EPIGENETIC INHERITANCE AND TRANSGENERATIONAL RESPONSE TO HEAT STRESS IN *ARABIDOPSIS THALIANA*

IVIE AYEMERE
Master of Science, Newcastle University (UK), 2010

A thesis submitted
in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

in

BIOLOGICAL SCIENCES

Department of Biological Sciences
University of Lethbridge
LETHBRIDGE, ALBERTA, CANADA

EPIGENETIC INHERITANCE AND TRANSGENERATIONAL RESPONSE TO HEAT STRESS IN *ARABIDOPSIS THALIANA*

IVIE AYEMERE

Date of Defence: December 17, 2019

Dr. Igor Kovalchuk Thesis Supervisor	Professor	Ph.D.
Dr. Gerlinde Metz Thesis Examination Committee Member	Professor	Ph.D.
Dr. Robert Laird Thesis Examination Committee Member	Associate Professor	Ph.D.
Dr. Anthony Russell Chair, Thesis Examination Committee	Associate Professor	Ph.D.

DEDICATION

To the loving memory of my father, S.U Ayemere whose words inspired me on this academic path and my forever supportive mother, Francisca O. Ayemere whose sacrifices lifts me up daily.

ABSTRACT

Plants experience abiotic stress and develop mechanisms to tolerate these stresses which preserves their progeny over time. Stress exposure induces physiological responses like accelerated growth rates; increased flowering/bolting time complementary to altered levels of genomic methylation. To understand the effects of epigenetic inheritance in the form of DNA methylation, changes triggered by heat stress in progeny genome, epigenetic variations in the genome of stressed and non-stressed progeny and parent plants over twenty-five generations were compared using whole genome bisulfite sequencing. Our study provides evidence that multigenerational exposure to heat stress resulted in physiological changes and impacted DNA methylation patterns across generations. Identification of SNPs gave insight to the broader genotypic diversity in stressed plants and showed a pattern of inheritance that reflects short-term memory of heat stress across twenty-five generations in *Arabidopsis thaliana*. Understanding transgenerational responses in plants can assist in implementing new techniques that produce non-transgenic stress tolerant crops.

ACKNOWLEDGEMENTS

I am especially thankful to my supervisor, Professor Igor Kovalchuk for providing the opportunity to work in his laboratory and eternally grateful for his patience in my work and the support he gave me throughout my graduate program. I would like to thank my committee members Dr. Gerlinde Metz and Dr. Robert Laird for their helpful advice and feedback.

In addition, I thank all my colleagues at the Kovalchuk Laboratories especially Aki Matsuoka, Nina Kepeshchuk, Andrey Golubov, Slava Ilnytskyy, Boseon Byeon, Rommy Rodriguez and everyone else for giving me valuable suggestions and assistance throughout my studies and research work.

I thank my friends for their love and support especially Folasade Adesina and Sabitri Dhakal for their help in figuring out challenges during my thesis writing and Samuel Ofori for his assistance in statistics analysis.

Finally, I would like to recognize the Niger Delta Development Commission for the financial support awarded me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

15d8	recombination reporter gene
5-azaC	5-azacytidine
5mC	5-methyl cytosine
ABA	abscisic acid
aba-1	ABA-insensitive mutant
ADH1	*alcohol dehydrogenase 1 gene*
AGO	ARGONAUTE
AGO4	ARGONAUTE4
AGO6	ARGONAUTE6
ANOVA	Analysis of Variance
APX2	ASCORBATE PEROXIDASE 2
Asr1 and *Asr2*	*ABA-stress ripening 1 and 2 gene respectively*
*At*HSP 101	*Arabidopsis thaliana* Hsp100/ClpB chaper*one*
atmbd9	*Arabidopsis thaliana methyl-CpG binding domain 9 gene*
ATP	adenosine triphosphate
avrRpt2	*cysteine protease avirulence protein*
bHLH	base helix-loop-helix
BRs	brassinosteroids
C-DMRs	differentially methylated regions in all three cytosine contexts
C.elegans	*Caenorhabditis elegans*
CG	cytosines at cytosine-guanine dinucleotide
CG-DMCs	differentially methylated cytosines at cytosine-guanine dinucleotide
CG-DMR	differentially methylated regions at cytosine-guanine dinucleotide
CHG	cytosines at cytosine- any nucleotide-guanine trinucleotide
CHH	cytosines at cytosine- any nucleotide-any nucleotide trinucleotide
Clp	CLP protease family of serine peptides
CMT3	CHROMOMETHYLASE3
CO	CONSTANS
Col wt	*Arabidopsis thaliana* Columbia wildtype
Col-0	*Arabidopsis thaliana* Columbia ecotype
CpG	cytosine and guanine separated by a phosphate
CTAB	cetyl trimethylammonium bromide
dCas9	endonuclease deficient Cas9 protein
DCL	Dicer-like protein
DCL1, 2, 3	Dicer-like protein1, 2, and 3 respectively
DDM1, 2	ATP-dependent DNA helicase DDM1
DDW	double distilled water
DEL	deletions
DMCs	differentially methylated cytosine
DME	DNA demethylase DEMETER
DML	DEMETER-LIKE
DML1, 3	DEMETER-LIKE PROTEIN 1 and 3 respectively
DMRs	differentially methylated regions
DNA	deoxyribonucleic acid

Dnak	chaperone protein dnak
Dnmt1, 3	DNA methyltransferase 1 and 3 respectively
dpg	days post germination
DRM1, 2	DOMAINS REARRANGED METHYLTRANSFERASE1 and 2 respectively
ds	double strand
EDTA	ethylenediaminetetraacetic acid
F25C	samples of Generation 25 grown in non-stressed conditions
F25H	samples of Generation 25 grown in heat stressed conditions
F2C	samples of Generation 2 grown in non-stressed conditions
Fe	iron
FLC	FLOWERING LOCUS C
FT	FLOWERING LOCUS T
GAs	gibberellins
GATK	Genome Analysis Toolkit
GC	guanine-cytosine content
gDNA	genomic DNA
GO	gene ontology
GUS	β-glucuronidase gene (uidA)
H2A, H2B	histone variants
H_2O_2	hydrogen peroxide
H3, H4	histone proteins in chromatin structure
H3K27me2, me3	di/tri-methylation of histone 3 lysine 27 respectively
H3K4me1, me2	mono/di- methylation of histone 3 lysine 4 respectively
H3K4me3	tri-methylation of histone 3 lysine 4
H3K9 and 27	histone 3 lysine 9 and 27 respectively
H3K9me2,	di-methylation of histone 3 lysine 9
H3K9me3	tri-methylation of histone 3 lysine 9
HCl	hydrochloric acid
HD-Zip	homeodomain-leucine zipper
HDA6, 19	histone deacetylase 6 and 19 respectively
HDAs	*Arabidopsis* histone deacetylases
HEN1	the methyltransferase HUA ENHANCER1
HRFs	heat stress transcription factors
HsfA1a	heat stress transcription factor A-1a
HsfA1b	*heat stress transcription factor A-1b* mutant
HsfA2, HsfA9	heat shock transcription factor A-2, 9 respectively
HsfB1	heat shock transcription factor B-1
Hsp100, 101	chaperone protein of ATPases
Hsp60, 70, 90	heat shock protein 60, 70 and 90 respectively
HSPs	heat shock proteins
huv-miR166a	*Hordeum vulgare microRNA166a*
HYL1	dsRNA-binding protein HYPONASTIC LEAVES1
IAA29	auxin-responsive gene 29
INDELs	insertions or deletions
INS	insertions

JA	jasmonic acid
LUC	*luciferase gene*
MBP	maltose-binding protein
mC	methylated cytosine
mCG	methylated CG island sites
mCHG	methylated CHG island sites
mCHH	methylated CHH island sites
MET1	METHYLTRANSFERASE1
miRNAs	micro RNAs
mRNA	messenger RNA
MS	Microsoft Excel
Msp1	a type II restriction endonuclease
NChIP	native chromatin immunoprecipitation
ncRNAs	non-coding RNAs
NRPD1	subunits of PolIV
NRPE1	subunits of PolV
NS	non-stressed
NS-NT	progeny of NS lineage in non-heat-treated group
NS-T	progeny of NS lineage in heat-treated group
NT	non-heat- stressed/treated group
nt	nucleotides
NtAlix1	pathogen-responsive gene
NtDPDL	tobacco *GLYCEROPHOSPHODIESTERASE-LIKE* gene
P5CDH	*DELTA-1-PYRROLINE-5-CARBOXYLATE DEHYDROGENASE*
PCA	Principal Component Analysis
PCR	polymerase chain reaction
PDC1	pyruvate decarboxylase 1 gene
PIF4	PHYTOCHROME INTERACTING FACTOR 4
Pol IV, V	polymerase IV and V respectively
PR	PATHOGENESIS-RELATED
PTGS	post translational gene silencing
PTM	post translational histone modification
RdDM pathway	RNA-dependent DNA methylation pathway
RDR	RNA-dependent polymerases
RDR2, 6	RNA-dependent polymerases 2 and 6 respectively
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNAi	RNA Interference
ROS1	REPRESSOR OF SILENCING1
rpm	revolution per minute
RPS2	*disease resistance protein 2*
S	stressed group
S-NT	progeny of S lineage in non-heat-treated group
S-T	progeny of S lineage in heat-treated group
SA	salicylic acid
SAHH1	S-ADENOSYL-HOMOCYSTEINE HYDROLASE1

SE	standard error
SGS3	Protein SUPPRESSOR OF GENE SILENCING 3
SHR	somatic homologous recombination
sHSP	small heat shock protein
siRNAs	small interfering RNA
SnpEff	program for annotated variants and coding effects
SNPs	single nucleotide polymorphism
sRNAs	small RNAs
SRO5	*SIMILAR TO RCD ONE5*
T	heat-stressed/treated group
ta-siRNAs	tans-acting siRNAs
TAIR10	The *Arabidopsis* Information Resource 10
TALE	transcription activator-like effector
TEs	transposable elements
TFIIS	transcription elongation factor
TFs	transcription Factors
TGS	transcriptional gene silencing
TMV	tobacco mosaic virus
TSS	transcription start site
UV-B	ultraviolet radiation-B rays
UV-C	ultraviolet radiation-C rays
VRN-A1	*Triticum aestivum* gene
WGBS	whole genome bisulfite sequencing
WGS	whole genome sequencing
ZF	zinc finger
zmGST	*Zea mays glutathione-S-transferase* genes
zmPP2C	*Zea mays protein phosphatase 2C* gene

CHAPTER 1: INTRODUCTION

The lifecycle of plants involves continuous interactions with their environment with impacts on growth and crop production. These interactions result in physiological and biomolecular responses including changes in plant phenotype and modulations in gene expressions through epigenetic modifications (Zheng et al., 2013). Plant response to its environment has been seen to bring about tolerance or adaptation to varying conditions and ensures the survival of the plant itself and that of its progeny (Huang et al., 2012). This phenomenon is feasible due to the plasticity of the plant genome, an ability to steadily maintain its genotype and evolve through environmentally induced alterations in its phenotype (Kovalchuk & Kovalchuk, 2012; Nicotra et al., 2010). As a result, phenotypic plasticity in reaction to environmental conditions allows for accelerated assimilation to wavering growth conditions leading to responses that are coordinated through stress perception, signalling pathways with cross-talks and the activation of response genes (Chinnusamy, Schumaker, & Zhu, 2004; Zheng et al., 2013).

Although, adaptation to an environment is not restricted to changes in the physiology of the plant but also to the existence of adequate mechanisms that influence its genome and epigenome in the presence of substantial stress (Gutzat & Scheid, 2012). The processes involved in the integration of exogenous environmental stimuli with endogenous developmental mechanisms in plants are yet to be completely understood. However, substantial number of studies in certain plant species demonstrate the formation of transgenerational memory; the ability to maintain the memory of stress exposure throughout ontogenesis and stably transmit this to the following generation (Bilichak & Kovalchuk, 2016; Boyko et al., 2010; Boyko & Kovalchuk, 2010; Suter & Widmer, 2013). For example, stress memory of transgenerational iron (Fe) deficiency was reported in *Arabidopsis thaliana* where the frequency of somatic homologous

recombination events, DNA strands breaks and TFIIS-like gene expression increased for one generation when plants were grown under Fe deficiency (Murgia et al., 2015). In rice, Ou et al., (2012) reported transgenerational inheritance of modified DNA methylation patterns in three successive generations with enhanced tolerance to heavy metal stress. Also, transgenerational inheritance of stress memory has been shown in multiple off-spring generations of *Arabidopsis thaliana* with or without the maintenance of the stress condition (Groot et al., 2016). Higher reproductive outputs such as flower and seed number per plant were reported in multiple generation exposures to heat treatment when compared to offspring grown in unheated environments (Whittle, Otto, Johnston, & Krochko, 2009). Similarly, transgenerational effects in response to heat stress were reported for three generations in other models like *Artemia* (Norouzitallab et al., 2014) and *Caenorhabditis ele*gans (Gouvêa, Aprison, & Ruvinsky, 2015).

The underlying mechanism of transgenerational memory is believed to be epigenetic; phenotype traits transferred to progeny without modification in gene sequence (Bilichak & Kovalchuk, 2016; Bird, 2007; Eichten, Schmitz, & Springer, 2014). Epigenetic marks result in variations within the genome evident by reversible enzyme-mediated modification of DNA or histones which controls transcription activity of genes, repetitive sequences and transposable elements (Allis, Jenuwein, & Reinberg, 2007; Hauser, Aufsatz, Jonak, & Luschnig, 2011). Epigenetic events are also characterized by the structural remodelling of chromosomal regions preserved through DNA repair phases and securely maintained across multiple cell generations (Bird, 2007). A contributing factor in the establishment of inheritance of epigenetic marks can be attributed to the occasional escape of the reprogramming process during gametogenesis (Bilichak & Kovalchuk, 2016) and the late differentiation of the germline in plants which can allow perception of stressor during vegetative growth (Mirouze & Paszkowski, 2011). Specific

epigenetic marks are associated with particular stress factor leading to the control of gene expression.

Epigenetic regulation consists of three primary mechanisms; DNA methylation, histone modification and small RNAs (sRNAs) expressions. DNA methylation is the most established gene expression modulations utilized by plants to change their phenotype in stressful environments (Mirouze & Paszkowski, 2011; Schmitz et al., 2011). The dynamic changes in chromatin structure and biogenesis of sRNAs contribute to transcriptional and post-transcriptional regulation of gene expression for stress response and is an active force behind epimutation (Mirouze & Paszkowski, 2011; Zheng et al., 2017). DNA methylation is suggested to act as a complementary mode of transferring heritable material that contributes to phenotypic variations (Heard & Martienssen, 2014; Molinier, Ries, Zipfel, & Hohn, 2006). The regulatory effects of DNA methylation in plants as seen in scientific studies refer to the prevention of DNA transcription, which can occur by DNA methylation inhibiting the binding of transcription factor(s) (TFs) (Moore, Le, & Fan, 2013) or by combining specific proteins (methyl CpG binding proteins) with methylated DNA acting as competitors for TFs (Liang et al., 2014).

Studying the genomic patterns of DNA methylation within angiosperms species can give insight on natural variations and epigenomic diversity that exist within this group (Niederhuth et al., 2016), this has been achieved by comparing single-base resolution of several DNA methylomes which depicts an extensive phylogenetic profile of any existing variations within the genome of study. DNA methylation being well studied has been reported to be an ancient form of defense against foreign DNA, as transposons and repeats are uniformly methylated in humans, mouse, *Arabidopsis thaliana* and *Neurospora crassa* (Zemach, McDaniel, Silva, & Zilberman, 2010). As such, variations in gene body DNA methylation, euchromatic silencing of transposons and repeats

or the silencing of heterochromatic transposons can reveal changes in DNA methylation of different plant groups. Applying these molecular analyses to plants exposed to stressing conditions allow scientists to distinguish if an exhibited phenotype is as a result of genetic or epigenetic variation or how each epigenetic modifications contribute to phenotypic differences (Lane, Niederhuth, Ji, & Schmitz, 2014), if at all.

Single Nucleotide Polymorphism (SNPs) is the difference at a single nucleotide base within individuals of same species which may cause phenotypic diversity such as flowering time adaptation, colour of plant fruit or tolerance to various biotic and biotic factors (Jang et al., 2015; Vidal, do Nascimento, Mondego, Pereira, & Carazzolle, 2012). The effects of SNPs in the genome occurs with different frequencies in different species and can cause changes in amino acids in the exon of a gene (asynonymous) or be silent, that is present in coding regions but synonymous. And also occur in non-coding regions with no effect (Huq et al., 2016). SNPs may access promoter activity for gene expression and can result in producing functional protein through transcription and translation (LeVan et al., 2001). For these reasons, identification and classification of functional SNPs in genes are essential as the distribution and frequency of polymorphism in a genome can tell the pattern of inheritance.

In rice, naturally occurring DNA methylation variations (epimutations) is well studied as it is a food crop, epimutations in single gene loci have been shown to result in heritable morphological variations without altering the DNA sequence of rice genes (Miura et al., 2009; Zheng et al., 2017), hence, revealing the role epimutation plays in plant growth and development. Epimutations can be modified by environmental stress and are reversible thereby serving as a critical function in the adaptation of plants to their environment. Differentially methylated regions (DMRs) are specific epimutations associated with epigenetic inheritance of a stress response.

Exposure to environmental factors can induce changes in the pattern of DNA methylation or specific DMRs which suggest the existence of exposure-specific DMRs (Haque, Nilsson, Holder, & Skinner, 2016; Manikkam, Guerrero-Bosagna, Tracey, Haque, & Skinner, 2012). Examples include DMRs associated with differentially expressed genes in the regulation of *Arabidopsis thaliana* immune system against *Pseudomonas syringae* (Dowen et al., 2012) and differences in location-specific DNA methylation pattern observed at specific cis-regulatory sites in tree crop *Hevea brasiliensis* indicating direct impact of cold and drought stress on the genome (Uthup, Ravindran, Bini, & Thakurdas, 2011). All of which support the proposition of site-specific stress-induced DNA methylation. Clustering analysis of changes in DNA methylation patterns allows areas of the genome with epimutations that are overrepresented or underrepresented to be identified which are achieved by comparing locations of DMR clusters to gene clusters within the genome and between test genomes.

Previous research from our laboratory demonstrated that the progeny of plants exposed to various abiotic and biotic stressors exhibit changes in phenotypic traits such as leaf size and number, flowering time and seed size. Theses progeny of stressed plants also displayed a certain degree of stress tolerance and cross-tolerance lasting for one to two generations without maintenance of the stress factor (Migicovsky, Yao, & Kovalchuk, 2014; Rahavi & Kovalchuk, 2013). These studies suggest differential DNA methylation and gene expression of non-coding RNAs (ncRNAs) in the gametes and in the progeny to be the underlying epigenetic mechanism for the establishment of transgenerational epigenetic inheritance reported (Bilichak et al., 2015; Boyko et al., 2010; Boyko & Kovalchuk, 2010).

We hypothesize that continuous exposure to stress will result in the establishment of short-term stress memory transmitted over multiple generations. We also hypothesize that

plant population propagated under stress will show a broader genotypic diversity when compared to the population of plants propagated without stress. We predict plants will exhibit stress–specific features of epigenetic inheritance (DNA methylation).

The primary goals of our work were to show the ability of plants to form transgenerational stress memory using repeated heat stress over consecutive generations. Also, to exhibit the stress-specific features of genetic and epigenetic changes (DNA methylation variations) in plant progeny and the possible inheritance of these effects. As a result, the following questions were raised; How long does transgenerational memory last? What are the related stable epigenetic changes associated with long-term transgenerational memory? Do genetic or epigenetic changes occur higher in the progeny of stressed plants when compared to the progeny of non-stressed plants? And if this happens, does a random or directional trend exist in DNA methylation pattern induced by heat stress? Are the changes in DNA methylation induced by heat stress preserved in subsequent generation of G_{25}. Overall, what are the specific indications of epigenetic stress adaptation?

Two different aims were established with the following objectives to achieve these primary goals and possibly assist in providing answers to the stated questions:

- To analyze multigenerational effects and epigenetic inheritance triggered by heat stress and interpret genetic and epigenetic changes that occur after multi-generational exposure to heat stress.
- Describe the role of the epigenetic component; DNA methylation in transgenerational inheritance, multigenerational exposure effects and plant adaptation to heat stress.

1. <u>Physiological Changes in Multi-Generational Stress Exposure</u>

- To perform detailed analysis of physiological changes in the progeny of heat-stressed Arabidopsis thaliana plants propagated for twenty-five generations.

- To examine evidence of stress response in leaf number, bolting rate and seed length of selected progeny under normal condition and in response to heat stress to identify stress-specific changes in plants.
- To study the influence of heat stress on the physiology of progeny of heat exposed plant, elucidate the role of continuous stress exposure and ascertain if such changes are persistent across the twenty-five generations.

2. Epigenetic Inheritance of Heat Stress

- To perform whole genome sequencing analysis on parental and progeny plants.
- To compare genetic and epigenetic variations due to heat-specific changes in parent and progeny plants using bioinformatics analysis and to identify stable and unstable genomic regions that show differences.
- To identify mutations and epimutations associated with tolerance to heat stress.

The influence of environment on plants as a driving force for variations in epigenome and genome of plants cannot be disregarded. It is essential to identify if epigenetic markers such as DNA methylation persist across generations and if they are associated with stress tolerance or cross-tolerance in the progeny. The characterization of these changes will enable the understanding of effects of multigenerational exposure on a molecular level and the resulting stress adaptation can facilitate the use of the non-transgenic approach to select plants that exhibit a better response to their shifting environment. This can be applied to sustainable plant breeding practices in plant seed production.

1.1 CHAPTER 1 REFERENCES

Allis, C. D., Jenuwein, T., & Reinberg, D. (2007). *Epigenetics*: CSHL Press.

Bilichak, A., Ilnytskyy, Y., Woycicki, R., Kepeshchuk, N., Fogen, D., & Kovalchuk, I. (2015). The elucidation of stress memory inheritance in brassica rapa plants. *Frontiers in Plant Science, 6*(5).

Bilichak, A., & Kovalchuk, I. (2016). Transgenerational response to stress in plants and its application for breeding. *Journal of experimental botany, 67*(7), 2081-2092.

Bird, A. (2007). Perceptions of epigenetics. *Nature, 447*(7143), 396-398.

Boyko, A., Blevins, T., Yao, Y., Golubov, A., Bilichak, A., Ilnytskyy, Y., . . . Kovalchuk, I. (2010). Transgenerational adaptation of arabidopsis to stress requires DNA methylation and the function of dicer-like proteins. *PloS one, 5*(3), e9514.

Boyko, A., & Kovalchuk, I. (2010). Transgenerational response to stress in *arabidopsis thaliana*. *Plant Signaling & Behavior, 5*(8), 995-998.

Chinnusamy, V., Schumaker, K., & Zhu, J. K. (2004). Molecular genetic perspectives on cross-talk and specificity in abiotic stress signalling in plants. *Journal of experimental botany, 55*(395), 225-236.

Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Dowen, J. M., Nery, J. R., . . . Ecker, J. R. (2012). Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences, 109*(32), E2183-E2191.

Eichten, S. R., Schmitz, R. J., & Springer, N. M. (2014). Epigenetics: Beyond chromatin modifications and complex genetic regulation. *Plant physiology, 165*(3), 933-947.

Gouvêa, D. Y., Aprison, E. Z., & Ruvinsky, I. (2015). Experience modulates the reproductive response to heat stress in c. Elegans via multiple physiological processes. *PloS one, 10*(12), e0145925.

Groot, M. P., Kooke, R., Knoben, N., Vergeer, P., Keurentjes, J. J., Ouborg, N. J., & Verhoeven, K. J. (2016). Effects of multi-generational stress exposure and offspring environment on the

expression and persistence of transgenerational effects in arabidopsis thaliana. *PloS one, 11*(3), e0151566.

Gutzat, R., & Scheid, O. M. (2012). Epigenetic responses to stress: Triple defense? *Current opinion in plant biology, 15*(5), 568-573.

Haque, M. M., Nilsson, E. E., Holder, L. B., & Skinner, M. K. (2016). Genomic clustering of differential DNA methylated regions (epimutations) associated with the epigenetic transgenerational inheritance of disease and phenotypic variation. *BMC Genomics, 17*(1), 418.

Hauser, M.-T., Aufsatz, W., Jonak, C., & Luschnig, C. (2011). Transgenerational epigenetic inheritance in plants. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1809*(8), 459-468.

Heard, E., & Martienssen, R. A. (2014). Transgenerational epigenetic inheritance: Myths and mechanisms. *Cell, 157*(1), 95-109.

Huang, G.-T., Ma, S.-L., Bai, L.-P., Zhang, L., Ma, H., Jia, P., . . . Guo, Z.-F. (2012). Signal transduction during cold, salt, and drought stresses in plants. *Molecular biology reports, 39*(2), 969-987.

Huq, M. A., Akter, S., Nou, I. S., Kim, H. T., Jung, Y. J., & Kang, K. K. (2016). Identification of functional snps in genes and their effects on plant phenotypes. *Journal of Plant Biotechnology, 43*(1), 1-11.

Jang, S.-J., Sato, M., Sato, K., Jitsuyama, Y., Fujino, K., Mori, H., . . . Yamada, T. (2015). A single-nucleotide polymorphism in an endo-1, 4-β-glucanase gene controls seed coat permeability in soybean. *PloS one, 10*(6), e0128527.

Kovalchuk, I., & Kovalchuk, O. (2012). *Epigenetics in health and disease*: FT Press.

Lane, A. K., Niederhuth, C. E., Ji, L., & Schmitz, R. J. (2014). Pencode: A plant encyclopedia of DNA elements. *Annual review of genetics, 48*, 49-70.

LeVan, T. D., Bloom, J. W., Bailey, T. J., Karp, C. L., Halonen, M., Martinez, F. D., & Vercelli, D. (2001). A common single nucleotide polymorphism in the cd14 promoter decreases the affinity of sp protein binding and enhances transcriptional activity. *The Journal of Immunology, 167*(10), 5838-5844.

Liang, D., Zhang, Z., Wu, H., Huang, C., Shuai, P., Ye, C.-Y., . . . Wang, J. (2014). *Single-base-resolution methylomes of populus trichocarpa reveal the association between DNA methylation and drought stress*.

Manikkam, M., Guerrero-Bosagna, C., Tracey, R., Haque, M. M., & Skinner, M. K. (2012). Transgenerational actions of environmental compounds on reproductive disease and identification of epigenetic biomarkers of ancestral exposures. *PloS one, 7*(2), e31901.

Migicovsky, Z., Yao, Y., & Kovalchuk, I. (2014). Transgenerational phenotypic and epigenetic changes in response to heat stress inarabidopsis thaliana. *Plant Signaling & Behavior, 9*(2), e27971.

Mirouze, M., & Paszkowski, J. (2011). Epigenetic contribution to stress adaptation in plants. *Current opinion in plant biology, 14*(3), 267-274.

Miura, K., Agetsuma, M., Kitano, H., Yoshimura, A., Matsuoka, M., Jacobsen, S. E., & Ashikari, M. (2009). A metastable dwarf1 epigenetic mutant affecting plant stature in rice. *Proceedings of the National Academy of Sciences, 106*(27), 11218-11223.

Molinier, J., Ries, G., Zipfel, C., & Hohn, B. (2006). Transgeneration memory of stress in plants. *Nature, 442*(7106), 1046-1049.

Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology, 38*(1), 23-38.

Murgia, I., Giacometti, S., Balestrazzi, A., Paparella, S., Pagliano, C., & Morandini, P. (2015). Analysis of the transgenerational iron deficiency stress memory in arabidopsis thaliana plants. *Frontiers in Plant Science, 6*(745).

Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E., Mathesius, U., . . . Valladares, F. (2010). Plant phenotypic plasticity in a changing climate. *Trends in plant science, 15*(12), 684-692.

Niederhuth, C. E., Bewick, A. J., Ji, L., Alabady, M. S., Do Kim, K., Li, Q., . . . Udall, J. A. (2016). Widespread natural variation of DNA methylation within angiosperms. *Genome biology, 17*(1), 194.

Norouzitallab, P., Baruah, K., Vandegehuchte, M., Van Stappen, G., Catania, F., Bussche, J. V., . . . Bossier, P. (2014). Environmental heat stress induces epigenetic transgenerational

inheritance of robustness in parthenogenetic artemia model. *The FASEB Journal, 28*(8), 3552-3563.

Ou, X., Zhang, Y., Xu, C., Lin, X., Zang, Q., Zhuang, T., . . . Liu, B. (2012). Transgenerational inheritance of modified DNA methylation patterns and enhanced tolerance induced by heavy metal stress in rice (oryza sativa l.). *PloS one, 7*(9), e41143.

Rahavi, S. M., & Kovalchuk, I. (2013). Changes in homologous recombination frequency in arabidopsis thaliana plants exposed to stress depend on time of exposure during development and on duration of stress exposure. *Physiol Mol Biol Plants, 19*(4), 479-488.

Schmitz, R. J., Schultz, M. D., Lewsey, M. G., O'Malley, R. C., Urich, M. A., Libiger, O., . . . Ecker, J. R. (2011). Transgenerational epigenetic instability is a source of novel methylation variants. *Science, 334*(6054), 369-373.

Suter, L., & Widmer, A. (2013). Phenotypic effects of salt and heat stress over three generations in arabidopsis thaliana. *PloS one, 8*(11), e80819.

Uthup, T. K., Ravindran, M., Bini, K., & Thakurdas, S. (2011). Divergent DNA methylation patterns associated with abiotic stress in hevea brasiliensis. *Molecular Plant, 4*(6), 996-1013.

Vidal, R. O., do Nascimento, L. C., Mondego, J. M. C., Pereira, G. A. G., & Carazzolle, M. F. (2012). Identification of snps in rna-seq data of two cultivars of glycine max (soybean) differing in drought resistance. *Genetics and Molecular Biology, 35*(1 Suppl), 331-334.

Whittle, C., Otto, S., Johnston, M. O., & Krochko, J. (2009). Adaptive epigenetic memory of ancestral temperature regime in arabidopsis thaliana. *Botany, 87*(6), 650-657.

Zemach, A., McDaniel, I. E., Silva, P., & Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science, 328*(5980), 916-919.

Zheng, X., Chen, L., Li, M., Lou, Q., Xia, H., Wang, P., . . . Luo, L. (2013). Transgenerational variations in DNA methylation induced by drought stress in two rice varieties with distinguished difference to drought resistance. *PloS one, 8*(11), e80253.

Zheng, X., Chen, L., Xia, H., Wei, H., Lou, Q., Li, M., . . . Luo, L. (2017). Transgenerational epimutations induced by multi-generation drought imposition mediate rice plant's adaptation to drought condition. *Scientific reports, 7*.

CHAPTER 2: LITERATURE REVIEW

2.1 Plant Reproductive Development and Life Cycle

Ontogenesis in plants involves stages of development that spans from egg cell fertilization, seed formation to maturation of vegetative tissues and flowering reproduction, which all comes full cycle and completion with senescence. The gradual transition from one stage of growth to the next is influenced not only by genetic factors within the plant but also by external environmental conditions such as day length, quality or intensity of light and temperature (Chailakhian, 1961).

The germline cells of plants are not distinguishable in early development but are differentiated from somatic cells later on in development (Slotkin et al., 2009). Plants alternate between a multicellular haploid state (the gametophyte) and a multicellular diploid state (the sporophyte). The sporophyte via meiosis produces microspores and megaspores that goes through cell proliferation and differentiation to form gametophytes. These gametophytes produce haploid gametes (pollen and embryo sac) which fuse to create the zygote and as such begin the formation of a diploid sporophyte that completes the haplodiplontic cycle (Yadegari & Drews, 2004).

Flowering plants (angiosperms) such as *Arabidopsis thaliana* have particular differentiating features; which is endosperm in their seeds and flowers. They also have both male and female gametophytes which arise from undifferentiated stem cells (gametophytes initials). These stem cells act as indistinct germlines that continuously produce vegetative tissues and organs. (Berger & Twell, 2011; Twell, 2011). The male gametophyte develops within the stamens and undergoes cell division to produce vegetative cell and germ cell. Then, the germ cell further divides to generate twin sperm cells utilized for double fertilization resulting in a male gamete composed of two sperm cells and a vegetative cell. On the other hand, the female gametophyte develops within the ovule consisting of seven cells and four different types; egg cells, antipodal

cells, accessory cells and synergid cells. These gametophytes are essential for the reproduction process, and the resulting cells developed during these processes are regulated through gene expression (Migicovsky & Kovalchuk, 2012; Twell, 2011; Yadegari & Drews, 2004).

2.1.1 Epigenetic Modifications in Angiosperms

Major epigenetic mechanisms such as DNA methylation, histone modification and small RNAs (sRNAs) control of transposons are epigenetic processes that function in plant gametes to regulate its epigenetic context. In *Arabidopsis,* previous studies propose that DNA methylation is reduced in the vegetative nucleus and central cell nuclei causing the activation of expression of transposons (Slotkin et al., 2009). Small interfering RNA (siRNAs) produced in response to the activation of transposon expression enforce silencing by travelling to sperm and egg cells from the vegetative nucleus and central cell (Migicovsky & Kovalchuk, 2012). It is crucial to note that changes occurring during plant development are dynamic processes which involve *de novo* methylation and demethylation resulting in differentiated cells with unique stable methylation patterns that regulate specific gene transcription (Law & Jacobsen, 2010). For instance, genome-wide hypomethylation was seen in *Arabidopsis* endosperm (Gehring, Bubb, & Henikoff, 2009; Hsieh et al., 2009; H. Zhang & Zhu, 2012); expression of DEMETER (DME), a DNA demethylase was detected in female gametophyte but absent after fertilization of central cell with sperm cell indicating endosperm hypomethylation may be due to demethylation of central cell nucleus (Y. Choi et al., 2002; H. Zhang & Zhu, 2012). As the transcriptional expression of MET1 was simultaneously repressed during female gametogenesis, it is suggested that endosperm hypomethylation likely occurs due to combined active and passive DNA demethylation (H. Zhang & Zhu, 2012).

In this review, epigenetic modifications in plants will be discussed as well as the extent to which stress-induced DNA methylation variations underlay transgenerational stress memory and multigenerational effects of stress exposure.

2.2 Epigenetic Modification in Plants

Chromatin is a complex structure of nucleoproteins that tightly pack DNA into the nucleus of a eukaryotic cell. Each unit of chromatin is made up of nucleosomes that consist of 147 base pairs of DNA wrapped around a core of eight histone protein molecules which affects the accessibility of DNA by transcription factors and RNA polymerases. As a result, nucleosome packing has to be dynamic and subject to alterations dependent on developmental and environmental cues (Grunstein, 1997; Khan, Shah, & Irshad, 2015; Narlikar, Fan, & Kingston, 2002). This dynamic nature of chromatin allows for alterations in cellular activities controlled through epigenetic modification and acts as a way to regulate corresponding genes which in turn affects biological processes like seed germination, flowering, embryo formation and responses to biotic and abiotic stresses (Khan et al., 2015; Roudier et al., 2011).

Epigenetic modifications involves heritable alterations in gene activity which are mitotically and/or meiotically transmitted without changes to DNA sequence (Holliday, 1994; Iwasaki & Paszkowski, 2014) and distinguishingly refers to heritable chromatin modifications that are reset in next generations (mitotically transmissible modifications) or transmitted to the progeny (meiotically transmissible modifications) (Khan et al., 2015; Lauria et al., 2004; Zemach, Kim, et al., 2010). Several epigenetic mechanisms involved in chromatin modifications include DNA methylation, post-translational histone modifications and non-coding RNAs as well as chromatin remodelling enzymes (Kouzarides, 2007; Rapp & Wendel, 2005).

2.2.1 DNA Methylation in Plants

DNA methylation is the chemical modification of genomic DNA involving the transfer of a methyl group onto a nucleotide base (cytosine or adenine) catalyzed by DNA methyltransferase. For this review, DNA methylation is interchangeable with cytosine methylation which refers strictly to the methylation of cytosine residues in DNA and has the significant function of defending against foreign DNA, maintaining genome stability and regulating the expression of transposable elements (TEs), repeat sequences and genes by inhibiting the binding of transcription factors to DNA. Also, DNA methylation is involved with many critical biological processes such as heterochromatin formation, genomic imprinting and regulation of endogenous genes. Cytosine methylation of promoter and coding sequences of genes can repress transcription and has been reported in major groups of eukaryotes - plants, animals and fungi (Chan, Henderson, & Jacobsen, 2005; Eichten, Schmitz, & Springer, 2014; Finnegan, Genger, Peacock, & Dennis, 1998; Henderson & Jacobsen, 2007; Law & Jacobsen, 2010; Moore, Le, & Fan, 2013; Vidalis et al., 2016; Xiaoguo Zheng et al., 2013).

Plants and animals show different patterns of DNA methylation, in plants methylation of cytosine residues in DNA occurs in three sequence context; strand-symmetrical CG and CHG (where H represents A, C or T) sites, and strand-asymmetrical CHH sites (Finnegan et al., 1998). While in animals, it is mostly reported in CG sites (Khan et al., 2015; Zemach, McDaniel, Silva, & Zilberman, 2010). CG methylation is frequently found within gene bodies (Law & Jacobsen, 2010) and non-CG methylation is enriched in transposon and repeat sequences (Dowen et al., 2012). Generally, symmetrical sequences are predominantly methylated while asymmetrical sequences show lesser frequency of cytosine methylation (Kovalchuk & Kovalchuk, 2012; Law & Jacobsen, 2010; Yaish, 2013).

Genome-wide studies indicate the overall level of cytosine methylation to be higher in plants than in animals and have shown several endogenous genes to be methylated within their promoter or transcribed regions with high correlation to transcription levels (Zhang et al., 2006). For example, analysis of genome-wide DNA methylation composition and distribution in *Arabidopsis thaliana* revealed 24% at CG sites, 6.7% at CHG sites and 1.7% at CHH sites (Cokus et al., 2008). However, these frequencies are sequence (gene coding or repeat sequence) context-dependent and as such gene coding sequence CG methylation sites have been seen to occur at 30% frequency while CHG and CHH sites have occurred at less than 1% frequency (Widman, Jacobsen, & Pellegrini, 2009). On the contrary, other genome sequencing reports of plants have reported varying frequencies in the occurrence of cytosine methylation at all three sequence context, as well as low frequencies in non-symmetrical sites indicating that the pattern of methylation at these sites are variable between molecules and organisms (Finnegan et al., 1998; Meyer, Niedenhof, & Lohuis, 1994). For example, global DNA methylome pattern found in *Populus trichocarpa* under normal conditions was 29% at CG sites, 29% at CHG sites and 42% at CHH sites. These frequencies were different under water stress; 26% at CG sites, 27% at CHG sites and 47% at CHH sites (Liang et al., 2014). Chinnusamy and Zhu (2009) reported approximately 50% DNA methylation observed in higher plants compared to 2% to 8% found in mammals.

Different enzymes are reported in plants to be involved in the establishment, maintenance and removal of DNA methylation. *De novo* methylation is catalyzed by DOMAINS REARRANGED METHYLTRANSFERASE 2 (DRM2), a homolog of the mammalian Dnmt3 family. There are three different pathways by which the maintenance of DNA methylation occurs depending on the sequence context; CG methylation is maintained by DNA METHYLTRANSFERASE 1 (MET1) by a plant homolog of Dnmt1, CHG methylation is

maintained by CHROMOMETHYLASE3 (CMT3) a plant-specific DNA methyltransferase and maintenance at CHH sites is by DRM2 through *de novo* methylation utilizing the plant-specific RNA-dependent DNA methylation (RdDM) pathway (Chan et al., 2005; Kovalchuk & Kovalchuk, 2012; Law & Jacobsen, 2010). Mutant plants lacking MET1 activity show a lack of CG methylation and passive loss of DNA methylation throughout generations suggesting that MET1 may not be required for establishing new methylation imprints. CMT3 mutant plants showed genome-wide loss of CHG methylation mostly at centromeric repeats and transposons and reduced CHH methylation in several genomic regions (Kovalchuk & Kovalchuk, 2012). However, overlapping functions between DNA methyltransferases have been reported indicating the ability of MET1 and CMT3 to establish *de novo* methylation and DRM1 and DRM2 to maintain symmetrical methylation (Lister et al., 2008).

The existence of multiple DNA methyltransferases in plants raises the notion that plant methyltransferases may differ in target specificity or stages of development as observed in the two pea methyltransferases which may have originated from a single gene but differed in target specificity. A similar difference was found in tobacco where methylation of cytosines in CG and CHG sequences exhibited differential sensitivity to methylation inhibitors indicating different enzymes was catalyzing methylation of these CG and CHG sites respectively (Finnegan et al., 1998; Kovařík, Koukalova, Holý, & Bezděk, 1994; Pradhan & Adams, 1995).

DNA methylation can be lost through passive or active demethylation processes. DNA methyltransferase dysfunction or deficit in methyl groups can lead to passive demethylation while the enzymatic replacement of 5mC with cytosine can result in active demethylation (H. Zhang & Zhu, 2012). These enzymes are known as DNA demethylases and contribute to active locus-specific and global DNA demethylation. Examples are (DEMETER (DME), REPRESSOR OF

SILENCING1 (ROS1) and DEMETER-LIKE (DML) PROTEIN DML2 and DML3 which all belong to a family of DNA glycosylases involved in DNA base pair excision repair (Saze, Tsugane, Kanno, & Nishimura, 2012; H. Zhang & Zhu, 2012). DME is engaged in global demethylation while ROS1 counteracts DNA methylation established by RdDM pathway (Agius, Kapoor, & Zhu, 2006; Gao et al., 2010; Gong et al., 2002; H. Zhang & Zhu, 2012); the loss of function of ROS1 in *ros1* mutants reported to cause DNA hypermethylation and enhanced transcriptional gene silencing (TGS) in *Arabidopsis.* The fusion of maltose-binding protein (MBP) with ROS1 and its DNA glycosylase domain was capable of incising plasmid DNA methylated with Msp1 methylase suggesting ROS1 to be a DNA demethylase (Agius et al., 2006; Gong et al., 2002). Mutations within coding regions of DDM1, DDM2, MET1 (Kakutani, Jeddeloh, & Richards, 1995; Yaish, Colasanti, & Rothstein, 2011) and ROS1 (Agius et al., 2006) have showed alterations in DNA methylation levels that resulted in inconsistent effects on phenotypes studied (Yaish et al., 2011). DNA methylation provides a mechanism of gene control, loss of DNA methylation disrupts normal plant development displaying phenotype abnormalities like altered leaf size and shape, reduced root length, loss of apical dominance and homeotic transformation of floral organs (Finnegan, Peacock, & Dennis, 1996; Kakutani et al., 1995; Ronemus, Galbiati, Ticknor, Chen, & Dellaporta, 1996).

The existence of a mechanism to reset methylation patterns between generations is significant as the methylation changes that occur during plant development contributes not only to tissue-specific gene expression but to the plant phenotype as well (Finnegan et al., 1998). In mammals, organ-specific changes in methylation do not affect germ cells as global demethylation occurs early in embryogenesis followed by remethylation during implantation (Finnegan et al., 1998; Razin & Cedar, 1993). On the contrary progamatic cells in plants inherit methylation

changes that have accumulated in its vegetative meristem and as such does not undergo global demethylation followed by remethylation during embryo development (Finnegan et al., 1996; Flavell, Pearce, & Kumar, 1994). As seen in methyltransferase antisense (*ddm1* mutant) *Arabidopsis* plants having reduced DNA methylation levels which were not restored to normal levels in progeny that had lost the antisense transgene through outcrossing (Finnegan et al., 1996; Kakutani et al., 1995). The methylation changes that occur during gametogenesis establish plant methylation patterns, studies in maize and *Arabidopsis* show methylation of transposable elements and repeated sequences increase during development which could ensure the packing of these sequences into heterochromatin (Finnegan et al., 1998).

The methylation of promoter and coding sequences can repress transcription while in other genomic regions it can control gene activities in different ways, depending on underlying nucleotide sequence. As a result, DNA methylation is also essential for regulation of tissue-specific gene expression (Finnegan et al., 1998; Moore et al., 2013). Global DNA methylation analysis of the *Arabidopsis* genome revealed about 20% of cytosines are methylated and located either in promoter regions or gene body with impacts on gene expression (Yaish, Al-Lawati, Al-Harrasi, & Patankar, 2018; X. Zhang et al., 2006). The regulation pattern of DNA methylation in genes is dependent on the regions within the genes it occurs. There is the negative correlation between the methylation of the promoter regions and transcription (X. Zhang et al., 2006). Methylation can also occur within the body of a gene; gene body methylation has either positive or negative effect on gene expression (Khan et al., 2015; X. Li et al., 2012; Zemach, McDaniel, et al., 2010). Reduction in global methylation does not consequently reduce DNA methylation of the target gene. Conversely, low level of global DNA methylation is frequently associated with hypermethylation at a specific gene. For example, the overexpression of an antisense

methyltransferase gene in *Arabidopsis* which caused an overall decrease in genomic cytosines also caused hypermethylation of the SUPERMAN gene (Jacobsen & Meyerowitz, 1997; Yaish, 2013).

2.2.1.2 DNA Methylation and Plant Stress Response

Epigenetic modifications are more dynamic than DNA sequence mutations and as such play an important role in an organism's initial response to changes in the environment to regulate gene expression. The variations in DNA methylation patterns due to environmental stress has been assessed with several empirical studies utilizing different crop plants and stress conditions (Table 2.1). Biotic and abiotic stress such as avirulent bacteria, bacterial pathogens, salicylic acid, drought, salinity, ultraviolet light, cold and heat can induce widespread dynamic changes in DNA methylation in plants (Bilichak, Ilnystkyy, Hollunder, & Kovalchuk, 2012; C.-S. Choi & Sano, 2007; Dowen et al., 2012; Migicovsky & Kovalchuk, 2013; W.-S. Wang et al., 2011); these changes typically correlate changes in replication, transcription, transposition, cell development and differentiation (Kovalchuk & Kovalchuk, 2012; Mirouze et al., 2012).

DNA demethylation and hypermethylation may occur in response to abiotic stress, for example in tobacco, *NtGPDL* gene was demethylated in the coding region by aluminum, low temperature and salt stress whereas promoter sequences were unmethylated irrespective of the stress. *NtDPDL* gene was not expressed under non-stressed conditions whereas upon exposure to biotic stress, its gene transcripts was induced and genomic locus was partly demethylated (Qiao & Fan, 2011). Plants may increase genome-wide DNA methylation in response to stress to maintain genome stability by decreasing transposon activity (Lukens & Zhan, 2007). On the other hand, genomic hypomethylation has been shown as a result of environmental stressors such as heavy metal and low temperature (Sano, 2010).

A typical plant response during exposure to stressing environmental factors is early flowering often associated with pleiotropic phenotypes; early flowering allows plants to complete their life cycle from seed to seed in less amount of time to reduce susceptibility to the stress. DNA methylation can control the expression of specific genes involved in flowering. Hence alterations in global DNA methylation can change flowering times in plants (Yaish, 2013). For instance, mutations within MET1 lead to late flowering in *Arabidopsis*, *met1* and *cmt3* lines resulted in abnormal cell division, embryonic malfunction and improper auxin gradient (Xiao et al., 2006). *atmbd9* mutants showed increased global DNA methylation that included the transcription factor FLC that represses flowering; these mutations led to early flowering (and an increase in multiple axillary branching which was alternatively restored to wild-type flowering time and normal FLC expression levels with 5-azaC (M. Peng, Cui, Bi, & Rothstein, 2006; Xiaoguo Zheng et al., 2017).

Several studies provide evidence of changes in DNA methylation as a stress response in plants; the role of DNA methylation in the inheritance of transgenerational memory in response to ultraviolet, UV-C cold and heat stress in *Arabidopsis* has been demonstrated by the work of Boyko et al., (2010a). Although there is a lack of direct consistency in the levels or type of methylation changes when plants are subjected to cold, heat or ultra-violet stress, DNA methylation changes in response to stress have been found to exist among different stress episodes in plant species. These methylation changes can either be short-term somatic response or be stably inherited and passed down for many generations (Uthup, Ravindran, Bini, & Thakurdas, 2011). Genome-wide methylation changes across generations are established when offspring of stressed and non-stressed plants are compared. The results indicate not only the occurrence of stress-induced changes but also determine if they are inherited by the offspring. Direct impacts of abiotic stress on the epigenome of plants via DNA methylation have been reported as epigenetic variations. For

example, the transmission of induced epigenetic variations to progeny as a result of various ecological stress was demonstrated in the study of asexual dandelions (*Taraxacum officinale*). The induced changes were faithfully transmitted to off-springs for two generation (Verhoeven, Jansen, van Dijk, & Biere, 2010) and in rice, modified DNA methylation patterns were inherited in a subsequent generation (Ou et al., 2012). Also, transgenerational changes in methylation profiles have been shown to be induced by biotic stress such as tobacco mosaic virus (TMV) in the form of increased rate of rearrangement in resistance gene-like loci (Boyko et al., 2007). DNA methylation especially CG methylation is a stable epigenetic mark as it is coupled with DNA replication and is significant in the coordination of epigenetic memory and transgenerational inheritance as a function of gene regulation in plants (Lauria & Rossi, 2011; Mathieu, Reinders, Čaikovski, Smathajitt, & Paszkowski, 2007).

Another example of related DNA methylation to stress response in plants is the induction of differentially methylated regions (DMRs). Profiles of DNA methylomes in plants exposed to bacterial pathogens, avirulent bacteria or salicylic (SA) hormones showed stress induced DMRs which were associated with differentially expressed genes. For instance, DNA methylation changes in repetitive sequence or transposons coupled with other epigenetic modifications regulated neighboring genes in response to SA stress in which transposon-associated DMRs were accompanied by upregulation of 21-nt siRNAs and changes of the proximal gene in response to SA (Dowen et al., 2012). Other studies revealed increase in global genomic methylation in pathogen-infected plants with locus-specific hypomethylation that allows for higher frequency of homologous recombination events (Boyko & Kovalchuk, 2011) and decrease in resistance related gene methylation after the viral infection (Boyko et al., 2007; Kathiria et al., 2010; Kovalchuk et al., 2003). These observations are explanatory by the view that an overall increase in methylation

promotes genome stability under viral attacks whereas a decrease in resistance gene methylation levels aids genetic recombination and consequently produces new genes that assist in resisting the pathogen (Engler, Weng, & Storb, 1993).

Table 2.1: DNA Methylation in Biotic and Abiotic Stresses in different Plant Species.

Stress	Plant Species	Genomic Region	Mode of Action/Effect	# of Generation studied	References
Heat/ Drought	*Arabidopsis*	Genome-wide	Higher HRF and Hypomethylation Higher reproductive output during exposure-adaptive memory.	Four	Bokyo 2010 Whittle 2009
	Tomatoes	*Asr2,* *Asr1*	CG hypermethylation and CHH hypomethylation. CHH hypomethylation in regulatory regions.	-	González et al., 2013
	Rice	Genome-wide	Differential methylation with genotypic and tissue specificity.	Three	Wang et al., 2011
Salinity/Salt stress	Wheat	24 genes Genome-wide	Stress-induced changes in global DNA methylation. Hypomethylation	-	Wang et al., 2014 L. Zhong et al., 2009
	Maize	Root zmPP2C Leaf zmGST	Hypermethylation Hypomethylation	-	Tan 2010
	Soybean	Glyma11g02400 (Promoter) Glyma16g27950 (Promoter)	-518 to -274 cytosines demethylated (exposure for 1–24 h). Hypomethylation at transcription start codon (-24 to -233).	-	Song et al., 2012
	Rice	Genome-wide	Gene body-specific methylation.	-	Karan et al., 2012
Cold	*Arabidopsis*	Genome-wide	Higher HRF and Hypomethylation	Four	Groot et al., 2006 Boyko 2010
	Maize	Genome-wide	Demethylation changes on fully methylated fragments causing a shift in global methylation.	-	Shan et al., 2013
	Wheat	VRN-A1	Hypermethylation (site-specific) in all cytosine context.	-	Khan et al., 2013
SA stress	Dandelions (apomictic) *Taraxacum officinale*	Genome-wide	Gain and loss methylation changes at AFLP markers under SA treatment.	Two	Verhoeven et al., 2010
Heavy metals	Rice	TE and protein coding genes	Hypomethylation at CHG sites.	Three	Ou et al., 2012
Osmosis	Maize	Transposon region	Hypermethylation	-	Tan, 2010
Iron deficiency	*Arabidopsis*	*TFIIS*-like gene	SHR frequency increase	Two	Murgia et al., 2015
Tobacco Mosaic Virus	Tobacco	NtGPDL NtAlix1 Genome-wide and disease resistance gene-like loci	Hypomethylation Hypermethylation Hypermethylation	-	Choi & Sano, 2007 Wada et al., 2004 Boyko et al, 2007

2.2.2 Histone Modifications in Plants

An alternative mechanism of controlling gene expression and the epigenetic process of retaining stress memory for expanded periods is the covalent modifications of histone proteins. Histones form the primary packing structure of chromosomal DNA made of two copies each of four different subunits (H2A, H2B, H3 and H4). The N-terminal tails on the surface of the nucleosome octamer allow for covalent modifications and undergo various post-translational modifications at different positions but mostly at lysine and arginine residues (Chinnusamy & Zhu, 2009; Feng & Jacobsen, 2011; Rapp & Wendel, 2005). Histone modifications and DNA methylation both play a role in determining chromatin structure thereby contributing to the transcriptional state and expression levels of genes (Chinnusamy & Zhu, 2009; Qiao & Fan, 2011). Changes and regulation of gene expression typically involve interlacing chromatin modifiers which are connected to post-translational histone modifications (PTM) and histone variant replacements. Histone tails can undergo an array of covalent modifications such as acetylation, methylation, phosphorylation, ubiquitination, sumoylation, biotinylation, carbonylation, glycosylation and ADP ribosylation catalyzed by many enzymes (Berr, Ménard, Heitz, & Shen, 2012; Lauria & Rossi, 2011; C. Liu, Lu, Cui, & Cao, 2010; Tariq & Paszkowski, 2004).

Histone modifications have been connected to several gene activities, for example, acetylation, certain phosphorylation and ubiquitination has been reported to enhance transcription (Sridhar et al., 2007; K. Zhang, Sridhar, Zhu, Kapoor, & Zhu, 2007) while others like biotinylation and sumoylation repress gene expression (Chinnusamy & Zhu, 2009; Nathan et al., 2006; Qiao & Fan, 2011). The modulation of gene activity by histone methylation is based on the lysine and arginine residue that is modified and the number of methyl groups attached to each lysine residue. For example, imprinted genes in *Arabidopsis* may be regulated via Histone 3 Lysine 27

Methylation (H3K27me2) through Polycomb group activation (Jullien, Katz, Oliva, Ohad, & Berger, 2006) which are the proteins linked to forming epigenetic memory and controlling global gene expression during development (Butenko & Ohad, 2011).

2.2.2.1 Histone Modifications and Plant Stress Response

Histone changes are essential forms of epigenetic regulation because they are involved in environmental stress-induced gene regulation (Boyko & Kovalchuk, 2008) and can exert non-genetic mediated effects on plant and its progeny. Although histone marks have different functions, the same histone mark can exert a different function in a different organism. Three types of histone methylation are most studied in plants as evolved epigenetic modification systems; H3K4 mono/di/trimethylation (H3K4me1, H3K4me2, H3K4me3), histone H3K27 trimethylation (H3K27me3) and histone H3K9 di-methylation (H3K9me2) (Feng & Jacobsen, 2011).

Genome-wide studies revealed H3K4me2 and H3K4me3 were predominantly enriched in actively transcribed gene sequences and were gene silencing markers (X. Zhang, Bernatavichute, Cokus, Pellegrini, & Jacobsen, 2009; X. Zhang et al., 2006) whereas H4K4me1 is correlated with CG methylation within transcribed regions of genes. H3K4me3 is associated with active transcription while H3K4me1 and H3K4me2 are directly connected with transcription (X. Zhang et al., 2009). Other forms of histone methylation have been reported to be associated with repressing transcription such as di-methylation of H3K9 and H3K27. Overall, H3K4 methylation is seen to be distributed in other organisms such as rice, yeast and humans indicating a highly conserved mechanism among species (K. Zhang et al., 2007).

In *Arabidopsis*, H3K4me1 H3K4me2 and H3K4me3 were seen in genes and promoters and absent in heterochromatic regions where transposons and repetitive DNA sequences are found

(Feng & Jacobsen, 2011). Also, dynamic changes to these histones were initiated by dehydration stress (van Dijk et al, 2010). In tomato plants, drought stress was found to induce linker histone variants H1-S connected to reducing stomatal conductance and minimize water loss (Scippa et al., 2004). In rice, water stress was reported to induce histone H3K4me3 and H3 acetylation in *ADH1* and *PDC1* genes correlated with enhanced expression of these genes under stress, although epigenetic modifications were restored to basal level after stress was relieved indicating the dynamic state of these changes (Chinnusamy & Zhu, 2009; Tsuji, Saika, Tsutsumi, Hirai, & Nakazono, 2006).

Environmental and endogenous cues can affect gene activity through altered levels of histone acetylation. For example, in rice and *Arabidopsis* histone deacetylases (HDAs); HDA6 and HDA19 catalyze histone deacetylation in response to abiotic and biotic stress which were both induced by jasmonic acid (JA) (Chinnusamy & Zhu, 2009). HDA6 is involved in transcriptional gene silencing (TGS) and RdDM (Aufsatz, Mette, van der Winden, Matzke, & Matzke, 2002; Chinnusamy & Zhu, 2009; Probst et al., 2004) while HDA19 possibly mediates pathogenic response in plants through JA-regulated PATHOGENESIS-RELATED (PR) genes (C. Zhou, Zhang, Duan, Miki, & Wu, 2005). Furthermore, modification in chromatin formation and histone H3 acetylation was responsible for the stress-mediated release of gene silencing in studied *Arabidopsis* transgene under temperature and UV-B stress. The impact on gene silencing was heritable and transmitted to two generation of non-stressed progeny (Lang-Mladek et al., 2010). Histone modifications add another level of epigenetic change that provides a source of heritable phenotypic variation, and by controlling the expression of stress-induced genes, histone modifications can affect epigenetic memory passed on by stressed plants (Xiaoguo Zheng et al., 2013).

2.2.3 Small RNAs and Plant Stress Response

Non-coding RNAs (ncRNA) play a regulatory role in the epigenome of a plant during early development, transcription and post-transcriptional gene silencing, imprinting and regulating heritable responses to varying environmental conditions (Kovalchuk & Kovalchuk, 2012). There are various ncRNAs classed based on size, biogenesis and their biological functions. Amongst these classes, small RNAs (sRNAs) are the most studied in plant epigenetic process because of its regulatory role and function in their response to environmental stress factors. As a result, ncRNAs has become relevant in understanding the dynamic relationship between sRNAs and other epigenetic modifications.

sRNAs are single-stranded ncRNAs that range from 20-24 nucleotide in length and are produced through RNA interference (RNAi) pathway involving plant-specific DNA -dependent RNA polymerase Pol IV and Pol V, RNA-dependent RNA polymerase RDR2, double-stranded DCL3 and Argonautes proteins AGO4 and AGO6 (Kasschau et al., 2007; Simon & Meyers, 2011; Xie et al., 2004; X. Zhang, Henderson, Lu, Green, & Jacobsen, 2007). The function of sRNAs in down-regulation of gene expression after transcription and mRNA maturation occurs through the binding of ncRNAs to mRNA molecule which inhibits translation and degrades mRNA transcript (Khraiwesh, Zhu, & Zhu, 2012; Kovalchuk & Kovalchuk, 2012). Groups of sRNAs have been characterized in plants like tobacco, tomato and *Arabidopsis* with the main function in post transcriptional gene silencing (PTGS) (Dalmay, Hamilton, Mueller, & Baulcombe, 2000; Hutvágner, Mlynarova, & Nap, 2000). Earlier example of the involvement of sRNAs in TGS and PTGS was established in RNA interference (RNAi) ability of dsRNAs to inactivate expression of genes in *C. elegans*. The discovery of *lin-4* RNA which controls the timing of larval development was the first miRNA to be identified. Over time a number of miRNAs was reported in plants and

found to regulate various biological processes through interaction with mRNA translation (Khraiwesh et al., 2012; R. C. Lee, Feinbaum, & Ambros, 1993).

In plants, the two major classes of endogenous sRNAs are micro RNAs (miRNAs), and small interfering RNA (siRNAs). The former is processed from single-stranded precursor RNAs that form a hairpin structure through the activity of Dicer-like enzyme and binds to RISC complex (Shukla, Chinnusamy, & Sunkar, 2008) while the latter can be generated from several sources including single-stranded RNAs converted by RNA-dependent polymerases (RDR), RNA transcribed from inverted repeats or natural cis-antisense transcript pairs (Khraiwesh et al., 2012; Ramachandran & Chen, 2008). miRNAs are involved in transcriptional silencing by binding to reverse complementary sequences which results in cleavage of target RNAs (Khraiwesh et al., 2012) while siRNAs play a role in heterochromatin formation and gene silencing by guiding sequence-specific DNA and histone methylation through RNA-directed DNA methylation (RdDM) pathway (Chen, 2012; Gao et al., 2010; Law & Jacobsen, 2010; Simon & Meyers, 2011). In general, sRNA are generated from partially double-stranded RNA precursors by the action of DICER proteins and the use of Argonaute (AGO) family which binds to a small RNA and confers regulatory functions on target genes (Ramachandran & Chen, 2008). Argonaute proteins act as the effectors of sRNA mediated regulation, for example, AGO4 or AGO6 recruit heterochromatic siRNA and assist in guiding chromatin modifications to homologous DNA sequences (Simon & Meyers, 2011; Xianwu Zheng, Zhu, Kapoor, & Zhu, 2007).

Furthermore, sRNAs have been reported in preserving genome integrity from transposons and assist in forming a long-term memory of stress (Boyko & Kovalchuk, 2010b). The loss of function mutations in miRNA biogenesis machinery DICER-LIKE 1-4 genes showed the epigenetic role of miRNA in controlling gene expression. DCL1 is involved in a process that

enhances DNA methylation which leads to the repression of transposons. DCL1 and DCL2 proteins are proposed to direct transgenerational memory of stress in plants as they control the production of miRNA (Boyko, Blevins, et al., 2010; Yaish et al., 2011). Small non-coding RNAs like miRNA and siRNA generated by somatic cells are believed to travel within organs to reach gametes and potentially influence phenotype appearance of the organism (Kovalchuk & Kovalchuk, 2012) and are considered as one of the mechanisms behind their ability to direct transgenerational memory. The unique ability to carry short- range signal from cells allows sRNAs to guide processes like DNA methylation (Molnar et al., 2010) which consolidates an association of sRNAs with environmental stress tolerance, changes in physiology and development such as flowering in plants.

Biotic and abiotic stress conditions cause plants to over- or under-express specific miRNA or synthesize new miRNAs to deal with stress which indicates miRNA is differentially regulated in response to stress but not necessarily involved in the adaptation response (Khraiwesh et al., 2012). Stress-induced sRNAs have been reported in several studies involving stress factors like salinity, drought or low temperature conditions. For instance, seventeen stress-induced miRNAs out of the known 117 miRNAs were detected in *Arabidopsis* under several abiotic stress used in the study (H.-H. Liu, Tian, Li, Wu, & Zheng, 2008). A library of small RNAs in seedlings exposed to cold, dehydration, high salt and ABA treatments were reported to have new miRNAs responsive to these conditions (Khraiwesh et al., 2012; Sunkar & Zhu, 2004). Genome-wide analysis of miRNAs in drought- challenged rice revealed about thirty novel miRNAs were induced across several developmental stages out of which nine were of different expression to that observed in *Arabidopsis* (L. Zhou et al., 2010). In another study, global gene expression analysis of rice revealed relative miRNAs expression levels, 32 miRNAs were found to be induced or suppressed

after drought, salt, cold or ABA treatments as well as stress-responsive cis-elements seen to be enriched in the promoters of stress-responsive miRNA genes (Shen, Xie, & Xiong, 2010).

Higher GC content can be indicative of the regulatory function of miRNAs sequences under stress conditions, and this can serve as a parameter for predicting their involvement as reported by the study of Mishra et al., (2009) which found higher GC content in known miRNA sequences. In addition to a role in environmental stress responses, miRNAs are involved in controlling flowering in *Arabidopsis*, in a study by Schmid et al, (2003) miRNAs were indicated to induce flowering precursors of miR172 microRNA and linked to mediate the effects of floral induction through signal transmission pathways that involve CO and FT genes (Schmid et al., 2003). Mutations within DCL1 and DCL3 resulted in deferred flowering due to high expression of FLC in these mutants (Robert J. Schmitz, Hong, Fitzpatrick, & Amasino, 2007; Yaish et al., 2011).

Similar to miRNAs, siRNAs are involved in the regulation of plant stress response. In wheat seedlings, different changes in the expression of four siRNAs were observed due to heat, salt or drought stresses (Khraiwesh et al., 2012; Yao et al., 2010). Salt stress in *Arabidopsis* revealed the expression of stress-regulated genes (*P5CDH* and *SRO5*); the presence of both transcripts formed a 24-nt siRNA through a DCL2- dependent pathway. Cleavage of *P5CDH* guided by the 24-nt siRNA leads to the subsequent generation of 21nt siRNAs and breakage of *P5CDH* transcripts which is constitutively expressed. The expression of *SRO5* gene was induced by salt stress which then initiated siRNA formation creating a gene pair with related functionality under salt stress. The downregulation of *P5CDH* leads to an accumulation of proline which is significant in plants ability to tolerate high salt stress (Borsani, Zhu, Verslues, Sunkar, & Zhu, 2005; Khraiwesh et al., 2012). Eukaryotes utilize siRNA- mediated RNA silencing as one of the defence mechanisms against invading viruses. A 22-nt siRNA, (nat-siRNAATGB2) was strongly induced in *P. syringae*

infected Arabidopsis plants carrying the avirulence gene, *avrRpt2.* The biogenesis of these siRNA requires DCL1, HYL1, HEN1, RDR6, NRPD1a and SGS3 and its induction is dependent on host resistance gene *RPS2* and *NDR1* gene. The nat-siRNAATGB2 functions by repressing a negative regulator of the RPS2 pathway (Katiyar-Agarwal et al., 2006; Shukla et al., 2008).

2.3 Plant Responses to Heat Stress

Plants exposed to adverse environmental conditions such as cold, drought, heat and pathogen infections have evolved refined genetic and epigenetic regulatory systems that allow for effective responses to their environment. The exposure to stress in plants can induce a rapid transition to reproductive development earlier than non-stressed treated plants (Yaish et al., 2011) at the expense of other features like yield. In the light of global warming, the diverse effects of temperature exposures on plant developments are well studied as the impact of climate change on crop production is of significant importance to food supply (J. Liu, Feng, Li, & He, 2015; S. Peng et al., 2004). A temperature increase of 1°C is estimated to implement reduction in global yield (approximately 8.9% drop in crop yield) of six majorly grown crops; barley, wheat, rice, soybeans, maize and sorghum (J. Liu, 2015; Lobell, Schlenker, & Costa-Roberts, 2011). Exposure to temperatures within 22°C - 27°C range is classed as warm ambient, within 27°C -30°C is considered high temperature while within 37°C - 50°C extremely high temperature and are all considered as heat stress for the model plant *Arabidopsis thaliana* (J. Liu et al., 2015). Temperatures that are above optimal growth range may exert irreversible damage to plant cellular function and alter metabolic processes. Heat stress can increase the rate of plant development, shorten lifecycle and decrease seed yield (Porter, 2005), all in an attempt to tolerate the environment, minimize damage and protect cellular homeostasis (Kotak, Larkindale, et al., 2007).

Acquired heat tolerance is different from basal heat tolerance, and the latter refers to an ability to survive temperature exposures that are above optimal growth. It can be argued that plants can quickly acquire heat tolerance to lethal temperature if they are pre-exposed to a moderate to high temperature or exposed to temperatures that gradually increase thereby conferring acquired heat tolerance (Kotak, Larkindale, et al., 2007; Larkindale, Mishkind, & Vierling, 2005).

In many cases, plant growth and development are regulated by interactive endogenous regulators and external factors. For instance, phytohormones mediate stress responses that aid in an adaptation response to the environmental change which is seen in phytohormone abscisic acid-mediated stomatal pore closure which results in reduced water loss during transpiration (Blatt & Thiel, 1993; Gray, Östin, Sandberg, Romano, & Estelle, 1998). Plant responses to heat stress has been documented to include morphological acclimation and accelerated flowering (Bokszczanin et al., 2013; Zinn, Tunc-Ozdemir, & Harper, 2010), regulated circadian clock, plant immunity to high temperatures and thermotolerance to heat stress (Bray, 2000; J. Liu et al., 2015; Sung, Kaplan, Lee, & Guy, 2003). The genetic mechanism of these responses has been reported to include several heat sensors, HSFs and HSPs (heat shock proteins) response pathways, network of phytohormones, chaperones and secondary metabolites (Bokszczanin et al., 2013; J. Liu, 2015; Qu, Ding, Jiang, & Zhu, 2013) as well as epigenetic mechanisms which spans through DNA methylation, histone modifications and differential expression of non-coding RNAs.

Other aspects of plant's response to heat stress worth mentioning are the effects of repeated or continuous exposure that may occur during their lifecycle which can lead to plant stress priming. The impact of a prior exposure can enhance an ability to respond to future events of exposure (Ganguly, Crisp, Eichten, & Pogson, 2017) and create an inherent memory of such event.

2.3.1 Plant Physiological Response to Heat Stress

Abiotic stress occurs in nature due to the amount or absence of physical or chemical components in the environment such as water, nutrients and temperature extremes. Several factors play a role in plant stress response to its environment. Some of which include an interaction between the type of stressor, severity of stress, duration, number of exposures, the stage of development of the plant and its genomic and epigenomic state all determine the physiological response observed (Bray, 2000; Sung et al., 2003). Resistance or susceptibility are the two possible outcomes from stress responses, susceptibility can result in considerable damage in development, physiology or cause death while resistance relies on avoidance to minimize the amount of stress damage or tolerance to maximize an ability to brace stress (Bray, 2000; Xiaoguo Zheng et al., 2013) leading to acclimation. However, acclimation can occur when a change in response to the stress can be achieved through avoidance or tolerance at an individual level and acclimation resulting from resistance mechanism that has evolved within a population can confer adaptation as a stress response in affected organisms (Bray, 2000; Puijalon et al., 2011; Vinocur & Altman, 2005) .

Heat mediated morphological acclimation and accelerated flowering were reported in plants under changing light and high temperature conditions. The involvement of basic helix-loop-helix (bHLH) transcription factor PHYTOCHROME INTERACTING FACTOR 4 (PIF4) and underlying signalling mechanism was indicated to control these responses. PIF4 was found to activate FLOWERING LOCUS T in a temperature dependent way by binding to FT promoter, an early flowering response was maintained in *pif4* mutants under high temperature indicating PIF4 controls early flowering at a warm temperature with the direct activation of FT (Kumar et al., 2012). Also, the expression of PIF4 was induced under increased ambient temperature which binds

to the promoters of auxin biosynthesis possibly triggered by auxin-responsive gene 29 (IAA29), *pif4* mutants did not show hyponasty or petiole elongation at high temperature, and HSP70 function of responding to measurable changes in cellular temperature was not affected (Koini et al., 2009; Proveniers & van Zanten, 2013). Early flowering is a reported plant response to heat stress and assists in completing the lifecycle at a quicker rate (Yaish et al., 2011), reports from 385 British plant species studied showed an advancement of 4.5 days in the average first flowering date with sensitivity to temperatures of the previous months (Fitter & Fitter, 2002).

The physiology of a plant is generally determined by the combination of its genotype and interactions with its surroundings, and to harness this relationship one must comprehend environmentally induced variations in the phenotype of individual plants (Nicotra et al., 2010). A response of hypocotyl and petiole elongation to warm temperature was reported in light-grown *Arabidopsis* plants as well as leaf hyponasty and early flowering (Gray et al., 1998; J. Liu et al., 2015). Gray et al., (1998) study on *Arabidopsis thaliana* (*etr1–1, ein2–1, aba-1*, and *ga4–1* mutant lines) reported warm temperatures promoted the accumulation of auxin which in turn activated gibberellins (GAs) and brassinosteroids (BRs) pathway resulting in hypocotyl elongation or growth. Seedlings were compared at 29°C and 20°C and the mutants in auxin response or transport pathways exhibited temperature-dependent reduced growth response (reduced levels auxin). Conversely, mutants deficient in gibberellin and abscisic acid biosynthesis and ethylene response were unaffected (Gray et al., 1998; J. Liu et al., 2015).

2.3.2 Mechanisms of Plant Response to Heat Stress

Heat stress response and the resulting acquired thermotolerance in plants and other organisms is centered around the accumulation of HSPs controlled by the binding of heat stress

transcription factors (HSFs) to heat shock elements in promoters (Kotak, Larkindale, et al., 2007; Mahmood, Safdar, Abbasi, & Naqvi, 2010). HSPs act as molecular chaperones in protein quality control by folding proteins after translation into structures suited for membrane transport, they also prevent aggregation of denatured proteins and resolubilize protein aggregates (Bösl, Grimminger, & Walter, 2006). As a result of these diverse functions, HSPs are not only linked to resistance to cold and heat stress but also to other abiotic stress like drought and salinity (Boston, Viitanen, & Vierling, 1996; Iba, 2002; Sun, Bernard, Van De Cotte, Van Montagu, & Verbruggen, 2001).

Identified HSPs based on chaperone activity and molecular weight include five major families; the Hsp60 (chaperonins), Hsp70 (DnaK) family, the Hsp90 family, Hsp100 (Clp) family and the small HSP (sHSP) family. With regards to function during heat stress, studies on Hsp100 and sHSPs are the most advanced with specific information on how they contribute to heat stress. The Hsp100/Clp family chaperones function in removing non-functional polypeptides that could potentially be harmful. These non-functional polypeptides arise from misfolding, denaturation or aggregation of proteins and their removal is important to the maintenance of cellular balance. The mechanism by which Hsp100 proteins rescue proteins from aggregation is reported to involve interaction with another ATP-Dependent chaperone system (W. Wang, Vinocur, Shoseyov, & Altman, 2004). For instance, heat-induced Hsp100/Clp proteins with the assistance of Hsp70/DnaK chaperone system was reported in *AtHSP101* mutant plants to be essential for resolubilizing protein aggregates and tolerance to high temperatures (U. Lee et al., 2005). The over-expression of *Athsp101* in rice plants increased protection against high temperature (Agarwal, Katiyar-Agarwal, & Grover, 2002). Although the expression of Hsp100/Clp family is crucial for plant tolerance to heat stress, they are expendable during normal growth, developmentally

regulated and induced by different environmental stressors (Hong & Vierling, 2001; W. Wang et al., 2004).

On the other hand, some sHSPs are expressed during developmental stages and triggered by abiotic stress (Boston et al., 1996; W. Wang et al., 2004). In plants, sHSPs are involved in the targeted protection of all cellular compartments (Hu et al., 2010; Sun, Van Montagu, & Verbruggen, 2002) and are located in the nucleus, plastids, ER and mitochondria (Bray, 2000; W. Wang et al., 2004). sHSPs cannot singularly refold non-native proteins but bind to them to stabilize and prevent non-native aggregation which facilitates subsequent refolding by ATP-dependent chaperones like the Dnak system (HSP70) (Al-Whaibi, 2011; W. Wang et al., 2004). The prevention of cellular protein aggregation allows HSPs to be a significant factor in thermotolerance and the diversification of plant sHSPs enables their function in molecular adaptation to stress. For example, in maize sHSPs found in mitochondria were shown to improve mitochondrial electron transport during salt stress (Hamilton & Heckathorn, 2001; W. Wang et al., 2004).

Plants have HSFs encoding genes which mediate the expression of HSPs and other heat stress-induced transcripts, twenty-one members of these genes have been identified in *Arabidopsis*. HsfA9 is expressed in late stages of seed development alongside HSPs and accumulated in response to abscisic acid. In tomato, HsfA1a, HsfA2 and HsfB1 have been identified as a part of the regulatory network behind the expression of heat stress response, HsfA1a was constitutively expressed and found to regulate the expression HsfA2 and HsfB1 (Kotak, Larkindale, et al., 2007; S. K. Mishra et al., 2002). In *Arabidopsis,* analysis of *HsfA1a*, *HsfA1b* and *HsfA2* knockout mutants demonstrated that HsfA2 is involved in prolonged heat stress and recovery conditions while HsfA1a and HsfA1b are important in the initial phase of heat stress responsive gene expression (Busch, Wunderlich, & Schöffl, 2005; Kotak, Larkindale, et al., 2007; Kotak, Vierling,

Bäumlein, & von Koskull-Döring, 2007; Schramm et al., 2006). HsfA2 has been seen as the major HSF in thermotolerant cells and can be induced under other abiotic stress as well. For instance, HsfA2 was induced by light and H_2O_2 and involved in the regulation of the enzyme APX2 expressed during oxidative stress (Kotak, Larkindale, et al., 2007; Nishizawa et al., 2006).

2.3.3 Epigenetic Regulation of Heat Response

The influence of epigenetic regulations on chromatin structure is cued by developmental and environmental signals, and as a result, the single genome in a plant cell can have multiple epigenomes in response to developmental and environmental cues (Chinnusamy & Zhu, 2009; Zhu, 2008). Epigenetic modifications regulate the expression of heat-responsive genes to prevent heat damage to the plant. Methylation differences also contribute to natural heritable variations, DNA sequence polymorphisms such as the presence or absence of transposable elements and repeats with a target gene regulate expression of heat responsive genes. Also, global methylation is affected by heat stress in different ways compared to the regular inhibition of transcription initiation within promoter regions and transcript elongation within gene bodies (J. Liu, 2015). For instance, in *Arabidopsis* heat stress has increased global methylation and homologous recombination frequency (Boyko, Blevins, et al., 2010) and the increased genome methylation may have been attributed to the up-regulation of DRM2 and subunits of PolIV and PolV (NRPD1 and NRPE1 respectively) in response to heat stress (Naydenov et al., 2015). However, in cotton anthers, high temperatures decreased the expression of S-ADENOSYL-HOMOCYSTEINE HYDROLASE1 (SAHH1), DRM1 and DRM2 resulting in global hypomethylation (Min et al., 2014; Naydenov et al., 2015). Also, heat stress-induced transcriptional activation of several transgenes that was initially silenced via GUS without detectable changes to DNA methylation

levels (Lang-Mladek et al., 2010; Pecinka et al., 2010). In a similar study, DNA methylation defective mutants showed the RdDM pathway is required for basal thermotolerance (Popova, Dinh, Aufsatz, & Jonak, 2013). All of these several responses leads to the proposition that changes in DNA methylation under heat stress in different species may not follow a singular consistent trend and is indirectly involved in the regulation of heat responses (J. Liu et al., 2015).

Similarly, histone modifications can be affected differently by heat stress with histone acetylation associated with the heat-induced release of silencing. Levels of H3K9me2, H3K4me3 and H3K27me1 at transcriptionally silenced GUS transgene showed minor changes while H3Kac1 and H3K9/14ac2 were significantly increased in response to heat (Lang-Mladek et al., 2010; J. Liu, 2015). On the contrary, H3K9me2, H3K4me3, H3K27me2 and H3K27me3 were not affected by temperature increase, but H3K9ac-K14ac was moderately enriched at rDNA (J. Liu, 2015; Tittel-Elmer et al., 2010).

Heat-responsive miRNAs have been reported and identified to be differently regulated in various species (reviewed in J. Liu (2015). For instance, miR159 was down-regulated in wheat (Kumar & Wigge, 2010; Y. Wang et al., 2012), *Arabidopsis* (S.-H. Zhong et al., 2013) and cassava (Ballén-Taborda et al., 2013). The up-regulation of *huv-miR166a* and down- regulation of its target genes were heat -induced and found to possibly affect leaf morphology in barley (Kruszka et al., 2014; J. Liu, 2015). Another example of heat responsive miRNA activity is miR166 which targets HD-Zip transcription factor and was up-regulated in *Arabidopsis* (S.-H. Zhong et al., 2013). The central targets of miR159 are MYB transcription factors (J. Liu, 2015), Wang et al., 2012 demonstrated the direct cleavage of TaGAMYB1 and TaGAMYB2 was by *tae-miR159.* Rice line overexpressing tae-mi159 and *Arabidopsis* double mutants (*myb33myb65*) were found to be more sensitive to heat stress than wild types. The down-regulation of miR159 and up-regulation of its

targets after heat stress indicates tae-mi159 is involved in a heat stress-related signalling pathway and contributes to heat stress tolerance (J. Liu, 2015; Y. Wang et al., 2012).

Endogenous and exogenous siRNAs are also affected by heat, Zhong and colleagues showed an increase in growth temperatures from 22°C to 30°C effectively inhibited transgene induced posttranscriptional gene silencing (PTGS) in *Arabidopsis* which they suggested likely occurred during the formation of stable dsRNAs that involves RDR6 and SGS3 steps (S.-H. Zhong et al., 2013). tasiRNAs were significantly reduced by 30°C growth accompanied by increased transcript levels of TAS and tasiRNA-target genes related to morphology acclimation which occurs due to reduced amount of protein abundance of SGS3 from the temperature increase which in turn attenuated the formation of dsRNAs (J. Liu et al., 2015; S.-H. Zhong et al., 2013).

Responses to other abiotic stress have also been linked to an epigenetic control mechanism as changes in both plant physiology and genome allow for many genes to be stress-inducible resulting in an immediate plant response (Thomashow, 1999) and long-term stress adaption (Iwasaki & Paszkowski, 2014). Repeated or continuous exposure to stress can lead to plant stress priming and prepare plants to better cope with future stress encounters (Ganguly, Crisp, Eichten, & Pogson, 2017) or even lead to the inheritance of the memory of such event.

2.4 Epigenetic Inheritance of Stress Response in Plants

Evolution and adaptive responses are driven by phenotypic diversity underlined by the selection of randomly generated genetic variations or heritable epigenetic variations under defined environmental conditions which allow some organisms to be more fit than others in various environments (Hauser, Aufsatz, Jonak, & Luschnig, 2011; Jiang et al., 2014). Environmental and genetic stimuli that induce methylation changes within the lifecycle of a plant could also create

novel epigenetic variations and affect subsequent generations. Since plant reproductive cell lineage is derived from somatic tissue late in development, genomic changes and environmental genome methylation status in response to stress that occur during a plant's lifespan can be transmitted to its progeny (Lukens & Zhan, 2007; Madlung & Comai, 2004). While genetic mechanisms of plant responses to stress target morphological acclimations, epigenetic control of plant heat responses has been linked to changes in genome stability via an increased frequency of somatic HR events, soft inheritance (changes in gene expression caused by epigenetic changes to environmental cues) and transgenerational adaptive plasticity of plant phenotype (Boyko & Kovalchuk, 2011). All of these responses show a provisional genome-wide roadmap as to how genetic and epigenetic changes accumulate over successive generations.

Also, stress-induced epigenetic changes have been proposed to include an evolutionary perspective that involves sRNAs as mobile stress generated signals that assess the gametes and influence DNA methylation patterns where epimutation events are represented by a loss or gain of DNA methylation accompanied by repressive or active chromatin marks (Boyko & Kovalchuk, 2011). The stable propagation of DNA methylation states makes possible the formation of plant stress memory (Boyko & Kovalchuk, 2010b). Retention of stress memory is an evident form of acclimation response made possible through stress-induced proteins, RNAs and metabolites (Iba, 2002; Thomashow, 1999). The duration of memory depends on if the processes involve reprogramming in phenology and morphology or stably inherited DNA methylation and histone modifications (Chinnusamy & Zhu, 2009).

Epigenetic inheritance requires passage of epigenetic marks through germline without erasure from the mechanisms that establish cellular totipotency at early ontogenesis (Hauser et al., 2011). An empirical example is paramutation in maize, the switch between epiallelic states of *r1*

and *b1* loci affected kernel pigmentation and was heritably passed on to the next generation (Chandler, Eggleston, & Dorweiler, 2000). The regulatory cross-talk between alleles allows for inherited alterations in the expression status of one of the alleles. sRNAs, locus-specific DNA repeats and their methylation status play a role in the underlying mechanism of this phenomenon (Hauser et al., 2011; Simon & Meyers, 2011). Transgenerational epigenetic inheritance on the other hand also considers the number of generations that stably inherits the epigenetic marks and the transmission of environmental effects into subsequent generations. For example, the frequency of TS-GUS reactivation increased in two generations after heat, cold and UV-B exposure (Hauser et al., 2011; Lang-Mladek et al., 2010). Also, homologous recombination increased for four generations after exposure to UV-C (Molinier, Ries, Zipfel, & Hohn, 2006).

The effects of DNA methylation on gene expression has heightened the premise that it could be complementary to genetic variations and be a mode of transfer for heritable information since DNA methylation states can be retained faithfully over mitotic and meiotic cell divisions by an array of pathways and enzymes (Law & Jacobsen, 2010; Probst, Dunleavy, & Almouzni, 2009). Although, it remains uncertain the extent of which genome-wide patterns of DNA methylation in plants are reset, however, it seems the parental methylome is reestablished and propagated during gametogenesis and spermatogenesis (Calarco et al., 2012; Ganguly et al., 2017; Slotkin et al., 2009). As these processes occur within postembryonic growth in plants, acquired variations in the DNA methylome (epiallele) which are either spontaneous or induced by the environment can be potentially carried over generations (Dubin et al., 2015; Ganguly et al., 2017; Q. Li, Eichten, Hermanson, & Springer, 2014). Stable inheritance of DNA methylation state has also been seen with epialleles in the form of differentially methylated regions (DMRs) occurring at frequencies proportionate to genetic polymorphism (Becker et al., 2011; Robert J Schmitz et al., 2011).

A study of heritable genetic and epigenetic variations in *Arabidopsis* showed elevated frequencies of accumulated epimutations under high salinity stress, the stressed lineage accumulated about 45% more differentially methylated cytosines (DMCs) than control plants with about 75% of these DMCs inherited and some lost over subsequent generation (Jiang et al., 2014)**.** With most results of DNA methylome variation attributed to an underlying genetic difference rather than an absolute epigenetic difference (Dubin et al., 2015; Eichten et al., 2013; Eichten et al., 2014; Q. Li et al., 2014), it is somewhat challenging for investigations to entirely tie transgenerational methylation changes to the observed adapted plant phenotype traits. As a result, a study by Ganguly and colleagues (2017) attempted to show the potential of drought-induced changes in DNA methylome of *Arabidopsis* being conveyed as transgenerational stress memory by correlating heritable epialleles and drought-responsive gene expressions in five generations of stressed plants. There was no evidence of conserved DMRs between treatments, only DMRs induced by the onset of drought within a single generation. The formation of drought stress memory was measured by enhanced drought tolerance and physiological traits in lineages. Overall, elevated seed dormancy was highlighted in the direct seed of drought-stressed parents and the seed from the first progeny of drought-exposed lineages grown in the absence of stress. The investigators concluded that transgenerational drought stress-induced epialleles were not acting as a mechanism to convey any form of transgenerational stress memory but the variations observed within and across generations seemed to be stochastic with highly stable DNA methylation states (Ganguly et al., 2017).

2.4.1 Transgenerational and Multi-generational Effects: Epimutations and Plant Adaptation

Adaptability of a genome is tied to the information on the DNA sequence and its expression in response to stress (Chinnusamy & Zhu, 2009). Transgenerational changes broadly indicate transmission of heritable changes in an organism occurring across multiple successive generations from parental to offspring in the absence of direct exposure to the stressor experienced by the parent organism (Boyko et al., 2007; Migicovsky, Yao, & Kovalchuk, 2014). A benefit to plants of these transmitted changes is acclimation to abiotic and biotic stressors which when propagated to the next generation confers transgenerational acquired tolerance (Bilichak et al., 2015; Boyko, Blevins, et al., 2010). The mechanism of action is proposed to include the various forms of epigenetic modifications in a different capacity.

The stress-induced transgenerational response in *Arabidopsis* depends on DNA methylation, and sRNA silencing pathways, the stress treatment of parental plants can lead to progeny with higher tolerance to stress (Boyko & Kovalchuk, 2010b). However, as a different environment has been involved in the parental lifecycle, the off-spring gametes or seeds are directly exposed to the parental environment (Blödner, Goebel, Feussner, Gatz, & Polle, 2007; Suter & Widmer, 2013a) and because their germ cells develop during stress treatment, related changes that persist in the next generations are referred to as transgenerational effects (Molinier et al., 2006). Another reference in distinguishing the effects of environmentally induced plant properties or persistent change in epigenetic states is multigenerational effects which bring about epigenetic variations along lineages.

Heritable epigenetic variations allow organisms to cope with different environments, reports of heritable variations in DNA methylation exist for at least eight generations in

Arabidopsis, in the absence of DNA sequence polymorphisms and without selection (Johannes et al., 2009; Xiaoguo Zheng et al., 2013). The environment or phenotype of the parents (also known parental effects) has been seen to determine plant phenotype in addition to its genotype and interactions with the environment (Groot et al., 2016). In *A. thaliana* mutation accumulation (MA) lineages, the changes in the properties of genome-wide accumulated de novo variants due to soil-salinity exposure in ten successive generations were seen as increased frequency accumulated de novo mutations and epimutations (differentially methylated cytosine positions) which differed from mutations accrued in non-stressed controlled MA lineages (Jiang et al., 2014). Epigenetic inheritance, in this case, was seen with multigenerational exposure to soil salinity.

The genetic alteration of CG methylation levels induces epialleles, as plants have no hard reset of CG methylation, transgenerational effects can be seen in the formation and transmission of heritable epialleles which corresponds to the different levels and distributions of CG methylation contributed by the activities of MET1 and DDM1 maintaining the epigenetic changes and inducing epigenetic inheritance (Paszkowski & Grossniklaus, 2011). In cases where traits are inherited for some generations (beyond the second generation), transgenerational inheritance is arguably declared in the absence of a stressor as only parent plants and their somatic cells were directly influenced.

Due to environment-induced transgenerational effects and correlated heritable epialleles that establish persistent changes in DNA methylation, studies aim to identify the induced persistent changes. A well-studied example of environmentally induced epigenetic states is the acceleration of flowering by long periods of cold. This response is epigenetically inherited over several cell divisions, and the memory of cold exposure is propagated through the lifecycle in winter annuals even after the initial stimulus is passed. However, the progeny of the vernalized plants were

observed to have lost the memory of the stress and resets their epigenetic state (Henderson, Shindo, & Dean, 2003; Paszkowski & Grossniklaus, 2011).

Acquired traits induced by environmental conditions and the possibility of these traits being heritable dates back to Lamarck theory (Suter & Widmer, 2013a), in an attempt to validate this concept, several studies have tried to demonstrate how parental environment influences the phenotype of off-springs. For example, Zheng et al, (2013) using two rice varieties demonstrated transgenerational changes in DNA methylation were induced by drought stress and, demethylated status was maintained in generation six without the induction of drought stress which indicates methylation status was carried along the plant lineage under drought stress (Xiaoguo Zheng et al., 2013). Interestingly, the environment or phenotype of parent plants (also known parental effects) has been reported to determine plant phenotype in addition to genotype and interactions with the environment (Groot et al., 2016). For example the offspring of *Arabidopsis* exposed to multiple generations of heat treatment showed higher fitness and adaptive memory (Whittle et al., 2009), ancestral exposure to salt stress caused improved growth under salt stress (Boyko, Golubov, Bilichak, & Kovalchuk, 2010), larger leaves and rosette diameter in some offspring from salt-treated plants when grown under salt stress (Suter & Widmer, 2013a, 2013b).

Although, concerns of consistency in the effects of transgenerational responses observed in empirical studies are raised. The difficulty in reproducing same or similar results in complementary studies (coined as context-dependency) is thought to depend on the genotype and environment context (Groot et al., 2016). It should be noted that due to plant sessile nature, its resulting dynamic biological processes ensure vastly varied responses to its ever-changing environment and cannot allow for an exert same response to stress factors, especially in diverse experimental methodologies and context.

However, the representations of epigenetic effects of the environment can be seen in related stress-induced transposon activation, communication between cells and organs, evidence of transgenerational changes and non-targeted mutagenesis (Kovalchuk & Kovalchuk, 2012). Epigenetic inheritance and its contribution to stress adaptation in plant can be two-sided, the late differentiation of the germline could perceive stresses during vegetative growth and memorize them by epigenetic mechanism in cell lineages and later contribute to germline as the stress memory which is possibly passed to progeny whose germline needs to be protected from environmental effects (Mirouze & Paszkowski, 2011). This occurrence highlights the importance of stability of transgenerational inheritance of the genome and epigenome under stress conditions.

2.5 CHAPTER 2 REFERENCES

Agarwal, M., Katiyar-Agarwal, S., & Grover, A. (2002). Plant hsp100 proteins: Structure, function and regulation. *Plant Science, 163*(3), 397-405.

Agius, F., Kapoor, A., & Zhu, J.-K. (2006). Role of the arabidopsis DNA glycosylase/lyase ros1 in active DNA demethylation. *Proceedings of the National Academy of Sciences, 103*(31), 11796-11801.

Al-Whaibi, M. H. (2011). Plant heat-shock proteins: A mini review. *Journal of King Saud University-Science, 23*(2), 139-150.

Aufsatz, W., Mette, M. F., van der Winden, J., Matzke, M., & Matzke, A. J. (2002). Hda6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded rna. *The EMBO journal, 21*(24), 6832-6841.

Ballén-Taborda, C., Plata, G., Ayling, S., Rodríguez-Zapata, F., Becerra Lopez-Lavalle, L. A., Duitama, J., & Tohme, J. (2013). Identification of cassava micrornas under abiotic stress. *International journal of genomics, 2013*.

Becker, C., Hagmann, J., Muller, J., Koenig, D., Stegle, O., Borgwardt, K., & Weigel, D. (2011). Spontaneous epigenetic variation in the arabidopsis thaliana methylome. *Nature, 480*(7376), 245-249.

Berger, F., & Twell, D. (2011). Germline specification and function in plants. *Annual review of plant biology, 62*, 461-484.

Berr, A., Ménard, R., Heitz, T., & Shen, W. H. (2012). Chromatin modification and remodelling: A regulatory landscape for the control of arabidopsis defence responses upon pathogen attack. *Cellular microbiology, 14*(6), 829-839.

Bilichak, A., Ilnystkyy, Y., Hollunder, J., & Kovalchuk, I. (2012). The progeny of arabidopsis thaliana plants exposed to salt exhibit changes in DNA methylation, histone modifications and gene expression. *PloS one, 7*(1), e30515.

Bilichak, A., Ilnytskyy, Y., Woycicki, R., Kepeshchuk, N., Fogen, D., & Kovalchuk, I. (2015). The elucidation of stress memory inheritance in brassica rapa plants. *Frontiers in Plant Science, 6*(5).

Blatt, M. R., & Thiel, G. (1993). Hormonal control of ion channel gating. *Annual review of plant biology, 44*(1), 543-567.

Blödner, C., Goebel, C., Feussner, I., Gatz, C., & Polle, A. (2007). Warm and cold parental reproductive environments affect seed properties, fitness, and cold responsiveness in arabidopsis thaliana progenies. *Plant, cell & environment, 30*(2), 165-175.

Bokszczanin, K. L., Fragkostefanakis, S., Bostan, H., Bovy, A., Chaturvedi, P., Chiusano, M. L., . . . Klaczynskid, K. (2013). Perspectives on deciphering mechanisms underlying plant heat stress response and thermotolerance. *Frontiers in Plant Science, 4*, 315.

Borsani, O., Zhu, J., Verslues, P. E., Sunkar, R., & Zhu, J.-K. (2005). Endogenous sirnas derived from a pair of natural cis-antisense transcripts regulate salt tolerance in arabidopsis. *Cell, 123*(7), 1279-1291.

Bösl, B., Grimminger, V., & Walter, S. (2006). The molecular chaperone hsp104—a molecular machine for protein disaggregation. *Journal of Structural Biology, 156*(1), 139-148.

Boston, R. S., Viitanen, P. V., & Vierling, E. (1996). Molecular chaperones and protein folding in plants. In *Post-transcriptional control of gene expression in plants* (pp. 191-222): Springer.

Boyko, A., Blevins, T., Yao, Y., Golubov, A., Bilichak, A., Ilnytskyy, Y., . . . Kovalchuk, I. (2010). Transgenerational adaptation of arabidopsis to stress requires DNA methylation and the function of dicer-like proteins. *PloS one, 5*(3), e9514.

Boyko, A., Golubov, A., Bilichak, A., & Kovalchuk, I. (2010). Chlorine ions but not sodium ions alter genome stability of arabidopsis thaliana. *Plant Cell Physiol, 51*(6), 1066-1078.

Boyko, A., Kathiria, P., Zemp, F. J., Yao, Y., Pogribny, I., & Kovalchuk, I. (2007). Transgenerational changes in the genome stability and methylation in pathogen-infected plants: (virus-induced plant genome instability). *Nucleic Acids Res, 35*(5), 1714-1725.

Boyko, A., & Kovalchuk, I. (2008). Epigenetic control of plant stress response. *Environmental and molecular mutagenesis, 49*(1), 61-72.

Boyko, A., & Kovalchuk, I. (2010a). Transgenerational response to stress in *arabidopsis thaliana*. *Plant Signaling & Behavior, 5*(8), 995-998.

Boyko, A., & Kovalchuk, I. (2010b). Transgenerational response to stress in arabidopsis thaliana. *Plant signaling & behavior, 5*(8), 995-998.

Boyko, A., & Kovalchuk, I. (2011). Genome instability and epigenetic modification--heritable responses to environmental stress? *Curr Opin Plant Biol, 14*(3), 260-266.

Bray, E. A. (2000). Response to abiotic stress. *Biochemistry and molecular biology of plants*, 1158-1203.

Busch, W., Wunderlich, M., & Schöffl, F. (2005). Identification of novel heat shock factor-dependent genes and biochemical pathways in arabidopsis thaliana. *The Plant Journal, 41*(1), 1-14.

Butenko, Y., & Ohad, N. (2011). Polycomb-group mediated epigenetic mechanisms through plant evolution. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1809*(8), 395-406.

Calarco, J. P., Borges, F., Donoghue, M. T., Van Ex, F., Jullien, P. E., Lopes, T., . . . Becker, J. D. (2012). Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small rna. *Cell, 151*(1), 194-205.

Chailakhian, M. K. (1961). Principles of ontogenesis and physiology of flowering in higher plants. *Canadian Journal of Botany, 39*(7), 1817-1841.

Chan, S. W.-L., Henderson, I. R., & Jacobsen, S. E. (2005). Gardening the genome: DNA methylation in arabidopsis thaliana. *Nature Reviews Genetics, 6*(5), 351.

Chandler, V. L., Eggleston, W. B., & Dorweiler, J. E. (2000). Paramutation in maize. In *Plant gene silencing* (pp. 1-25): Springer.

Chen, X. (2012). Small rnas in development–insights from plants. *Current opinion in genetics & development, 22*(4), 361-367.

Chinnusamy, V., & Zhu, J.-K. (2009). Epigenetic regulation of stress responses in plants. *Current opinion in plant biology, 12*(2), 133-139.

Choi, C.-S., & Sano, H. (2007). Abiotic-stress induces demethylation and transcriptional activation of a gene encoding a glycerophosphodiesterase-like protein in tobacco plants. *Molecular Genetics and Genomics, 277*(5), 589-600.

Choi, Y., Gehring, M., Johnson, L., Hannon, M., Harada, J. J., Goldberg, R. B., . . . Fischer, R. L. (2002). Demeter, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in arabidopsis. *Cell, 110*(1), 33-42.

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., . . . Jacobsen, S. E. (2008). Shotgun bisulphite sequencing of the arabidopsis genome reveals DNA methylation patterning. *Nature, 452*(7184), 215-219.

Dalmay, T., Hamilton, A., Mueller, E., & Baulcombe, D. C. (2000). Potato virus x amplicons in arabidopsis mediate genetic and epigenetic gene silencing. *The Plant Cell, 12*(3), 369-379.

Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Dowen, J. M., Nery, J. R., . . . Ecker, J. R. (2012). Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences, 109*(32), E2183-E2191.

Dubin, M. J., Zhang, P., Meng, D., Remigereau, M.-S., Osborne, E. J., Casale, F. P., . . . Vilhjálmsson, B. (2015). DNA methylation in arabidopsis has a genetic basis and shows evidence of local adaptation. *Elife, 4*, e05255.

Eichten, S. R., Briskine, R., Song, J., Li, Q., Swanson-Wagner, R., Hermanson, P. J., . . . Tiffin, P. (2013). Epigenetic and genetic influences on DNA methylation variation in maize populations. *The Plant Cell, 25*(8), 2783-2797.

Eichten, S. R., Schmitz, R. J., & Springer, N. M. (2014). Epigenetics: Beyond chromatin modifications and complex genetic regulation. *Plant physiology, 165*(3), 933-947.

Engler, P., Weng, A., & Storb, U. (1993). Influence of cpg methylation and target spacing on v (d) j recombination in a transgenic substrate. *Molecular and cellular biology, 13*(1), 571-577.

Feng, S., & Jacobsen, S. E. (2011). Epigenetic modifications in plants: An evolutionary perspective. *Current opinion in plant biology, 14*(2), 179-186.

Finnegan, E., Genger, R., Peacock, W., & Dennis, E. (1998). DNA methylation in plants. *Annual review of plant biology, 49*(1), 223-247.

Finnegan, E., Peacock, W. J., & Dennis, E. S. (1996). Reduced DNA methylation in arabidopsis thaliana results in abnormal plant development. *Proceedings of the National Academy of Sciences, 93*(16), 8449-8454.

Fitter, A., & Fitter, R. (2002). Rapid changes in flowering time in british plants. *Science, 296*(5573), 1689-1691.

Flavell, A. J., Pearce, S. R., & Kumar, A. (1994). Plant transposable elements and the genome. *Current opinion in genetics & development, 4*(6), 838-844.

Ganguly, D., Crisp, P. A., Eichten, S. R., & Pogson, B. J. (2017). The arabidopsis DNA methylome is stable under transgenerational drought stress. *Plant physiology*.

Gao, Z., Liu, H.-L., Daxinger, L., Pontes, O., He, X., Qian, W., . . . Zhang, S. (2010). An rna polymerase ii-and ago4-associated protein acts in rna-directed DNA methylation. *Nature, 465*(7294), 106.

Gehring, M., Bubb, K. L., & Henikoff, S. (2009). Extensive demethylation of repetitive elements during seed development underlies gene imprinting. *Science, 324*(5933), 1447-1451.

Gong, Z., Morales-Ruiz, T., Ariza, R. R., Roldán-Arjona, T., David, L., & Zhu, J.-K. (2002). Ros1, a repressor of transcriptional gene silencing in arabidopsis, encodes a DNA glycosylase/lyase. *Cell, 111*(6), 803-814.

González, R. M., Ricardi, M. M., & Iusem, N. D. (2013). Epigenetic marks in an adaptive water stress-responsive gene in tomato roots under normal and drought conditions. *Epigenetics, 8*(8), 864-872.

Gray, W. M., Östin, A., Sandberg, G., Romano, C. P., & Estelle, M. (1998). High temperature promotes auxin-mediated hypocotyl elongation in arabidopsis. *Proceedings of the National Academy of Sciences, 95*(12), 7197-7202.

Groot, M. P., Kooke, R., Knoben, N., Vergeer, P., Keurentjes, J. J., Ouborg, N. J., & Verhoeven, K. J. (2016). Effects of multi-generational stress exposure and offspring environment on the expression and persistence of transgenerational effects in arabidopsis thaliana. *PloS one, 11*(3), e0151566.

Grunstein, M. (1997). Histone acetylation in chromatin structure and transcription. *Nature, 389*(6649), 349.

Hamilton, E. W., & Heckathorn, S. A. (2001). Mitochondrial adaptations to nacl. Complex i is protected by anti-oxidants and small heat shock proteins, whereas complex ii is protected by proline and betaine. *Plant physiology, 126*(3), 1266-1274.

Hauser, M.-T., Aufsatz, W., Jonak, C., & Luschnig, C. (2011). Transgenerational epigenetic inheritance in plants. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1809*(8), 459-468.

Henderson, I. R., & Jacobsen, S. E. (2007). Epigenetic inheritance in plants. *Nature, 447*(7143), 418-424.

Henderson, I. R., Shindo, C., & Dean, C. (2003). The need for winter in the switch to flowering. *Annual review of genetics, 37*(1), 371-392.

Holliday, R. (1994). Epigenetics: An overview. *genesis, 15*(6), 453-457.

Hong, S. W., & Vierling, E. (2001). Hsp101 is necessary for heat tolerance but dispensable for development and germination in the absence of stress. *The Plant Journal, 27*(1), 25-35.

Hsieh, T.-F., Ibarra, C. A., Silva, P., Zemach, A., Eshed-Williams, L., Fischer, R. L., & Zilberman, D. (2009). Genome-wide demethylation of arabidopsis endosperm. *Science, 324*(5933), 1451-1454.

Hu, X., Li, Y., Li, C., Yang, H., Wang, W., & Lu, M. (2010). Characterization of small heat shock proteins associated with maize tolerance to combined drought and heat stress. *Journal of Plant Growth Regulation, 29*(4), 455-464.

Hutvágner, G., Mlynarova, L., & Nap, J.-P. (2000). Detailed characterization of the posttranscriptional gene-silencing-related small rna in a gus gene-silenced tobacco. *Rna, 6*(10), 1445-1454.

Iba, K. (2002). Acclimative response to temperature stress in higher plants: Approaches of gene engineering for temperature tolerance. *Annual review of plant biology, 53*(1), 225-245.

Iwasaki, M., & Paszkowski, J. (2014). Epigenetic memory in plants. *EMBO J, 33*(18), 1987-1998.

Jacobsen, S. E., & Meyerowitz, E. M. (1997). Hypermethylated superman epigenetic alleles in arabidopsis. *Science, 277*(5329), 1100-1103.

Jiang, C., Mithani, A., Belfield, E. J., Mott, R., Hurst, L. D., & Harberd, N. P. (2014). Environmentally responsive genome-wide accumulation of de novo arabidopsis thaliana mutations and epimutations. *Genome research*, gr. 177659.177114.

Johannes, F., Porcher, E., Teixeira, F. K., Saliba-Colombani, V., Simon, M., Agier, N., . . . Audigier, P. (2009). Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS genetics, 5*(6), e1000530.

Jullien, P. E., Katz, A., Oliva, M., Ohad, N., & Berger, F. (2006). Polycomb group complexes self-regulate imprinting of the polycomb group gene medea in arabidopsis. *Current Biology, 16*(5), 486-492.

Kakutani, T., Jeddeloh, J. A., & Richards, E. J. (1995). Characterization of an arabidopsis thaliana DNA hypomethylation mutant. *Nucleic acids research, 23*(1), 130-137.

Karan, R., DeLeon, T., Biradar, H., & Subudhi, P. K. (2012). Salt stress induced variation in DNA methylation pattern and its influence on gene expression in contrasting rice genotypes. *PloS one, 7*(6), e40203.

Kasschau, K. D., Fahlgren, N., Chapman, E. J., Sullivan, C. M., Cumbie, J. S., Givan, S. A., & Carrington, J. C. (2007). Genome-wide profiling and analysis of arabidopsis sirnas. *PLoS biology, 5*(3), e57.

Kathiria, P., Sidler, C., Golubov, A., Kalischuk, M., Kawchuk, L. M., & Kovalchuk, I. (2010). Tobacco mosaic virus infection results in an increase in recombination frequency and resistance to viral, bacterial, and fungal pathogens in the progeny of infected tobacco plants. *Plant Physiol, 153*(4), 1859-1870.

Katiyar-Agarwal, S., Morgan, R., Dahlbeck, D., Borsani, O., Villegas, A., Zhu, J.-K., . . . Jin, H. (2006). A pathogen-inducible endogenous sirna in plant immunity. *Proceedings of the National Academy of Sciences, 103*(47), 18002-18007.

Khan, A., Enjalbert, J., Marsollier, A.-C., Rousselet, A., Goldringer, I., & Vitte, C. (2013). Vernalization treatment induces site-specific DNA hypermethylation at the vernalization-a1 (vrn-a1) locus in hexaploid winter wheat. *BMC plant biology, 13*(1), 209.

Khan, A., Shah, S., & Irshad, M. (2015). Immediate and transgenerational regulation of plant stress response through DNA methylation. *Journal of Agricultural Science, 7*(4), 144.

Khraiwesh, B., Zhu, J.-K., & Zhu, J. (2012). Role of mirnas and sirnas in biotic and abiotic stress responses of plants. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms, 1819*(2), 137-148.

Koini, M. A., Alvey, L., Allen, T., Tilley, C. A., Harberd, N. P., Whitelam, G. C., & Franklin, K. A. (2009). High temperature-mediated adaptations in plant architecture require the bhlh transcription factor pif4. *Current Biology, 19*(5), 408-413.

Kotak, S., Larkindale, J., Lee, U., von Koskull-Döring, P., Vierling, E., & Scharf, K.-D. (2007). Complexity of the heat stress response in plants. *Current opinion in plant biology, 10*(3), 310-316.

Kotak, S., Vierling, E., Bäumlein, H., & von Koskull-Döring, P. (2007). A novel transcriptional cascade regulating expression of heat stress proteins during seed development of arabidopsis. *The Plant Cell, 19*(1), 182-195.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell, 128*(4), 693-705.

Kovalchuk, I., & Kovalchuk, O. (2012). *Epigenetics in health and disease*: FT Press.

Kovalchuk, I., Kovalchuk, O., Kalck, V., Boyko, V., Filkowski, J., Heinlein, M., & Hohn, B. (2003). Pathogen-induced systemic plant signal triggers DNA rearrangements. *Nature, 423*(6941), 760.

Kovařík, A., Koukalova, B., Holý, A., & Bezděk, M. (1994). Sequence-specific hypomethylation of the tobacco genome induced with dihydroxypropyladenine, ethionine and 5-azacytidine. *FEBS letters, 353*(3), 309-311.

Kruszka, K., Pacak, A., Swida-Barteczka, A., Nuc, P., Alaba, S., Wroblewska, Z., . . . Szweykowska-Kulinska, Z. (2014). Transcriptionally and post-transcriptionally regulated micrornas in heat stress response in barley. *Journal of experimental botany, 65*(20), 6123-6135.

Kumar, S. V., Lucyshyn, D., Jaeger, K. E., Alós, E., Alvey, E., Harberd, N. P., & Wigge, P. A. (2012). Transcription factor pif4 controls the thermosensory activation of flowering. *Nature, 484*, 242.

Kumar, S. V., & Wigge, P. A. (2010). H2a. Z-containing nucleosomes mediate the thermosensory response in arabidopsis. *Cell, 140*(1), 136-147.

Lang-Mladek, C., Popova, O., Kiok, K., Berlinger, M., Rakic, B., Aufsatz, W., . . . Luschnig, C. (2010). Transgenerational inheritance and resetting of stress-induced loss of epigenetic gene silencing in arabidopsis. *Molecular Plant, 3*(3), 594-602.

Larkindale, J., Mishkind, M., & Vierling, E. (2005). Plant responses to high temperature. *Plant abiotic stress, 100*, 144.

Lauria, M., & Rossi, V. (2011). Epigenetic control of gene regulation in plants. *Biochim Biophys Acta, 1809*(8), 369-378.

Lauria, M., Rupe, M., Guo, M., Kranz, E., Pirona, R., Viotti, A., & Lund, G. (2004). Extensive maternal DNA hypomethylation in the endosperm of zea mays. *The Plant Cell, 16*(2), 510-522.

Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics, 11*(3), 204-220.

Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The c. Elegans heterochronic gene lin-4 encodes small rnas with antisense complementarity to lin-14. *Cell, 75*(5), 843-854.

Lee, U., Wie, C., Escobar, M., Williams, B., Hong, S.-W., & Vierling, E. (2005). Genetic analysis reveals domain interactions of arabidopsis hsp100/clpb and cooperation with the small heat shock protein chaperone system. *The Plant Cell, 17*(2), 559-571.

Li, Q., Eichten, S. R., Hermanson, P. J., & Springer, N. M. (2014). Inheritance patterns and stability of DNA methylation variation in maize near-isogenic lines. *Genetics, 196*(3), 667-676.

Li, X., Zhu, J., Hu, F., Ge, S., Ye, M., Xiang, H., . . . Wang, W. (2012). Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics, 13*(1), 300.

Liang, D., Zhang, Z., Wu, H., Huang, C., Shuai, P., Ye, C.-Y., . . . Wang, J. (2014). *Single-base-resolution methylomes of populus trichocarpa reveal the association between DNA methylation and drought stress*.

Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, A. H., & Ecker, J. R. (2008). Highly integrated single-base resolution maps of the epigenome in arabidopsis. *Cell, 133*(3), 523-536.

Liu, C., Lu, F., Cui, X., & Cao, X. (2010). Histone methylation in higher plants. *Annual review of plant biology, 61*, 395-420.

Liu, H.-H., Tian, X., Li, Y.-J., Wu, C.-A., & Zheng, C.-C. (2008). Microarray-based analysis of stress-regulated microrнas in arabidopsis thaliana. *Rna, 14*(5), 836-843.

Liu, J. (2015). Genetic and epigenetic control of plant heat responses. *Frontiers in Plant Science, 6*, 267.

Liu, J., Feng, L., Li, J., & He, Z. (2015). Genetic and epigenetic control of plant heat responses. *Frontiers in Plant Science, 6*.

Lobell, D. B., Schlenker, W., & Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science, 333*(6042), 616-620.

Lukens, L. N., & Zhan, S. (2007). The plant genome's methylation status and response to stress: Implications for plant improvement. *Current opinion in plant biology, 10*(3), 317-322.

Madlung, A., & Comai, L. (2004). The effect of stress on genome regulation and structure. *Annals of Botany, 94*(4), 481-495.

Mahmood, T., Safdar, W., Abbasi, B. H., & Naqvi, S. S. (2010). An overview on the small heat shock proteins. *African Journal of Biotechnology, 9*(7), 927-939.

Mathieu, O., Reinders, J., Čaikovski, M., Smathajitt, C., & Paszkowski, J. (2007). Transgenerational stability of the arabidopsis epigenome is coordinated by cg methylation. *Cell, 130*(5), 851-862.

Meyer, P., Niedenhof, I., & Lohuis, M. t. (1994). Evidence for cytosine methylation of non-symmetrical sequences in transgenic petunia hybrida. *The EMBO journal, 13*(9), 2084-2088.

Migicovsky, Z., & Kovalchuk, I. (2012). Epigenetic modifications during angiosperm gametogenesis. *Frontiers in Plant Science, 3*.

Migicovsky, Z., & Kovalchuk, I. (2013). Changes to DNA methylation and homologous recombination frequency in the progeny of stressed plants. *Biochemistry and Cell Biology, 91*(1), 1-5.

Migicovsky, Z., Yao, Y., & Kovalchuk, I. (2014). Transgenerational phenotypic and epigenetic changes in response to heat stress inarabidopsis thaliana. *Plant Signaling & Behavior, 9*(2), e27971.

Min, L., Li, Y., Hu, Q., Zhu, L., Gao, W., Wu, Y., . . . Zhang, X. (2014). Sugar and auxin signaling pathways respond to high-temperature stress during anther development as revealed by transcript profiling analysis in cotton. *Plant physiology, 164*(3), 1293-1308.

Mirouze, M., Lieberman-Lazarovich, M., Aversano, R., Bucher, E., Nicolet, J., Reinders, J., & Paszkowski, J. (2012). Loss of DNA methylation affects the recombination landscape in arabidopsis. *Proceedings of the National Academy of Sciences, 109*(15), 5880-5885.

Mirouze, M., & Paszkowski, J. (2011). Epigenetic contribution to stress adaptation in plants. *Current opinion in plant biology, 14*(3), 267-274.

Mishra, A. K., Agarwal, S., Jain, C. K., & Rani, V. (2009). High gc content: Critical parameter for predicting stress regulated mirnas in arabidopsis thaliana. *Bioinformation, 4*(4), 151-154.

Mishra, S. K., Tripp, J., Winkelhaus, S., Tschiersch, B., Theres, K., Nover, L., & Scharf, K.-D. (2002). In the complex family of heat stress transcription factors, hsfa1 has a unique role as master regulator of thermotolerance in tomato. *Genes & development, 16*(12), 1555-1567.

Molinier, J., Ries, G., Zipfel, C., & Hohn, B. (2006). Transgeneration memory of stress in plants. *Nature, 442*(7106), 1046-1049.

Molnar, A., Melnyk, C. W., Bassett, A., Hardcastle, T. J., Dunn, R., & Baulcombe, D. C. (2010). Small silencing rnas in plants are mobile and direct epigenetic modification in recipient cells. *Science, 328*(5980), 872-875.

Moore, L. D., Le, T., & Fan, G. (2013). DNA methylation and its basic function. *Neuropsychopharmacology, 38*(1), 23-38.

Murgia, I., Giacometti, S., Balestrazzi, A., Paparella, S., Pagliano, C., & Morandini, P. A. (2015). Analysis of the transgenerational iron deficiency stress memory in arabidopsis thaliana plants. *Frontiers in Plant Science, 6*, 745.

Narlikar, G. J., Fan, H.-Y., & Kingston, R. E. (2002). Cooperation between complexes that regulate chromatin structure and transcription. *Cell, 108*(4), 475-487.

Nathan, D., Ingvarsdottir, K., Sterner, D. E., Bylebyl, G. R., Dokmanovic, M., Dorsey, J. A., . . . Meluh, P. B. (2006). Histone sumoylation is a negative regulator in saccharomyces cerevisiae and shows dynamic interplay with positive-acting histone modifications. *Genes & development, 20*(8), 966-976.

Naydenov, M., Baev, V., Apostolova, E., Gospodinova, N., Sablok, G., Gozmanova, M., & Yahubyan, G. (2015). High-temperature effect on genes engaged in DNA methylation and affected by DNA methylation in arabidopsis. *Plant Physiology and Biochemistry, 87*, 102-108.

Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E., Mathesius, U., . . . Valladares, F. (2010). Plant phenotypic plasticity in a changing climate. *Trends in plant science, 15*(12), 684-692.

Nishizawa, A., Yabuta, Y., Yoshida, E., Maruta, T., Yoshimura, K., & Shigeoka, S. (2006). Arabidopsis heat shock transcription factor a2 as a key regulator in response to several types of environmental stress. *The Plant Journal, 48*(4), 535-547.

Ou, X., Zhang, Y., Xu, C., Lin, X., Zang, Q., Zhuang, T., . . . Liu, B. (2012). Transgenerational inheritance of modified DNA methylation patterns and enhanced tolerance induced by heavy metal stress in rice (oryza sativa l.). *PloS one, 7*(9), e41143.

Paszkowski, J., & Grossniklaus, U. (2011). Selected aspects of transgenerational epigenetic inheritance and resetting in plants. *Current opinion in plant biology, 14*(2), 195-203.

Pecinka, A., Dinh, H. Q., Baubec, T., Rosa, M., Lettner, N., & Scheid, O. M. (2010). Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in arabidopsis. *The Plant Cell, 22*(9), 3118-3129.

Peng, M., Cui, Y., Bi, Y. M., & Rothstein, S. J. (2006). Atmbd9: A protein with a methyl-cpg-binding domain regulates flowering time and shoot branching in arabidopsis. *The Plant Journal, 46*(2), 282-296.

Peng, S., Huang, J., Sheehy, J. E., Laza, R. C., Visperas, R. M., Zhong, X., . . . Cassman, K. G. (2004). Rice yields decline with higher night temperature from global warming. *Proceedings of the National Academy of Sciences of the United States of America, 101*(27), 9971-9975.

Popova, O. V., Dinh, H. Q., Aufsatz, W., & Jonak, C. (2013). The rddm pathway is required for basal heat tolerance in arabidopsis. *Molecular Plant, 6*(2), 396-410.

Porter, J. R. (2005). Rising temperatures are likely to reduce crop yields. *Nature, 436*(7048), 174.

Pradhan, S., & Adams, R. L. (1995). Distinct cg and cng DNA methyltransferases in pisum sativum. *The Plant Journal, 7*(3), 471-481.

Probst, A. V., Dunleavy, E., & Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. *Nature reviews Molecular cell biology, 10*(3), 192-206.

Probst, A. V., Fagard, M., Proux, F., Mourrain, P., Boutet, S., Earley, K., . . . Furner, I. (2004). Arabidopsis histone deacetylase hda6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rdna repeats. *The Plant Cell, 16*(4), 1021-1034.

Proveniers, M. C. G., & van Zanten, M. (2013). High temperature acclimation through pif4 signaling. *Trends in plant science, 18*(2), 59-64.

Puijalon, S., Bouma, T. J., Douady, C. J., van Groenendael, J., Anten, N. P., Martel, E., & Bornette, G. (2011). Plant resistance to mechanical stress: Evidence of an avoidance–tolerance trade-off. *New Phytologist, 191*(4), 1141-1149.

Qiao, W., & Fan, L. (2011). Epigenetics, a mode for plants to respond to abiotic stresses. *Frontiers in Biology, 6*(6), 477-481.

Qu, A.-L., Ding, Y.-F., Jiang, Q., & Zhu, C. (2013). Molecular mechanisms of the plant heat stress response. *Biochemical and Biophysical Research Communications, 432*(2), 203-207.

Ramachandran, V., & Chen, X. (2008). Small rna metabolism in arabidopsis. *Trends in plant science, 13*(7), 368-374.

Rapp, R. A., & Wendel, J. F. (2005). Epigenetics and plant evolution. *New Phytologist, 168*(1), 81-91.

Razin, A., & Cedar, H. (1993). DNA methylation and embryogenesis. In *DNA methylation* (pp. 343-357): Springer.

Ronemus, M. J., Galbiati, M., Ticknor, C., Chen, J., & Dellaporta, S. L. (1996). Demethylation-induced developmental pleiotropy in arabidopsis. *Science, 273*(5275), 654-657.

Roudier, F., Ahmed, I., Bérard, C., Sarazin, A., Mary-Huard, T., Cortijo, S., . . . Al-Shikhley, L. (2011). Integrative epigenomic mapping defines four main chromatin states in arabidopsis. *The EMBO journal, 30*(10), 1928-1938.

Sano, H. (2010). Inheritance of acquired traits in plants: Reinstatement of lamarck. *Plant Signaling & Behavior, 5*(4), 346-348.

Saze, H., Tsugane, K., Kanno, T., & Nishimura, T. (2012). DNA methylation in plants: Relationship to small rnas and histone modifications, and functions in transposon inactivation. *Plant and Cell Physiology, 53*(5), 766-784.

Schmid, M., Uhlenhaut, N. H., Godard, F., Demar, M., Bressan, R., Weigel, D., & Lohmann, J. U. (2003). Dissection of floral induction pathways using global expression analysis. *Development, 130*(24), 6001-6012.

Schmitz, R. J., Hong, L., Fitzpatrick, K. E., & Amasino, R. M. (2007). Dicer-like 1 and dicer-like 3 redundantly act to promote flowering via repression of flowering locus c in arabidopsis thaliana. *Genetics, 176*(2), 1359-1362.

Schmitz, R. J., Schultz, M. D., Lewsey, M. G., O'Malley, R. C., Urich, M. A., Libiger, O., . . . Ecker, J. R. (2011). Transgenerational epigenetic instability is a source of novel methylation variants. *Science, 334*(6054), 369-373.

Schramm, F., Ganguli, A., Kiehlmann, E., Englich, G., Walch, D., & von Koskull-Döring, P. (2006). The heat stress transcription factor hsfa2 serves as a regulatory amplifier of a subset of genes in the heat stress response in arabidopsis. *Plant molecular biology, 60*(5), 759-772.

Scippa, G. S., Di Michele, M., Onelli, E., Patrignani, G., Chiatante, D., & Bray, E. A. (2004). The histone-like protein h1-s and the response of tomato leaves to water deficit. *Journal of experimental botany, 55*(394), 99-109.

Shan, X., Wang, X., Yang, G., Wu, Y., Su, S., Li, S., . . . Yuan, Y. (2013). Analysis of the DNA methylation of maize (zea mays l.) in response to cold stress based on methylation-sensitive amplified polymorphisms. *Journal of Plant Biology, 56*(1), 32-38.

Shen, J., Xie, K., & Xiong, L. (2010). Global expression profiling of rice microrna by one-tube stem-loop reverse transcription quantitative pcr revealed important roles of microrna in abiotic stress responses. *Molecular Genetics and Genomics, 284*(6), 477-488.

Shukla, L. I., Chinnusamy, V., & Sunkar, R. (2008). The role of microrna and other endogenous small rna in plant stress responses. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms, 1779*(11), 743-748.

Simon, S. A., & Meyers, B. C. (2011). Small rna-mediated epigenetic modifications in plants. *Current opinion in plant biology, 14*(2), 148-155.

Slotkin, R. K., Vaughn, M., Borges, F., Tanurdžić, M., Becker, J. D., Feijó, J. A., & Martienssen, R. A. (2009). Epigenetic reprogramming and small rna silencing of transposable elements in pollen. *Cell, 136*(3), 461-472.

Song, Y., Ji, D., Li, S., Wang, P., Li, Q., & Xiang, F. (2012). The dynamic changes of DNA methylation and histone modifications of salt responsive transcription factor genes in soybean. *PloS one, 7*(7), e41274.

Sridhar, V. V., Kapoor, A., Zhang, K., Zhu, J., Zhou, T., Hasegawa, P. M., . . . Zhu, J.-K. (2007). Control of DNA methylation and heterochromatic silencing by histone h2b deubiquitination. *Nature, 447*(7145), 735.

Sun, W., Bernard, C., Van De Cotte, B., Van Montagu, M., & Verbruggen, N. (2001). At-hsp17. 6a, encoding a small heat-shock protein in arabidopsis, can enhance osmotolerance upon overexpression. *The Plant Journal, 27*(5), 407-415.

Sun, W., Van Montagu, M., & Verbruggen, N. (2002). Small heat shock proteins and stress tolerance in plants. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression, 1577*(1), 1-9.

Sung, D.-Y., Kaplan, F., Lee, K.-J., & Guy, C. L. (2003). Acquired tolerance to temperature extremes. *Trends in plant science, 8*(4), 179-187.

Sunkar, R., & Zhu, J.-K. (2004). Novel and stress-regulated microrna and other small rna from arabidopsis. *The Plant Cell, 16*(8), 2001-2019.

Suter, L., & Widmer, A. (2013a). Environmental heat and salt stress induce transgenerational phenotypic changes in arabidopsis thaliana. *PloS one, 8*(4), e60364.

Suter, L., & Widmer, A. (2013b). Phenotypic effects of salt and heat stress over three generations in arabidopsis thaliana. *PloS one, 8*(11), e80819.

Tan, M.-p. (2010). Analysis of DNA methylation of maize in response to osmotic and salt stress based on methylation-sensitive amplified polymorphism. *Plant Physiology and Biochemistry, 48*(1), 21-26.

Tariq, M., & Paszkowski, J. (2004). DNA and histone methylation in plants. *TRENDS in Genetics, 20*(6), 244-251.

Thomashow, M. F. (1999). Plant cold acclimation: Freezing tolerance genes and regulatory mechanisms. *Annual review of plant biology, 50*(1), 571-599.

Tittel-Elmer, M., Bucher, E., Broger, L., Mathieu, O., Paszkowski, J., & Vaillant, I. (2010). Stress-induced activation of heterochromatic transcription. *PLoS genetics, 6*(10), e1001175.

Tsuji, H., Saika, H., Tsutsumi, N., Hirai, A., & Nakazono, M. (2006). Dynamic and reversible changes in histone h3-lys4 methylation and h3 acetylation occurring at submergence-inducible genes in rice. *Plant and Cell Physiology, 47*(7), 995-1003.

Twell, D. (2011). Male gametogenesis and germline specification in flowering plants. *Sexual Plant Reproduction, 24*(2), 149-160.

Uthup, T. K., Ravindran, M., Bini, K., & Thakurdas, S. (2011). Divergent DNA methylation patterns associated with abiotic stress in hevea brasiliensis. *Molecular Plant, 4*(6), 996-1013.

Verhoeven, K. J., Jansen, J. J., van Dijk, P. J., & Biere, A. (2010). Stress-induced DNA methylation changes and their heritability in asexual dandelions. *New Phytologist, 185*(4), 1108-1118.

Vidalis, A., Zivkovic, D., Wardenaar, R., Roquis, D., Tellier, A., & Johannes, F. (2016). Methylome evolution in plants. *Genome biology, 17*(1), 264.

Vinocur, B., & Altman, A. (2005). Recent advances in engineering plant tolerance to abiotic stress: Achievements and limitations. *Current opinion in biotechnology, 16*(2), 123-132.

Wada, Y., Miyamoto, K., Kusano, T., & Sano, H. (2004). Association between up-regulation of stress-responsive genes and hypomethylation of genomic DNA in tobacco plants. *Molecular Genetics and Genomics, 271*(6), 658-666.

Wang, M., Qin, L., Xie, C., Li, W., Yuan, J., Kong, L., . . . Liu, S. (2014). Induced and constitutive DNA methylation in a salinity-tolerant wheat introgression line. *Plant and Cell Physiology, 55*(7), 1354-1365.

Wang, W., Vinocur, B., Shoseyov, O., & Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends in plant science, 9*(5), 244-252.

Wang, W.-S., Pan, Y.-J., Zhao, X.-Q., Dwivedi, D., Zhu, L.-H., Ali, J., . . . Li, Z.-K. (2010). Drought-induced site-specific DNA methylation and its association with drought tolerance in rice (oryza sativa l.). *Journal of experimental botany, 62*(6), 1951-1960.

Wang, W.-S., Zhao, X., Pan, Y., Zhu, L., Fu, B., & Li, Z. (2011). DNA methylation changes detected by methylation-sensitive amplified polymorphism in two contrasting rice genotypes under salt stress. *Journal of Genetics and Genomics, 38*(9), 419-424.

Wang, Y., Sun, F., Cao, H., Peng, H., Ni, Z., Sun, Q., & Yao, Y. (2012). Tamir159 directed wheat tagamyb cleavage and its involvement in anther development and heat response. *PloS one, 7*(11), e48445.

Whittle, C., Otto, S., Johnston, M. O., & Krochko, J. (2009). Adaptive epigenetic memory of ancestral temperature regime in arabidopsis thaliana. *Botany, 87*(6), 650-657.

Widman, N., Jacobsen, S. E., & Pellegrini, M. (2009). Determining the conservation of DNA methylation in arabidopsis. *Epigenetics, 4*(2), 119-124.

Xiao, W., Custard, K. D., Brown, R. C., Lemmon, B. E., Harada, J. J., Goldberg, R. B., & Fischer, R. L. (2006). DNA methylation is critical for arabidopsis embryogenesis and seed viability. *The Plant Cell, 18*(4), 805-814.

Xie, Z., Johansen, L. K., Gustafson, A. M., Kasschau, K. D., Lellis, A. D., Zilberman, D., . . . Carrington, J. C. (2004). Genetic and functional diversification of small rna pathways in plants. *PLoS biology, 2*(5), e104.

Yadegari, R., & Drews, G. N. (2004). Female gametophyte development. *The Plant Cell, 16*(suppl 1), S133-S141.

Yaish, M. W. (2013). DNA methylation-associated epigenetic changes in stress tolerance of plants. In *Molecular stress physiology of plants* (pp. 427-440): Springer.

Yaish, M. W., Al-Lawati, A., Al-Harrasi, I., & Patankar, H. V. (2018). Genome-wide DNA methylation analysis in response to salinity in the model plant caliph medic (medicago truncatula). *BMC Genomics, 19*(1), 78.

Yaish, M. W., Colasanti, J., & Rothstein, S. J. (2011). The role of epigenetic processes in controlling flowering time in plants exposed to stress. *Journal of experimental botany, 62*(11), 3727-3735.

Yao, Y., Ni, Z., Peng, H., Sun, F., Xin, M., Sunkar, R., . . . Sun, Q. (2010). Non-coding small rnas responsive to abiotic stress in wheat (triticum aestivum l.). *Functional & Integrative Genomics, 10*(2), 187-190.

Zemach, A., Kim, M. Y., Silva, P., Rodrigues, J. A., Dotson, B., Brooks, M. D., & Zilberman, D. (2010). Local DNA hypomethylation activates genes in rice endosperm. *Proceedings of the National Academy of Sciences, 107*(43), 18729-18734.

Zemach, A., McDaniel, I. E., Silva, P., & Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science, 328*(5980), 916-919.

Zhang, H., & Zhu, J.-K. (2012). *Active DNA demethylation in plants and animals*.

Zhang, K., Sridhar, V. V., Zhu, J., Kapoor, A., & Zhu, J.-K. (2007). Distinctive core histone post-translational modification patterns in arabidopsis thaliana. *PloS one, 2*(11), e1210.

Zhang, X., Bernatavichute, Y. V., Cokus, S., Pellegrini, M., & Jacobsen, S. E. (2009). Genome-wide analysis of mono-, di-and trimethylation of histone h3 lysine 4 in arabidopsis thaliana. *Genome biology, 10*(6), R62.

Zhang, X., Henderson, I. R., Lu, C., Green, P. J., & Jacobsen, S. E. (2007). Role of rna polymerase iv in plant small rna metabolism. *Proceedings of the National Academy of Sciences, 104*(11), 4536-4541.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S. W.-L., Chen, H., . . . Jacobsen, S. E. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell, 126*(6), 1189-1201.

Zheng, X., Chen, L., Li, M., Lou, Q., Xia, H., Wang, P., . . . Luo, L. (2013). Transgenerational variations in DNA methylation induced by drought stress in two rice varieties with distinguished difference to drought resistance. *PloS one, 8*(11), e80253.

Zheng, X., Chen, L., Xia, H., Wei, H., Lou, Q., Li, M., . . . Luo, L. (2017). Transgenerational epimutations induced by multi-generation drought imposition mediate rice plant's adaptation to drought condition. *Scientific reports, 7*.

Zheng, X., Zhu, J., Kapoor, A., & Zhu, J. K. (2007). Role of arabidopsis ago6 in sirna accumulation, DNA methylation and transcriptional gene silencing. *The EMBO journal, 26*(6), 1691-1701.

Zhong, L., Xu, Y.-h., & Wang, J.-b. (2009). DNA-methylation changes induced by salt stress in wheat triticum aestivum. *African Journal of Biotechnology, 8*(22).

Zhong, S.-H., Liu, J.-Z., Jin, H., Lin, L., Li, Q., Chen, Y., . . . Qi, Y.-J. (2013). Warm temperatures induce transgenerational epigenetic release of rna silencing by inhibiting sirna biogenesis in arabidopsis. *Proceedings of the National Academy of Sciences, 110*(22), 9171-9176.

Zhou, C., Zhang, L., Duan, J., Miki, B., & Wu, K. (2005). Histone deacetylase19 is involved in jasmonic acid and ethylene signaling of pathogen response in arabidopsis. *The Plant Cell, 17*(4), 1196-1204.

Zhou, L., Liu, Y., Liu, Z., Kong, D., Duan, M., & Luo, L. (2010). Genome-wide identification and analysis of drought-responsive microrna s in oryza sativa. *Journal of experimental botany, 61*(15), 4157-4168.

Zhu, J.-K. (2008). Epigenome sequencing comes of age. *Cell, 133*(3), 395-397.

Zinn, K. E., Tunc-Ozdemir, M., & Harper, J. F. (2010). Temperature stress and plant sexual reproduction: Uncovering the weakest links. *Journal of experimental botany, 61*(7), 1959-1968.

3. MATERIALS AND METHODS

3.1 Parental Generation and Progeny Plants

Plant seeds used for parental generation (P) were obtained from single homozygous *Arabidopsis thaliana* (Columbia ecotype, Col-0) transgenic for *luciferase* (LUC) recombination reporter gene (15*d*8 line) carrying a copy of a direct repeat of *luciferase* recombination transgene. All parental plants were allowed to flower, and seeds collected from approximately 16-20 plants. Subsequent progeny plants were obtained from these seeds with ~ 100 seeds sown, ~ 20 plants were heat treated during development, grown to maturity and seeds collected, this made up a generation. The same process was repeated consecutively until twenty -five generations representing progeny plant lines were obtained. Seeds were kept in storage under dry conditions at room temperature.

Progeny plants obtained from progeny seeds described above for every generation had two groups; stressed (S) and non-stressed groups (NS, which acted as the control group within each generation). The S group included heat-stressed plants to assist in determining the potential role of multi-generational heat stress response while NS group included plants grown under normal conditions to act as control.

All progeny plants representing the S group were heat treated during development at 42°C for four hours per day for four consecutive days starting at day 10 post germination and allowed to flower and mature. Seeds were then collected and propagated to obtain the next successive generation; creating 25 generations of stressed S group with three biological replicates in each generation. On the other hand, progeny plants representing the NS group were not heat stressed but grown under normal conditions, allowed to flower and mature. Seeds were then collected and

propagated to obtain the next successive generation; creating 25 generations of non- stressed NS group with two biological replicates in each generation.

In summary, each generation had two replicates of NS and three replicates of S plant lines totaling five samples per generation.

3.1.1 Sampling Generations

Sampling generations were selected to allow comprehensive representation of time intervals within the twenty-five generations. Sampling generation chosen were generations two (G_2) acting as parental, generations ten (G_{10}), twenty (G_{20}) and twenty-five (G_{25}) representing sampled progenies. As a result, there were in total fifteen samples from the sampled progeny generations consisting of both S and NS plants while for the parental generation two samples consisting of NS plants.

3.2 Plant Seed Preparation and Growth

Seeds from aforementioned parental and progeny sampling generations were kept at 4°C for seven days to initiate stratification and then planted in all-purpose potting soil prepared with water containing generic fertilizer (Miracle-Gro, Scotts Canada.) made to field capacity in 4 x 4 pots. The moist soil containing plant seeds was further stratified for 24 hours at 4°C. Plants were grown in a growth chamber (BioChambers) at 22°C under extended day conditions of 16 hours light and 8 hours dark (18°C). Approximately three days post germination (dpg) plant seedlings were transplanted into individual pots containing 9:1 soil to vermiculite composition prepared with fertilizer. Pots containing plants were placed in trays measuring ~ 25 cm x 50 cm and watered from below. This gave a total of eight (8) plants per pot and two pots per sample group resulting in approximately sixteen (16) plants per treatment group.

3.3 Plant Experimental Conditions

Two experimental conditions were established; heat-treated (T) and non-heat-treated (NT) groups. NT plants were placed out of growth chambers at room temperature of approximately 22°C for the duration of which T group were under heat stress treatment (see description in 3.4). Experiments on all sampling generations were not simultaneously carried out but were staggered over time due to the magnitude of the total of individual experimental plants involved under each experimental condition (~ 8100 plants). A second independent repeat for all sampling generations was carried out simultaneously for physiological analysis to confirm phenotype changes observed from the initial staggered group of experiments. Plants from which tissues were collected were killed to prevent bias in progeny due to mechanical or chemical damage.

3.4 Heat Stress Treatment

Stress treatment was applied to plants at 10_dpg for four consecutive days. Plants belonging to T group were incubated at 50°C for 4 hours, 42°C for 6 hours and 37°C for 12 hours while plants belonging to NT group were kept alongside its corresponding T group at approximately 22°C for 4hours, 6 hours and 12hours respectively.

3.5 Physiological Profiling and Measurements

Measurements were taken to create a physiological profile of plants response to heat stress treatment by observing plant changes at certain development stages according to criteria described by Boyes et al., (2001). Phenotype changes measured were number of leaves at 21_dpg (approximate time to reach full rosette stage), percentage bolting rate (flowering), fresh and dry weight of mature plant and seed length.

The number of leaves on each plant was counted (excluding cotyledons) at 21_dpg. At 30_dpg bolting was assessed on each plant to determine percentage of plants in each group that had bolted and were bolting**.** Biomass measurements of fresh weight and dry weight at 30_dpg were also measured. For fresh weight, individual plants were uprooted and weighed on five-digit balance (Mettler Teledo) while for the dry weight same particular plant was incubated at 37°C for four days, cooled and weighed to obtain dry weight. For seed length measurements, seeds were collected from plants photographed using a microscope (HUND WETZLAR, Germany) and the diameter of ~ 20seeds were measured per heat treatment using Image J to determine seed length.

3.5.1 Measurement of Multigenerational Effects of Heat Stress

Progeny of stressed plants was analyzed for the possible acquisition of higher tolerance to heat stress and whether the stress response varied among generations by measuring same physiological profile at same day post germination in progeny generations and then comparing with parental generation.

3.6 Molecular Techniques Used

Parental G_2 plants and progeny G_{25} plants were grown to about 21dpg, leaf tissue was harvested for molecular analysis. G_2 samples consisted of two replicates of NS group while G_{25} samples consisted of S and NS group making up five samples (two replicates of NS and three replicates of S). Also, duplicate leaf tissues representing all samples were harvested as well resulting in a total of fifteen samples used for molecular analysis.

3.6.1 DNA Isolation

Whole rosette leaves of G_2 and G_{25} plants were collected, frozen in liquid Nitrogen and stored at -80°C for DNA extraction. Total genomic DNA was extracted from approximately 100mg of leaf tissue homogenized in liquid Nitrogen using CTAB protocol with some modifications. DNA extraction buffer consisted of 31.8 g Sorbitol, 6 g Trizma base (Tris), 0.84 g EDTA in 500ml made with DDW, pH adjusted to 7.5 with HCl. Nucleic lysis buffer was prepared using 30.29g Tris, 23.27g EDTA, 73.05g NaC, 5g CTAB dissolved in ~ 250ml DDW, pH adjusted to 7.5. Total extraction buffer used was prepared with Na-biSulfite (38mg/10mL) added before use, 10 ml of nucleic lysis buffer and 4 ml of 5% Sarkosyl. 700µl of total extraction buffer was used per sample.

Samples were incubated at 65°C for 1 hour and inverted periodically. 700µl of chloroform was added to samples and shaken by hand for 5 minutes. Samples were centrifuged at 16,000xg for 10 minutes at 4°C with the supernatant phase transferred to a new tube, this chloroform step was repeated. Two-third volume of Isopropanol was added and incubated at room temperature for about 24 hours to precipitate DNA. Samples were centrifuged at 12,000xg for 15 minutes at 4°C, pellet of precipitated gDNA was then rinsed twice with 70% ethanol and once in 100% ethanol, then air-dried at room temperature for about 10 minutes. 100 µl of P1 buffer (Qiagen Kit**)** mixed with RNAase was added, and samples were incubated for 10 minutes at 37°C.

Repeat extraction was performed at this stage by adding 100 µl of Phenol: Chloroform mixture, centrifuged at 16,000xg for 10 minutes at 4°C with 90 µl the supernatant phase transferred to a new tube. 9 µl of Sodium acetate and 250 µl of 100% ethanol was added and incubated at room temperature to precipitate DNA. Samples were centrifuged at 12,000xg for 10 minutes at 4°C to obtain pellets of precipitated DNA and washed twice with 1ml of 70% ethanol and once in

1ml of 100% ethanol, then air-dried at room temperature. DNA pellets were dissolved in distilled water and quantified using NanoDrop 2000C spectrophotometer (Thermo Fisher Scientific Inc.). Also, agarose gel electrophoresis was performed to verify integrity of DNA sample.

3.7 Whole Genome Sequencing and Whole Genomic Bisulfite Sequencing

Genomic DNA isolated was used for whole genome sequencing (WGS) and whole genome bisulfite sequencing (WGBS) (Illumina) to assist in identifying genomic and epigenomic (associated with changes in DNA methylation) profile and variations. Data obtained were analyzed using several toolkits found in *methylKit* package.

WGBS allows for the investigation of genome-wide patterns of DNA methylation at a single base resolution. It involves sodium bisulfite conversion of unmethylated cytosine to uracil, the resulting cytosine residues in the sequence represent methylated cytosine in the genome which is then mapped to a reference genome (Susan, Harrison, Paul, & Frommer, 1994). Binomial tests was applied and used to determine observed methylation frequency against bisulfite conversion reaction and percentage methylation (%methylation) levels were calculated at each base (Schultz, Schmitz, & Ecker, 2012).

3.7.1 WGBS Data Computation and Analysis

Raw sequencing reads were quality controlled and trimmed using Trim Galore software (version 0.4.4). Trimmed reads were then aligned to the TAIR10 reference genome using Bismark bisulfite mapping tool (Krueger & Andrews, 2011). Methylated cytosines (Cs) were extracted from aligned reads with Bismark methylation extractor on default parameters, followed by the computation of methylation frequency using the R package software; *methylKit*. The

%methylation was calculated by counting the ratio of frequency of Cs divided by reads with C or a T at each base and computed at bases with coverage ≥ 10 (Akalin et al., 2012).

$$\%\text{Methylation} = \{Frequency\ of\ C \div read\ coverage\} \times 100$$

Common bases covered across all samples were extracted and compared, differential hyper- and hypo- methylated bases in each chromosome were extracted. Differentially methylated cytosines (DMCs) overlapping with genomic regions were assessed (in the preference of promotor > exon > intron), and the average %methylation of DMCs around genes with the distances of DMCs to nearest transcription start sites (TSSs) were also calculated.

Annotation analysis was performed with *genomation* package within *methylKit* to obtain biological understanding of genomic intervals over pre-defined functional regions like promoters, exons, and introns (Akalin, Franke, Vlahoviček, Mason, & Schübeler, 2014). Functional commentary of generated gene expression profiles was performed using SuperViewer tool with Bootstrap to show difference between samples (Provart & Zhu, 2003). Hierarchical clustering of samples was used to analyze for similarities and detect sample outliners based on %methylation scores and possible molecular signature. Also, Principal Component Analysis (PCA) was utilized for variations and any biological relevant clustering of samples. Scatterplots, bar plots showing percentage of hyper-/hypo- methylated bases overall chromosome and heat maps were used to visualize similarities and dissimilarities between DNA methylation profiles.

3.7.1.1 Differentially Methylated Regions (DMRs)

Comparison of differential DNA methylation levels between samples reveals significant locations of different changes in the epigenome**.** The direction (random or non-random) and location can be correlated. DMRs information obtained were investigated over predefined regions for all context; CG, CHG, and CHH on 100bp and 1000bp tiles across the genome to identify stochastic and treatment–associated DMRs (Akalin et al., 2012). Differential hyper-/hypo-methylated regions were extracted as well and compared across samples. By default, DMRs were extracted with q-values < 0.01 and %methylation difference > 25%. The differential methylation patterns between sample groups and the methylation events of these differences per chromosome were extracted as well. Methylation profiles of sample groups used were F25H versus F2C, F25H versus F25C and F25C versus F2C.

In summary, sliding windows of 100bp and 1000bp were considered for both DMRs and DMCs, and extractions were extracted based on at least 25% and 50% differences (q-values > 0.01) to assess significant differences among samples.

3.7.2 WGS Data Computation and Analysis

Raw sequencing reads were trimmed and aligned to the TAIR10 reference genome as described in 3.7.1; the only difference was that duplicates were marked using Picard tool. Local realignments around SNPs and INDELs were performed using GATK (genome analysis toolkit) which accounts for genome aligners, mapping errors and gives consistent regions that contain SNPs and INDELs. The resulting reads were quality controlled with Haplotype scores and sample variant sites were called individually and jointly using Haplotypercaller with GATK. Sites marked

as low-quality score by GATK were filtered out and used. The effects of variants in the genome sequences were classified using SnpEff program (Cingolani et al., 2012).

Similar to WBGS computation analysis, toolkits used included *genomation* to obtain biological understanding of genomic intervals and Functional Classification SuperViewer to create gene expression profiles and show difference between samples. The genes nearest to the non-overlapping SNPs and INDELs sites were annotated.

3.8 Statistical Analysis and Quality Control Values

Mapped reads were obtained with quality score of < 30, differential hyper- and hypomethylated bases were extracted with q-values < 0.01 and %methylation difference larger than 25% in *methylKit.* Heatmaps of differentially methylated bases were quantified at q-values < 0.01 and %methylation difference more significant than 50%. Distances of DMCs to nearest TSSs obtained from *genomation* were run at >25% and >50%. TSSs distance to DMCs was extracted within +/- 1000bp and annotated at DMCs >50%. DNA methylation profiles obtained from *melthylKit* used pairwise correlation coefficients of the %methylation levels obtained and 1-Pearson's correlation distance for hierarchical clustering of samples. Logistic regression and Fisher's exact test were used in the determination of differential methylation with calculations of q-values and Benjamini-Hochberg for p-values corrections. T-test for mean difference between groups was calculated and extracted with at least p-values < 0.05. Global genome methylation results were graphed using Microsoft Excel (MS) and output graphs from each corresponding program used.

Statistical analyses on physiological measurements data were performed using R (with RStudio tools) and MS program, standard errors or standard deviations were calculated and

significant differences between the means were compared using three-way ANOVA and Tukey-Kramer test. Significant differences between experimental groups were analyzed using p-values.

3.9 Plant Genealogy

In this study, multi-generational response to heat stress at different temperatures were evaluated by measuring different physiological response of plants during development at certain growth stages. All of which enabled the creation of a response profile with interpretations of phenotypic differences that occurred within plants grown in stress-induced environment and ambient conditions, as well as allowed analysis for changes within and across generations. Seedlings used for this study originated from a single generation of *Arabidopsis thaliana* (Col wt -*15d8* line) and had undergone generational heat stress treatment. Our experiment began with G_2 seeds from non-stressed lineage (control, i.e., NS), seeds of all sampling progeny generations were from NS lineage and stressed lineage (i.e., S) of that particular progeny generation. Generations selected for comparison were parental generation $G_{2,}$ and progeny generations were $G_{10,}$ G_{20}, and G_{25} as described in Figure 1. The number of leaves at 21_dpg**,** bolting percentage at 30dpg, fresh and dry weight of mature plants and average seed length were recorded. To identify whether heat stress response of the offspring generations improved over time and across generations, the physiological differences between parental and progeny generations under ambient and stress conditions were compared.

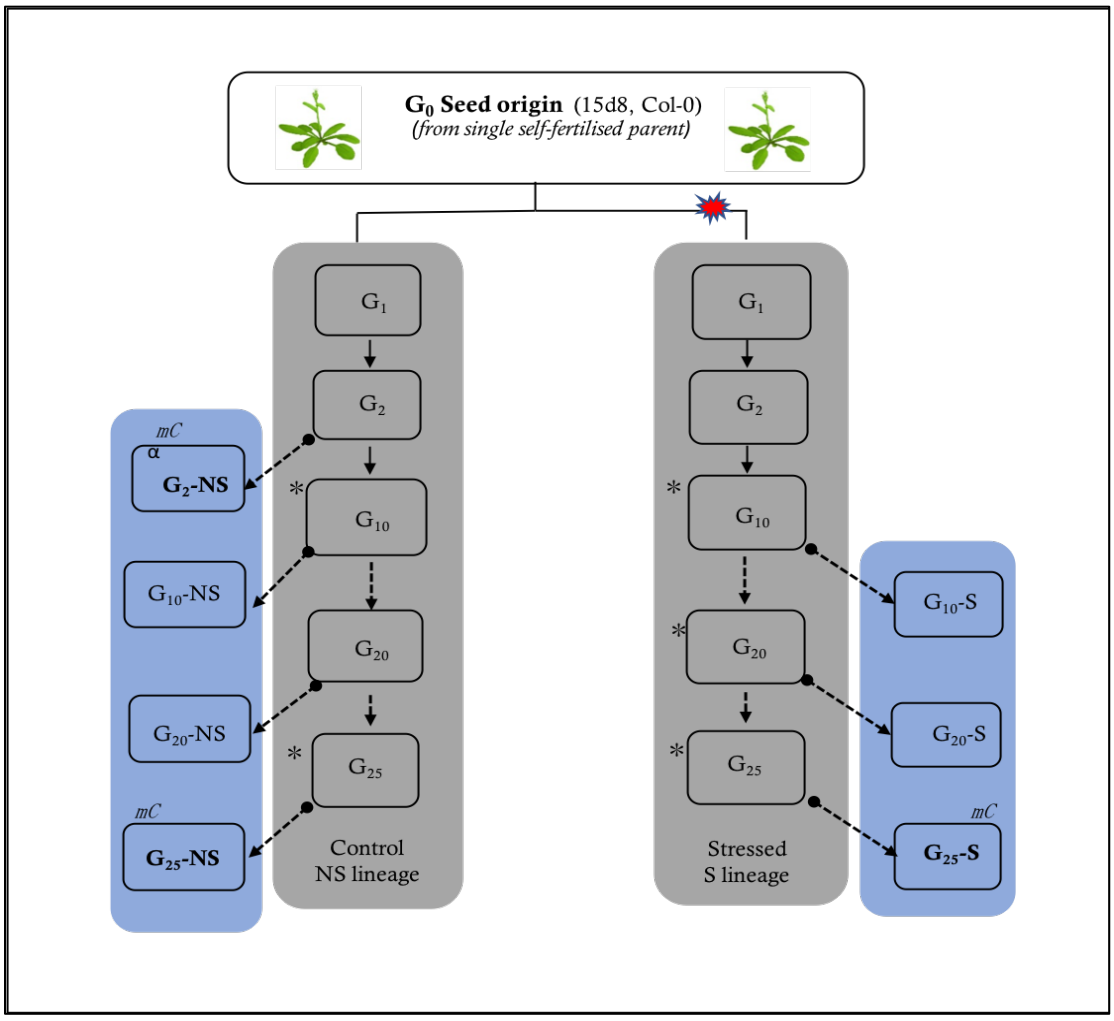

Figure 1. Origin of plant genealogy used to form experimental groups. Single *Arabidopsis thaliana* plant (Col-wt-*15d8,*) was seed origin. Plants were grown for twenty-five generations either exposed to temperature stress or in a control ambient environment. Seeds of second-generation G_2 (indicated by an alpha sign) were taken as experimental parent generation for comparison with other generations (progeny). Progeny generations used in this study were G_{10}, G_{20} and G_{25} and tested for effects of multigenerational heat stress exposure. Whole genome bisulfite sequencing and whole genome sequencing were performed with seeds grown independently from G_2 and G_{25} (indicated by mC).

3.10 CHAPTER 3 REFERENCES

Akalin, A., Franke, V., Vlahoviček, K., Mason, C. E., & Schübeler, D. (2014). Genomation: A toolkit to summarize, annotate and visualize genomic intervals. *bioinformatics, 31*(7), 1127-1129.

Akalin, A., Kormaksson, M., Li, S., Garrett-Bakelman, F. E., Figueroa, M. E., Melnick, A., & Mason, C. E. (2012). Methylkit: A comprehensive r package for the analysis of genome-wide DNA methylation profiles. *Genome biology, 13*(10), R87.

Boyes, D. C., Zayed, A. M., Ascenzi, R., McCaskill, A. J., Hoffman, N. E., Davis, K. R., & Görlach, J. (2001). Growth stage–based phenotypic analysis of arabidopsis a model for high throughput functional genomics in plants. *The Plant Cell, 13*(7), 1499-1510.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., . . . Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, snpeff: Snps in the genome of drosophila melanogaster strain w1118; iso-2; iso-3. *Fly, 6*(2), 80-92.

Krueger, F., & Andrews, S. R. (2011). Bismark: A flexible aligner and methylation caller for bisulfite-seq applications. *bioinformatics, 27*(11), 1571-1572.

Provart, N., & Zhu, T. (2003). A browser-based functional classification superviewer for arabidopsis genomics. *Currents in Computational Molecular Biology, 2003*, 271-272.

Schultz, M. D., Schmitz, R. J., & Ecker, J. R. (2012). 'Leveling'the playing field for analyses of single-base resolution DNA methylomes. *TRENDS in Genetics, 28*(12), 583-585.

Susan, J. C., Harrison, J., Paul, C. L., & Frommer, M. (1994). High sensitivity mapping of methylated cytosines. *Nucleic acids research, 22*(15), 2990-2997.

4. RESULTS

4.1 Phenotypic Analysis of *Arabidopsis thaliana* After Multi-Generational Exposure to Heat Stress

4.1.1 Changes in Leaf Number in Parental and Progeny Generations

There was an overall significant difference ($p<0.05$) in the number of leaves across generations when progeny plants are compared to parental plants. G_{25} and G_{20} plants showed a significant difference ($p<0.05$) in the number of leaves when compared to G_2 while G_{10} plants had the most number of leaves at 21dpg across all progeny generations with a considerable difference ($p<0.05$) when compared to G_{20} and G_{25}.

At 37°C within generations, leaf number on non-stressed plants that were progeny of NS lineage (NS-NT) were overall significantly higher than the stressed plants of the same origin (NS-T), except for G_{10} which was consistently the same for G_{10}-NS grown under stressed and normal conditions (Figure 2). Also, within G_{20} the leaf number on the progeny of plants exposed to multigenerational heat, S lineage that was grown under stressed condition (S-T) was higher when compared to those grown under normal conditions (S-NT), while in G_{25} the opposite was observed. There were response differences between progeny generation with G_{10} having a significant difference ($p<0.001$) from G_{20} and G_{25}.

At 42°C temperature, the number of leaves on plants varied based on experimental treatment and generational heat exposure of the plants. Progeny plants of stressed lineage had high number of leaves across all three progeny generations which was similar to heat stressed parental plants (NS-T) (Figure 3). There were significant differences between progeny generations ($p<0.05$). Although, within advanced generation G_{25}, both heat exposed (S), and non-heat (NS) exposed plants irrespective of the lineage had similar number of leaves ($p<0.05$).

At 50°C parental generation showed no difference in leaf number of plants grown under stress or normal conditions. Within G_{20} there was significant difference between plants exposed to heat stress and those grown under normal conditions ($p<0.05$; Figure 4). While within G_{25} non-stressed plants that were exposed to heat (NS-T) were lower than stressed plants that were exposed to heat (S-T; $p<0.05$).

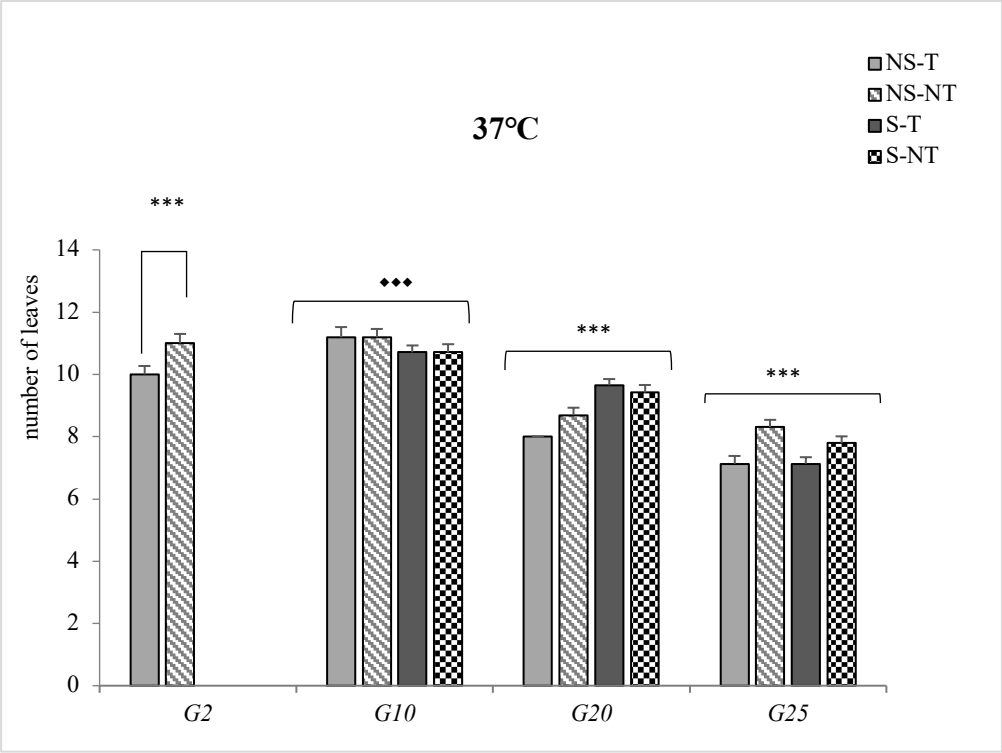

Figure 2. The response of leaf number on parental and progeny generations exposed to 37°C heat and ambient conditions. NS-T are plants of non-stressed lineage that was heat treated, NS-NT are non-stressed lineage that were non-stressed, S-T are plants of stressed lineage that was heat treated and S-NT are plants stressed lineage that was non-heat treated. Leaves were counted at 21dpg, (n=32, SE). The asterisk above bracket (*) show a significant difference compared to parental generation where * indicates $p < 0.05$, ** indicates $p < 0.01$ and *** indicates $p < 0.001$. The diamond (◆) symbol indicates the significant difference between progeny generations where ◆ indicates $p < 0.05$, ◆◆ indicates $p < 0.01$ and ◆◆◆ indicates $p < 0.001$(Tukey-Kramer test).

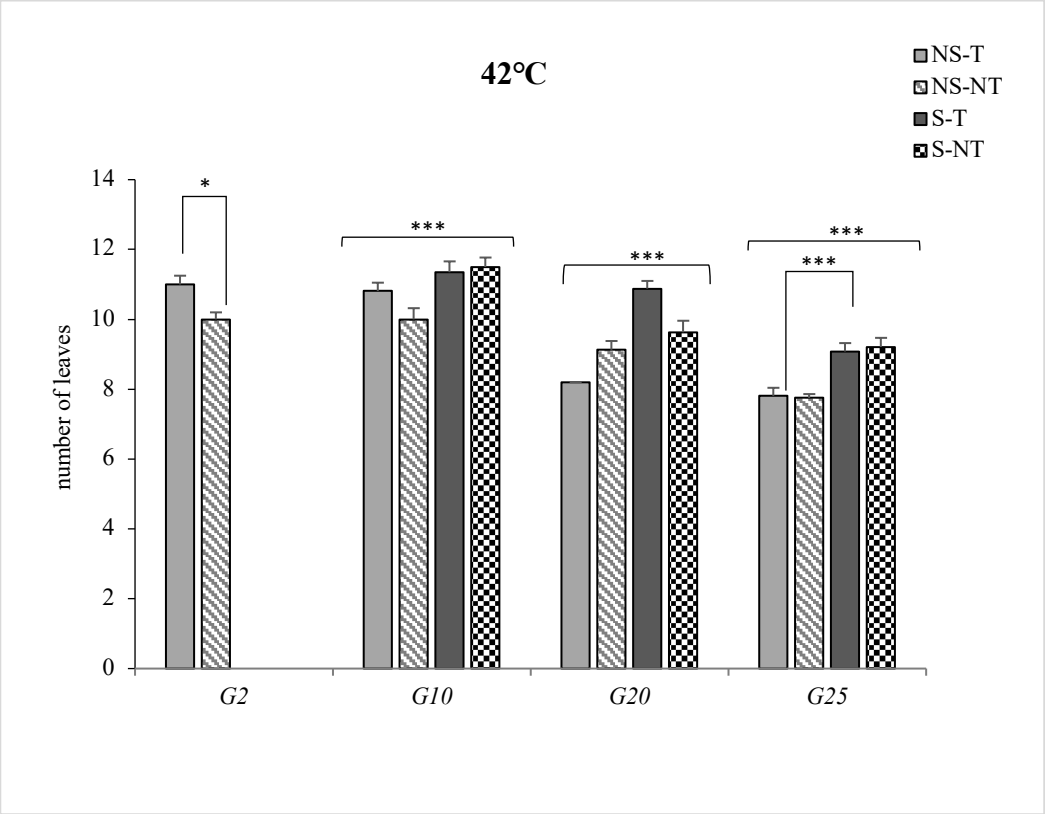

Figure 3. The response of leaf number on parental and progeny generations exposed to 42°C heat and ambient conditions. NS-T are plants of non-stressed lineage that was heat treated, NS-NT are non-stressed lineage that were non-stressed, S-T are plants of stressed lineage that was heat treated and S-NT are plants stressed lineage that was non-heat treated. Leaves were counted at 21dpg, (n=32, SE). The asterisk above bracket (*) show a significant difference compared to parental generation where * indicates $p < 0.05$, ** indicates $p < 0.01$ and *** indicates $p < 0.001$.

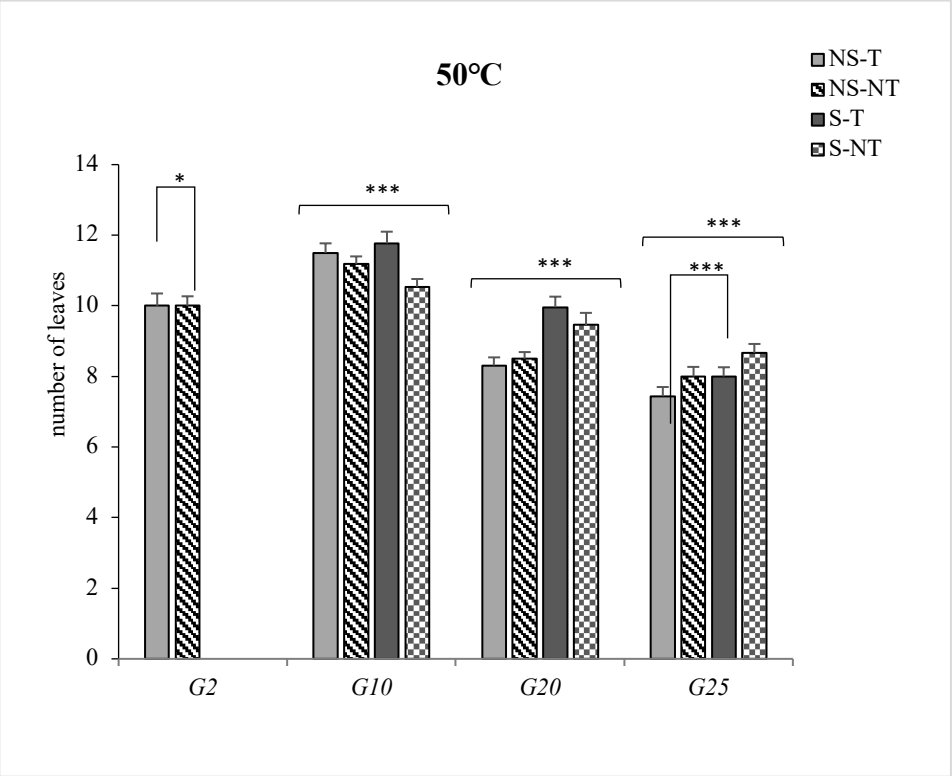

Figure 4. The response of leaf number on parental and progeny generations exposed to 50°C heat and ambient conditions. NS-T are plants of non-stressed lineage that was heat treated, NS-NT are non-stressed lineage that were non-stressed, S-T are plants of stressed lineage that was heat treated and S-NT are plants stressed lineage that was non-heat treated. Leaves were counted at 21dpg, (n=32, SE). The asterisk above bracket (*) show a significant difference compared to parental generation where * indicates $p < 0.05$, ** indicates $p < 0.01$ and *** indicates $p < 0.001$.

4.1.2 Changes in Bolting Time

The percentage of plants that bolted at 30_dpg in response to heat stress at various temperature treatments was calculated. Heat stress did result in increased percentage of plants that bolted which was observed in both progeny of stressed and non-stressed plants. Bolting time varied across generations and at the different temperatures studied, NS-T plants of parental G_2 had bolting at 37°C and 42°C but not at 50°C (Figure 5-7). Parental treatment i.e. multigenerational effect also impacted the bolting rate of plants under normal growth and

stressed conditions as there were differences in bolting based on if plants were from NS or S group. NS-NT and S-NT plants had less bolting rates than NS-T and S-T plants.

At 37°C, progeny generations G_2, G_{20} and G_{25} had increased bolting rates in plants that were progeny of S and had under gone multi-generational heat stress (S-T) when compared to S-NT which had reduced percentage of bolted plants in G_{10} and G_{20} but none in G_{25} at 30_dpg (Figure 5). Interestingly, plants within progeny generations that were progeny of heat stressed plants showed changes in bolting rate when compared to progeny of non-stressed with S-T plants having the most average of bolted plants**.** However, in G_{25} heat stressed plants showed bolting at 37°C while the plants that were not stressed had no bolting at 30_dpg.

At 42°C, G_{25} had an extensive response in bolting percentages (Figure 6); all treatment groups showed a bolting response within all the tested plant groups of G_{25}. Parental G_2 and progeny G_{10} and G_{20} had similar response except for the bolting response of NS-NT experimental group which was noticeably absent. Heat stressed plants had higher percentages of bolting in each population and this change was overall less substantial in plants that were grown under normal conditions irrespective of their lineage.

At 50°C, the impact of heat stress on progeny generation ten and twenty showed larger percentages in bolting when compared to parental G_2 which showed no impact. Advanced G_{25} showed similar effects of heat stress in bolting rate but had the lowest bolting percentage between progeny generation (Figure 7).

In summary, the response of bolting in tested plant groups were not consistent across generations, but each progeny displayed response based on treatment and temperature. These observations indicate generation twenty-five showed extensive bolting response at 42°C and

50°C than at 37°C as all experimental groups within G_{25} had bolting at different rates and were represented at both temperatures.

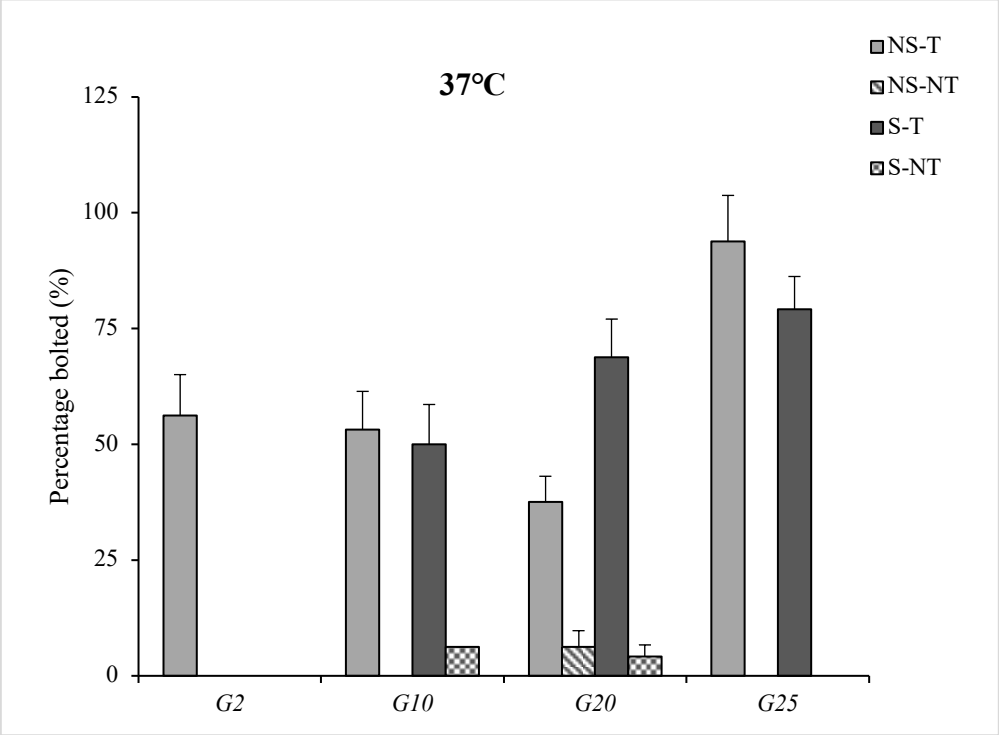

Figure 5. Percentage of plants bolted at 37°C. Plants were grown under normal and stressed conditions. Treatment groups within generations are labelled according to ancestral stress exposure and experiment treatment. Percentage averages were calculated for approximately 32 plants for NS and 48 plants for S, error bars represent SE.

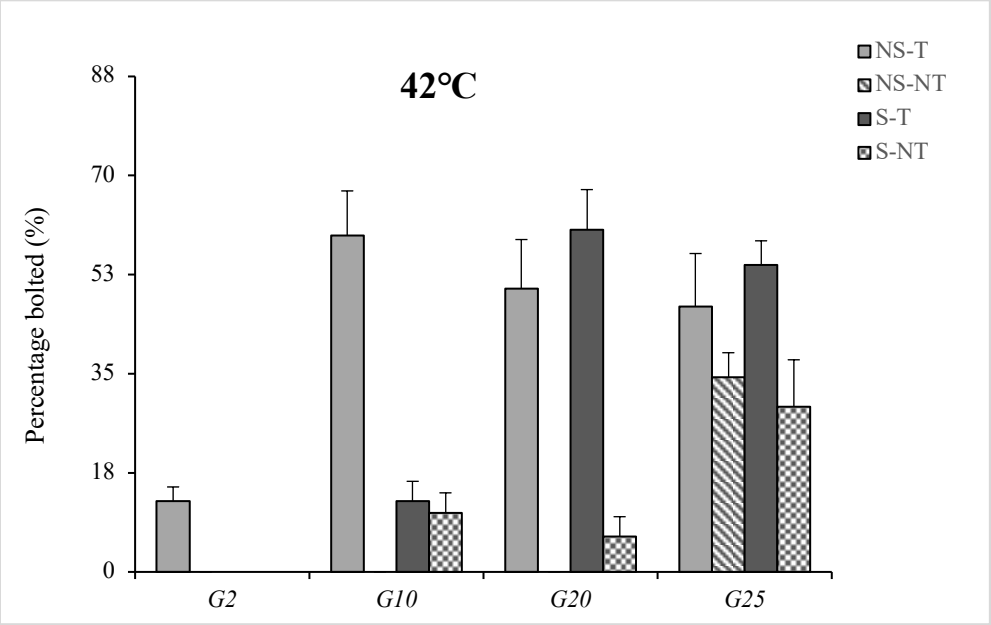

Figure 6. Percentage of plants bolted at 42°C. Plants were grown under normal and stressed conditions. Treatment groups within generations are labelled according to ancestral stress exposure and experiment treatment. Percentage averages were calculated for approximately 32 plants for NS and 48 plants for S, error bars represent SE.

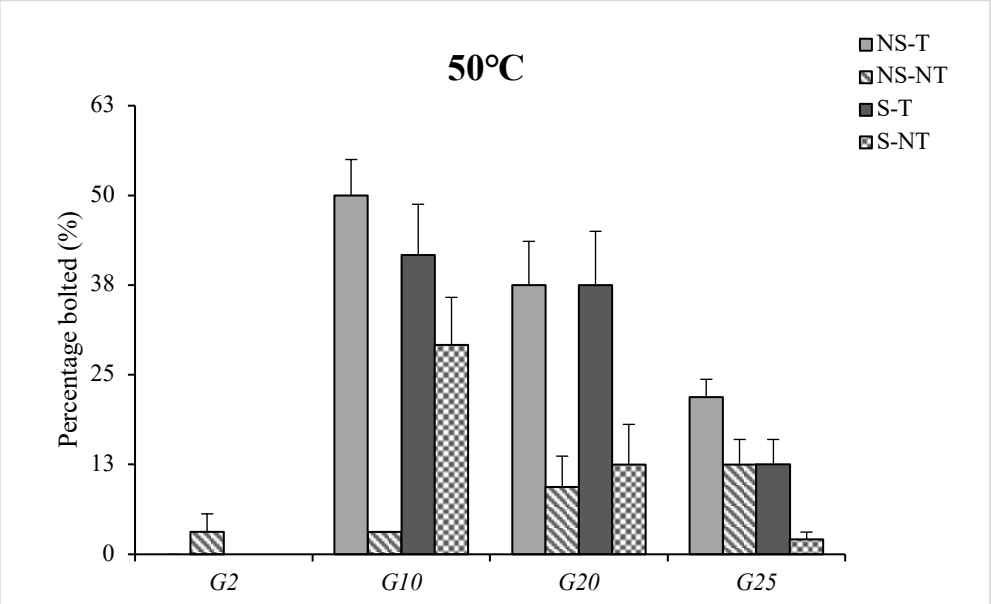

Figure 7. Percentage of plants bolted at 37°C. Plants were grown under normal and stressed conditions. Treatment groups within generations are labelled according to ancestral stress exposure and experiment treatment. Percentage averages were calculated for approximately 32 plants for NS and 48 plants for S, error bars represent SE.

4.1.3 Changes in Fresh and Dry Weight

The fresh and dry weight of mature plants were measured for all generations. There were variations between progeny and parental generations as well as within progeny generations. There were significant variations in plant weight ($p<0.05$) when progeny plants were compared to parental plants. Also, $G_{25,}$ G_{20} and G_{10} plants showed significant difference ($p<0.05$) based on temperature treatments, when weight of plants (fresh and dry) within each progeny generation were compared to parental G_2 plants, G_{25} plants showed the most change in weight of plants with significant difference ($p<0.05$) when all temperature treatments are considered.

At 37°C, within progeny generations, fresh weight of S-NT plant group was significantly different from S-T ($p<0.001$, Figure 8). There was no significant difference when progeny generations and parental generation were compared or within progeny generations at this temperature. For dry weight of plants measured at 37°C, there was a difference ($p<0.05$) within progeny generations (Figure 11).

At 42°C temperature, the fresh weight of plants was significantly different ($p<0.01$, Figure 9) between progeny plants and parental plants. The average fresh weight of G_{25} was significantly higher ($p<0.001$) from G_2 plants. Each generation had variations in plant dry weight, a comparison of stressed and non-stressed plants across all generations showed a significant difference with $p<0.05$ (Figure 12).

At 50°C, the fresh weight of G_{25} plants was significantly ($p<0.001$) higher in weight when compared to parental G_2 and G_{10} plants (Figure 10). G_{20} plants were also significantly different ($p<0.05$) in fresh weight from G_2 and G_{10} plants. There was significant difference

($p<0.01$) in dry weight only within progeny generation G_{10} when compared to parental G_2 (Figure 13), G_{20} and G_{25} revealed no significant difference in comparison with parental G_2.

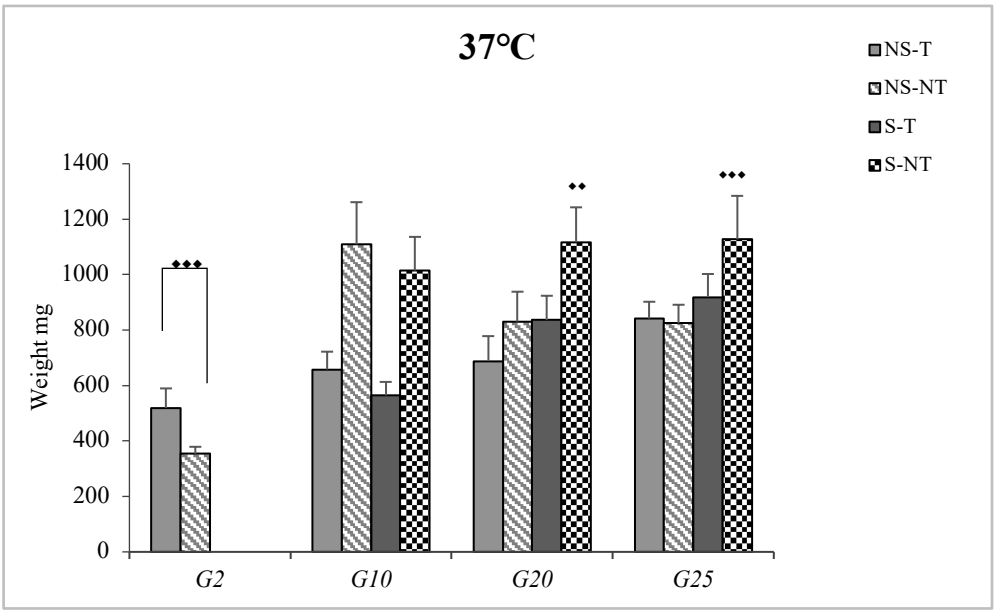

Figure 8. Fresh weight of mature plants at 37°C. Plants were grown under normal and stressed conditions. Treatment groups within generations are labelled according to ancestral stress exposure and experiment treatment. Averages were calculated for 16 plants for NS and 24 plants for S, error bars represent SE. The diamond (◆) symbol indicates the significant difference in plant treatment (T and NT) within generations where ◆ indicates $p < 0.05$, ◆◆ indicates $p < 0.01$ and ◆◆◆ indicates $p < 0.001$ (Tukey-Kramer test).

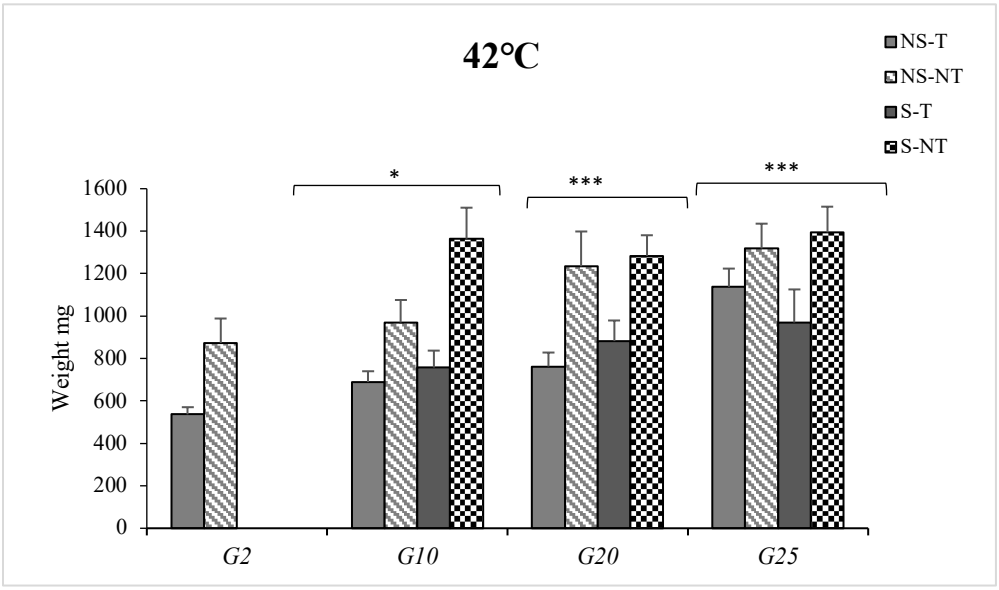

Figure 9. Fresh weight of mature plants at 42°C. Plants were grown under normal and stressed conditions. Treatment groups within generations are labelled according to ancestral stress exposure and experiment treatment. Averages were calculated for 16 plants for NS and 24 plants for S, error bars represent SE. The asterisk above bracket (*) show a significant difference compared to parental generation where * indicates $p < 0.05$, ** indicates $p < 0.01$ and *** indicates $p < 0.001$(Tukey-Kramer test).

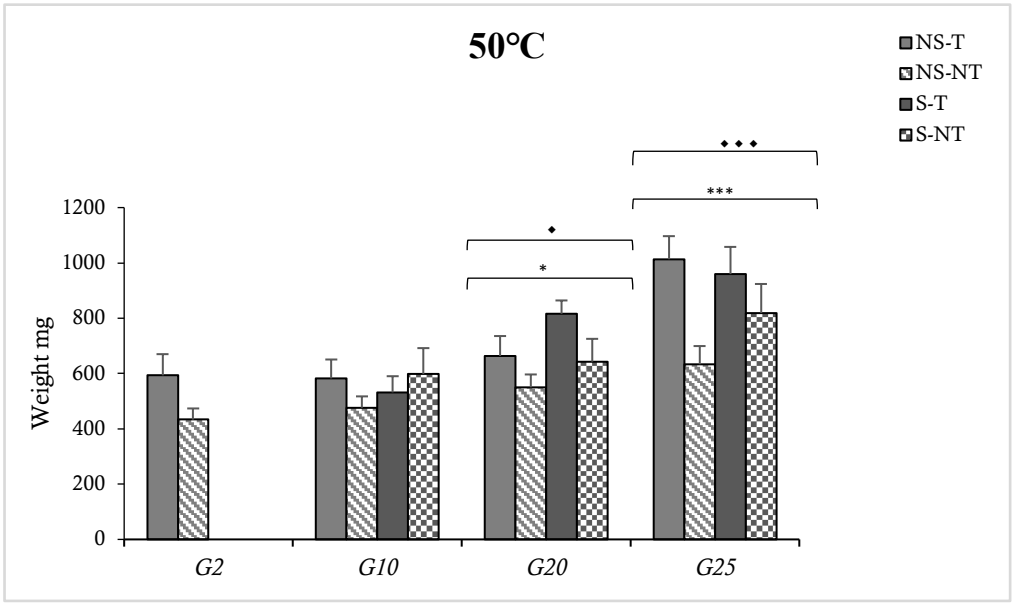

Figure 10. Fresh weight of mature plants at 50°C. Plants were grown under normal and stressed conditions. Treatment groups within generations are labelled according to ancestral stress exposure and experiment treatment. Averages were calculated for 16 plants for NS and 24 plants for S, error bars represent SE. The asterisk above bracket (*) show a significant difference compared to parental generation where ** indicates $p < 0.01$ *** indicates $p < 0.001$. The diamond (◆) symbol indicates the significant difference between progeny generations where ◆◆ indicates $p < 0.01$ ◆◆◆ indicates $p < 0.001$ (Tukey-Kramer test).

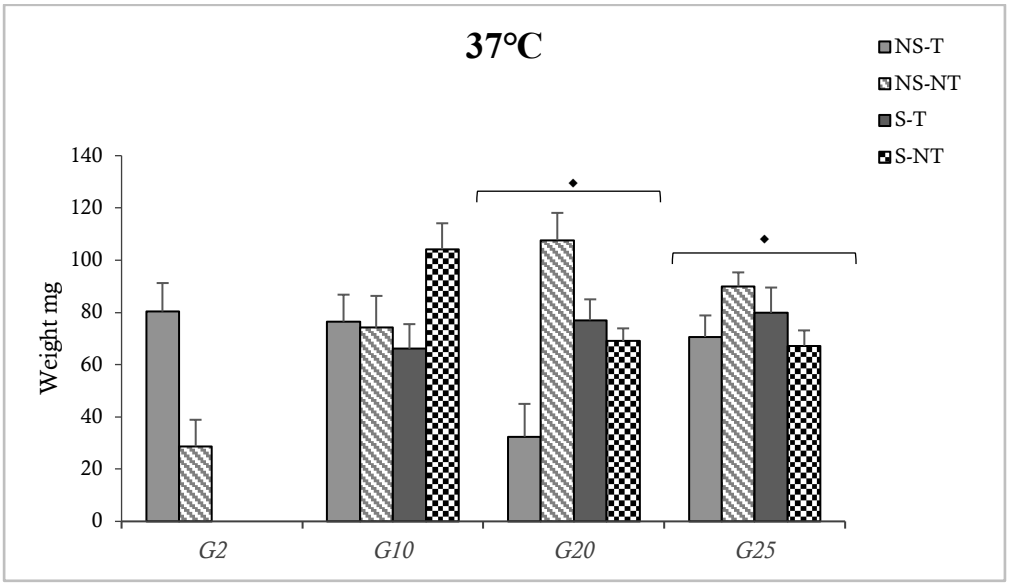

Figure 11. Dry weight of mature plants at 37°C. Plants were grown under normal and stressed conditions. Treatment groups within generations are labelled according to ancestral stress exposure and experiment treatment. Averages were calculated for 16 plants for NS and 24 plants for S, error bars represent SE. The diamond (◆) symbol indicates the significant difference between progeny generations where ◆ indicates $p < 0.05$ (Tukey-Kramer test).

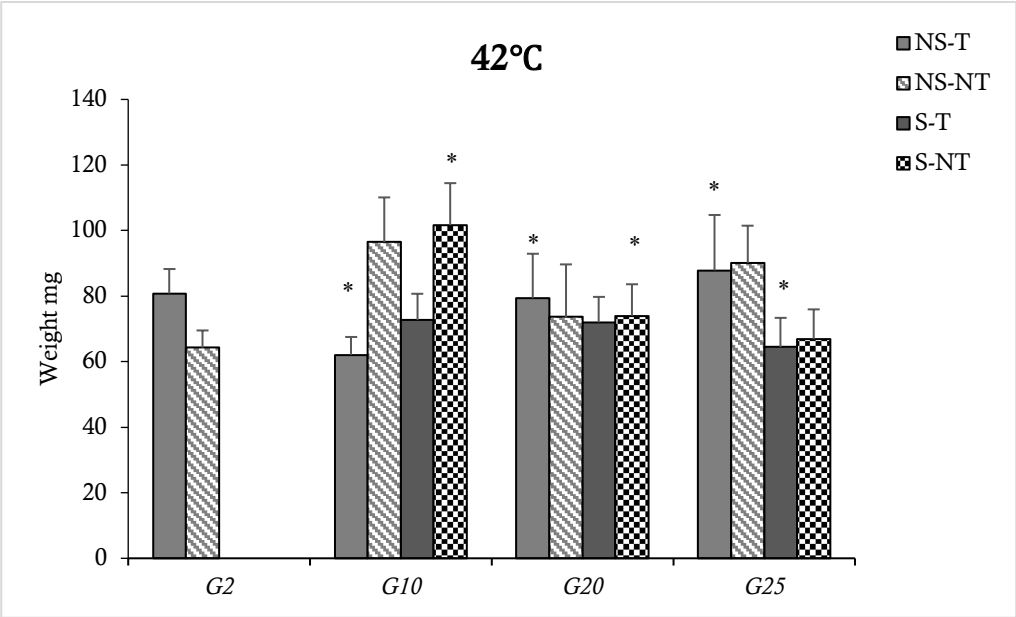

Figure 12. Dry weight of mature plants at 42°C.Plants were grown under normal and stressed conditions. Treatment groups within generations are labelled according to ancestral stress exposure and experiment treatment. Averages were calculated for 16 plants for NS and 24 plants for S, error bars represent SE. The asterisk above bracket (*) show a significant difference between stressed and non-stressed treatment where ** indicates $p < 0.01$ *** indicates $p < 0.001$ (Tukey-Kramer test).

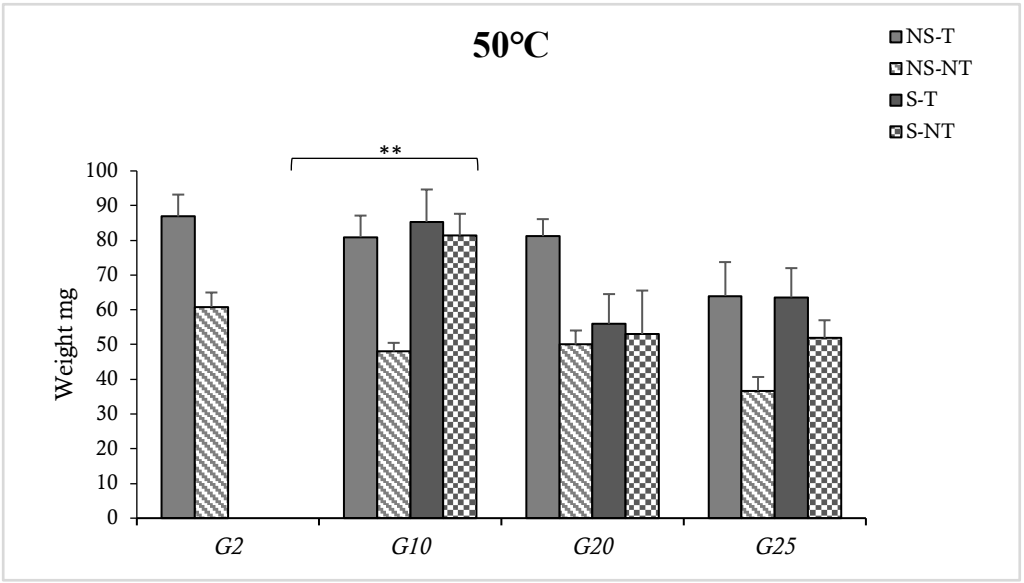

Figure 13. Dry weight of mature plants at 50°C.Plants were grown under normal and stressed conditions. Treatment groups within generations are labelled according to ancestral stress exposure and experiment treatment. Averages were calculated for 16 plants for NS and 24 plants for S, error bars represent SE. The asterisk above bracket (*) show a significant difference compared to parental generation where ** indicates $p < 0.01$ (Tukey-Kramer test).

4.1.4 Changes in Seed Length

Seeds were collected from both stressed and non-stressed plant groups, each seed length was measured and compared from twenty randomly selected seeds to represent population of sampled generation. Variations in the size of seeds produced by parental generations varied mostly based on longest temperature treatment (37°C). There were variations of seed size within progeny generations based on whether the plant was stressed or non-stressed (Figure 14-16)

At 37°C, progeny generations $G_{25,}$ G_{20} and G_{10} showed significant difference ($p<0.001$) in seed length when compared to parental G_2 (Figure 14). There was no significant difference within progeny generations, G_{10} and G_{20} were similar in average seed length. Generally, plants that were heat stressed (S-T) produced seeds that were significantly smaller ($p<0.01$) in length than seeds produced from plants grown under normal conditions (NS-T). There was also significant difference ($p< 0.01$) when each generation and their seed lengths were considered and compared based on heat treatment of each group, especially NS-NT seeds of G_{25} which were smaller than S-NT of G_{25} or NS-T seeds of G_{10} which were larger than S-T seeds G_{10}.

At 42°C, there was no significant difference between G_{25} and G_2 seed lengths, and the other progeny generations G_{20} and G_{10} showed similar trend when compared to G_2. However, there was significant variations ($p<0.001$) between G_{25} and G_{10} (Figure 15). The seed length from non-stressed and heat stressed plant groups did not highlight a difference based on group treatment in seed length either. As within each generation average seed length produced from NS-T and S-T plants were similar to those produced by their corresponding NS-NT and S-NT plant groups.

At 50°C the significant difference ($p<0.01$) was within progeny generation G_{25} and G_{10}, while significant difference ($p<0.05$) was between G_{20} and G_{10}, comparison of progeny generations with parental generation G_2 showed no significant variation (Figure 16).

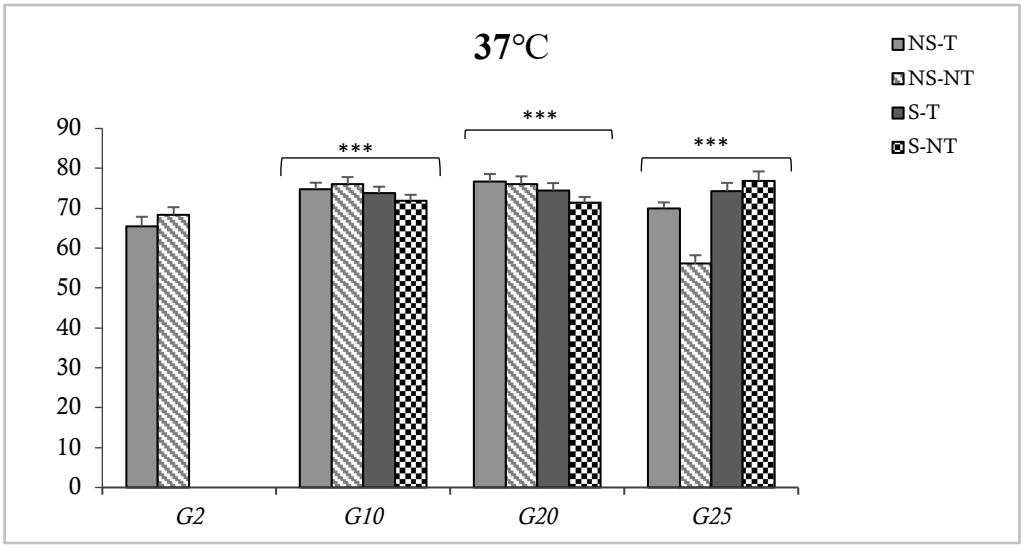

Figure 14. Length of seeds produced by parental and progeny generations at 37°C. Plants were grown under normal and stressed conditions. Approximately 20 seeds were measured for each group. Averages were calculated for 40 seeds for NS and 80 seeds for S, error bars represent SE. The asterisk above bracket (*) show a significant difference compared to parental generation where *** indicates $p < 0.001$ (Tukey-Kramer test)

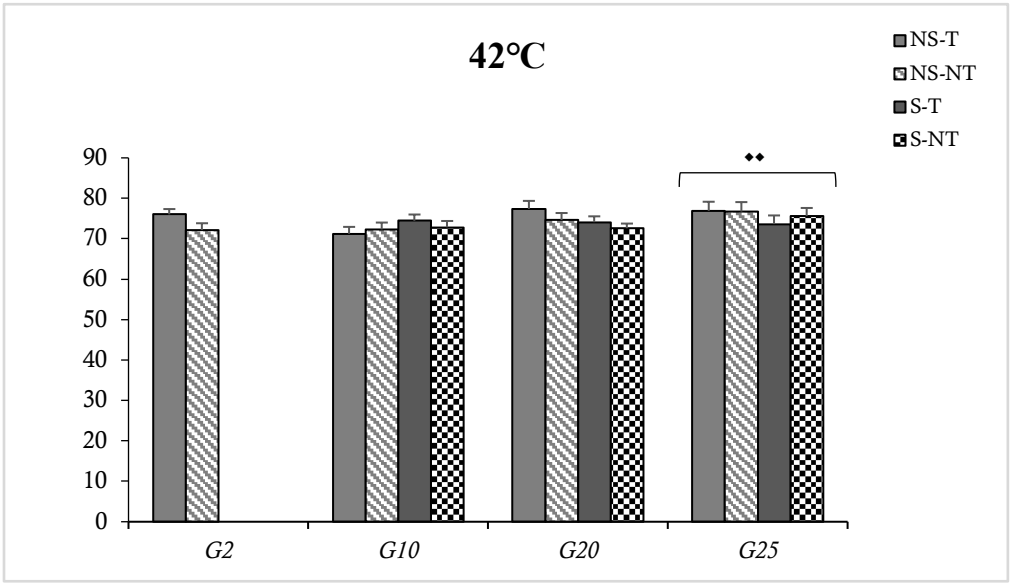

Figure 15. Length of seeds produced by parental and progeny generations at 42°C. Plants were grown under normal and stressed conditions. Approximately 20 seeds were measured for each group. Averages were calculated for 40 seeds for NS and 80 seeds for S, error bars represent SE. The diamond above bracket (◆) symbol indicates the significant difference between progeny generations where ◆◆ indicates $p < 0.01$ (Tukey-Kramer test).

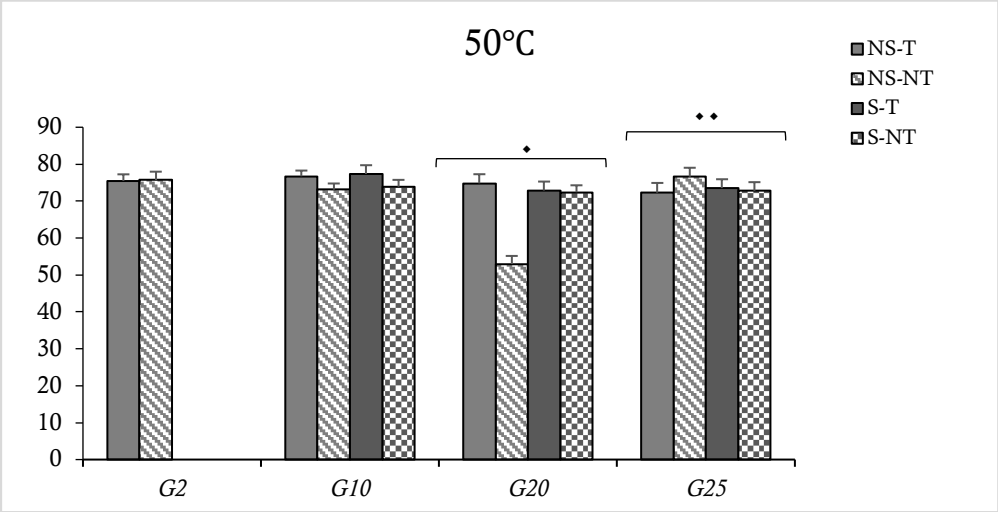

Figure 16. Length of seeds produced by parental and progeny generations at 50°C. Plants were grown under normal and stressed conditions. Approximately 20 seeds were measured for each group. Averages were calculated for 40 seeds for NS and 80 seeds for S, error bars represent SE. The diamond (◆) symbol indicates the significant difference between progeny generations where ◆ indicates $p < 0.05$ (Tukey-Kramer test).

4.2 Genetic and Epigenetic Changes in G_2 and G_{25}

Genetic changes in independently soil-grown plants of G_2-NS (designated here as F2C), G_{25}-S (designated here as F25H) and G_{25}-NS (designated here as F25C) were investigated by whole genome sequencing. Epigenetic changes were examined as the extent of DNA methylation patterns using high-throughput bisulfite sequencing of the whole genome. The distribution of epimutations induced by multi-generational heat stress investigated in this study included differentially methylated cytosines (DMCs) and differentially methylated regions (DMRs). In total fifteen methylomes were analyzed which included five biological replicates (sibling plants) of F2C, F25C, and F25H each.

4.2.1 Global DNA Methylation and Changes in G_2 and G_{25}

Sequencing data mapped into reference *Arabidopsis* genome revealed an average of 16,081,856, 16,095,139 and 15,869,955 methylated cytosine (mC) residues for F2C, F25C, and F25H respectively, representing all three-sequence context. Total global methylation levels with minimum coverage ≥10 extracted were similar for G_2 and G_{25} (Figure 17). The proportions of methylation percentage in the three sequence contexts (CG, CHG, and CHH) were 31%, 35%, and 34% respectively in F25H, F25C and F2C, regardless of progeny analyzed.

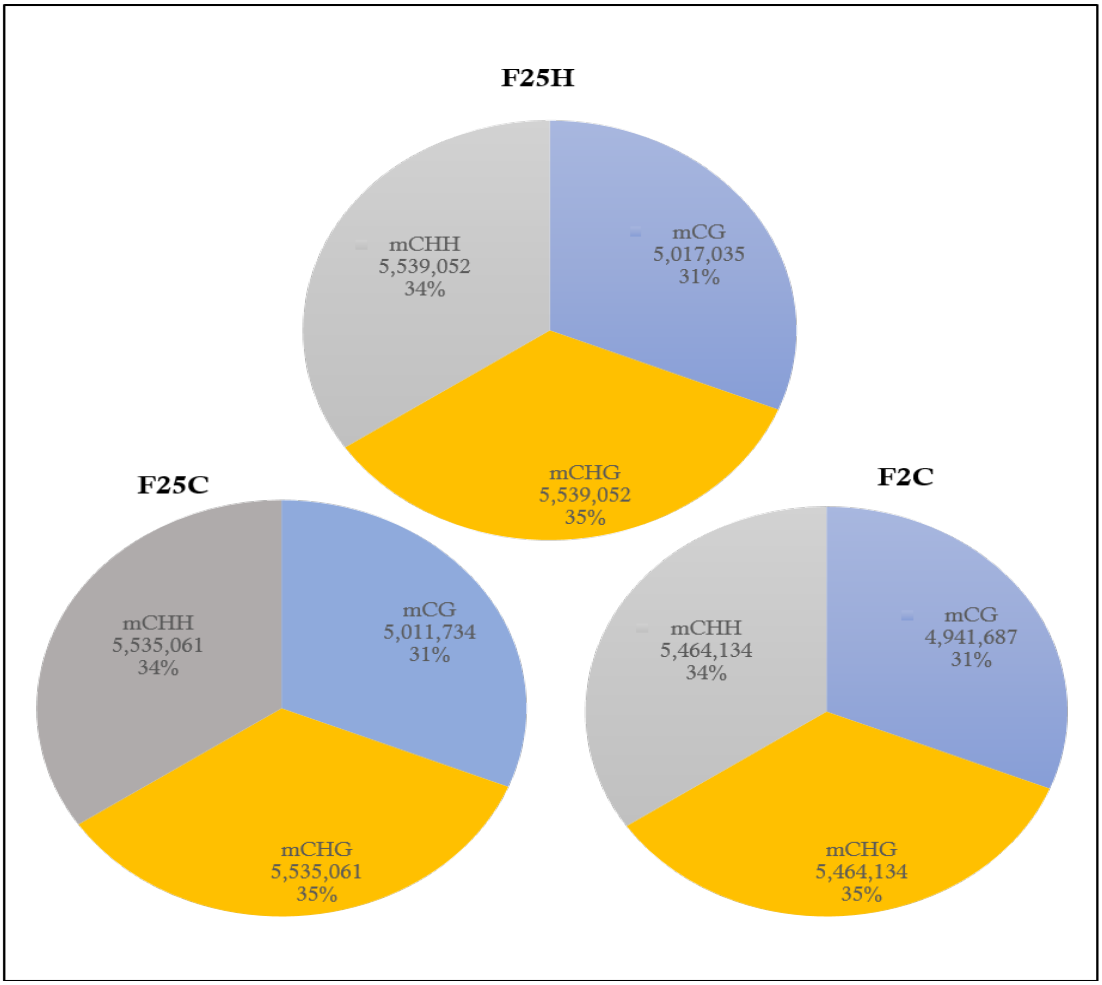

Figure 17. The average number of methylated cytosine (bases) identified with corresponding percentages in F25H, F25C and F2C for each sequence context (H=A, T, C and mean, n=5). Methylation levels were from reads with minimum coverage ≥ 10 mapped to TAIR 10 reference and called with Bismark.

The frequency distribution of methylation percentages (%methylation) in G_2 and G_{25} samples showed most bases to have either high or low global methylation levels mostly in methylated CG (mCG) and methylated CHG (mCHG) sequences. When each sequence context was considered, there was no difference between groups. F25H, F25C and F2C showed similar distribution of methylation percentages (Figure 18-20). Cytosine methylation levels per base showed higher frequencies distributed mostly at mCG (Figure 18) and mCHG sites (Figure 19) than methylated CHH (mCHH) sites (Figure 20). Cytosine methylation in CHG context was lower in frequency when compared to CG context with most bases having higher %methylation distributed in CG. The corresponding reads coverage show a similar pattern with lowest %methylation frequency in CHH context (Appendix 1).

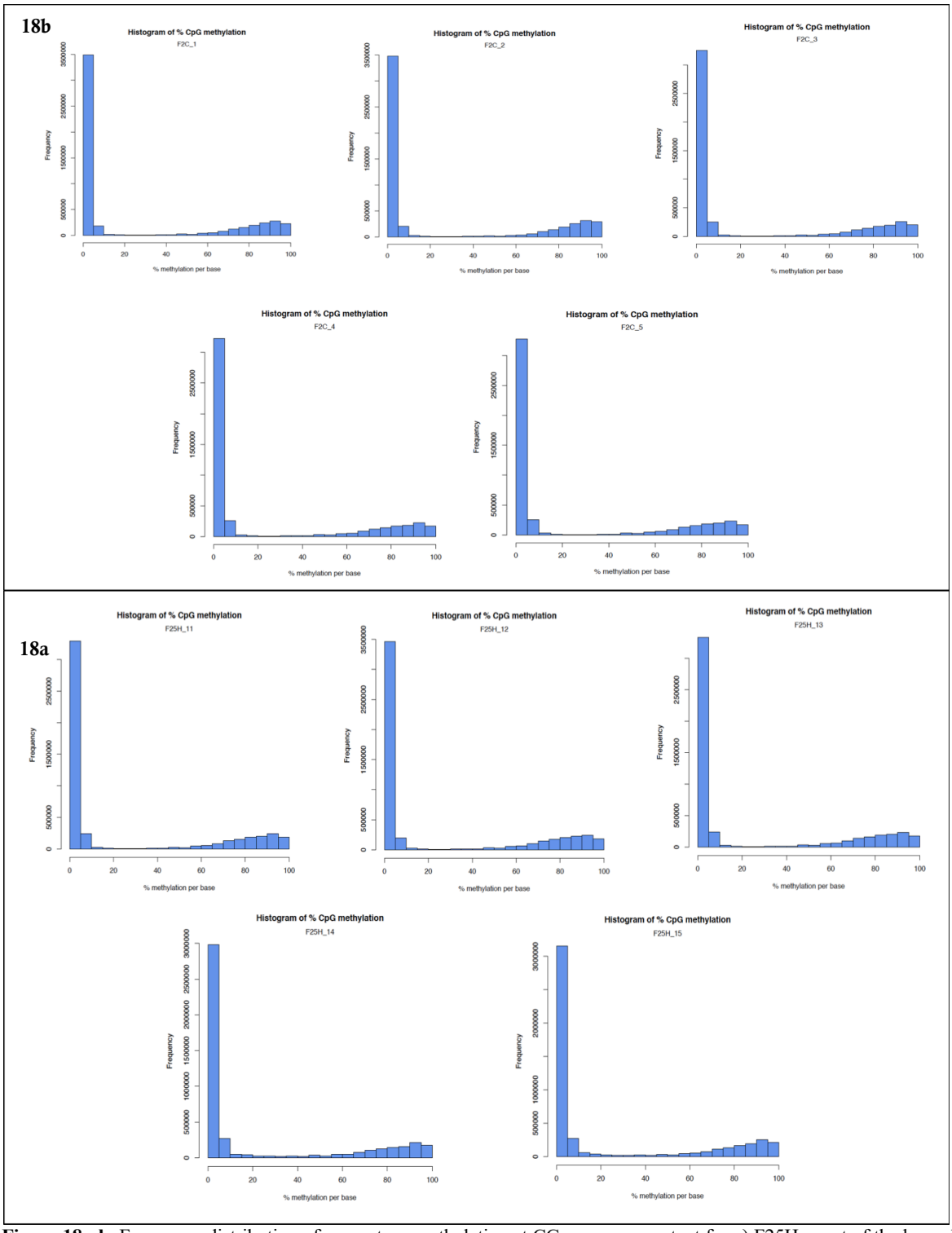

Figure 18a-b. Frequency distribution of percentage methylation at CG sequence context for a) F25H, most of the bases have either high or low methylation based on bimodal distribution. Data were normalized in *Methylkit* to account for clonal reads (PCR duplication bias), and each histogram represents a biological replicate where n=5 for F25H.

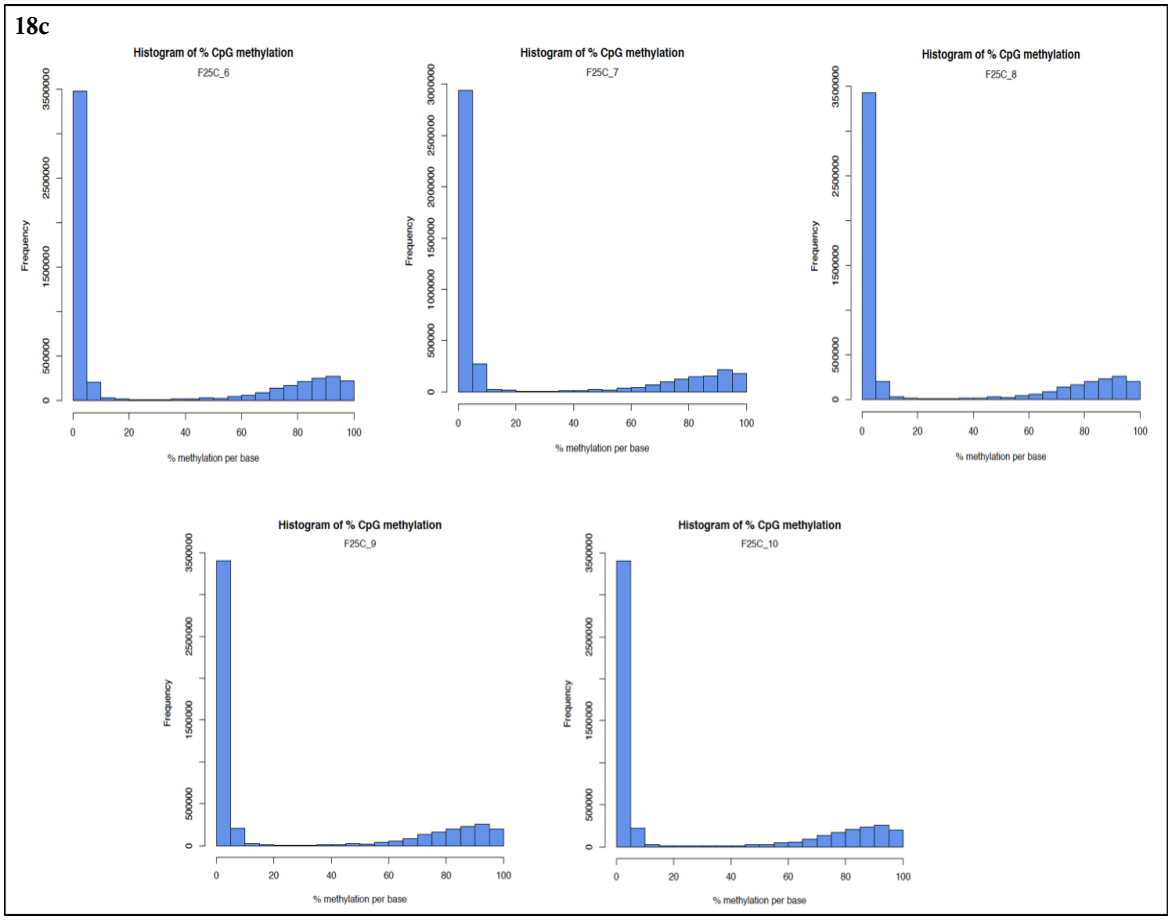

Figure 18c. Frequency distribution of percentage methylation at CG sequence context for b) F2C and c) F25C samples, most of the bases have either high or low methylation based on bimodal distribution. Data were normalized in *Methylkit* to account for clonal reads (PCR duplication bias), and each histogram represents a biological replicate, n=5 for F2C and F25H.

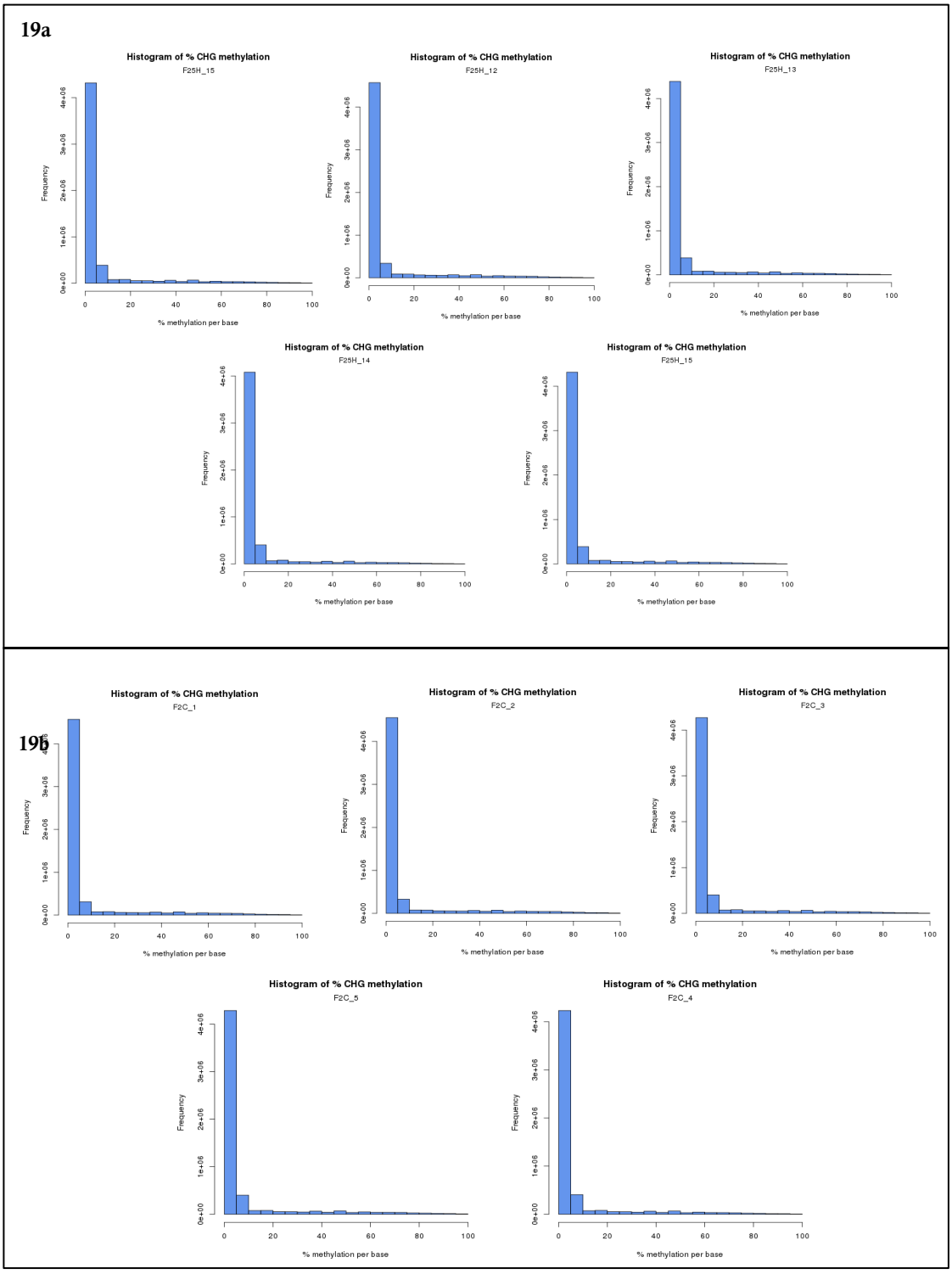

Figure 19a-b. Frequency distribution of percentage methylation at CHG sequence context for a) F25H and b) F2C samples, most of the bases have either high or low methylation based on bimodal distribution. Data were normalized in *Methylkit* to account for clonal reads (PCR duplication bias), and each histogram represent a biological replicate mean, n=5 for F25H and F2C.

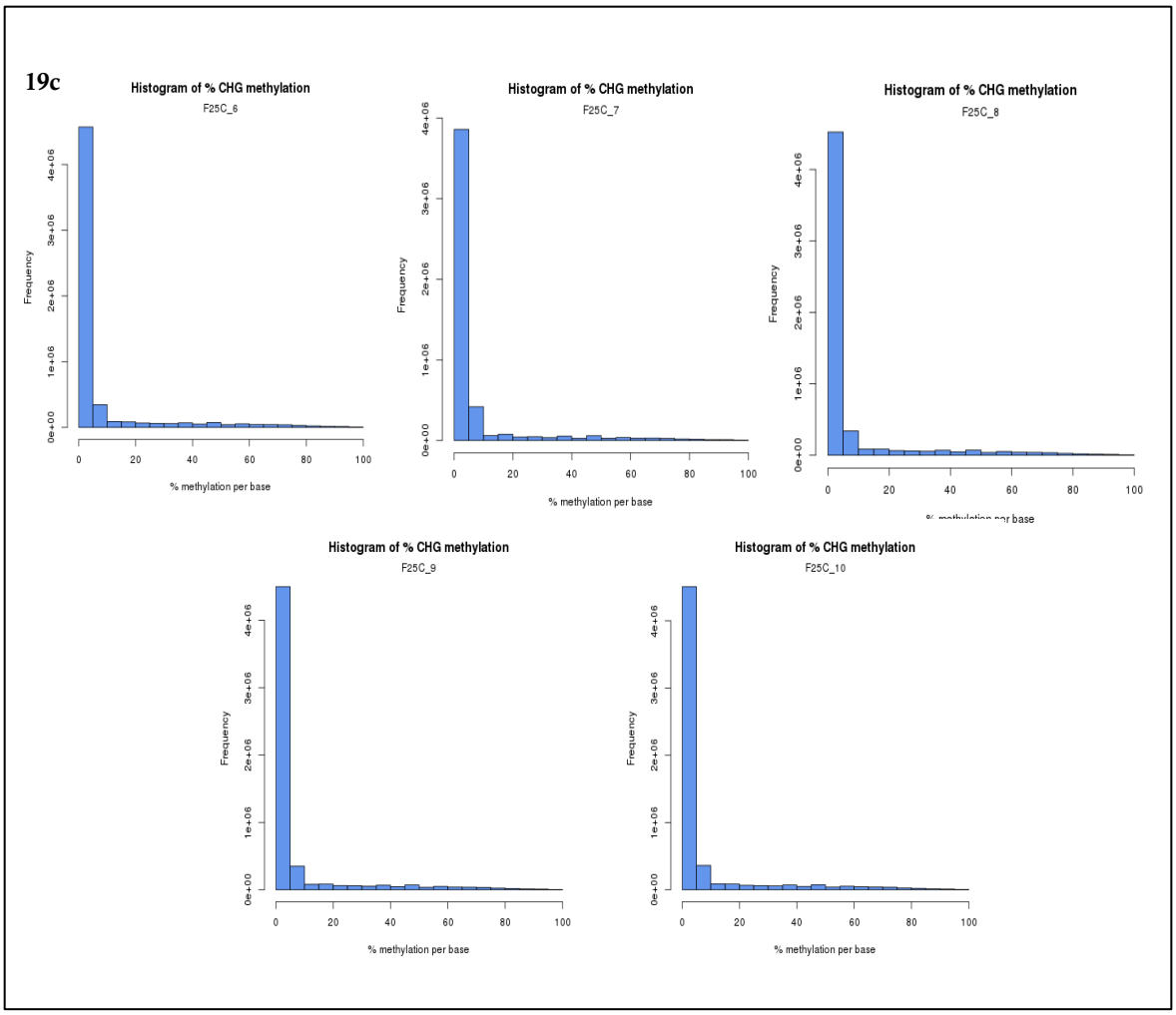

Figure 19c. Frequency distribution of percentage methylation at CHG sequence context for c) F25C samples, most of the bases have either high or low methylation based on bimodal distribution. Data were normalized in *Methylkit* to account for clonal reads (PCR duplication bias), and each histogram represents a biological replicate n=5 for F25C, (H=A, T, C).

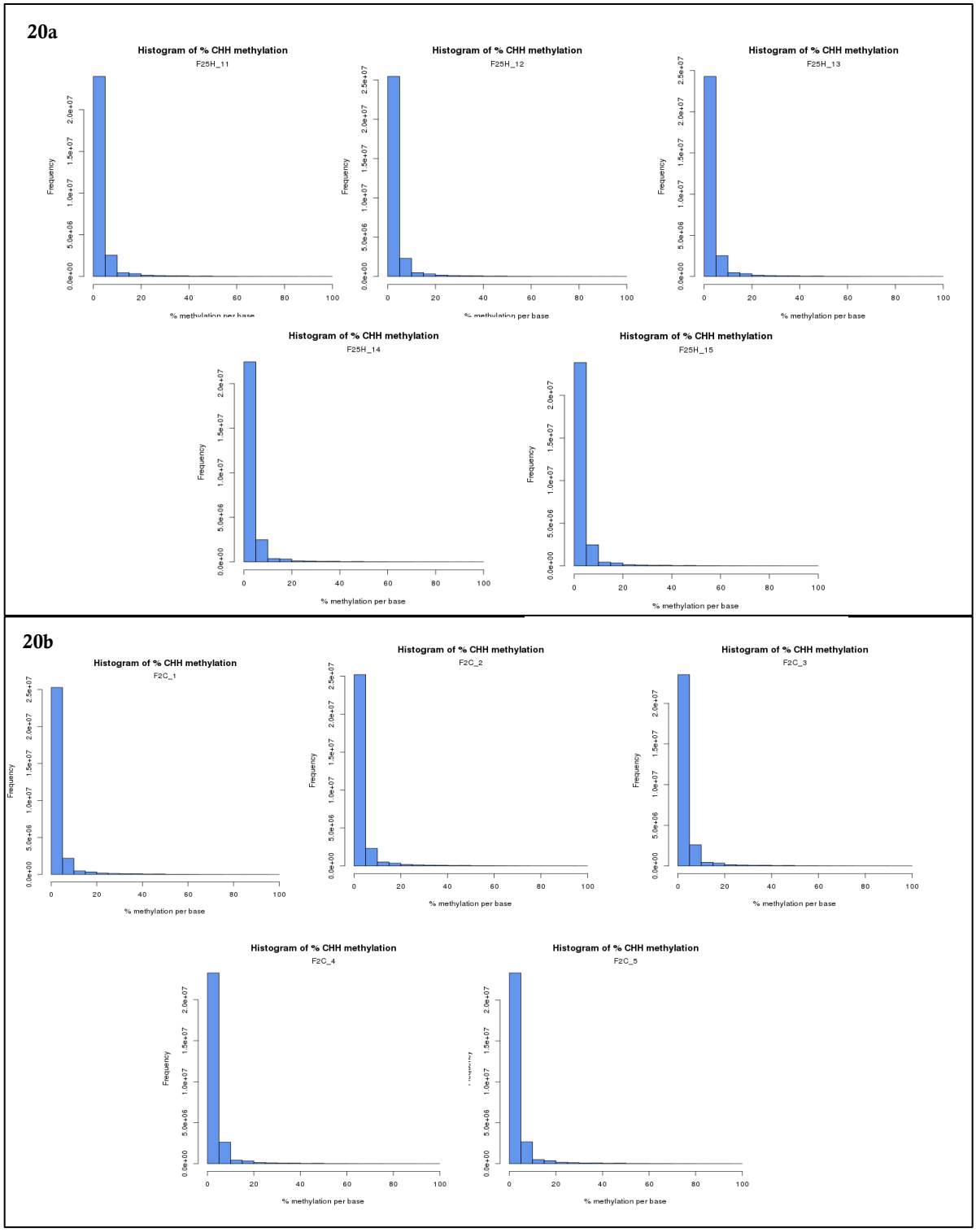

Figure 20a-b. Frequency distribution percentage methylation at CHH sequence context (H=A, T, C) for a) F25H and b) F2C samples, most of the bases have either high or low methylation based on bimodal distribution. Data were normalized in *Methylkit* to account for clonal reads (PCR duplication bias), and each histogram represents a biological replicate n=5 for F25H and F2C.

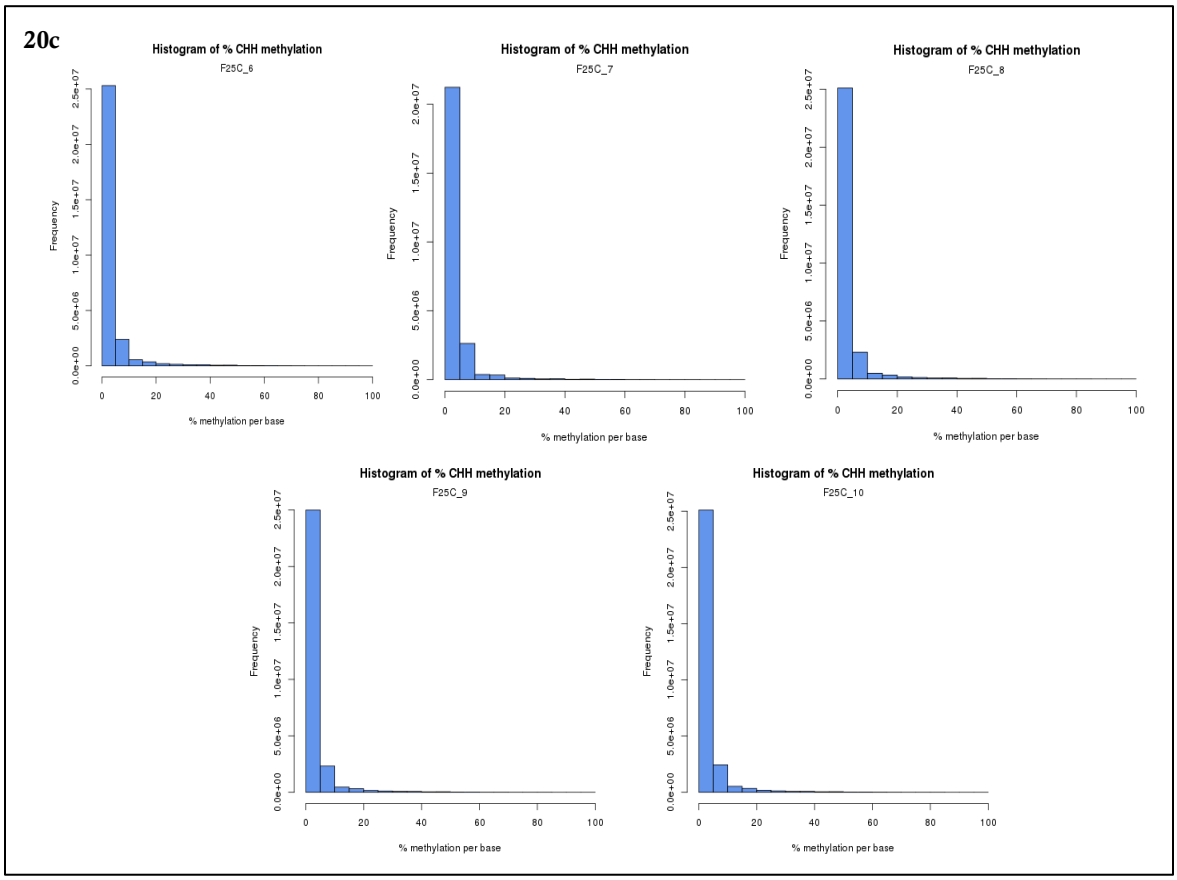

Figure 20c. Frequency distribution percentage methylation at CHH sequence context (H=A, T, C) for c) F25C samples, most of the bases have either high or low methylation based on bimodal distribution. Data were normalized in *Methylkit* to account for clonal reads (PCR duplication bias) and each histogram represents a biological replicate n=5 for F25C.

4.2.1.1 Hierarchical Clustering within Generations

In general, the DNA methylation profiles obtained for each sample group were matching to each other, similarities in DMCs at CG methylation profiles between G_2 and G_{25} samples showed most of the non- stressed advanced generation (F25C) clustered with parental generation (F2C) (Figure 21). The descendant plants segregate as an independent cluster from ancestral plants, three out of five samples of stressed twenty-fifth progeny generation (F25H) showed association with each other (Figure 21a) and the remaining two samples showed similarities within all three groups. The pairwise comparisons and Principal Component Analysis (PCA) plots derived (Figure 21b) estimated by similarities in global methylation levels on CG sites confirms similarities outlined in Figure 21a which also outlines the biological denotation and homogeneity of all tested samples. Hierarchical clustering of DMCs at CHG (Figure 22) and CHH (Figure 23) methylation profiles displayed similar clusters of sample groupings as seen in their CG profile with F25C and F2C mainly clustering together. This was also the case for DMRs at CG (Figure 24), CHG (Figure 25) and CHH (Figure 26).

In summary, hierarchical clustering of samples based on DNA methylation percentages of differential methylated cytosine (DMCs) (Figure 21a-b) and differential methylated regions (DMRs) (Figure 24a-b) at CG sites separated the stressed plants and non-stressed plants into distinct groups by their ancestral generational treatment suggesting that multi-generational heat stress impacts the DNA methylation patterns of these positions or regions in the *Arabidopsis* genome.

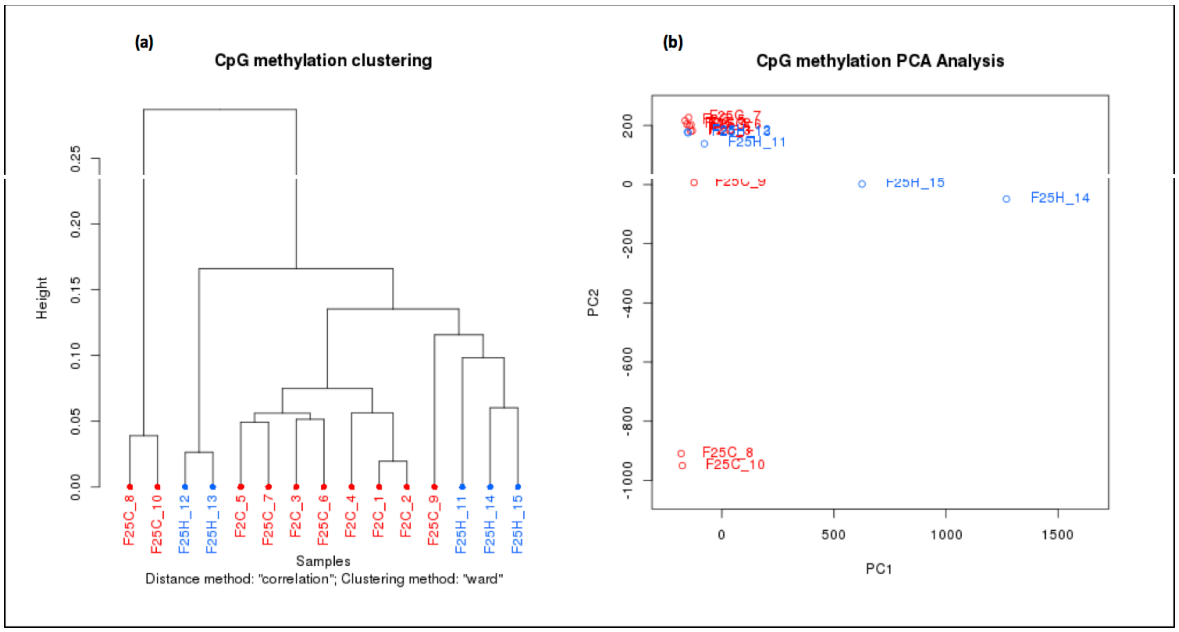

Figure 21. Global methylation clustering of DMCs at CG (a) Hierarchical clustering of all fifteen methylomes using 1-Pearson's correlation distance. (b) Principal Component Analysis (PCA) of all samples in F25H vs. F2C vs. F25C methylation profiles. The plot shows principal component 1 and principal component 2 of each sample. Samples closer to each other in principal component space are similar in methylation profiles.

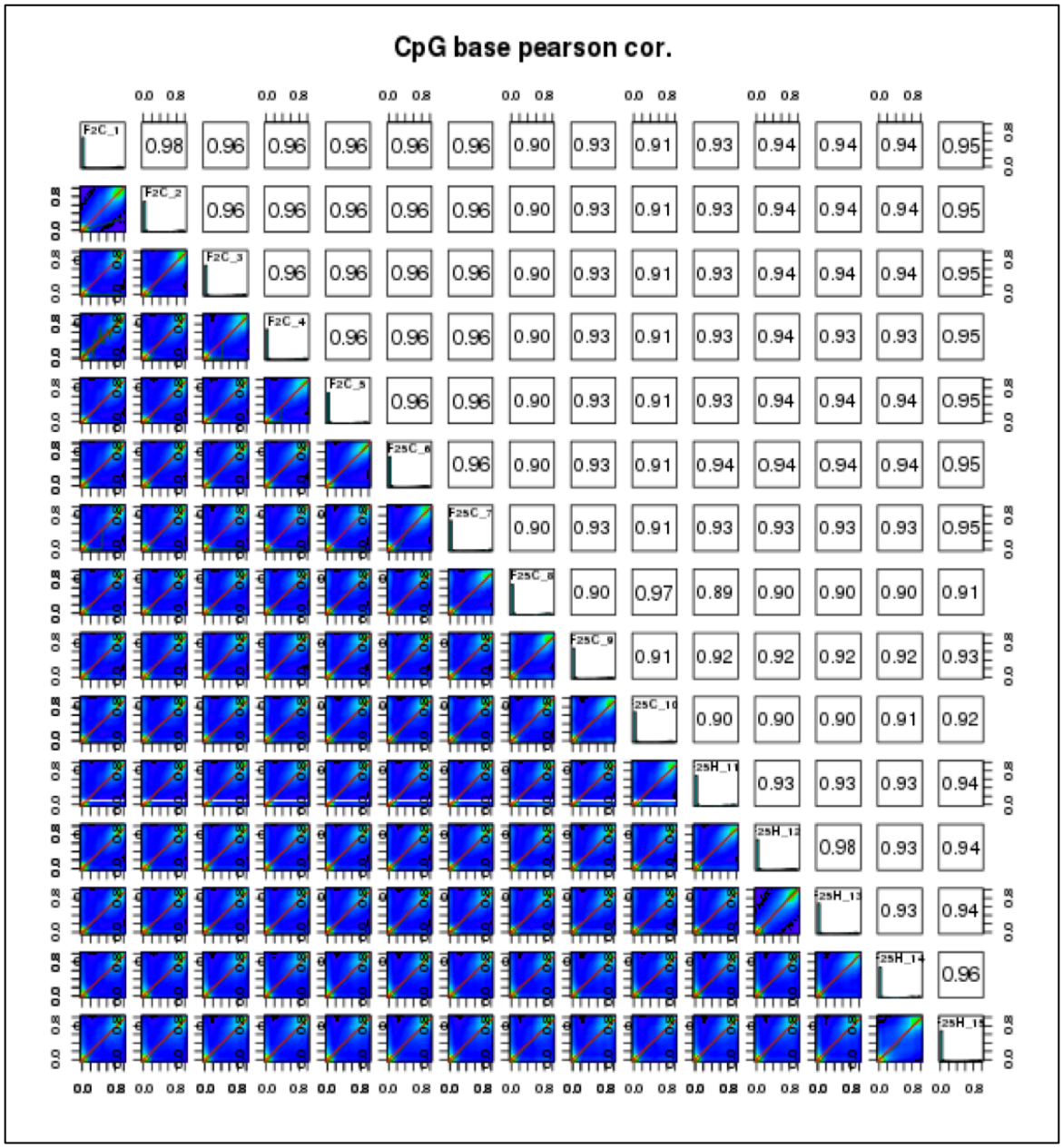

Figure 21. **(c)** Correlation matrix of scatter plots of methylation percentage values for samples in F25H, F2C, and F25C at CG bases. Numbers on upper right corner denotes pairwise Pearson's correlation scores. Histograms on the diagonal are methylation percentage histogram similar to Figure 14-16 for each sample profiles.

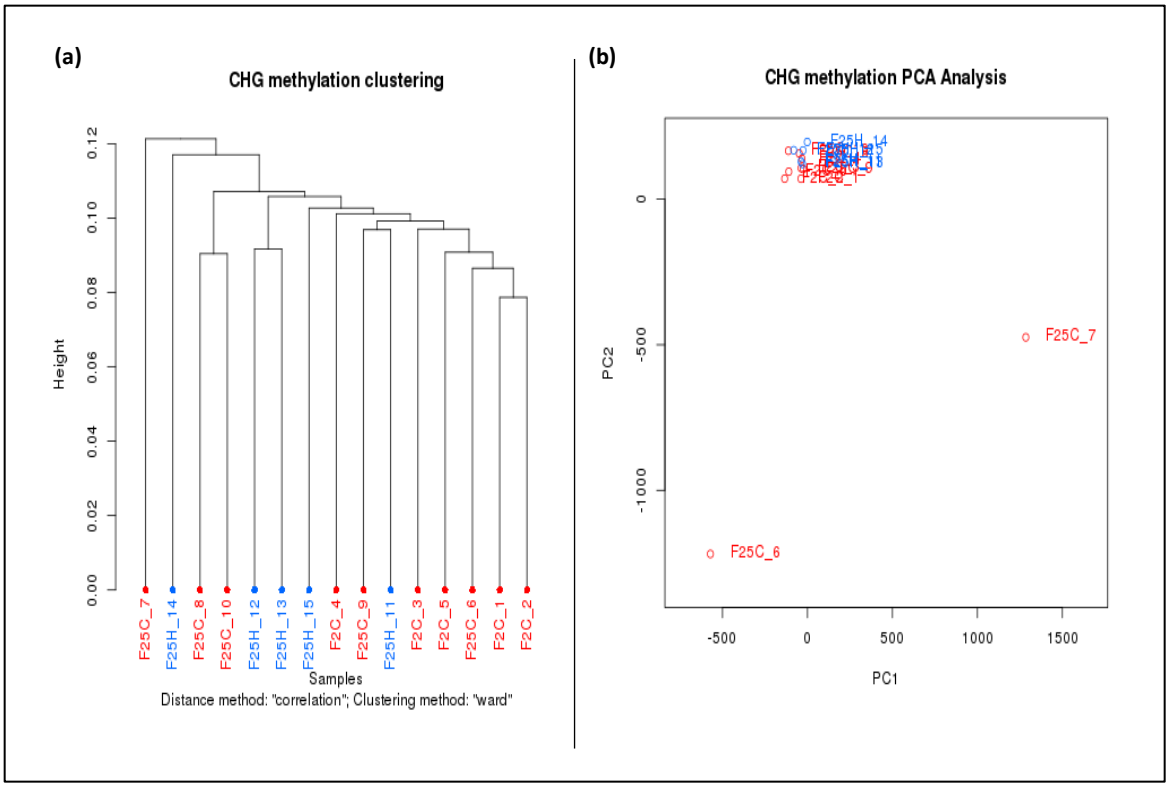

Figure 22. Global methylation clustering of DMCs at CHG (H=A, T, C) (a) Hierarchical clustering of all fifteen methylomes using 1-Pearson's correlation distance. (b) Principal Component Analysis (PCA) of all samples in F25H vs. F2C vs. F25C methylation profiles. The plot shows principal component 1 and principal component 2 of each sample. Samples closer to each other in principal component space are similar in methylation profiles.

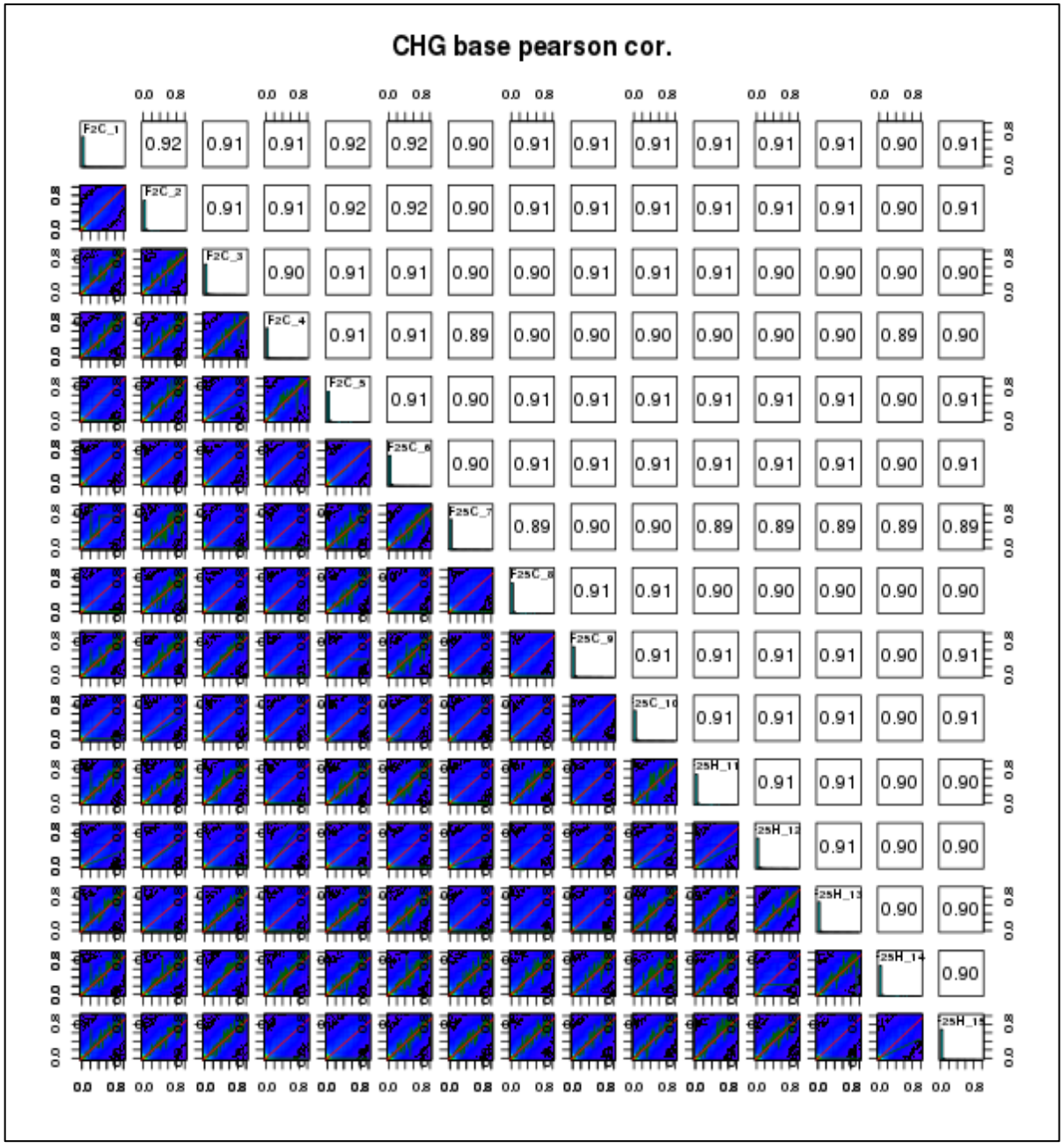

Figure 22c. Correlation matrix of scatter plots of methylation percentage values for samples in F25H, F2C, and F25C at CHG bases. Numbers on upper right corner denotes pairwise Pearson's correlation scores. Histograms on the diagonal are methylation percentage histogram similar to Figure 14-16 for each sample profiles.

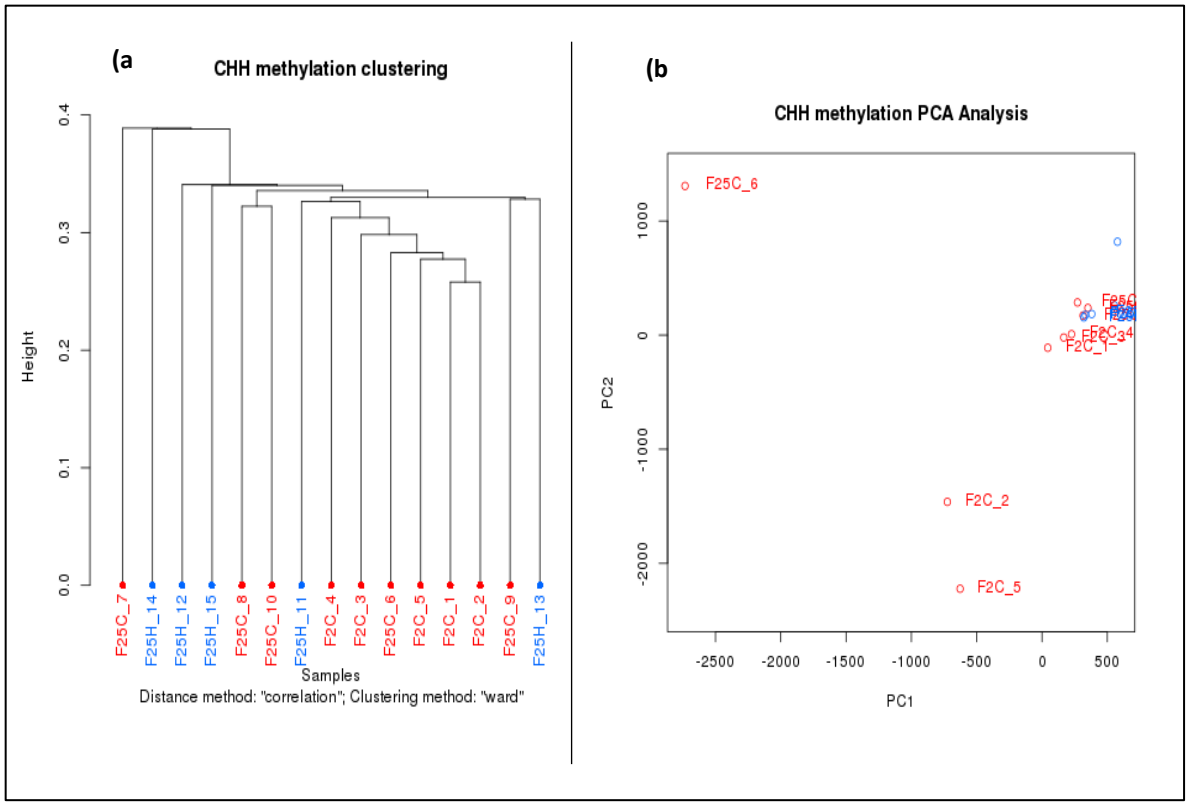

Figure 23. Global methylation clustering of DMCs at CHH (H=A, T, C) (a) Hierarchical clustering of all fifteen methylomes using 1-Pearson's correlation distance. (b) Principal Component Analysis (PCA) of all samples in F25H, F2C, and F25C methylation profiles. The plot shows principal component 1 and principal component 2 of each sample. Samples closer to each other in principal component space are similar in methylation profiles.

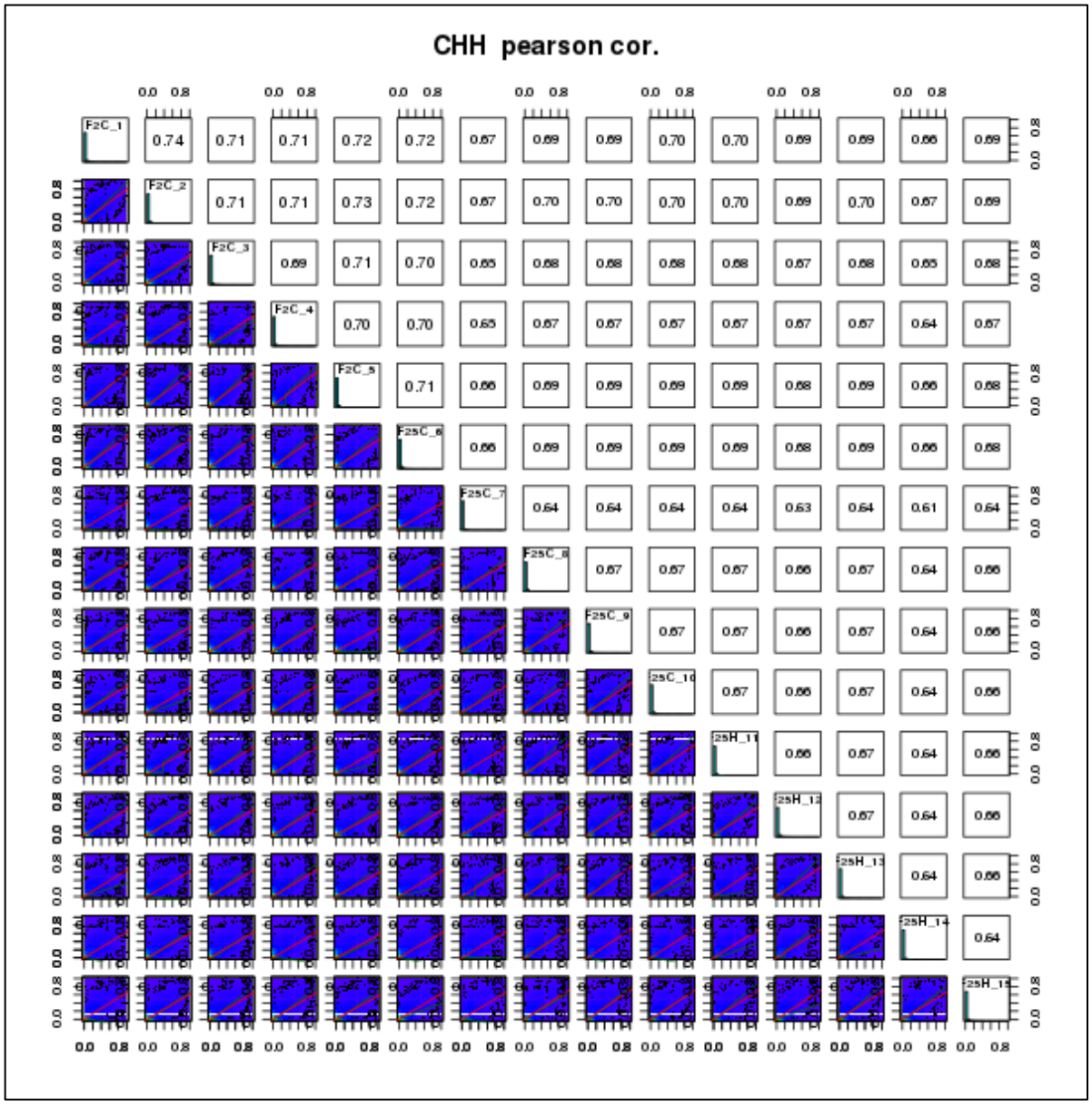

Figure 23c. Correlation matrix of scatter plots of methylation percentage values for samples in F25H, F2C, and F25C at CHH bases. Numbers on upper right corner denotes pairwise Pearson's correlation scores. Histograms on the diagonal are methylation percentage histogram similar to Figure 14-16 for each sample profiles.

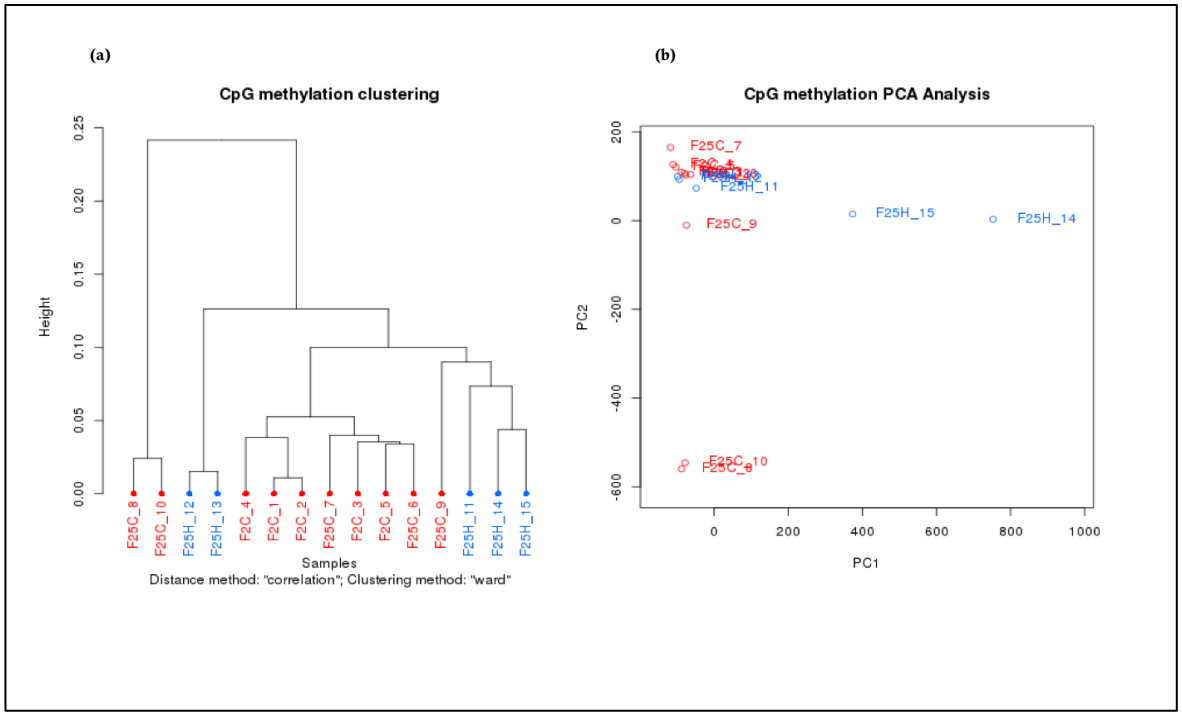

Figure 24. CpG methylation clustering of DMRs (a) Hierarchical clustering of all fifteen methylomes using 1-Pearson's correlation distance. (b) Principal Component Analysis (PCA) of all samples in F25H vs. F2C vs. F25C methylation profiles. The plot shows principal component 1 and principal component 2 of each sample. Samples closer to each other in principal component space are similar in methylation profiles.

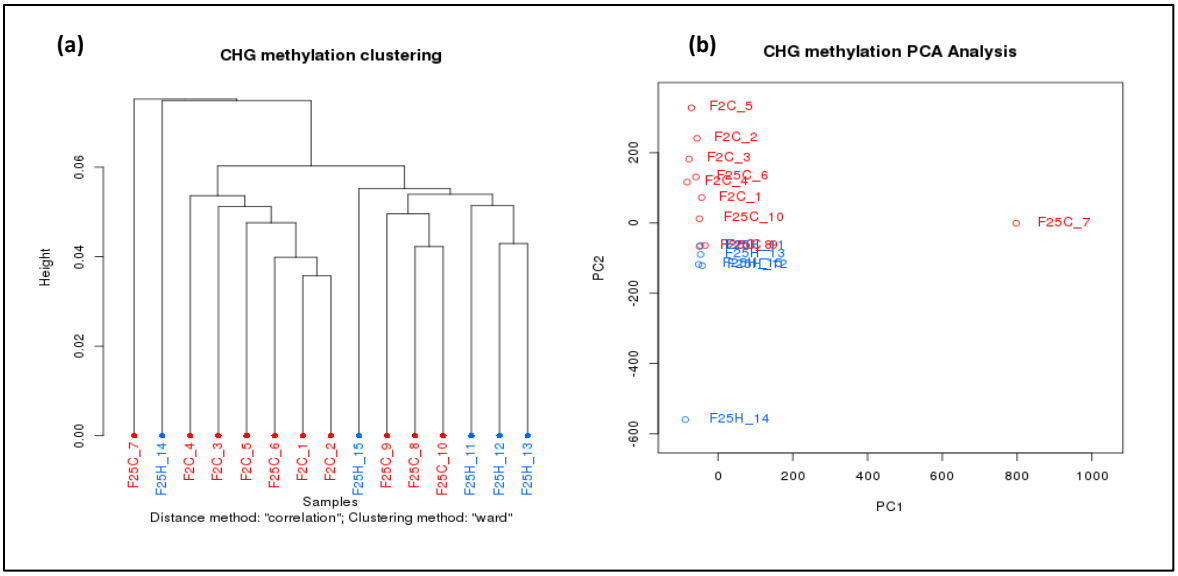

Figure 25. CHG methylation clustering of DMRs (a) Hierarchical clustering of all fifteen methylomes using 1-Pearson's correlation distance. (b) Principal Component Analysis (PCA) of all samples in F25H vs. F2C vs. F25C methylation profiles. The plot shows principal component 1 and principal component 2 of each sample. Samples closer to each other in principal component space are similar in methylation profiles.

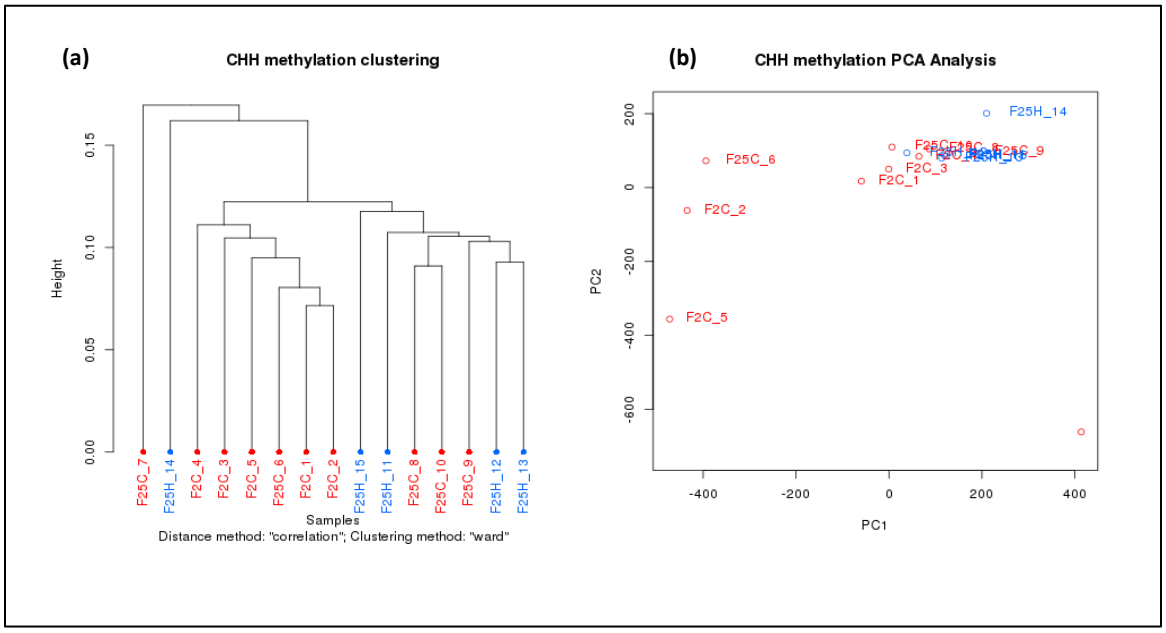

Figure 26. CHH methylation clustering of DMRs (a) Hierarchical clustering of all fifteen methylomes using 1-Pearson's correlation distance. (b) Principal Component Analysis (PCA) of all samples in F25H vs. F2C vs. F25C methylation profiles. The plot shows principal component 1 and principal component 2 of each sample. Samples closer to each other in principal component space are similar in methylation profiles.

Direct comparison of global CG methylation profiles of individual samples of F2C and F25H groups show similaries in each biological replicates as well as an increase in cluster height indicating a difference between the two generations G_2 and G_{25} (Figure 27).

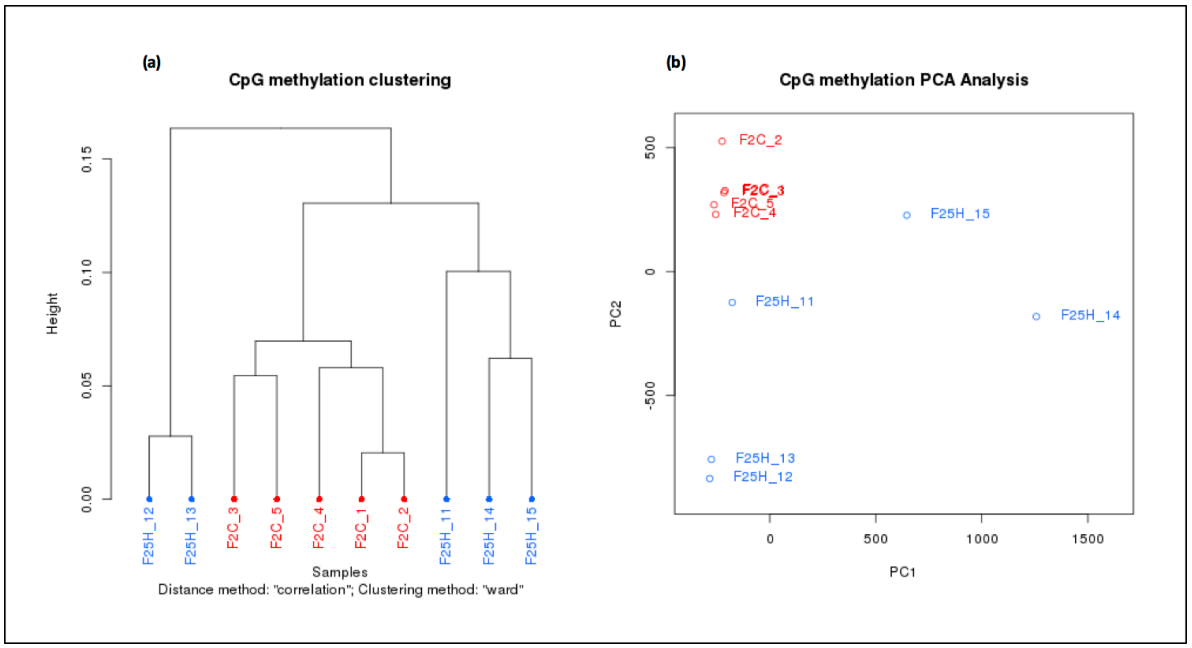

Figure 27. Global CG methylation clustering (a) Hierarchical clustering of parental and twenty-fifth progeny methylomes using 1-Pearson's correlation distance. (b) Principal Component Analysis (PCA) comparing F25H and F2C methylation profiles. The plot shows principal component 1 and principal component 2 of each sample. Samples closer to each other in principal component space are similar in methylation profiles.

4.2.2 Differentially Methylated Bases and Regions

Methylation levels between test (F25H and F25C) and control (F2C) samples were compared to identify locations of important divergent changes between parental and progeny generations and further characterize genomic methylated bases and regions.

4.2.2.1 Distribution of DMCs and Generational Changes

The percentage of hypermethylated and hypomethylated DMCs per chromosome in CG context showed an almost equal proportion of hyper- and hypomethylation across all five chromosomes (Figure 28). Although the average number of methylated bases in CHG and CHH profiles were slightly higher than those in CG as reported in 3.3.1, the distribution of

methylation levels varied across the different chromosomes for each sequence context. Hypermethylation and hypomethylation was apparent in mCG (Figure 28) while mCHG context (Figure 29) had more of hypermethylated bases distributed at chromosome level for all comparison groups. Hypomethylation was mostly distributed in mCHH context (Figure 30) and less in mCHG context for all comparison groups.

F25C vs. F2C had more hypermethylated bases per chromosome than hypo-methylated bases in CHG sequence context. At CHH contest, more hypomethylated bases per chromosome were found in all sample comparison except for chromosome three which had an equal distribution of hyper- and hypomethylated bases and a higher percentage than the other chromosomes (Figure 30).

The proportions of DMCs per chromosome observed in F25H vs. F2C revealed more of hyper-methylated DMCs at CHG (Figure 29b) while at CHH DMCs per chromosome were more hypo-methylated (Figure 30b).

The analysis of hyper/ hypo methylated events per chromosome indicate that F25H has a significant proportion ($p<0.01$) of hyper methylation in CHG context and hypo-methylation in CHH when compared to the non-stressed parental generation, F2C. The non-stressed progeny F25C also shows significant proportion ($p<0.01$) of hyper-methylation when compared to parental generation F2C (Figure 30c).

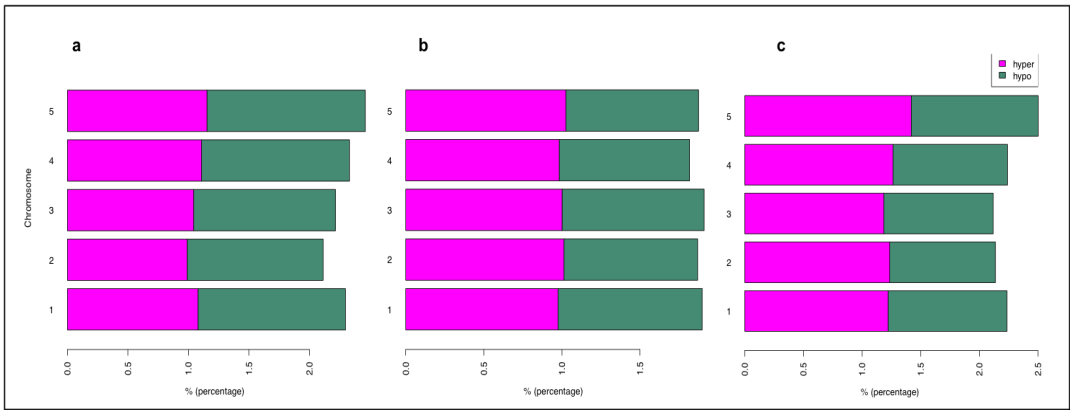

Figure 28. Percentage of Hyper- and Hypo- methylated DMCs at chromosome level. Differentially methylated bases in CpG context for each chromosome was identified in a 100bp window a) F25H vs. F25C b) F25H vs. F2C and c) F25C vs. F2C samples. Horizontal bar plots show the number of methylated events per chromosome as a percent of the sites with minimum coverage and differential. Pink section indicates the percentage of hypermethylation and green indicates hypomethylation, q-value <0.01 and methylation difference > 25%.

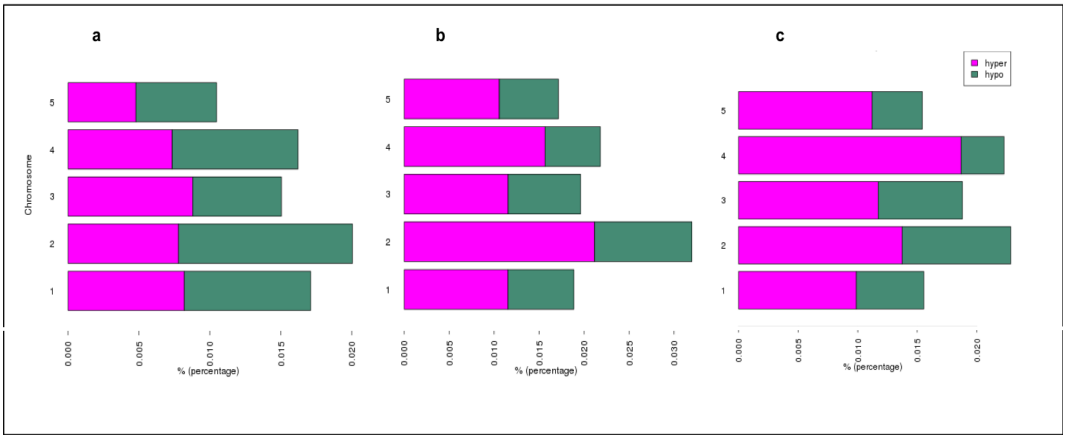

Figure 29. Percentage of Hyper- and Hypo- methylated DMCs at chromosome level. Differentially methylated bases in CHG context for each chromosome was identified in a 100bp window a) F25H vs. F25C b) F25H vs. F2C and c) F25C vs. F2C samples. Horizontal bar plots show the number of methylated events per chromosome as a percent of the sites with minimum coverage and differential. Pink section indicates the percentage of hypermethylation and green indicates hypomethylation, q-value <0.01 and methylation difference > 25%, (H=A, T, C).

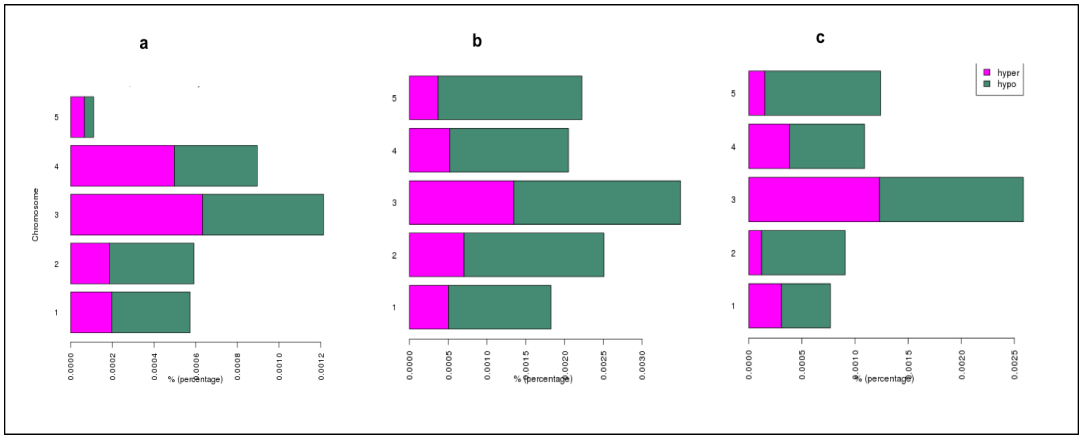

Figure 30. Percentage of Hyper- and Hypo- methylated DMCs at chromosome level. Differentially methylated bases in CHH context for each chromosome was identified in a 100bp window a) F25H vs. F25C b) F25H vs. F2C and c) F25C vs. F2C samples. Horizontal bar plots show the number of methylated events per chromosome as a percent of the sites with minimum coverage and differential. Pink section indicates the percentage of hypermethylation and green indicated hypomethylation, q-value <0.01 and methylation difference > 25%, (H=A, T, C).

Analysis of changes in DMCs with changes in methylation larger than 50% revealed that F25H group on heat map shows clear separation of all five samples from five samples of F2C group, whereas comparison of F25C group with F2C group or with F25H groups shows that only three samples clearly separate, whereas the other two show similarity to both F2C and F25H groups; this observation was similar for both, hypermethylated (Figure 31) and hypomethylated (Figure 32) regions.

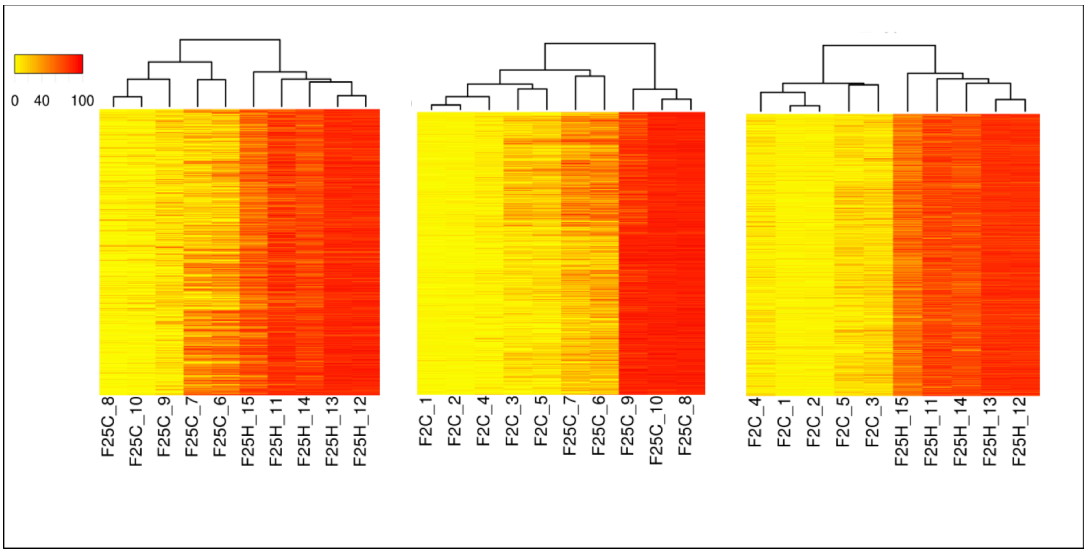

Figure 31. Heat maps of DMCs for hypermethylated cytosines. Differentially methylated cytosine in the genome with differences > 50% showing percentage of methylation in F25H vs. F25C (left), F25C vs. F2C (middle) and F25H vs. F2C (left). Red section indicates a larger percentage of methylation and yellow indicates the lower percentage, q-value <0.01. Y-axis represents height of samples, and the proportion of its position on the genome, x-axis represent each sample.

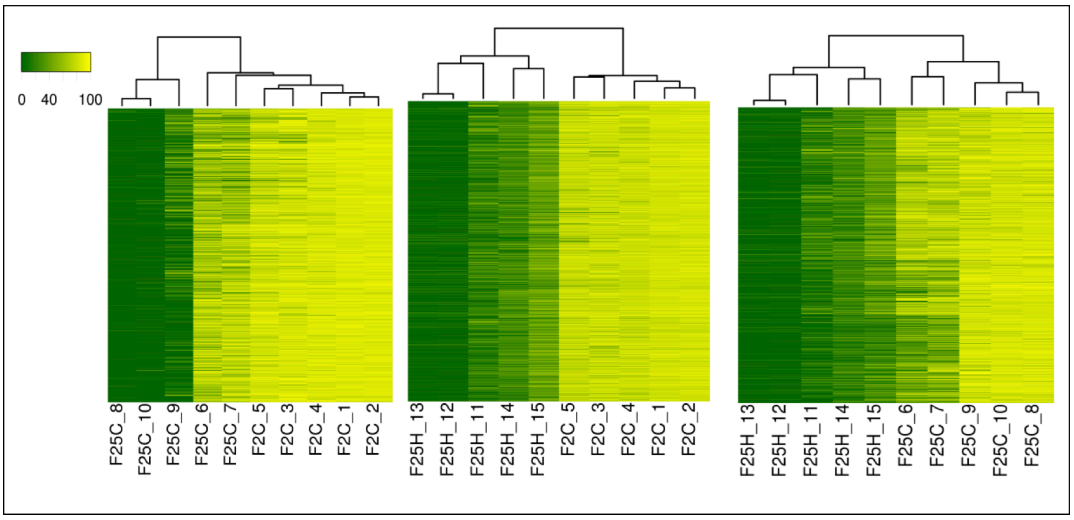

Figure 32. Heat maps of DMCs for hypomethylated differential methylated bases. Differentially methylated cytosine in the genome with differences > 50% showing percentage of methylated bases in F25H vs. F25C (left), F25C vs. F2C (middle) and F25H vs. F2C (left). Green section indicates a larger percentage of methylation and yellow indicates the lower percentage, q-value <0.01. Y-axis represents height of samples, and the proportion of its position on the genome, x-axis represent each sample.

4.2.2.2 Distribution of DMRs and Generational Changes

Differentially methylated regions (DMRs) extracted show similar trends to differentially methylated bases (DMCs) where DMRs were either hyper- or hypomethylated in various proportions on the chromosome. The distribution of methylated regions highlights more hypermethylation in CG and CHG context when F25H vs. F2C are compared (Figure 33), and the percentage methylation (x-axis) reduces from CG to CHG and CHH context over a 100_bp window. Also, when the sliding window was increased to 1000_bp, the proportions of hypermethylated DMRs somewhat increased per chromosome in CG while at CHG context the distribution varied with complete hypermethylation at chromosome two. DMRs at CHH context for 1000bp sliding window highlighted no DMRs per chromosome (Figure 34).

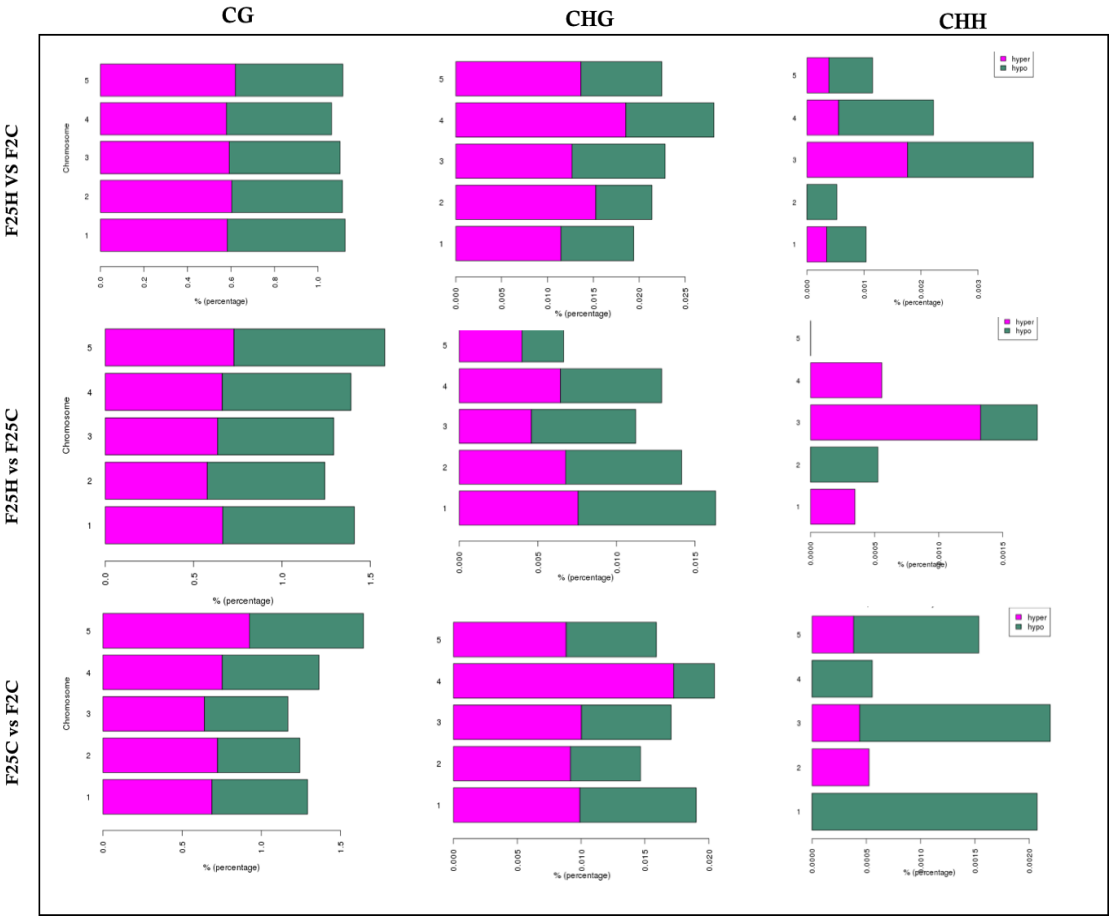

Figure 33. Distribution of DMRs and percentage of Hyper- and Hypo- methylated bases at chromosome level. Differentially methylated regions per chromosome per sequence context over a 100bp window. Global DMRs with differences > 25% showing percentage of hyper- and hypomethylated regions pink section indicates the percentage of hypermethylation and green sections indicates percentage of hypomethylation, q-value <0.01.

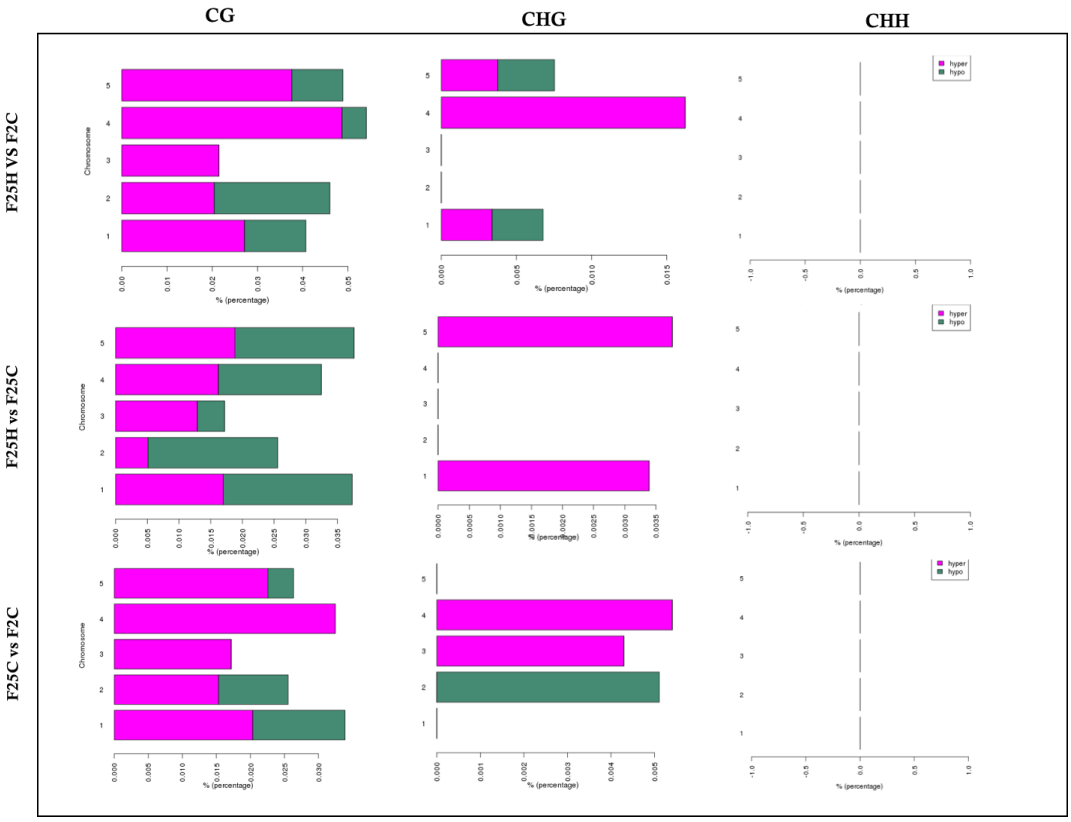

Figure 34. Distribution of DMRs and percentage of Hyper- and Hypo- methylated bases at chromosome level Differentially methylated regions per chromosome per sequence context over a 1000bp window. Global DMRs with differences > 25% showing percentage of hyper- and hypomethylated regions pink section indicates the percentage of hypermethylation and green sections indicate the percentage of hypomethylation, q-value <0.01.

DNA methylation changes (epimutations) between F25H vs. F2C and F25H vs. F25C were considered to be potentially induced by heat stress and epimutations between F25C and F2C were considered as the difference in the genome of parental and advanced progeny. Total number of DMRs (8

1) revealed more differential hypermethylated regions existing in F25C vs. F2C and the most differential hypomethylated regions existing in F25H vs. F25C which suggests that the multigenerational heat stress exposure triggered methylation changes in F25H plants.

Table 4.1: Number of Differentially Methylated Regions

Base Pair window	Sample Groups	Hypermethylated	Hypomethylated
	F25H vs F2C	5560	4754
100	F25C vs F2C	6991	5628
	F25H vs F25C	6109	6831
	F25H vs F2C	36	13
1000	F25C vs F2C	25	7
	F25H vs F25C	17	19

Regions that were differentially methylated between samples were defined relative to a control group (F25C and F2C), and those larger than 25% differences make up the numbers in Table 4.1.

Based on the 5560 hyper-methylated DMRs uncovered for F25H vs. F2C, heatmaps show F25H clustering separately from F2C. Similar comparison between F25H vs. F25C show similar trend but less obvious clustering. Also, DMRs for F25C vs. F2C showed less obvious clustering (Figure 35). Similar pattern was observed for the group of hypomethylated DMRs; clear clustering to two groups was observed between F25H vs. F2C, and less obvious between F25H and F25C and F25C and F2C group comparison (Figure 36).

In summary, analysis of DMRs confirms the trend observed for DMCs, F25H group of plants clearly segregates in clustering analysis from F2C group, whereas differences between F2C group and F25C groups as well as between F25H and F25C are less obvious.

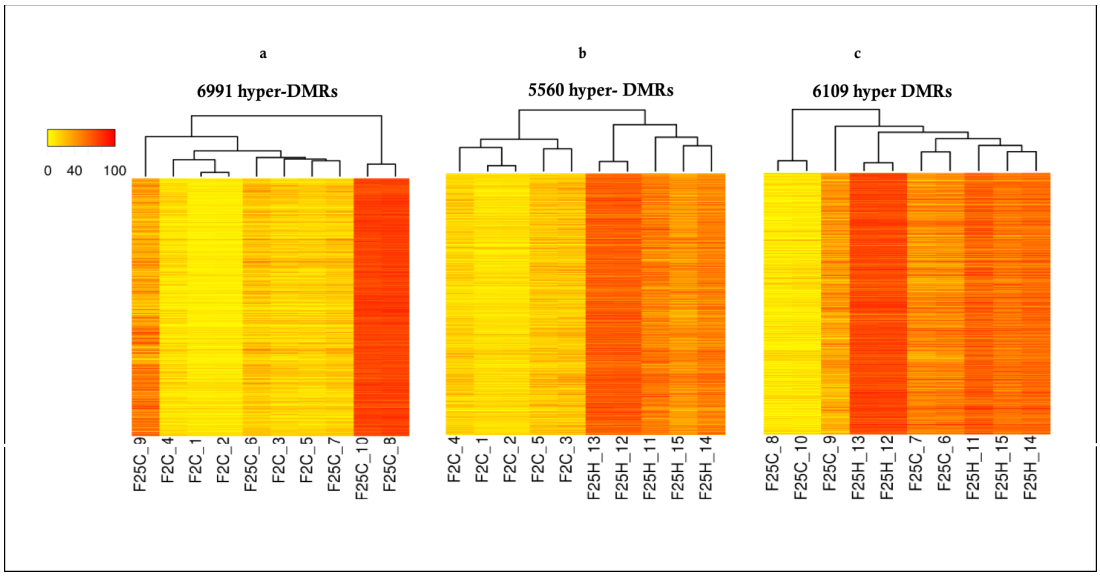

Figure 35. Heat maps of differentially hypermethylated regions at 100_bp. Differentially hypermethylated cytosine in the genome with differences > 50% showing percentage of methylated bases in F25C vs F2C (left), F25H vs F2C (middle) and F25H vs F25C (right) comparison groups. Red section indicates a larger percentage of methylation and yellow indicate the lower percentage, q-value <0.01. Y-axis represents height of samples, and the proportion of its position on the genome, x-axis represent each sample.

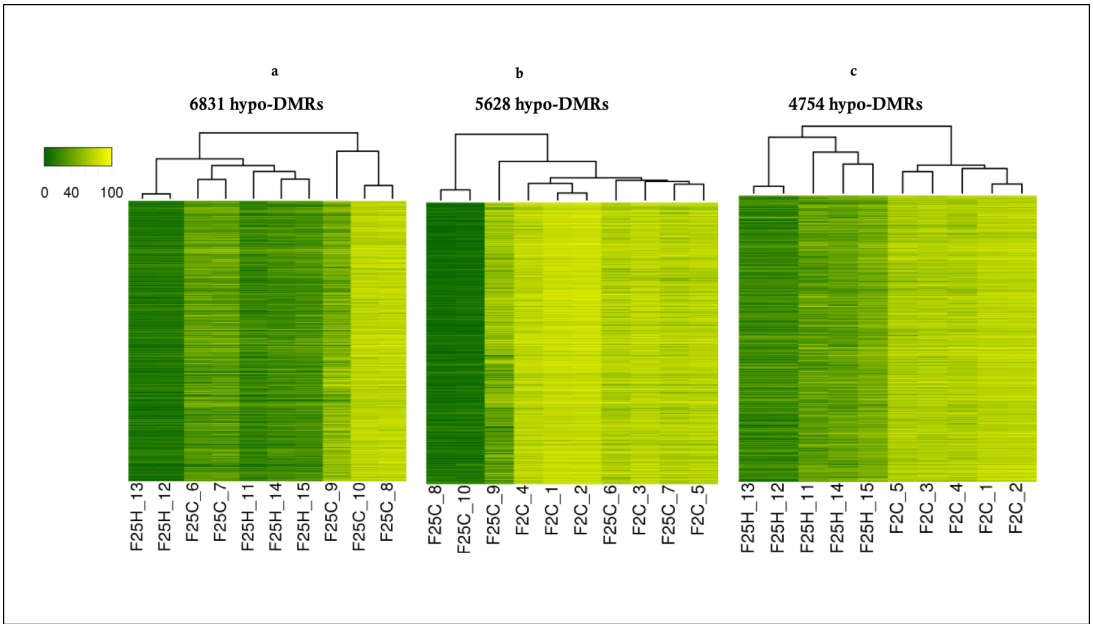

Figure 36. Heat maps of differentially hypomethylated regions at 100bp. Differentially hypomethylated cytosine in the genome with differences > 50% showing percentage of methylated bases in F25H vs F25C (left), F25C vs. F2C (middle) and F25H vs. F2C (right) comparison groups. Green section indicates a higher percentage of methylation and yellow indicates the lower percentage, q-value <0.01. Y-axis represents height of samples, and the proportion of its position on the genome, x-axis represent each sample.

4.2.3 Annotated Differential Methylation Events

The biological impact of all observed differential methylation event was put into genomic context with subsequent analysis detailing variable regions and positions of methylation within gene structure and sequence islands. To obtain insights on the relationship of DMCs with promoter regions, we calculated their distances to the nearest transcription start site (TSS).

The distribution collected by absolute distance of individual DMCs in sample replicates revealed methylated DMCs in CG context were located approximately within 0 to 2 kb to TSS (Figure 37). The overall nucleotide distance to TSS dropped for DMCs in CHG and CHH context (Figure 38) but was similar to known ranges within 0 to 1kb distance to TSS. At a 50% overlapping difference, the DMCs present if any, were insufficient to be binned for CHG and CHH context. Furthermore, the binned proximity of DMRs nearest to TSS in the stress- induced methylated CG sites as seen Figure 39 showed a similar distribution as DMCs in Figure 37.

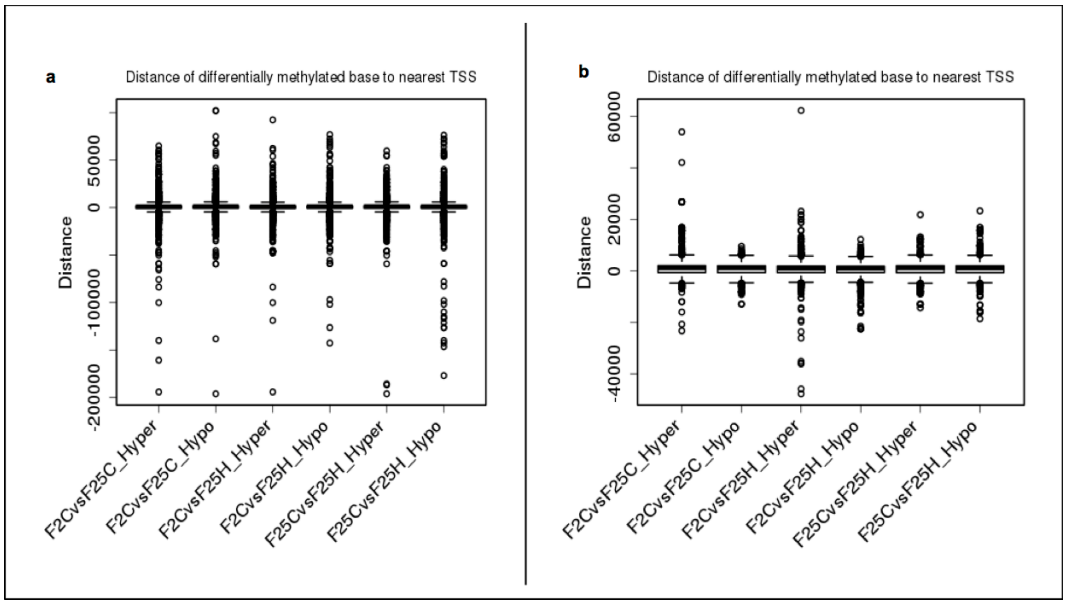

Figure 37. The distance of DMCs to nearest TSS. Differentially methylated bases in mCG context shores overlapping by (a) >25% and (b) >50%. Distance to TSS for differentially methylated CGs are plotted from F25H vs. F2C, F25C vs. F2C and F25H vs. F25C hyper/hypomethylation analysis. Y-axis is nucleotide base distance to a transcription start site.

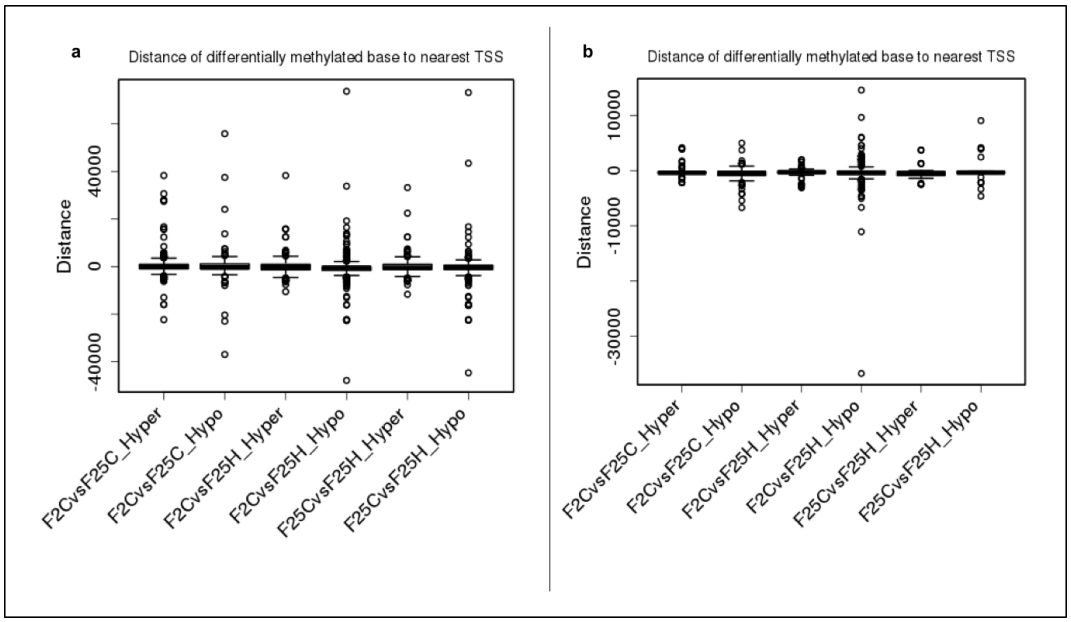

Figure 38. The distance of DMCs to nearest TSS. Differentially methylated bases in (a) CHG (b) CHH context shores overlapping by >25%. Distance to TSS for differentially methylated CGs are plotted from F25H vs. F2C, F25C vs. F2C and F25H vs. F25C hyper/hypomethylation analysis. Y-axis is nucleotide base distance to a transcription start site.

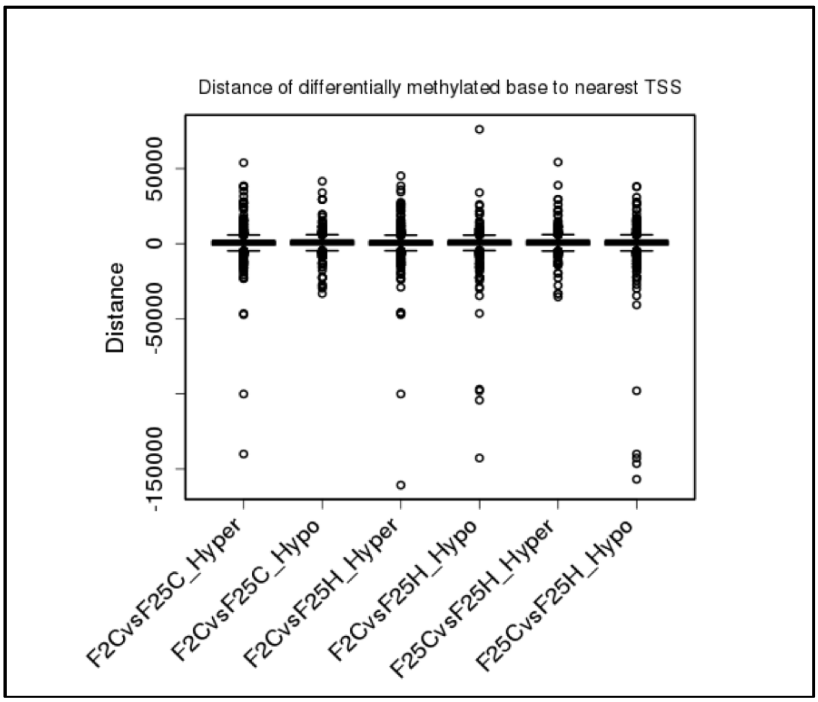

Figure 39. The distance of DMRs to nearest TSS. Differentially methylated regions in CG context shores overlapping by >25%. Distance to TSS for differentially methylated CGs are plotted from F25H vs. F2C, F25C vs. F2C and F25H v. F25C hyper/hypomethylation analysis. Y-axis is nucleotide base distance to a transcription start site.

DMRs and DMCs were characterized to determine whether they were preferably located near genes. The location of hypo or hyper methylated DMRs was compared to annotated *Arabidopsis* genes using genomation (Akalin, Franke, Vlahoviček, Mason, & Schübeler, 2014). Gene components of DMCs and DMRs overlapping with exon/intron/promoters in the preference of promoter > exon > intron showed DNA methylation events were concentrated mostly in the promoter regions within all sequence context.

Differentially methylated bases in F25H vs. F2C were highest in promoter region followed by intergenic region for CHH sites (Figure 40 and 41). DMCs on CG sites were methylated in gene bodies especially at exon components (Figure 42, 44). DMCs in F25H vs. F25C were hypomethylated mostly at promoter and intergenic regions for CHG sites (Figure

43). Additionally, hypomethylated DMCs were concentrated in intergenic regions (Figure 41, 45). DMRs overlapped with exons most at CG sites followed by promoter regions.

Differentially methylated regions based on the overlap of sequence shores within gene components showed the highest percentage of hyper and hypomethylated DMRs overlapping promoter regions were at non-CG sites for all sample groups (Figure 46-47). DMRs overlapped with exons most at CG context followed by promoter regions. Hypermethylated DMRs in F25H vs F2C comparison group (Figure 46) were most present in exons similar to what was found in F25C vs F2C.

F25H vs. F2C had higher percentage of DMRs overlapping with promoter and intergenic regions than in F25C vs. F2C while F25C vs. F2C had higher percentage of DMRs overlapping with exon and intron regions than in F25H vs. F2C irrespective of methylation changes (hyper/hypo) (Figure 46-47).

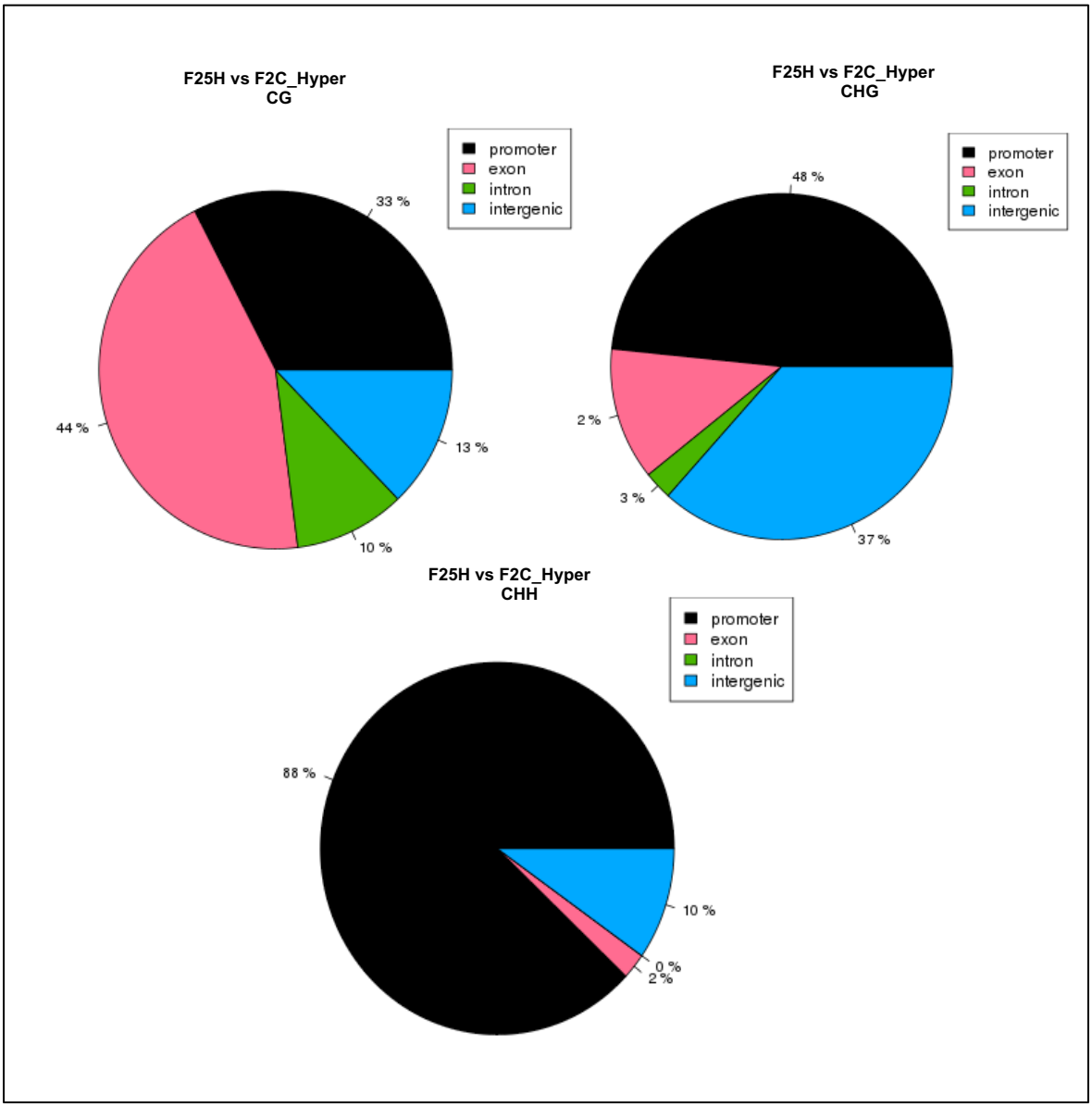

Figure 40. Gene components of differentially hypermethylated DMCs in F25H vs. F2C. Percentages of differentially methylated bases of all sequence context overlapping with the promoter, exon, intron and intergenic regions. Percentages plotted are averages of hyper/hypomethylation from F25H vs. F2C, DMCs analysis was with > 25% differences and around 10bins.

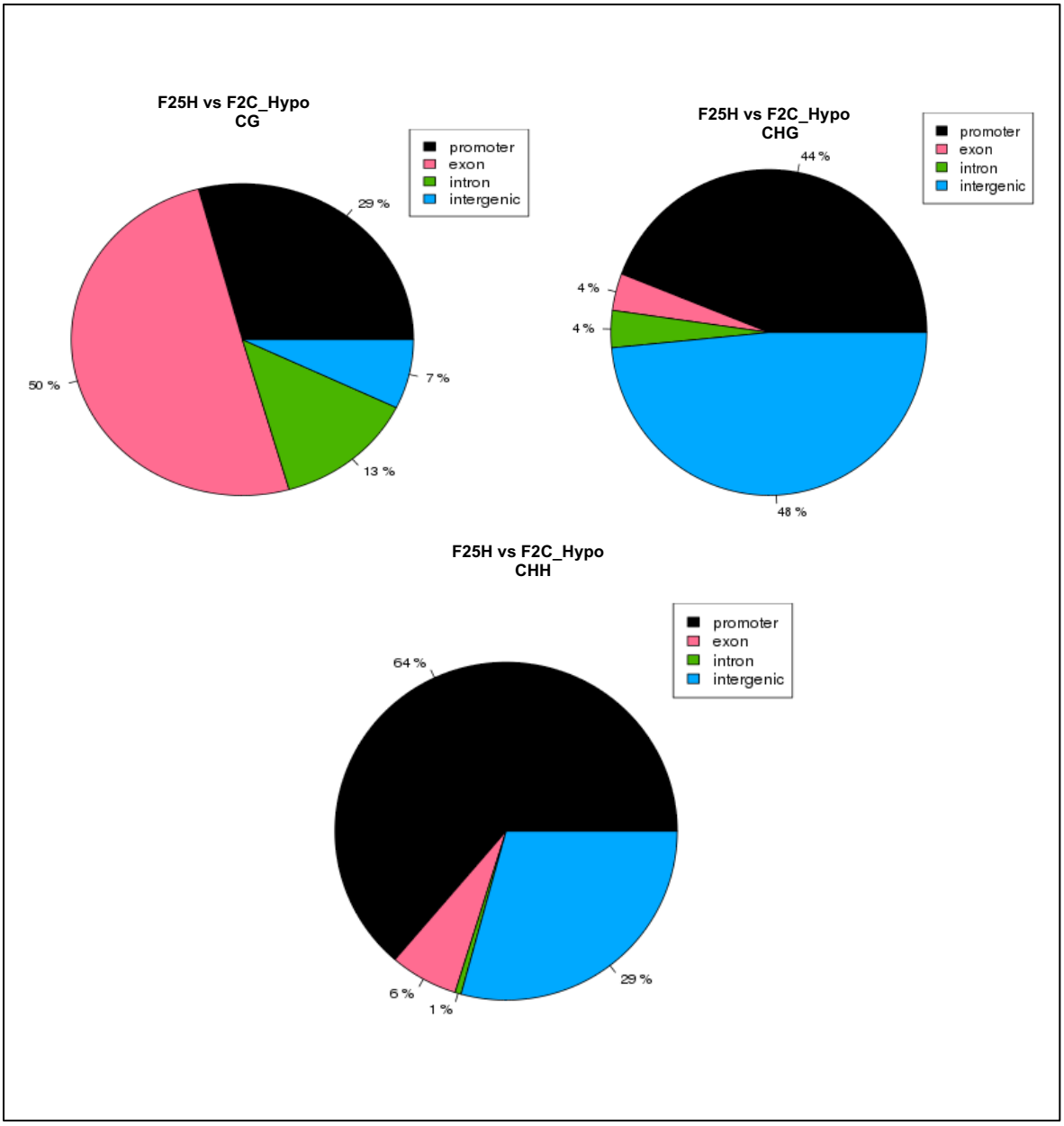

Figure 41. Gene components of differentially hypomethylated DMCs in F25H vs. F2C. Each bar represents percentages of differentially methylated bases of all sequence context overlapping with the promoter, exon, intron and intergenic regions. Percentages plotted are averages of hyper/hypomethylation from F25H versus F2C, DMCs analysis was with > 25% differences and around 10bins.

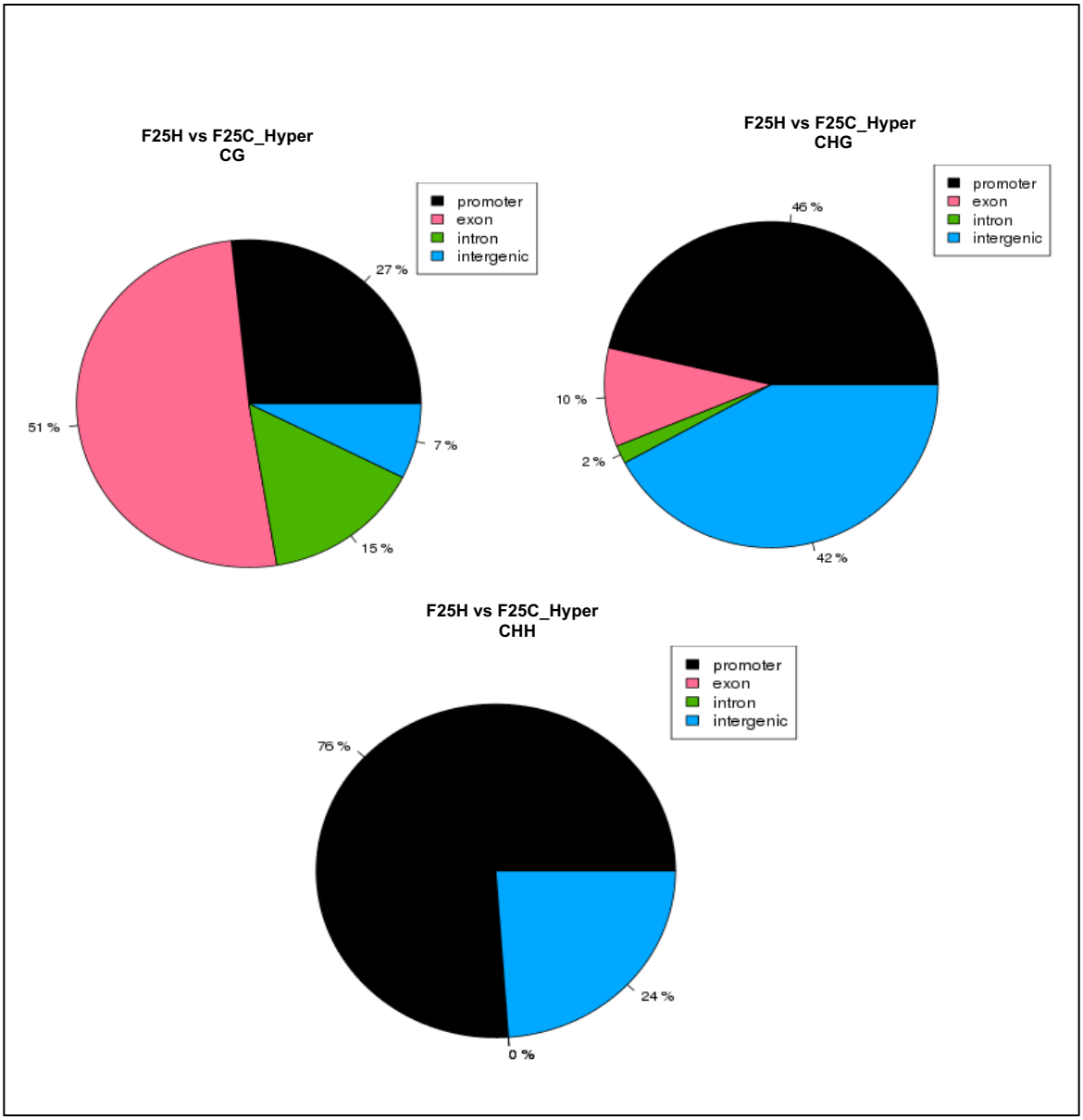

Figure 42. Gene components of differentially hypermethylated DMCs in F25H vs F25C. Percentages of differentially methylated bases of all sequence context overlapping with the promoter, exon, intron and intergenic regions. Percentages plotted are averages of hyper/hypomethylation from F25H versus F25C, DMCs analysis was with > 25% differences and around 10bins.

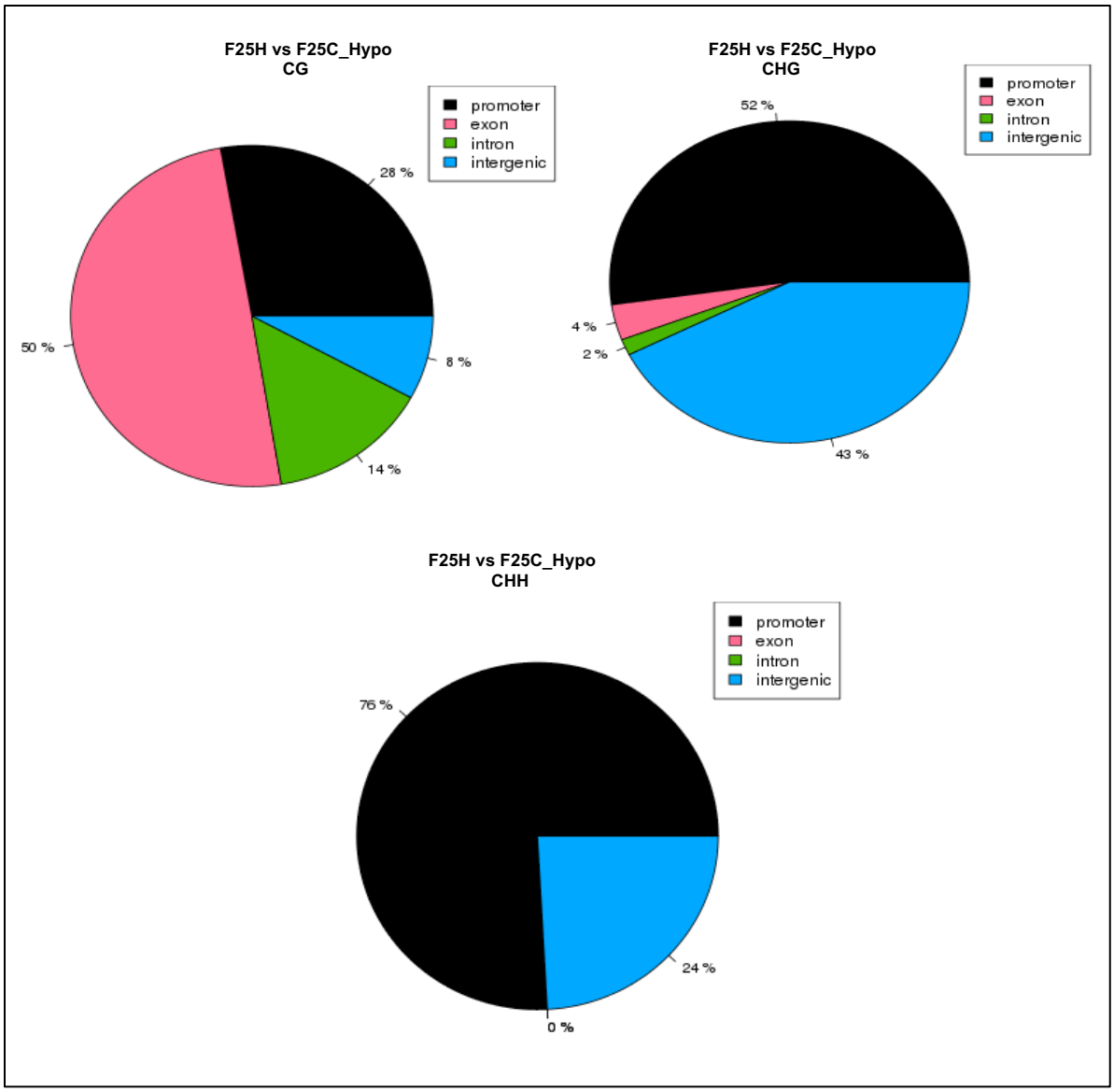

Figure 43. Gene components of differentially hypomethylated DMCs in F25H vs F25C. Percentages of differentially methylated bases of all sequence context overlapping with the promoter, exon, intron and intergenic regions. Percentages plotted are averages of hyper/hypomethylation from F25H versus F25C, DMCs analysis was with > 25% differences and around 10bins.

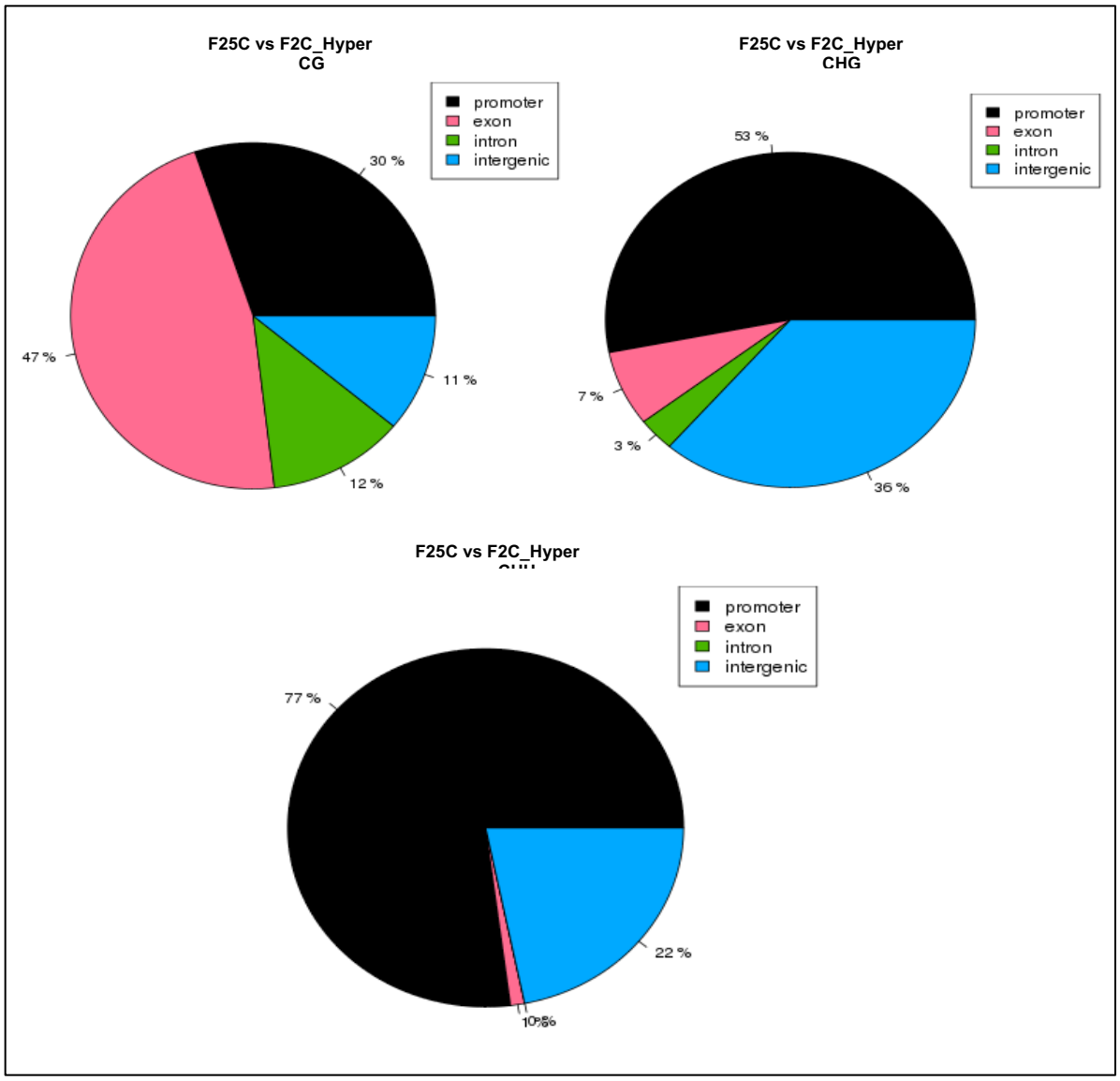

Figure 44. Gene components of differentially hyper-methylated DMCs in F25C vs. F2C. Percentages of differentially methylated bases of all sequence context overlapping with the promoter, exon, intron and intergenic regions. Percentages plotted are averages of hyper/hypomethylation from F25C versus F2C, DMCs analysis was with > 25% differences and around 10bins.

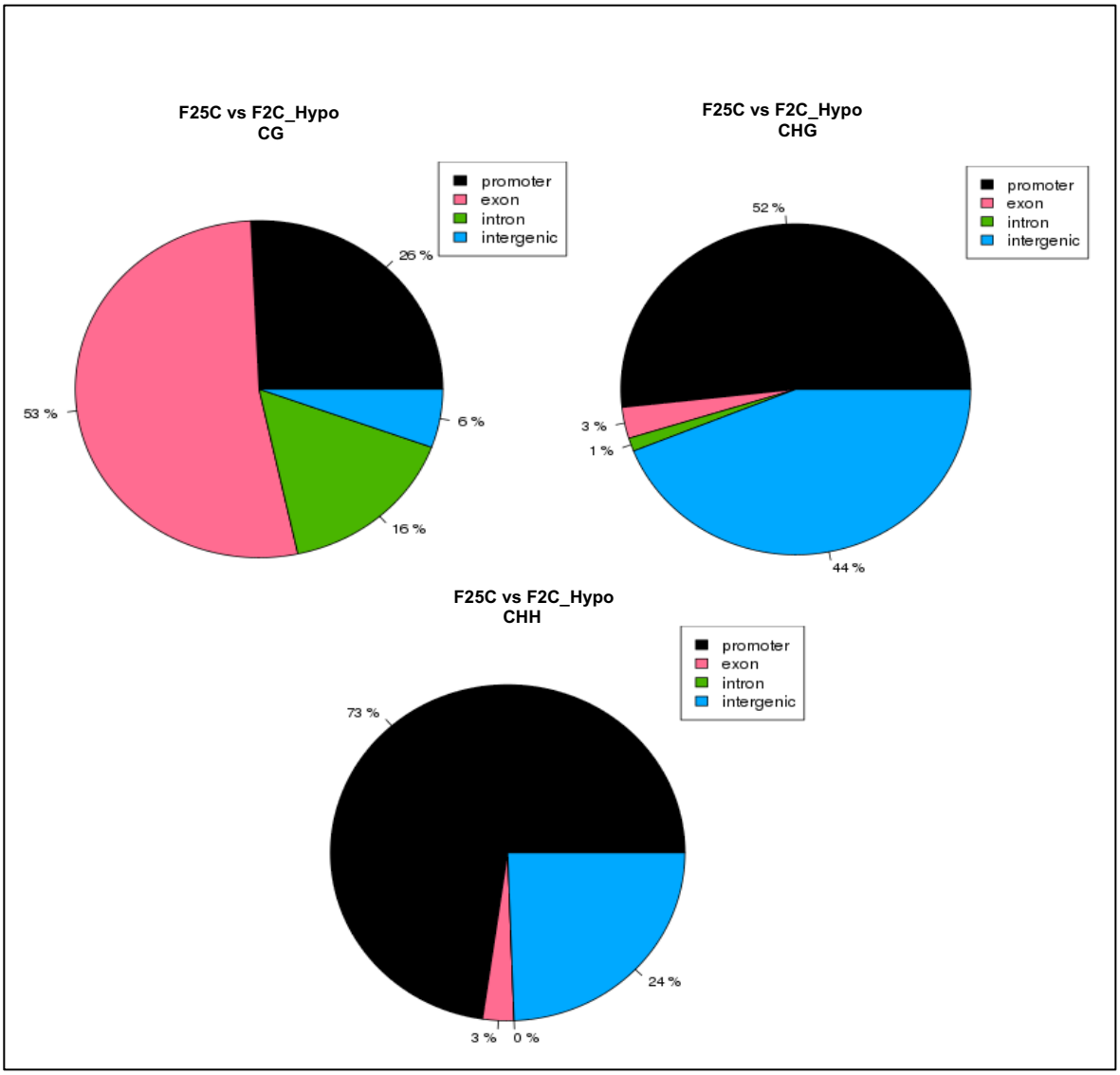

Figure 45. Gene components of differentially hypomethylated DMCs in F25C vs. F2C. Percentages of differentially methylated bases of all sequence context overlapping with the promoter, exon, intron and intergenic regions. Percentages plotted are averages of hyper/hypomethylation from F25C vs F2C, DMCs analysis was with > 25% differences and around 10bins.

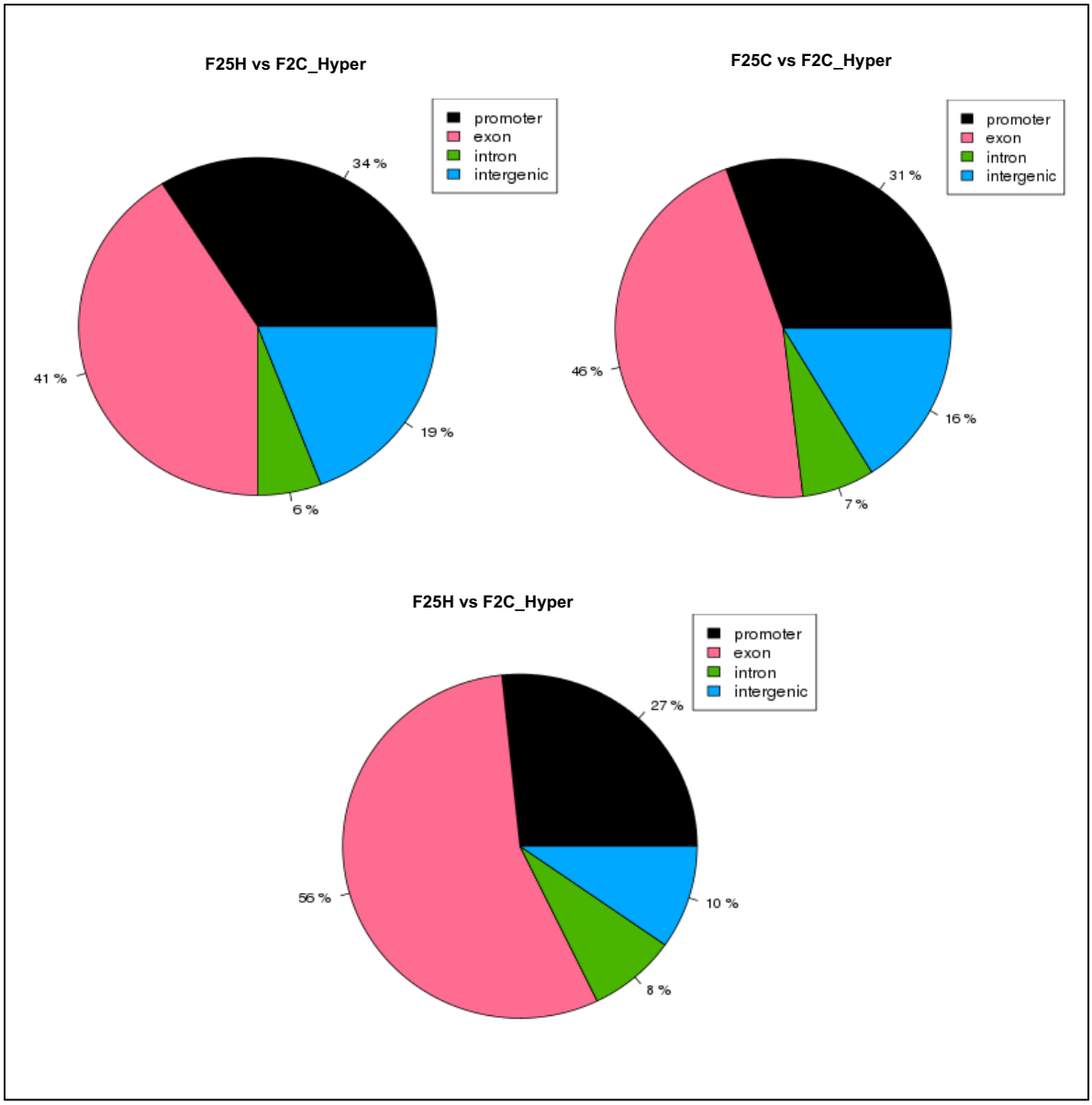

Figure 46. Gene components of differentially hypermethylated DMRs. Percentages of differentially methylated bases of all sequence context overlapping with the promoter, exon, intron and intergenic regions. Percentages plotted are averages of hyper/hypomethylation from F25C vs. F2C, DMRs analysis was with > 25% differences and around 10bins.

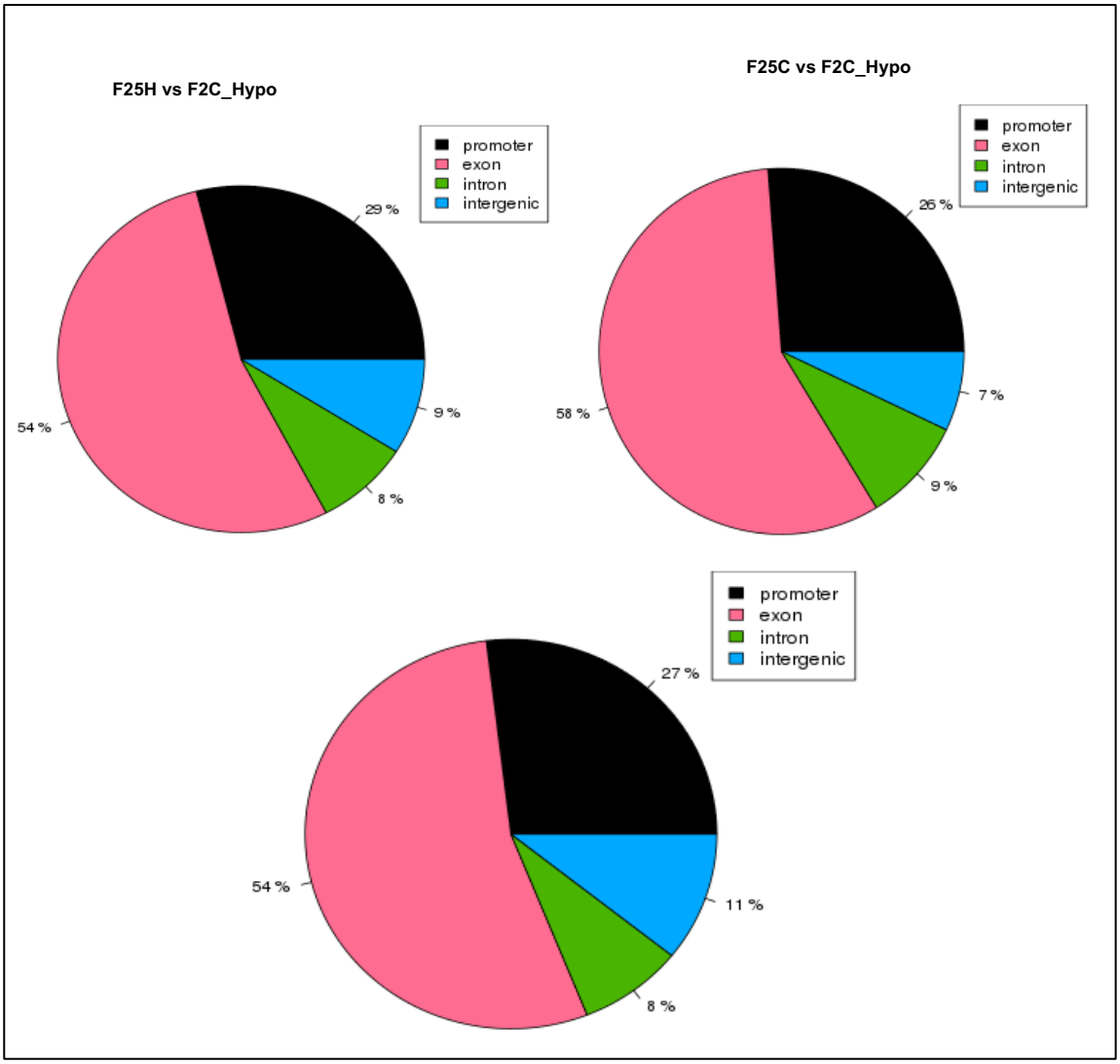

Figure 47. Gene components of differentially hypomethylated DMRs. Percentages of differentially methylated bases of all sequence context overlapping with the promoter, exon, intron and intergenic regions. Percentages plotted are averages of hyper/hypomethylation from F25C vs F2C, DMRs analysis was with > 25% differences and around 10bins.

4.3 Stress Associated Sequence Variations in G_{25} vs G_2 Comparison

Genomic variants comprising of single nucleotide polymorphisms (SNPs), insertions and deletions (INDELs) were explored. Their corresponding genomic location such as intronic, untranslated regions (5' UTR or 3'UTR), upstream and downstream of gene regions and intergenic regions were identified as well as predicted coding effects in respect to protein-coding genes such as synonymous or non- synonymous mutations. According to the SnpEff program (Cingolani et al., 2012) utilized in this study, upstream was defined as 5 kb upstream of the distal transcription start site and downstream was defined as 5 kb downstream of the most distal polyA addition site. Variants affecting non-coding regions were expounded and biotypes identified with available information after comparing with reference *Arabidopsis* genome.

Variant rates for F25H were 1 variant for every ~ 1700 bases, F25C was 1 variant for every ~ 5000 bases, and F2C was 1 variant for every ~ 5000 bases suggesting that F25H had a higher rate of variance in its genome when all samples were jointly considered.

4.3.1 Classification of Sequence Variants by Type and Regions

The variation profiles of coding sequences, introns and untranslated regions investigated revealed higher number of variant effects in generational heat stressed sample F25H when compared to generational non-heat stressed F25C and parental F2C (Table 4.2). Classification of variants by regions revealed the largest number of total variants in F25H to be upstream (110,837) followed by those found downstream (103,301) and intergenic (40,941) regions accordingly. Variants were also found within the 5' untranslated (2312) and 3' untranslated regions (1919) of the genome.

Table 4.2 Number of effects by region of genome

	Downstream	Upstream	Intergenic	5’ UTR	3’ UTR
F25H	103,301	110,837	40,941	1919	2312
F2C	27,947	28,403	15,058	442	442
F25C	28,035	28,762	15,184	438	461

Table 4.2 summarizes jointly called variant effects by region highlighting only regions with the most number of coding effects in F25H, F25C and F2C. Variant types represented here are downstream and upstream gene variants, intergenic regions and 5’ and 3’ UTR. Figure 49-51 represents the total number of effects in F25H, F25C and F2C respectively.

Variant effects in F25H were mostly found in upstream regions (Table 4.2) and F25H having the highest total variant number than F2C and F25C**.** Interestingly, the insertion and deletion lengths of variants in F25C and F2C had similar numbers in all regions when compared to each other while F25H was significantly different when compared to F25C and F2C (Figure 48), suggesting that heat stress did trigger the difference when F25H is compared to F25C and F2C.

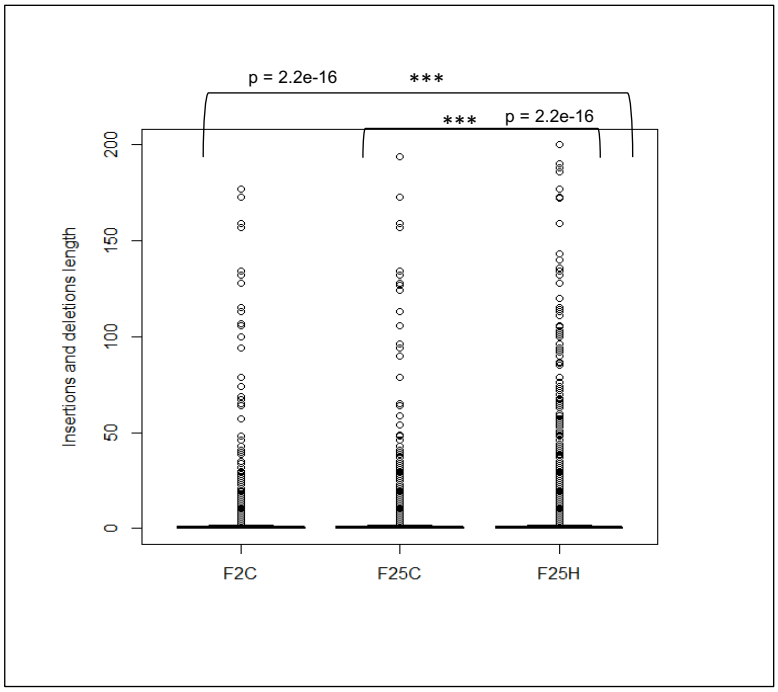

Figure 48. Box plot for insertions and deletions length of variants within F2C, F25C and F25H. The asterisk above bracket (*) show a significant difference between F25H vs. F2C and F25H vs. F25C where *** indicates $p < 0.001$. Wilcoxon rank sum text was used to determine statistical differences ($p<0.05$).

Other regions that showed some variants are represented in Figure 49-51. Also, the number of variants by type as described in Table 4.3 shows F25H group having a larger number of SNPs and INDELs when compared to F25C and F2C. The number of SNPs found in F25H were 53,678, 8715 insertions (INS) and 8954 deletions (DEL). F25C and F2C showed a similar range with a slight difference between their number of SNPs and INDELs (Table 4.3).

Table 4.3 Number of Variants by Type

	SNPs	INS	DEL
		INDELs	
F25H	53,678	8,715	8,954
F2C	15,599	4,188	3,798
F25C	15,526	4,340	3,803

In F25H, 18 SNPs representing missense mutations converting stop codon to amino acid coding triplet and 66 SNPs representing nonsense mutations generating stop codon sequence were identified (Figure 49). 2,312 synonymous coding variants that causes no change in the amino acid sequence were identified. F25C and F2C had similar variants, 313 synonymous-coding variants (Figure 50) were classed in F25C while parental F2C had 316 synonymous-coding variants (Figure 51). The different SNPs frequencies detected in all sample groups highlighted major classes outlined in Table 4.2 as well as other variant classes such as non-coding transcript loss, start- loss and exon- loss.

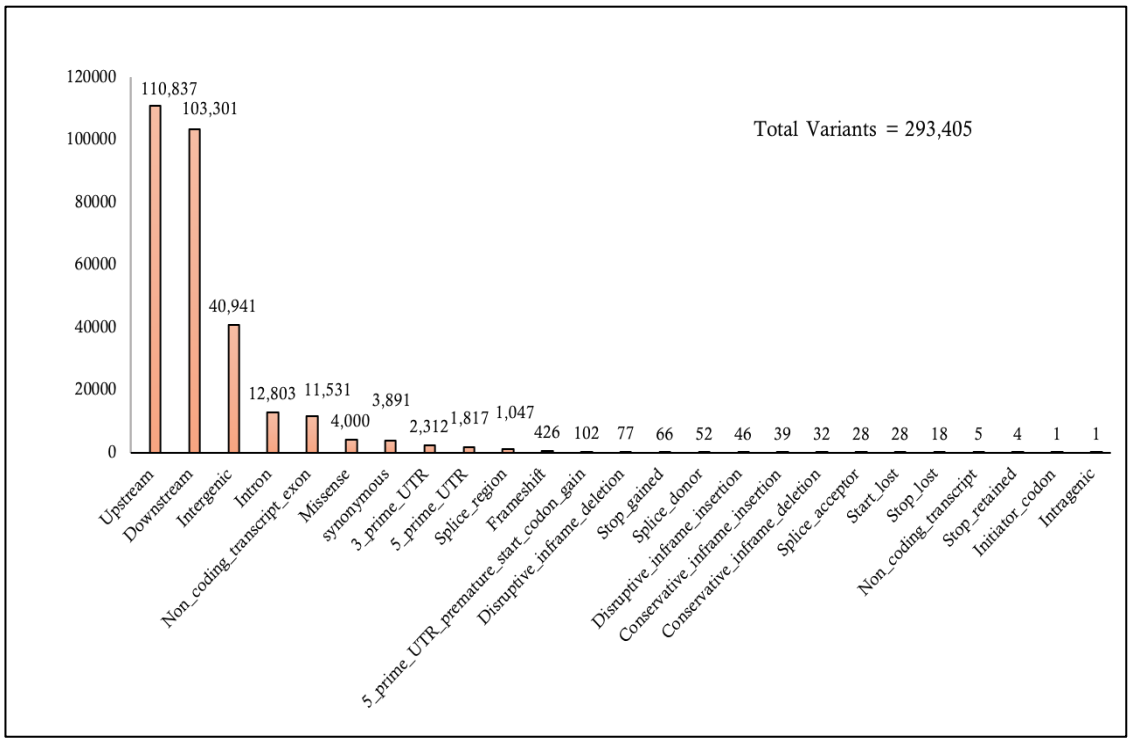

Figure 49. Classifications of variants in F25H. The number of variants is shown above each bar and effects from SnpEff listed on x-axis.

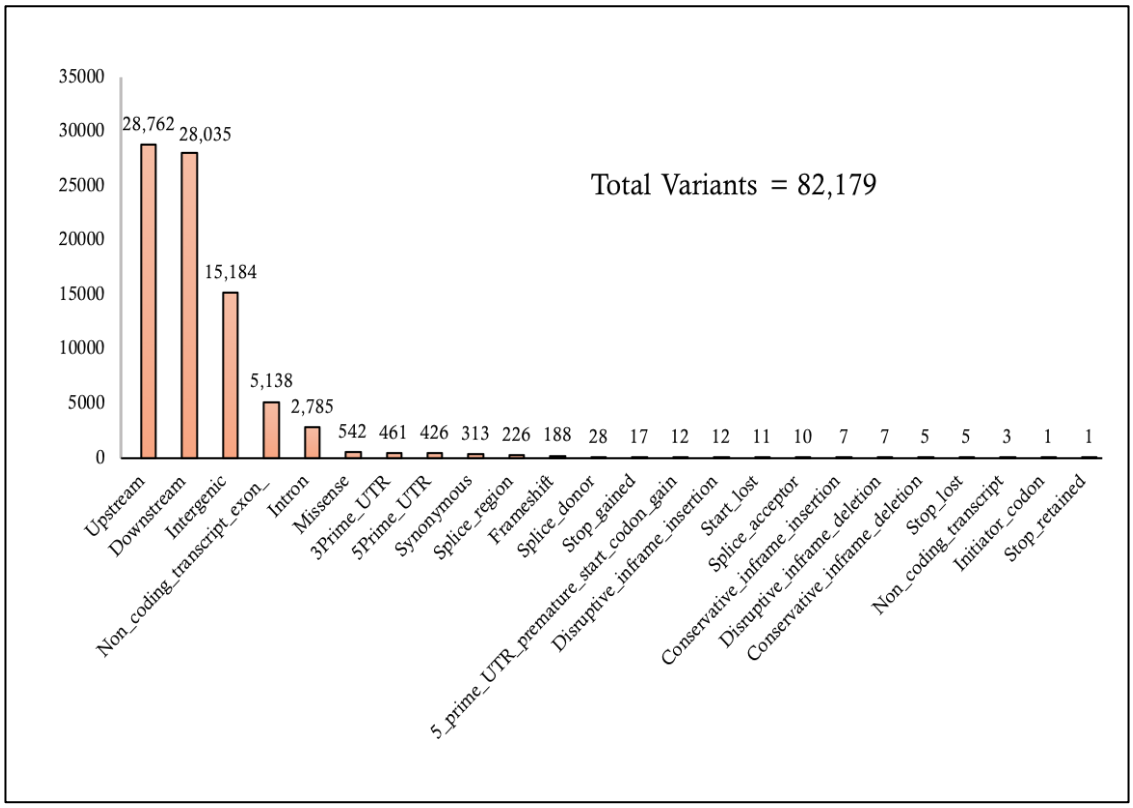

Figure 50. Classifications of variants in F25C. The number of variants is shown above each bar and effects from SnpEff listed on x-axis.

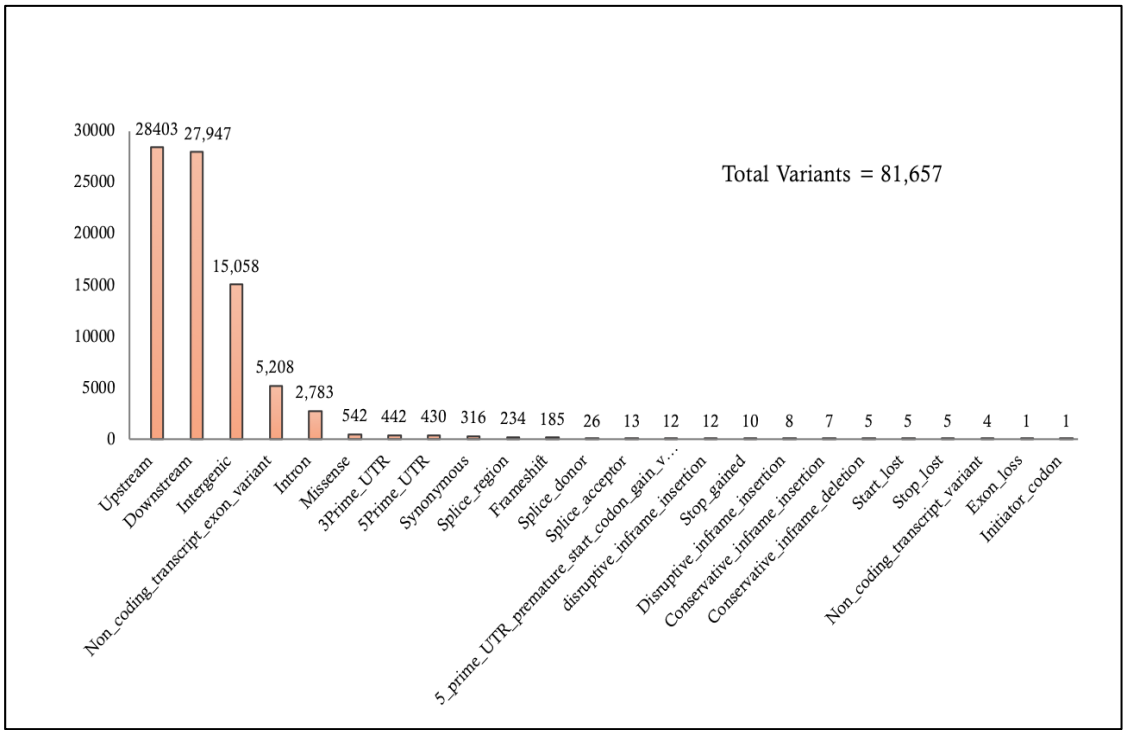

Figure 51. Classifications of variants in F2C by regions. The number of variants is shown above each bar and effects from SnpEff listed on x-axis.

GATK Haplotype caller calls SNPs and INDELs by an assembly of haplotypes in an active region of the genome which enabled the identification of variants unique to each sample genome by extracting the non-overlapping sites specific for F25H, F25C, and F2C. The resulting variants showed genes nearest to non-overlapping SNPs sites to be 775, 740 and 590 for F2C, F25C, and F25H respectively. 159, 140 and 104 INDELs were identified for F2C, F25C and F25H respectively (Figure 52). SNPs and INDELS jointly called show F25H with most SNPs and INDELS (Figure 53).

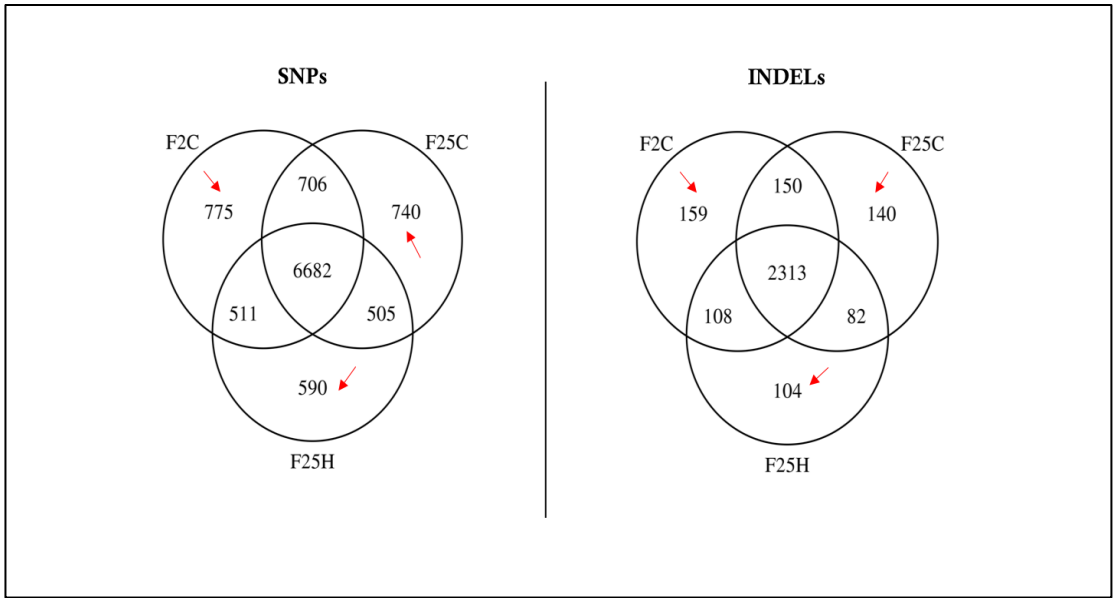

Figure 52. Venn diagram of variants. The red arrow indicates non- overlapping SNPs and INDELs unique to F25H, F25C and F2C and the Venn diagram identifies numbers unique to the various samples. Sample variants were jointly called by "GATK haplotypes Caller" with a low quality filter. Each sample was of five biological replicates.

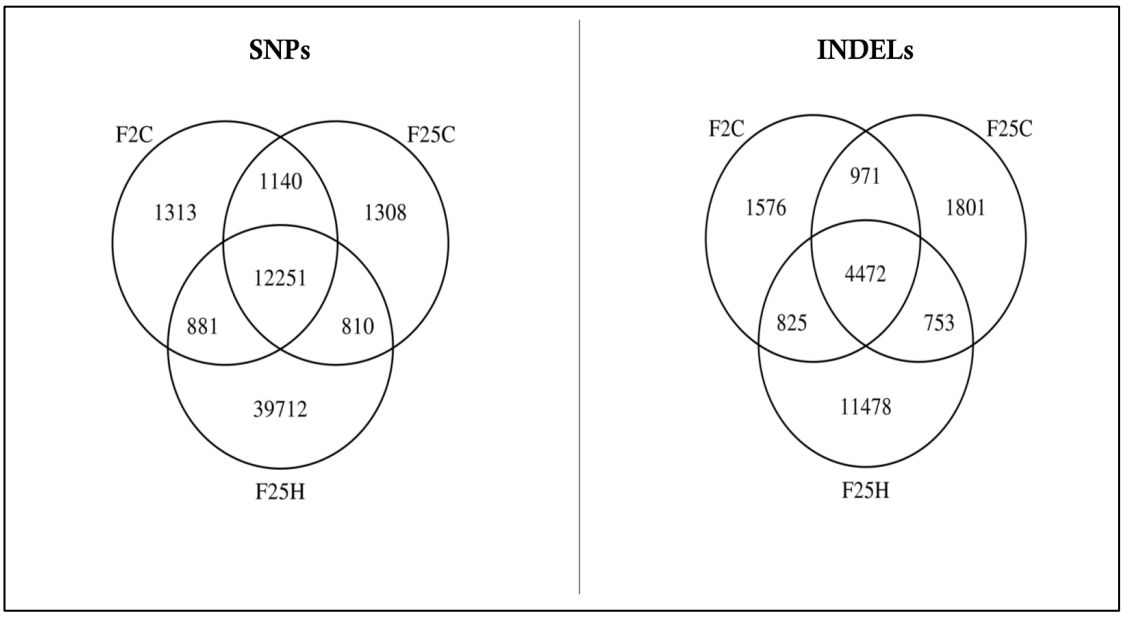

Figure 53. Venn diagram of jointly called variants, non- overlapping area of Venn indicates the number of SNPs and INDELs of each sample group. Sample variants were jointly called by "GATK haplotypes Caller" meaning variants of each condition (using the five samples together) were considered by GATK haplotypes.

4.3.2 Biological Enrichment Analysis

Functional classification of variants, DMRs and DMCs unique to each test group was interpreted using SuperViewer to identify regions with statistically over-represented numbers of genes and genomic features. Biological processes that might be enriched or under-represented within and between generations were assessed. All values were normalized by bootstrap x100 and p-values < 0.05 only were retrieved as significant.

4.3.2.1 Variants and Predicted Genomic Effects

Variants included genes nearest to the non -overlapping SNPs and INDELs sites represented in Figure 52. The most enriched biological process for SNPs was classed as unknown biological processes in all samples although processes such as response to stress and stimulus were commonly present for SNPs in F25H, F2C and F25C with F25H having the most enrichment in response to stimulus biological component. F25H-SNPs were underrepresented in processes like developmental, protein metabolism, cell organization, and biogenesis which were all present in F2C- SNPs (Figure 54).

104 INDELs in F25H were enriched in genes termed other cellular and metabolic processes. Analysis of the genes affected by the 159 INDELs in F2C showed transport as the most enriched biological process followed by protein metabolism and other cellular and metabolic processes (Figure 54, Table 4). Lastly, in F25C, 140 INDELs were specific to G_{25} representing variants that are different from parental plants F2C (Figure 52), these INDELs were enriched in the response of abiotic and biotic stimulus and protein metabolism.

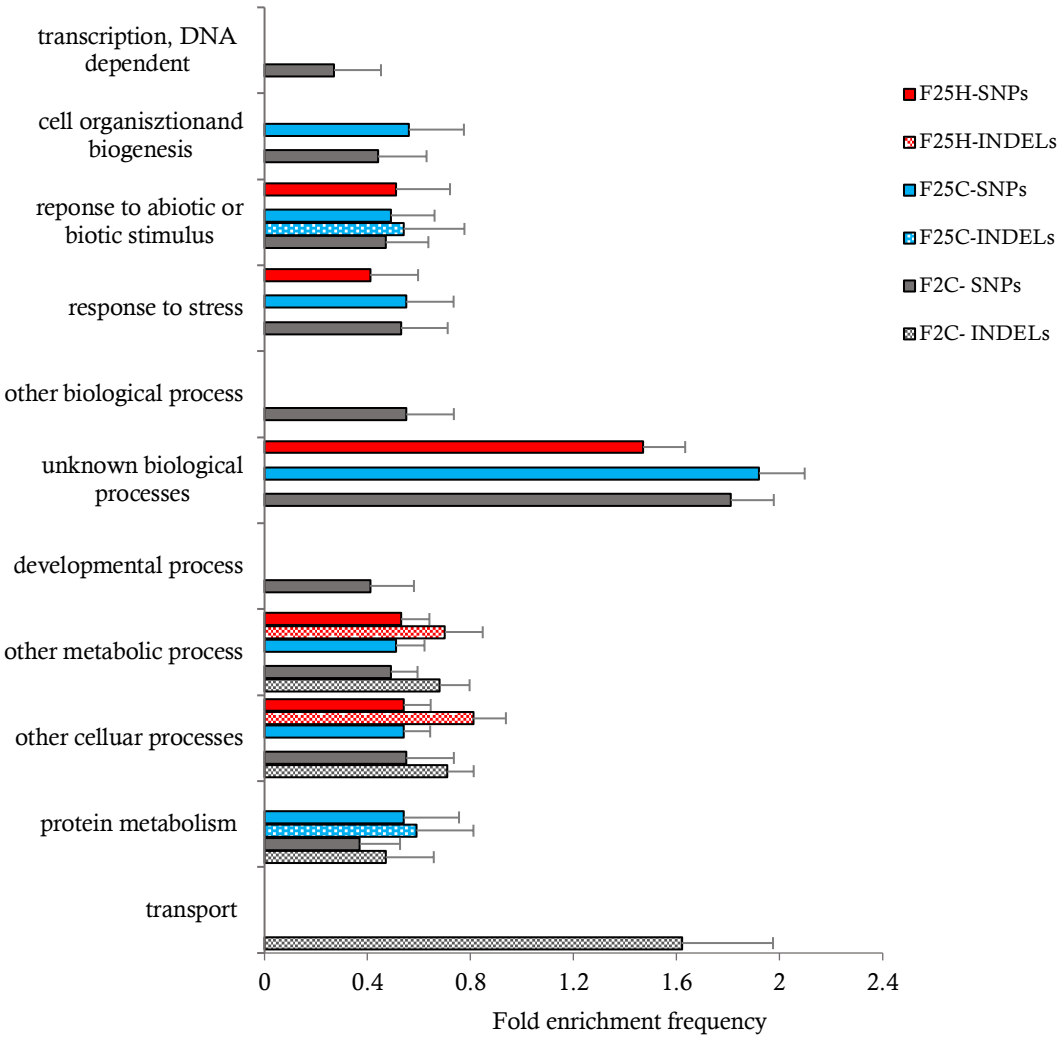

Figure 54. Enrichment analysis of SNPs and INDELs associated genes and their classification based on biological processes. y-axis is normalized frequency with binomial coefficients as calculated by SuperViewer. Genes are classed with p-values < 0.05, ±bootstrap StdDev.

Table 4.4: Classification of Annotated SNPs and INDELs

Biological Processes		F25H-SNPs	F25C – SNPs	F2C-SNPs	F25H-INDELs	F25C - INDELs	F2C-INDELs
Transport	Freq. ±SD p-value	-	-	-	-	-	1.62 ±0.345 0.013
Protein Metabolism	Freq. ±SD p-value	-	0.54 ±0.215 0.000	0.37 ±0.156 0.000	-	0.59 ±0.221 0.000	0.47 ±0.187 0.008
Other Cellular Processes	Freq. ±SD p-value	0.54 ±0.105 0.000	0.54 ±0.103 0.000	0.55 ±0.185 0.000	0.81 ±0.127 0.043	-	0.71 ±0.102 0.004
Other Metabolic Processes	Freq. ±SD p-value	0.53 ±0.11 0.000	0.51 ±0.111 0.000	0.49 ±0.104 0.000	0.7 ±0.147 0.017	-	0.68 ±0.116 0.004
Developmental processes	Freq. ±SD p-value	-	-	0.41 ±0.17 0.01	-	-	-
Unknown biological process	Freq. ±SD p-value	1.47 ±0.163 0.002	-	1.81 ±0.167 0.000	-	-	-
Other biological Processes	Freq. ±SD p-value	-	1.92 ±0.177 0.000	0.55 ±0.185 0.03	-	-	-
Response to stress	Freq. ±SD p-value	0.41 ±0.186 0.009	-	0.53 ±0.181 0.02	-	-	-
Response to abiotic or biotic stimulus	Freq. ±SD p-value	0.51 ±0.21 0.026	0.55 ±0.184 0.03	0.47 ±0.166 0.02	-	0.54 ±0.236 0.000	-
Cell organization	Freq. ±SD p-value	-	0.49 ±0.17 0.021	0.44 ±0.189 0.02	-	-	-
Transcription, DNA dependent	Freq. ±SD p-value	-	0.56 ±0.214 0.043	0.27 ±0.182 0.02	-	-	-

4.3.2.2 DNA Methylation and Predicted Genomic Effects

DMRs were mostly enriched at CG sites than non-CG sites with more annotated biological processes outlined for hypermethylated DMRs **(**Figure 55, Table 5) and hypomethylated DMRs (Figure 56, Table 6) at CG sites. The top five enriched biological processes annotated for both hypermethylated and hypomethylated DMRs at CG sites were DNA or RNA metabolism, Unknown biological processes, Developmental processes, Transcription, DNA-dependent process and cell organization and biogenesis (Table 5-6) for all compared groups; F25H vs. F2C, F25C vs. F2C, and F25H vs. F25C. Others such as response to stress and stimuli were also enriched. For hypomethylated DMRs in F25H vs F2C and F25H vs. F25C groups and hypermethylated DMRs in F25H vs. F2C group at CHG context (Table 7) were classed as "unknown" and "other" which could mean the function of these genes is yet to be identified in *Arabidopsis.* This also occurred in CHH context for F25H vs F25C for hypomethylated DMRs (Table 8)

At CG context, all comparison groups had hypermethylated and hypomethylated DMRs enriched in biological processes classes as "unknown" and "other" (Table 5-6). The variation was highlighted for hypomethylated DMRs in F25H vs F2C and F25H vs. F25C groups and hypermethylated DMRs in F25H vs. F2C group at CHG context (Table 7) which also had enriched biological processes classed as "unknown" which could mean the function of these genes is yet to be identified in *Arabidopsis.*

Methylated regions related to biological processes such as signal transduction was enriched in CHG regions that were hypermethylated in F25C vs F2C group (Figure 57, Table 7), while transport protein metabolism, other cellular, metabolic and unknown processes were enriched in regions that were hypomethylated in F25H vs F2C group (Figure 57). Protein

metabolism was the biological process different in hypermethylated DMRs for F25H vs F2C as it was outlined in hypomethylated DMRs for F25H vs F2C (Table 7).

Only hypomethylated DMRs at CHH were enriched significantly ($p<0.05$) in F25H vs. F25C, F25C vs. F2C and F25H vs. F2C comparison groups (Figure 58, Table 8), hypermethylated DMRs at CHH were not significantly enriched (p-values were >0.1). Transport, other metabolic and cellular processes and developmental processes were enriched for F25H vs. F2C (Figure 58). Transcription, DNA dependent, protein metabolism and other cellular processes were enriched in F25C vs F2C. Unknown biological process was the only annotated process enriched in F25H vs F25C. Hypermethylated DMRs seem to be underrepresented in any biological process at CHH site (Figure 58).

The differences outlined between samples when compared to whole *Arabidopsis* reference genome, especially for F25H group which represents multigenerational heat exposure in this study**,** indicates that methylation status of these genes regulating these outlined annotated biological processes play a role in heat stress response of *Arabidopsis.* Biological relevant DMRs may increase and persist especially when stress is frequently experienced within a generation and across generations (Ganguly, Crisp, Eichten, & Pogson, 2017) rendering insights to stress-responsive changes to DNA methylation.

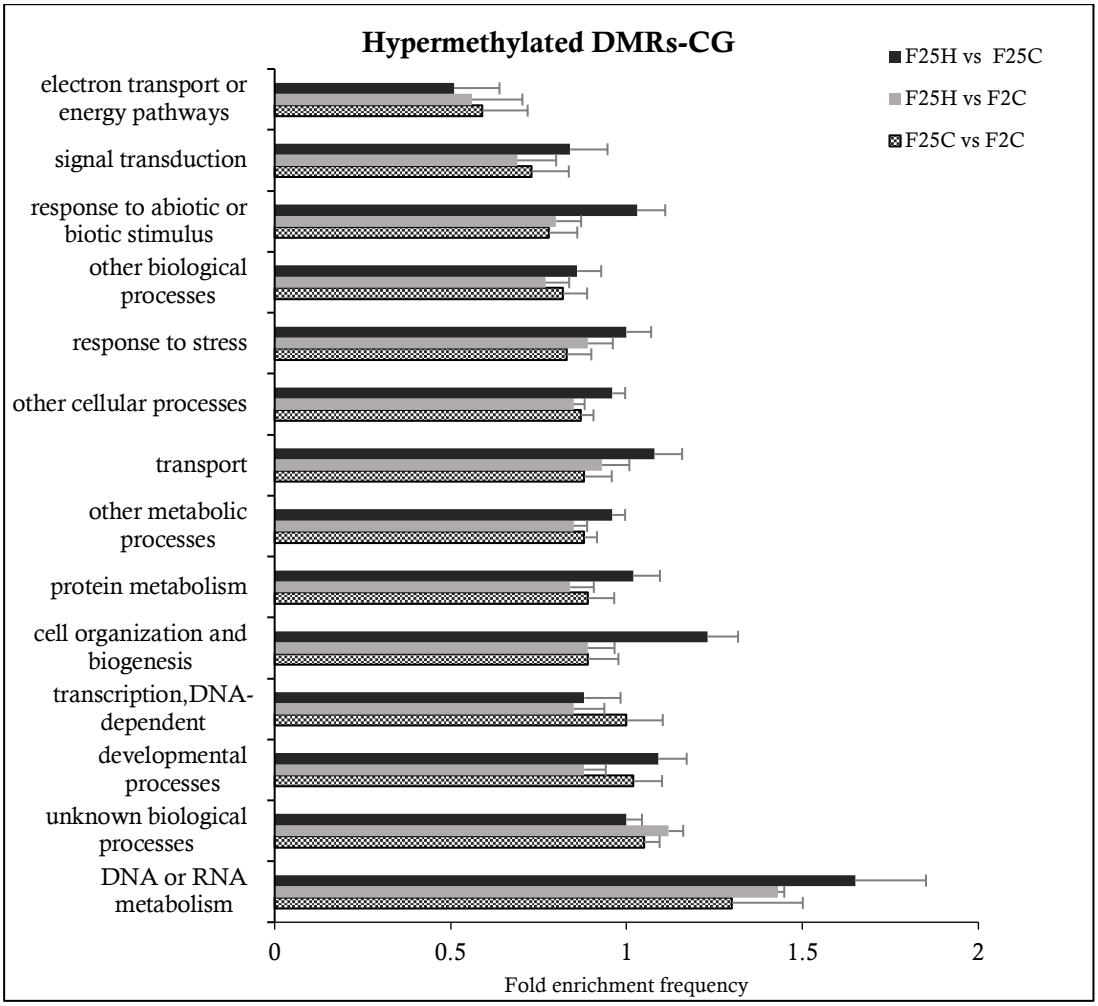

Figure 55. Enrichment analysis of hypermethylated DMRs on CG sites and their classification based on biological processes. y-axis is normalized frequency with binomial coefficients as calculated by SuperViewer. Genes are classed with p-values < 0.05, ±bootstrap StdDev.

Table 4.5: Hypermethylated DMRs enriched on CG Context

		F25C vs F2C	F25H vs F2C	F25H vs F25C
DNA or RNA Metabolism	Freq. ±SD p-value	1.3 ±0.203 0.009	*1.43* *0.018* *0.002*	*1.65* ±0.201 0.000
Unknown Biological Processes	Freq. ±SD p-value	1.05 ±0.044 0.008	*1.12* ±0.041 0.000	*1* ±0.044 0.026
Developmental Processes	Freq. ±SD p-value	1.02 ±0.072 0.031	*0.88* ±0.061 0.008	*1.09* ±0.081 0.015
Transcription, DNA-dependent	Freq. ±SD p-value	1 ±0.098 0.041	*0.85* ±0.087 0.014	*0.88* *±0.103* *0.024*
Cell Organization and Biogenesis	Freq. ±SD p-value	0.89 ±0.065 0.011	*0.89* ±0.076 0.076	*1.23* ±0.087 0.001
Protein Metabolism	Freq. ±SD p-value	0.89 ±0.066 0.007	*0.84* *±0.067* *0.002*	*1.02* *±0.075* *0.031*
Other Metabolic Processes	Freq. ±SD p-value	0.88 ±0.034 0.000	*0.85* ±0.038 0.000	*0.96* ±0.036 0.016
Transport	Freq. ±SD p-value	0.88 *±0.072* *0.012*	*0.93* ±0.078 0.028	*1.08* ±0.078 0.023
Other Cellular Processes	Freq. ±SD p-value	0.87 ±0.031 0.000	*0.85* ±0.031 0.000	*0.96* ±0.036 0.015
Response to Stress	Freq. ±SD p-value	0.83 ±0.055 0.001	*0.89* ±0.071 0.009	*1* ±0.07 0.034
Other Biological Processes	Freq. ±SD p-value	0.82 ±0.066 0.001	*0.77* *±0.067* *0.000*	*0.86* ±0.068 0.006
Response to Abiotic or Biotic Stimulus	Freq. ±SD p-value	0.78 ±0.057 0.000	*0.8* ±0.071 0.000	*1.03* ±0.08 0.032
Signal Transduction	Freq. ±SD p-value	0.73 ±0.095 0.001	*0.69* ±0.11 0.001	*0.84* ±0.106 0.022
Electron Transport or Energy Pathways	Freq. ±SD p-value	0.87 ±0.133 0.004	*0.56* ±0.144 0.003	*0.51* ±0.129 0.002

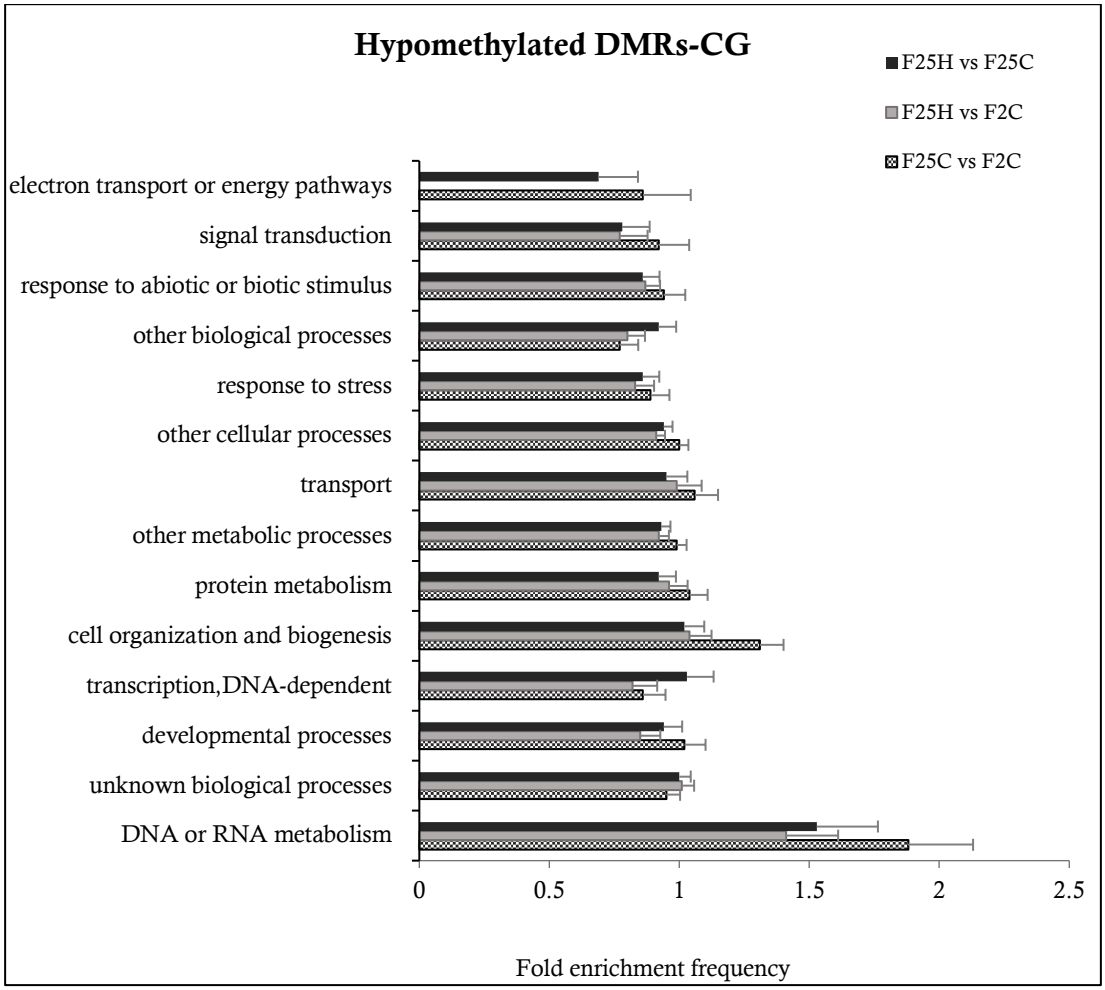

Figure 56. Enrichment analysis of hypomethylated DMRs on CG sites and their classification based on biological processes. y-axis is normalized frequency with binomial coefficients as calculated by SuperViewer. Genes are classed with p-values < 0.05, ±bootstrap StdDev.

Table 4.6: Hypomethylated DMRs enriched on CG Context

		F25C vs F2C	F25H vs F2C	F25H vs F25C
DNA or RNA Metabolism	Freq. ±SD p-value	1.88 ±0.25 0.000	1.41 ±0.201 0.006	1.53 ±0.234 0.000
Unknown Biological Processes	Freq. ±SD p-value	0.95 ±0.053 0.016	1.01 ±0.047 0.027	1 ±0.044 0.025
Developmental Processes	Freq. ±SD p-value	1.02 ±0.081 0.035	0.85 ±0.077 0.006	0.94 ±0.071 0.026
Transcription, DNA-dependent	Freq. ±SD p-value	0.86 ±0.087 0.02	0.82 ±0.095 0.012	1.03 ±0.102 0.039
Cell Organization and Biogenesis	Freq. ±SD p-value	1.31 ±0.091 0.000	1.04 ±0.084 0.034	1.02 ±0.076 0.033
Protein Metabolism	Freq. ±SD p-value	1.04 ±0.069 0.027	0.96 ±0.072 0.032	0.92 ±0.067 0.018
Other Metabolic Processes	Freq. ±SD p-value	0.99 ±0.038 0.025	0.92 ±0.04 0.005	0.93 ±0.036 0.004
Transport	Freq. ±SD p-value	1.06 ±0.089 0.029	0.99 ±0.096 0.041	0.95 ±0.081 0.033
Other Cellular Processes	Freq. ±SD p-value	1 ±0.035 0.025	0.91 ±0.035 0.002	0.94 ±0.034 0.005
Response to Stress	Freq. ±SD p-value	0.89 ±0.072 0.012	0.83 ±0.073 0.003	0.86 ±0.063 0.004
Other Biological Processes	Freq. ±SD p-value	0.77 ±0.072 0.00	0.8 ±0.068 0.001	0.92 ±0.068 0.02
Response to Abiotic or Biotic Stimulus	Freq. ±SD p-value	0.94 ±0.083 0.027	0.87 ±0.056 0.009	0.86 ±0.064 0.005
Signal Transduction	Freq. ±SD p-value	0.92 ±0.118 0.046	0.77 ±0.108 0.011	0.78 ±0.106 0.007
Electron Transport or Energy Pathways	Freq. ±SD p-value	0.86 ±0.184 0.000	-	0.69 ±0.151 0.02

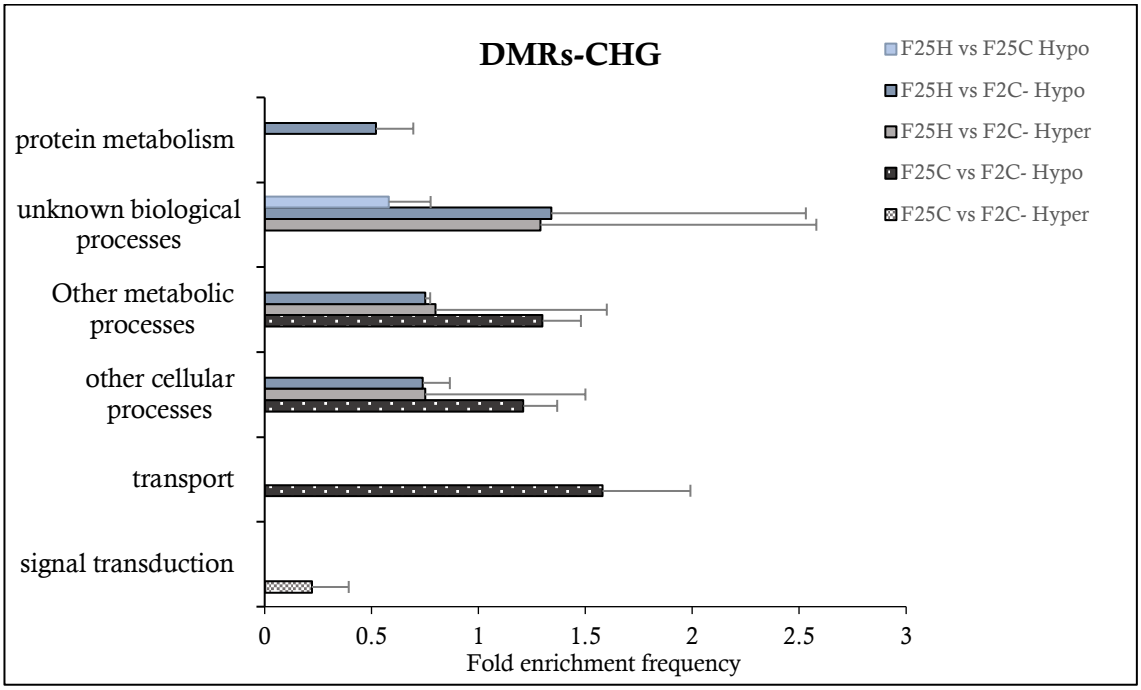

Figure 57. Enrichment analysis of hypermethylated and hypomethylated DMRs on CHG sites and their classification based on biological processes. y-axis is normalized frequency with binomial coefficients as calculated by SuperViewer. Genes are classed with p-values < 0.05, ±bootstrap StdDev.

Table 4.7: DMRs enriched on CHG Context

		Hypomethylated			Hypermethylated	
		F25C vs F2C	F25H vs F2C	F25H vs F25C	F25C vs F2C	F25H vs F2C
Signal Transduction	Freq. ±SD p-value	-	-	-	0.22 ±0.173 0.048	-
Transport	Freq. ±SD p-value	1.58 ±0.411 0.41	-	-	-	-
Other Cellular Processes	Freq. ±SD p-value	1.21 ±0.164 0.04	0.74 ±0.015 0.043	-	-	0.75 ±0.146 0.021
Other Metabolic Processes	Freq. ±SD p-value	1.3 ±0.179 0.02	0.75 ±0.024 0.024	-		0.8 ±0.164 0.043
Unknown Biological Processes	Freq. ±SD p-value	-	1.34 ±1.191 0.015	0.58 ±0.196 0.02	-	1.29 ±0.208 0.027
Protein Metabolism	Freq. ±SD p-value	-	0.52 ±0.175 0.029	-	-	-

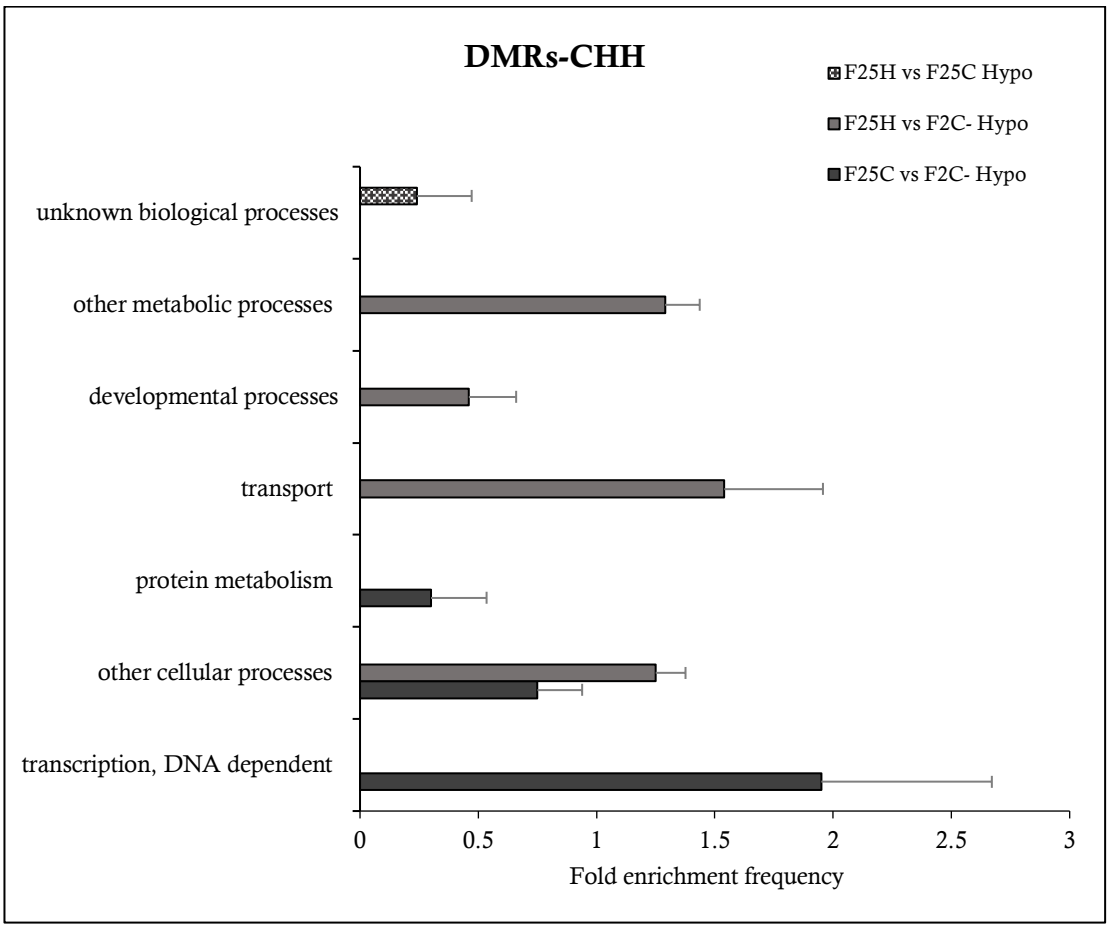

Figure 58. Enrichment analysis of hypomethylated DMRs on CHG sites and their classification based on biological processes. y-axis is normalized frequency with binomial coefficients as calculated by SuperViewer. Genes are classed with p-values < 0.05, ±bootstrap StdDev.

Table 4.8: Hypomethylated DMRs Enriched On CHH Context

		F25C vs F2C	F25H vs F2C	F25H vs F25C
Transcription, DNA Dependent	Freq. ±SD p-value	1.95 ±0.721 0.04	-	-
Other Cellular Processes	Freq. ±SD p-value	0.75 ±0.189 0.049	1.25 ±0.126 0.01	-
Protein Metabolism	Freq. ±SD p-value	0.3 ±0.235 0.24	-	-
Transport	Freq. ±SD p-value	-	1.54 ±0.417 0.03	-
Developmental Processes	Freq. ±SD p-value	-	0.46 ±0.2 0.02	-
Other Metabolic Processes	Freq. ±SD p-value	-	1.29 ±0.146 0.01	-
Unknown Biological Processes	Freq. ±SD p-value	-	-	0.24 ±0.232 0.05

4.4 CHAPTER 4 REFERENCES

Akalin, A., Franke, V., Vlahoviček, K., Mason, C. E., & Schübeler, D. (2014). Genomation: A toolkit to summarize, annotate and visualize genomic intervals. *bioinformatics, 31*(7), 1127-1129.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., . . . Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, snpeff: Snps in the genome of drosophila melanogaster strain w1118; iso-2; iso-3. *Fly, 6*(2), 80-92.

Ganguly, D., Crisp, P. A., Eichten, S. R., & Pogson, B. J. (2017). The arabidopsis DNA methylome is stable under transgenerational drought stress. *Plant physiology*.

5. DISCUSSION

5.1 Multigenerational Effects and Phenotypic Changes in Response to Heat Stress

Plant phenotype can be affected by parental environment (Groot et al., 2016). How long these parental effects last across off-spring generations is a question the science community thrives to clarify. Although it has been reported that multigenerational exposure changes parental effects compared to single generational exposures (otherwise known as transgenerational effects), multigenerational effects of stress exposure on off-spring phenotype elicit different outcomes depending on dose effects and off-spring environment (Groot et al., 2016). Though our study did not begin or observe the response of G_0 plants to heat stress, observations on changes in phenotypes traits that occurred in G_2 in comparison to its off-spring generations was achieved by comparing changes in heat-stressed plants to plants grown under normal conditions. This study explored parental effects on plant phenotypes and differences in response to multigenerational exposure to heat stress over twenty-five generations and can expand on the concept that expression of parental effects does not only depend on the off-spring environment but also on the number of past exposures.

5.1.1 Leaf Number

Changes in leaf number were observed in parent and progeny plants. These variations occurred with the parent, and off-spring plants grown under stress and ambient conditions, progeny plants that were generationally stressed (S lineage) or non-stressed (NS lineage) had differences based on temperature treatment. Heat stress can transition plants into an early and increased rate of development as well as significantly increasing number of leaves (Prasad, Boote, & Allen Jr, 2006; Prasad, Pisipati, Mutava, & Tuinstra, 2008), thereby decreasing growth interval, seed and

plant yield (Yaish, Colasanti, & Rothstein, 2011). Interestingly, progeny generation G_{10} had the most difference in leaf number compared to G_{20} and G_{25} at all three temperatures which may be attributed to G_{10} having the most transition to increased development due to stress. Also, it could be that G_{10} had the most expression of epigenetic memory of parental stress compared to the other progeny generations which are further down the ancestral timeline. Some reports have mentioned the effects of epigenetic inheritance or transference of transgenerational stress memory to be restricted to the number of generation without the maintenance of stress (Pecinka et al., 2009; Suter & Widmer, 2013a). It is possible that transgenerational effects can only extend to a particular generation or be induced by specific environmental factors experienced by the progeny.

Also, within G_{10}, parental effects (i.e. lineage treatment; NS and S) showed variations in leaf number across the different temperature treatment. Parental non-stressed plants (NS-NT and NS-T) and parental stressed (S-T and S-NT) plants had the same number of leaves at 37°C and varied at 42°C and 50°C for G_{10}. This trend was not consistent in other progeny generations and could be attributed to the intensity of temperature stress affecting the expression of multigenerational and transgenerational effects which can be sensitive to timing, duration and severity of the environmental factor. Moreover, the expression of transgenerational effects in the form of changes of phenotype responses strongly depends on offspring environment (Boyko & Kovalchuk, 2010; Groot et al., 2016) and varies between genotype and trait (Suter & Widmer, 2013b; Verhoeven & van Gurp, 2012). Changes in leaf number were not consistently impacted by parental treatment in all generations. In some cases, it was evident that parental heat stress impacted the response of off-spring generation providing evidence that heat stress accelerates leaf production in an attempt to speed up the development in the progeny of stressed plants (Liu, Feng, Li, & He, 2015; Porter, 2005). Also, under normal conditions, leaf number changes in the progeny

of non-stressed (NS) plants that were heat stressed may be under the control of heat-shock proteins. For example, HSP101 is present in mature seeds of plants grown in the absence of stress (Hong & Vierling, 2001) and reported to be essential for resolubilizing proteins following heat stress. Our previous studies showed that heat-induced modifications on gene expression might have a role in phenotype changes such as leaf number and flowering time (Migicovsky, Yao, & Kovalchuk, 2014).

5.1.2 Bolting Time

In *Arabidopsis*, the elongation of the first internode refers to bolting transition which initiates the first flower (floral transition) marking the change into the reproductive growth phase. However, reproductive phase change can lead to other morphological changes as well as be initiated by inductive treatments from environmental conditions such as photoperiodic and stress treatments (Pouteau & Albertini, 2009). As previously mentioned, heat increases plant growth reproductive development and, in this instance, most likely prompted plants to rapidly advance from rosette growth stage to bolting at higher degrees. In this study, the percentage of bolted plants in each group population was an indicator of bolting time. Earlier bolting in heat-stressed plants (T) in comparison to non-stressed plants (NT) may be attributed to the accelerated change to reproductive phase induced by the stressing condition. The specific mechanism behind the transgenerational inheritance of environmental stress is unknown, but possibly temperature treatments trigger changes in plant hormone regulatory growth cycle (Kotak et al., 2007), thereby initiating early development and reproductive growth phase as a way to ensure survival.

Heat elevated the rate of bolting in the progeny of stressed and non-stressed plants in progeny generation but not in parental generation G_2. Curiously at 50°C parental non-stressed

plants bolted earlier than stressed plants. Multigenerational effects on progeny generation affected bolting time of plants grown under normal and stressed conditions depending on parental treatment (NS or S). However, plants under heat-stressed conditions bolted more than plants grown under normal conditions. Similarly, Brachi et al., (2012) reported a selection for early bolting in the progeny of water-stressed plants. Also, shorter flowering duration and inheritance of early bolting were more efficient with increased stress severity (Brachi et al., 2012).

The response of early bolting was most extensive in general at 50°C as all progeny plants groups regardless of parental treatment, bolted earlier than the parental generation. Progeny of heat-stressed plants bolted even under normal conditions (S-NT) but not as much as S-T especially at 42°C, which attests to the theory that transgenerational inheritance is relational to severity of stress. Also, provides further evidence that parental heat stress may increase bolting even under normal growth conditions to portably compensate for reduced plant height which occurs in response to heat stress (Cicchino, Edreira, Uribelarrea, & Otegui, 2010).

5.1.3 Fresh and Dry Weight

The weight of a plant is useful in determining active growth and development of the plant (Dornbos Jr & Mullen, 1991). Progeny plants weighed more than parental plants and heat stress impacted the weight of progeny generations when compared to parental plants. G_{25} plants were the most impacted in weight change at all temperatures, progeny of stressed plants that were non-heat treated (S-NT) had more fresh weight at 37°C and 42°C while at 50°C, the progeny of stressed plants that were heat treated (S-T) had more fresh weight. The degree or severity of high temperature stress can impact cellular homeostasis (Kotak et al., 2007) and exposure to above

optimal growing temperature over time can result in the acquainted thermotolerance (Bray, 2000; Sung, Kaplan, Lee, & Guy, 2003).

5.1.4 Seed Length

Heat stress has been reported to decrease seed size which results in reduced yields (Prasad, Staggenborg, & Ristic, 2008; Sadras, 2007), similar findings of decreased seed size were reported in soybean under water stress and high air-temperature stress (Dornbos Jr & Mullen, 1991). Reduction in seed size has also been linked to decreasing seed germination (Dornbos Jr & Mullen, 1991). The opposite effect in seed size has also been reported, in that larger seeds increase seed survival as a selective advantage in nutrient deprivation (Krannitz, Aarssen, & Dow, 1991). In our study, seed length was measured as an indicator of seed size, parent and progeny seeds varied based on temperature treatment. For instance, seeds heat treated at 37°C that were from stressed linage (S-T) were smaller than seeds from non-stressed lineage (NS-T). Similar observation occurred in progeny generation ten at 50°C. In other instances, seeds from G_{25} plants grown under normal condition that were of NS lineage (NS-NT) were smaller in seed size when compared with seeds from S lineage (S-NT).

In some plant species, parental effects associated with temperature stress has included an increase in biomass and quantity of reserve in seeds (Blödner, Goebel, Feussner, Gatz, & Polle, 2007; Galloway, 2001). While a decrease in seed size is indicative of the plants' effort to conserve energy resources and complete its lifecycle rapidly in the expense of overall yield, increase in seed size can also be observed as an effort to ensure longevity and continuance into the next generation. It is also possible that the effects of multigenerational stress impacted the generation towards smaller seed length as an adaptive trait (Morrison & Stewart, 2002) and the minimal differences

between seed sizes irrespective of the lineage, whether it was a progeny with transgenerational stress or a progeny with multigenerational stress was also indicative of acquired tolerance over time.

In general, comparing transgenerational and phenotype changes in response to heat stress, the progeny of stressed plants showed the various range of phenotype changes associated with heat stress responses especially in leaf number and bolting. Plants can benefit from a tremendous or minimal change to their growth cycle depending on what adaptive feature is needed for instance a decrease in leaf number can allocate more resources to other growth phrases thereby resulting in higher biomass and seeds after the heat stress has been removed (Migicovsky et al., 2014). The responses seen from the progeny of non-stressed plants under ambient conditions gave insights on transgenerational effects without the maintenance of stress and the extension of ancestral effects on progenies. It is possible that progeny of non-stressed plants maintained a similar phenotype to parental plants while stressing plants which maintained exposure to stress overtime had such exposure trigger the observed phenotype changes irrespective of if the stress was present or absent. However, under stress conditions the progeny of stressed plants tend to exhibit the phenotype that assisted in tolerating the stressing environmental factor, this likely occurred due to inherited epigenetic memory that allowed for specific favourable response to the heat stress. Although multigenerational exposure often reduces the expression of the parental effects compared to single generation exposure (Groot et al., 2016). In this study some level of high offspring performance under normal and heat conditions was observed as well as multigenerational effects, the length of exposure to generational stress also exerted strong effects.

5.2 Genome-wide Alterations in DNA Methylation

Transgenerational plant stress memory is often linked to DNA and chromatin alterations as the potential mechanism underpinning transmission of stressing experiences along plant lineage (Crisp, Ganguly, Eichten, Borevitz, & Pogson, 2016). Studies point to the existence of stress-induced variations in DNA methylomes and an adaptability of DNA methylome in plants responding to environmental conditions. Differential cytosine methylation, a form of heritable epigenetic polymorphism, underlies phenotypic variation. In order to understand epigenetic inheritance, the differences in plant epialleles and its effects on the expression of nearby genes is to be considered as a means of DNA adaptability in plants (Becker et al., 2011; Paszkowski & Grossniklaus, 2011; Richards, 2008)**.**

Analysis of whole genome cytosine methylation with the assistance of sequencing technologies has reported about one single-base-pair mutation per haploid genome and generation in *A. thaliana* mutation accumulation lines derived from a single ancestor (Ossowski et al., 2010). Genome-wide DNA methylation of 10 *A. thaliana* derived over 30 generations from a single-seed descent was examined and about 30,000 cytosines were found to be differentially methylated in each strain (Shaw, Byers, & Darmo, 2000). This study compared 31st generation individuals with two independent lines of 3rd generation (closer to the founder generation) to reflect differences accumulated by the 31th generation and reported CG sites to be over-represented among differentially methylated positions (Becker et al., 2011). In our study, we examined whole genome-cytosine methylation and compared 25th generation individuals with 2nd generation (because seeds from founders were no longer available); changes shared within these lines should reflect differences accumulated by the 25th generation. All cytosine residues among the three differentially methylated positions (sequence context) were represented with distribution

percentages consistent with observations reported in plants. The overall genome-wide DMCs averages of 31% CG, 35% CHG and 34% CHH methylation observed in our analysis were slightly different compared to previous studies (Cokus et al., 2008; Lister et al., 2008) which could be attributed to the average > 10 coverage depth used in sequencing analysis. Differences in methylation averages and distribution highlights the notion that each type of methylation is under specific genetic control (Cokus et al., 2008; Richards, 2008). Higher averages in genome wide methylation is expected within plant genome as large fractions of methylation can also be found in CHG sequence context (Chan, Henderson, & Jacobsen, 2005). The average number of global methylated cytosine bases were similar for both non-stressed parental (F2C) and advanced progeny (F25C) with a considerable difference for stressed advanced progeny (F25H). Although the methylation percentages of each sequence context were very similar in F25H, F25C and F2C, cytosine methylation was mostly distributed in symmetrical sequences which is consistent with patterns found in plants (Finnegan, Genger, Peacock, & Dennis, 1998; Law & Jacobsen, 2010; Niederhuth et al., 2016).

Similarities in genome showed patterns based on heat stress treatment as hierarchical clustering of parental plants and non-stressed progeny were aligned while stressed progeny were separate which affirmed the homogeneity of our seed line from the single-parent origin. Same clustering patterns were outlined by DMCs and DMRs which indicated that the accumulated methylation differences between progeny and parent were captured. Correlation was highest among three individuals of non-stressed progeny (F25C) out of the five tested. Each individual of parental generation (F2C) was more similar to these three F25C progeny lines. Generally, these patterns reflect stable whole-genome methylation patterns which are inheritable in *A. thaliana* and

also points to the gradual accumulation of differences in methylation status similar to genetic mutations (Becker et al., 2011).

Differentially methylated regions can occur in CG context (CG-DMR) or CG, CHG and CHH context (C-DMRs) (Niederhuth & Schmitz, 2014) and as such the inheritance of methylation variant or epimutations in progenies can be seen in either of these sequence contexts. Methylation patterns of DMRs studied in epigenetic mutants has linked DNA methylation to response to environmental signals like heat (Popova, Dinh, Aufsatz, & Jonak, 2013) and salt (Yao, Bilichak, Golubov, & Kovalchuk, 2012) stress. The significance of these studies suggest changes in DNA methylation specifically targets an epigenetic system that enhances gene expression profiles deviant from the wild type (Meyer, 2015; Popova et al., 2013). We identified more hyper-methylated regions between non-stressed progeny and parent generation and mostly hypo-methylated regions between stressed and non-stressed progeny generation. Of the 5 lines of stressed progeny, 2 lines had more DMRs than the remaining three (which where the 3 lines that correlated highest with non-stressed progeny). Comparison between stressed progeny generation and parent generation revealed mostly hyper-methylated regions. Plants with previous multiple exposures to drought stress convey enhanced ability to respond to future or new stress by rapid adaptive changes to gene expression patterns compared to plants with no previous exposure (Y. Ding, Fromm, & Avramova, 2012; Ganguly, Crisp, Eichten, & Pogson, 2017).

DMCs were widely distributed on chromosome arms and displayed hypermethylation and hypomethylation across all five chromosomes, with DMCs in CG context having higher percentage distribution which agrees with other plant studies that have found high gene density of CG-DMCs on chromosome arms (Becker et al., 2011; Zheng et al., 2013). Differential cytosine methylation between parent and progeny were considered to be induced by the heat stress;

comparison of DMRs in G_{25} and G_2 profiles identified a predominance of aberrant hypermethylation in F25H and F25C when both were compared to F2C and hypomethylation in F25H when compared to F25C which may indicate advance generation was hypermethylated. G_{25} showed variations in DNA methylation reflected mainly in F25H group and hierarchical clustering analysis of DMRs in G_{25} indicates a more considerable difference in methylation status in response to multigenerational heat stress exposure. Compared to G_2, the DNA methylation patterns of G_{25} changed after being cultivated successively for twenty-five generations under heat stress, DMRs were noticeably induced by heat stress within generations.

DNA methylation can occur non-randomly and cluster in specific segments of the genome (Becker et al., 2011; Vaughn et al., 2007; Zhang et al., 2006). The functional consequences of DNA methylation are influenced by genomic location, basal DNA sequence and site class. Gene body methylation is found within coding sequences with a potential to contribute to phenotypes (Bewick & Schmitz, 2017) and CG sequences are usually found to cluster around regulatory region of genes; in promoters and first exons (Zheng et al., 2013). DMCs were concentrated in promoter and exon regions for CG sites while in CHH sites, DMCs were mostly distributed in the promoter and intergenic regions. Methylated CG sites were enriched within gene body for all groups, F25H and F25C were mostly methylated at gene body with higher average methylation percentile than F2C which is consistent with gene body methylation and correlates with low expression of genes (Bewick & Schmitz, 2017).

Expressed genes are likely to be unmethylated and transcription extremes are mostly depleted in mCG which proposes a dependence on RNA polymerase transit and suggests a link between patterns of DNA methylation and transcription elongation (Zilberman, Gehring, Tran, Ballinger, & Henikoff, 2007) and such reduced promoter methylation correlates with enhanced

expression (Zemach, McDaniel, Silva, & Zilberman, 2010). Most DNA methylated genes compared to unmethylated genes often consist of housekeeping, constitutively expressed long (bp) genes conserved within gene bodies among species (Bewick & Schmitz, 2017). It is possible that these bulky genes represent the higher percentages observed in intergenic regions in CHH sites. Enrichment of mCHG and depletion of mCHH within transcribed regions has been reported as a class of methylated genes in angiosperms resulting in genes in this area to be typically expressed at lower levels than mCG and mCHH sites indicating that mCG may or may not be enriched all the time (Bewick & Schmitz, 2017; Niederhuth et al., 2016). DNA methylation in stressed plants increased and stress-induced variations in the epigenetic landscape of stressed plants were apparent as well and are consistent with previous studies which found environmental stresses induced changes in rice (Zheng et al., 2013).

5.3 Stress-Associated Variations in DNA Methylation

Natural selection and consecutive phenotypic selection are the driving forces behind evolution which is why scientists use environmental changes to challenge selection forces on model organisms in search of adaptable phenotype traits. The differences or changes in DNA methylation patterns induced by specific stressors indicates an epigenetic response to various environmental conditions and supports the basis of stress-induced DNA methylation that is proposed to exist in plant stress-responsive mechanisms. Growing evidence points to the role of epigenetic mechanisms in coordinating inheritable and reversible changes in gene expression in *Arabidopsis* and other plants (Bilichak, Ilnystkyy, Hollunder, & Kovalchuk, 2012; Li et al., 2012). To observe the relationship between heat stress and DNA methylation variations, our study explored patterns of DNA methylation polymorphisms over twenty-five generations of

multigenerational stress exposure and considered genome-wide methylation profiles in *Arabidopsis* under temperature stress.

DNA methylation profiles for both parental and off-spring generations outlined numerous genes and found methylation in different genetic regions especially in the promoter and exon regions which were relative to cytosine sequence context suggesting the potential role of DNA methylation in downstream gene regulation. Among CHG sites, relative abundance of DMCs was found in promoter and intergenic regions while among CG sites it was mainly in the promoter and coding regions agreeing with genomic studies in *A. thaliana* which revealed endogenous genes to be methylated either within promoter or within their transcribed regions (Cokus et al., 2008). Previous studies have shown promoter methylation to correspond with negatively expressed genes and gene body methylation to correspond to positively expressed genes. For example, genes with promoter regions overlapping with DMRs in *Populus nigra* were found to repress the expression of downstream genes. These genes were shown to be involved in multiple biological processes which point to the significance of DNA methylation in plant cells (C.-J. Ding, Liang, Diao, Su, & Zhang, 2018).

The variations in DNA methylation included hypomethylation and hypermethylation events which is consistent with previous studies that found environmental stresses induced DNA methylation changes in other plants (Migicovsky & Kovalchuk, 2014; Suter & Widmer, 2013a; Uthup, Ravindran, Bini, & Thakurdas, 2011; Verhoeven & van Gurp, 2012), rice (Feng et al., 2012; Wang et al., 2011; Zheng et al., 2013) and other models (Norouzitallab et al., 2014). DMCs found in our studied methylomes were seen to cluster by treatment with the progeny of non-stressed plants clustering with parental plants suggesting that there was no predominant source of variation in the DNA methylome between F25C and F2C. However, the progeny of stressed plants

mostly clustered separately from the progeny of non-stressed plants suggesting that heat stress-induced conserved variations in F25H and indicates the malleability of the *Arabidopsis g*enome to abiotic stress (Ganguly et al., 2017). The different patterns observed could be because stress-induced modifications can be reset to basal level after the removal of the stress, and some can be transmitted as stress memory (Chinnusamy & Zhu, 2009).

The epimutations found between F25H vs F25C and F25H vs F2C were considered to be potentially induced by heat stress in G_{25}. Hierarchical clustering of these epimutations separated mutations based on the influence of heat stress on the genome where advanced generations had regions that were more methylated than parent generations. DMRs gives insight on the biodiversity of revealed regions that exhibit consistent differences in the most biological variations (Hansen, Langmead, & Irizarry, 2012). Although most studies with *Arabidopsis* show stress memory to a particular lineage, for example, DMRs increased and persisted over repeated stress experience within a generation and one trait (increased seed dormancy) out of six was seen in descendants of plants drought stressed for five generation (Ganguly et al., 2017). DMRs appeared after the onset of heat stress within multiple generations which is consistent with studies in *Arabidopsis*.

Genomic variants comprising of single nucleotide polymorphisms (SNPs) and insertions and deletions (INDELs) are helpful in identification of allele-specific epigenetic events and to quantify absolute methylation levels. The positions of SNPs may add to transcription, or translation effects on the genome depending on their corresponding genomic location such as intronic and untranslated regions as well as predict coding effects in respect to protein-coding genes such as synonymous or non- synonymous amino acid replacement (Cingolani et al., 2012). Variants in F25H were mostly found upstream and downstream and were significantly more in F25H than F25C or F2C. Although the number of SNPs were similar between F25C and F2C, the

polymorphisms between ancestral and heat stressed descendants were different which is in line with heritable variations found in *Arabidopsis* plant. For example, flowering time and plant height in *Arabidopsis* was due to stable inheritance of epialleles which included DNA insertion variants rather than accumulated new epimutation linked to a phenotype trait (Johannes et al., 2009).

Although effects of variants were represented mostly as synonymous coding variants, a few represented variants responsible in changing amino acid coding triplets were gained and could be related to enhancing a specific protein function as SNPs upstream or downstream might affect transcription or translation (Cingolani et al., 2012).

Genes with SNPs involved in a wide range of function was identified in progeny stressed and non-stressed plants some of which were involved in the response of abiotic and biotic stimulus in plants according to the GO analysis. DMRs were enriched mostly at CG sites and in processes like transcription and DNA dependent processes, DNA or RNA metabolism but SNPs were enriched mostly at unknown biological processes which probably could indicate that variants impact are yet to be identified as at the time of analysis.

Possible variations could be due to seed age storage and or selection for seed survival as DNA methylation is predominantly static over relatively long periods time and changes in cytosine methylation occur at a frequency higher than that of mutations observed at DNA sequence levels (Schmitz et al., 2011). Heat stress had a significant cumulative effect on *Arabidopsis* methylome, the DNA methylation patterns of G_{25} changed after being cultivated successively for twenty-five generations under heat stress and was significantly enriched in biological processes that are relevant to plant responses to stress.

5.4 Directional/Non-Random Changes of Heat-Induced Epimutations

Epimutations (DMRs) between G_{25} and G_2 were considered to be potentially induced by stress and accumulated after twenty-five generation of heat stress. Hierarchical clustering of these epimutations separated F25H plants and F2C into distinct groups suggesting that long-term heat stress impacted the DNA methylation patterns of these regions. Such influence is considered directional as the epimutations in offspring generation grouped siblings together and was not random, similar patterns were observed in rice varieties that were drought stressed for eleven successive generations (Zheng et al., 2017). Epimutations separated parent G0 rice plants and other generations in discrete clusters for both varieties tested suggesting the DNA methylation patterns in these regions were specific in occurrence. Also, these epimutations were comparable to drought induced site-specific DNA methylation in rice and associated with drought tolerance (Wang et al., 2010a; Zheng et al., 2017).

DNA methylation is established to cluster in specific segments of the genome and non-random patterns (Becker et al., 2011; Tran et al., 2005; Vaughn et al., 2007; Zhang et al., 2006). Methylation changes shared within a line should broadly reflect differences that had accumulated over generations (Becker et al., 2011). In our study, the accumulated differences show methylation changes in the twenty-fifth generation triggered by heat treatment (from comparing F25H versus F2C). We identified 5560 hypermethylated DMRs that were at least 100bp long between F25H and F2C and 4754 hypomethylated DMRs. Hierarchical clustering according to DMRs distinctly separated parent and progeny lines into groups which allows us to consider methylation status in 2nd generation individuals to be the reflection of their ancestral pattern and the direction of heat-induced epimutations.

Non-random distribution of DNA methylation and enrichment levels were linked with high expression percentiles within *Arabidopsis thaliana* genes in the genome-wide study conducted by Zilberman and colleagues (2007). In this study, the relationship between DNA methylation and transcription was highlighted and showed gene transcription to be influenced by DNA methylation where highest-transcribed genes were found to be undermethylated and methylation within longer genes were likely to be sustained imposing a non-random relationship between gene expression and size (bp percentiles) overtime as the short methylated genes were poorly expressed (Zilberman et al., 2007).

DNA methylation variants and its effects on the expression of nearby genes in transgenerational inheritance of affected traits is as significant as DNA sequence mutations when materials for evolutionary change is being considered (Paszkowski & Grossniklaus, 2011; Richards, 2008). Importantly, our genome-wide observation was consistent with previous reports linking stress-induced methylation changes to transgenerational inheritance in plants system. Differential methylation events and their genomic effects indicated nucleic metabolic pathways and responsive pathways to be mostly enriched. As Gene Ontology analysis of methylated DMRs revealed the product of these genes in *Arabidopsis* to be involved in DNA or RNA metabolism, responses to abiotic and biotic stress, transport, response to stimulus and cell organization and biogenesis. These pathways are associated with plant development and stress response. There is the possibility of a change in individual stress response genes between progeny and parent generation which is an interesting concept to analyze in future studies.

DNA methylation changes accompanied with an environmental stimuli usually result in a decrease or increase in cytosine methylation within the entire genome or at specific loci (Gao et al., 2014; Lukens & Zhan, 2007; Tan, 2010; Wang et al., 2010b). Certain studies have indicated a

relative stress tolerance in genotypes studied by showing that differential DNA methylation occurs among single genotypes as well as specific genotypes with stress-tolerant or sensitive traits. For instance, in rice the relationship between DNA methylation and abiotic stress tolerance revealed genotype dependent responses to salt stress with gene expressions that varied with genotypes irrespective of level of tolerance of rice genotype (Karan, DeLeon, Biradar, & Subudhi, 2012). In cotton, methylation changes were seen to be tissue specific under different salt stresses, methylation levels of salt tolerant cultivar roots differed from leaves and methylation levels of salt sensitive cultivar increased continuously under salt stress (Lu et al., 2015). In another study, the seedlings of heat-sensitive (Fengyou 1) and heat-tolerant (Huyou 2) *B. napus* (Gao et al., 2010) showed changes in genomic methylation patterns under heat stress and showed seedlings from heat-sensitive genotype had relatively increased cytosine methylation events occurring (under heat stress and control conditions) than those of heat-tolerant genotype (Gao et al., 2014). The methylation differences outlined indicate that differential DNA methylation occurs within specific genotypes in certain species and highlights an association between methylation changes and stress tolerance (Gao et al., 2014).

A definite correlation between changes in methylation and heat tolerance that occur from direct stress treatment or as a result of an association with the incidence of stress is hard to ignore. Plants use collective effects of assorted metabolic pathways to survive various stresses (like salt, heat and drought) as well as induce changes in DNA methylation levels (Lu et al., 2015). However, thermotolerance in plants as a means of adaption has been extensively explored to establish an association between DNA methylation changes and conferred tolerance as a relative mechanism of plant adaptation (Ganguly et al., 2017; Gao et al., 2014; Karan et al., 2012).

Our study points to DNA methylation changes at a genome-wide scale and reflects differences accumulated and faithfully transmitted by 25th generation indicating offspring tolerance to temperature change. Our data support a model whereby DNA methylation not only impacts persistent control over some stress defense genes during non-stressful conditions but also in response to environmental stimuli which can change dynamically to alter gene expression (Dowen et al., 2012; Zilberman et al., 2007) and allow for an adaptation to the environment. The potential for heat-induced formation of heritable epialleles that would correlate with heat-responsive gene expression is the next direction for our study where the stress responses of individual plants from parent and progeny generation would be explored and identified based on the genomic variants found in this report.

5.5 CHAPTER 5 REFERENCES

Becker, C., Hagmann, J., Muller, J., Koenig, D., Stegle, O., Borgwardt, K., & Weigel, D. (2011). Spontaneous epigenetic variation in the arabidopsis thaliana methylome. *Nature, 480*(7376), 245-249.

Bewick, A. J., & Schmitz, R. J. (2017). Gene body DNA methylation in plants. *Current opinion in plant biology, 36*, 103-110.

Bilichak, A., Ilnystkyy, Y., Hollunder, J., & Kovalchuk, I. (2012). The progeny of arabidopsis thaliana plants exposed to salt exhibit changes in DNA methylation, histone modifications and gene expression. *PloS one, 7*(1), e30515.

Blödner, C., Goebel, C., Feussner, I., Gatz, C., & Polle, A. (2007). Warm and cold parental reproductive environments affect seed properties, fitness, and cold responsiveness in arabidopsis thaliana progenies. *Plant, cell & environment, 30*(2), 165-175.

Boyko, A., & Kovalchuk, I. (2010). Transgenerational response to stress in *arabidopsis thaliana*. *Plant Signaling & Behavior, 5*(8), 995-998.

Brachi, B., Aimé, C., Glorieux, C., Cuguen, J., & Roux, F. (2012). Adaptive value of phenological traits in stressful environments: Predictions based on seed production and laboratory natural selection. *PloS one, 7*(3), e32069.

Bray, E. A. (2000). Response to abiotic stress. *Biochemistry and molecular biology of plants*, 1158-1203.

Chan, S. W.-L., Henderson, I. R., & Jacobsen, S. E. (2005). Gardening the genome: DNA methylation in arabidopsis thaliana. *Nature Reviews Genetics, 6*(5), 351.

Chinnusamy, V., & Zhu, J.-K. (2009). Epigenetic regulation of stress responses in plants. *Current opinion in plant biology, 12*(2), 133-139.

Cicchino, M., Edreira, J. I. R., Uribelarrea, M., & Otegui, M. E. (2010). Heat stress in field-grown maize: Response of physiological determinants of grain yield *Crop Science, 50*, 1438-1448.

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., . . . Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide

polymorphisms, snpeff: Snps in the genome of drosophila melanogaster strain w1118; iso-2; iso-3. *Fly, 6*(2), 80-92.

Cokus, S. J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C. D., . . . Jacobsen, S. E. (2008). Shotgun bisulphite sequencing of the arabidopsis genome reveals DNA methylation patterning. *Nature, 452*(7184), 215-219.

Crisp, P. A., Ganguly, D., Eichten, S. R., Borevitz, J. O., & Pogson, B. J. (2016). Reconsidering plant memory: Intersections between stress recovery, rna turnover, and epigenetics. *Science advances, 2*(2), e1501340.

Ding, C.-J., Liang, L.-X., Diao, S., Su, X.-H., & Zhang, B.-Y. (2018). Genome-wide analysis of day/night DNA methylation differences in populus nigra. *PloS one, 13*(1), e0190299.

Ding, Y., Fromm, M., & Avramova, Z. J. N. C. (2012). Multiple exposures to drought'train'transcriptional responses in arabidopsis. *3*, 740.

Dornbos Jr, D., & Mullen, R. (1991). Influence of stress during soybean seed fill on seed weight, germination, and seedling growth rate. *Canadian Journal of Plant Science, 71*(2), 373-383.

Dowen, R. H., Pelizzola, M., Schmitz, R. J., Lister, R., Dowen, J. M., Nery, J. R., . . . Ecker, J. R. (2012). Widespread dynamic DNA methylation in response to biotic stress. *Proceedings of the National Academy of Sciences, 109*(32), E2183-E2191.

Feng, Q., Yang, C., Lin, X., Wang, J., Ou, X., Zhang, C., . . . Liu, B. (2012). Salt and alkaline stress induced transgenerational alteration in DNA methylation of rice ('oryza sativa'). *Australian Journal of Crop Science*, 6(5), 877.

Finnegan, E., Genger, R., Peacock, W., & Dennis, E. (1998). DNA methylation in plants. *Annual review of plant biology, 49*(1), 223-247.

Galloway, L. F. (2001). Parental environmental effects on life history in the herbaceous plant campanula americana. *Ecology, 82*(10), 2781-2789.

Ganguly, D., Crisp, P. A., Eichten, S. R., & Pogson, B. J. (2017). The arabidopsis DNA methylome is stable under transgenerational drought stress. *Plant physiology*.

Gao, G., Li, J., Li, H., Li, F., Xu, K., Yan, G., . . . Wu, X. (2014). Comparison of the heat stress induced variations in DNA methylation between heat-tolerant and heat-sensitive rapeseed seedlings. *Breeding science, 64*(2), 125-133.

Gao, G., Wu, X., Lv, X., Chen, B., Xu, K., & Yan, G. (2010). Genotype differences of seed viability in rapeseed during storage at different temperature. *Chinese Journal of Oil Crop Sciences, 32*(4), 495-499.

Groot, M. P., Kooke, R., Knoben, N., Vergeer, P., Keurentjes, J. J., Ouborg, N. J., & Verhoeven, K. J. (2016). Effects of multi-generational stress exposure and offspring environment on the expression and persistence of transgenerational effects in arabidopsis thaliana. *PloS one, 11*(3), e0151566.

Hansen, K. D., Langmead, B., & Irizarry, R. A. (2012). Bsmooth: From whole genome bisulfite sequencing reads to differentially methylated regions. *Genome biology, 13*(10), R83.

Hong, S. W., & Vierling, E. (2001). Hsp101 is necessary for heat tolerance but dispensable for development and germination in the absence of stress. *The Plant Journal, 27*(1), 25-35.

Johannes, F., Porcher, E., Teixeira, F. K., Saliba-Colombani, V., Simon, M., Agier, N., . . . Audigier, P. (2009). Assessing the impact of transgenerational epigenetic variation on complex traits. *PLoS genetics, 5*(6), e1000530.

Karan, R., DeLeon, T., Biradar, H., & Subudhi, P. K. (2012). Salt stress induced variation in DNA methylation pattern and its influence on gene expression in contrasting rice genotypes. *PloS one, 7*(6), e40203.

Kotak, S., Larkindale, J., Lee, U., von Koskull-Döring, P., Vierling, E., & Scharf, K.-D. (2007). Complexity of the heat stress response in plants. *Current opinion in plant biology, 10*(3), 310-316.

Krannitz, P. G., Aarssen, L. W., & Dow, J. M. (1991). The effect of genetically based differences in seed size on seedling survival in arabidopsis thaliana (brassicaceae). *American Journal of Botany, 78*(3), 446-450.

Law, J. A., & Jacobsen, S. E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics, 11*(3), 204-220.

Li, X., Zhu, J., Hu, F., Ge, S., Ye, M., Xiang, H., . . . Wang, W. (2012). Single-base resolution maps of cultivated and wild rice methylomes and regulatory roles of DNA methylation in plant gene expression. *BMC Genomics, 13*(1), 300.

Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, A. H., & Ecker, J. R. (2008). Highly integrated single-base resolution maps of the epigenome in arabidopsis. *Cell, 133*(3), 523-536.

Liu, J., Feng, L., Li, J., & He, Z. (2015). Genetic and epigenetic control of plant heat responses. *Frontiers in Plant Science, 6*.

Lu, X., Zhao, X., Wang, D., Yin, Z., Wang, J., Fan, W., . . . Wuwei, Y. (2015). Whole-genome DNA methylation analysis in cotton (gossypium hirsutum l.) under different salt stresses. *Turkish Journal of Biology, 39*(3), 396-407.

Lukens, L. N., & Zhan, S. (2007). The plant genome's methylation status and response to stress: Implications for plant improvement. *Current opinion in plant biology, 10*(3), 317-322.

Meyer, P. (2015). Epigenetic variation and environmental change. *Journal of experimental botany, 66*(12), 3541-3548.

Migicovsky, Z., & Kovalchuk, I. (2014). Transgenerational changes in plant physiology and in transposon expression in response to uv-c stress in arabidopsis thaliana. *Plant Signal Behav, 9*(11), e976490.

Migicovsky, Z., Yao, Y., & Kovalchuk, I. (2014). Transgenerational phenotypic and epigenetic changes in response to heat stress inarabidopsis thaliana. *Plant Signaling & Behavior, 9*(2), e27971.

Morrison, M. J., & Stewart, D. W. (2002). Heat stress during flowering in summer brassica ecorc contribution no. 02-41. *Crop Science, 42*(3), 797-803.

Niederhuth, C. E., Bewick, A. J., Ji, L., Alabady, M. S., Do Kim, K., Li, Q., . . . Udall, J. A. (2016). Widespread natural variation of DNA methylation within angiosperms. *Genome biology, 17*(1), 194.

Niederhuth, C. E., & Schmitz, R. J. (2014). Covering your bases: Inheritance of DNA methylation in plant genomes. *Molecular Plant, 7*(3), 472-480.

Norouzitallab, P., Baruah, K., Vandegehuchte, M., Van Stappen, G., Catania, F., Bussche, J. V., . . . Bossier, P. (2014). Environmental heat stress induces epigenetic transgenerational inheritance of robustness in parthenogenetic artemia model. *The FASEB Journal, 28*(8), 3552-3563.

Ossowski, S., Schneeberger, K., Lucas-Lledó, J. I., Warthmann, N., Clark, R. M., Shaw, R. G., . . . Lynch, M. (2010). The rate and molecular spectrum of spontaneous mutations in arabidopsis thaliana. *Science, 327*(5961), 92-94.

Paszkowski, J., & Grossniklaus, U. (2011). Selected aspects of transgenerational epigenetic inheritance and resetting in plants. *Current opinion in plant biology, 14*(2), 195-203.

Pecinka, A., Rosa, M., Schikora, A., Berlinger, M., Hirt, H., Luschnig, C., & Mittelsten Scheid, O. (2009). Transgenerational stress memory is not a general response in arabidopsis. *PloS one, 4*(4), e5202.

Popova, O. V., Dinh, H. Q., Aufsatz, W., & Jonak, C. (2013). The rddm pathway is required for basal heat tolerance in arabidopsis. *Molecular Plant, 6*(2), 396-410.

Porter, J. R. (2005). Rising temperatures are likely to reduce crop yields. *Nature, 436*(7048), 174.

Pouteau, S., & Albertini, C. (2009). The significance of bolting and floral transitions as indicators of reproductive phase change in arabidopsis. *Journal of experimental botany, 60*(12), 3367-3377.

Prasad, P., Boote, K. J., & Allen Jr, L. H. (2006). Adverse high temperature effects on pollen viability, seed-set, seed yield and harvest index of grain-sorghum [sorghum bicolor (l.) moench] are more severe at elevated carbon dioxide due to higher tissue temperatures. *Agricultural and forest meteorology, 139*(3-4), 237-251.

Prasad, P., Pisipati, S., Mutava, R., & Tuinstra, M. (2008). Sensitivity of grain sorghum to high temperature stress during reproductive development. *Crop Science, 48*(5), 1911-1917.

Prasad, P., Staggenborg, S., & Ristic, Z. (2008). Impacts of drought and/or heat stress on physiological, developmental, growth, and yield processes of crop plants. *Response of crops to limited water: Understanding and modeling water stress effects on plant growth processes*(responseofcrops), 301-355.

Richards, E. J. (2008). Population epigenetics. *Current opinion in genetics & development, 18*(2), 221-226.

Sadras, V. O. (2007). Evolutionary aspects of the trade-off between seed size and number in crops. *Field Crops Research, 100*(2-3), 125-138.

Schmitz, R. J., Schultz, M. D., Lewsey, M. G., O'Malley, R. C., Urich, M. A., Libiger, O., . . . Ecker, J. R. (2011). Transgenerational epigenetic instability is a source of novel methylation variants. *Science, 334*(6054), 369-373.

Shaw, R. G., Byers, D. L., & Darmo, E. (2000). Spontaneous mutational effects on reproductive traits of arabidopsis thaliana. *Genetics, 155*(1), 369-378.

Sung, D.-Y., Kaplan, F., Lee, K.-J., & Guy, C. L. (2003). Acquired tolerance to temperature extremes. *Trends in plant science, 8*(4), 179-187.

Suter, L., & Widmer, A. (2013a). Environmental heat and salt stress induce transgenerational phenotypic changes in arabidopsis thaliana. *PloS one, 8*(4), e60364.

Suter, L., & Widmer, A. (2013b). Phenotypic effects of salt and heat stress over three generations in arabidopsis thaliana. *PloS one, 8*(11), e80819.

Tan, M.-p. (2010). Analysis of DNA methylation of maize in response to osmotic and salt stress based on methylation-sensitive amplified polymorphism. *Plant Physiology and Biochemistry, 48*(1), 21-26.

Tran, R. K., Henikoff, J. G., Zilberman, D., Ditt, R. F., Jacobsen, S. E., & Henikoff, S. J. C. B. (2005). DNA methylation profiling identifies cg methylation clusters in arabidopsis genes. *15*(2), 154-159.

Uthup, T. K., Ravindran, M., Bini, K., & Thakurdas, S. (2011). Divergent DNA methylation patterns associated with abiotic stress in hevea brasiliensis. *Molecular Plant, 4*(6), 996-1013.

Vaughn, M. W., Tanurdžić, M., Lippman, Z., Jiang, H., Carrasquillo, R., Rabinowicz, P. D., . . . Bulski, A. J. P. b. (2007). Epigenetic natural variation in arabidopsis thaliana. *5*(7), e174.

Verhoeven, K. J., & van Gurp, T. P. (2012). Transgenerational effects of stress exposure on offspring phenotypes in apomictic dandelion. *PloS one, 7*(6), e38605.

Wang, W.-S., Pan, Y.-J., Zhao, X.-Q., Dwivedi, D., Zhu, L.-H., Ali, J., . . . Li, Z.-K. (2010a). Drought-induced site-specific DNA methylation and its association with drought tolerance in rice (oryza sativa l.). *Journal of experimental botany, 62*(6), 1951-1960.

Wang, W.-S., Pan, Y.-J., Zhao, X.-Q., Dwivedi, D., Zhu, L.-H., Ali, J., . . . Li, Z.-K. J. J. o. e. b. (2010b). Drought-induced site-specific DNA methylation and its association with drought tolerance in rice (oryza sativa l.). *62*(6), 1951-1960.

Wang, W.-S., Zhao, X., Pan, Y., Zhu, L., Fu, B., & Li, Z. (2011). DNA methylation changes detected by methylation-sensitive amplified polymorphism in two contrasting rice genotypes under salt stress. *Journal of Genetics and Genomics, 38*(9), 419-424.

Yaish, M. W., Colasanti, J., & Rothstein, S. J. (2011). The role of epigenetic processes in controlling flowering time in plants exposed to stress. *Journal of experimental botany, 62*(11), 3727-3735.

Yao, Y., Bilichak, A., Golubov, A., & Kovalchuk, I. (2012). Ddm1 plants are sensitive to methyl methane sulfonate and nacl stresses and are deficient in DNA repair. *Plant cell reports, 31*(9), 1549-1561.

Zemach, A., McDaniel, I. E., Silva, P., & Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science, 328*(5980), 916-919.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S. W.-L., Chen, H., . . . Jacobsen, S. E. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. *Cell, 126*(6), 1189-1201.

Zheng, X., Chen, L., Li, M., Lou, Q., Xia, H., Wang, P., . . . Luo, L. (2013). Transgenerational variations in DNA methylation induced by drought stress in two rice varieties with distinguished difference to drought resistance. *PloS one, 8*(11), e80253.

Zheng, X., Chen, L., Xia, H., Wei, H., Lou, Q., Li, M., . . . Luo, L. (2017). Transgenerational epimutations induced by multi-generation drought imposition mediate rice plant's adaptation to drought condition. *Scientific reports, 7*.

Zilberman, D., Gehring, M., Tran, R. K., Ballinger, T., & Henikoff, S. (2007). Genome-wide analysis of arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. *Nature genetics, 39*(1), 61.

CONCLUSION

Environmental stimuli including heat, salt and pathogen infection on *Arabidopsis thaliana* initiate a cascade of metabolic pathways to optimize physiological traits from leaf amount, weight of mature plants to the size of reproductive element; seed production and induces DNA methylation changes that allow for acclimation to these stressors especially temperature changes. The results of this study suggest that heat stress impacted DNA methylation patterns across generations and induced multigenerational inheritance of DNA methylation changes.

This study has also shown that the frequency of methylated cytosine revealed stress-triggered DNA methylation patterns with differential hypermethylation and hypomethylation in progeny generation than in parent generation. DNA alterations across *Arabidopsis* genome can occur due to long term adverse environmental condition as the plant over time strives to achieve improved balance between surviving and completing a lifecycle. These features manifested in physiological traits like number of leaves, flowering time and seed yield as well as differential changes in DNA methylation.

Methylation patterns of the genomic DNA changed under heat stress over multiple generations highlighted the role DNA methylation plays in environmentally induced phenotypic variations and plant epigenetic changes in response to abiotic stresses like drought and heat. Our results suggest that DNA methylation is a directed response to abiotic stress and DNA methylation variants act to convey heritable stress-induced epimutations that conduct heat stress memory for physiological traits and possible gene expressions. We found multi-generational heat stress to expend considerable influence on the methylome of *Arabidopsis* plants and because the lines used were derived from a single individual strain, the observed methylation variations in the off-spring of stressed plants were associated with parental environments and not with genetic variation among

the plants. Thus, the results reflect transgenerational inheritance of a single genotype in response to continuous heat stress.

Genome-wide sequencing analysis in this experiment indicated methylation changes were transmitted over multiple generation of plants under continuous stress and short-term memory was established. The identification of SNPs did point to a pattern of inheritance that reflected a broader genotypic diversity in stressed plants than plants without stress. DMRs showed specific epimutation associated with epigenetic inheritance of a stress response.

This research expands our understanding of the *Arabidopsis thaliana* genome to drought and heat tolerance. In a broader term, it gives an insight to an epigenetic mechanism of inheritance in the form of DNA methylation and expands the ideas of DNA methylation as the epigenetic modification that acts as the complementary transfer of phenotypic differences to influence plant response to heat stress. As well as create roadmaps to the molecular changes behind patterns of inheritance and adaptations to heat tolerance over generations.

DNA methylation has the ability to alter gene expression and change the phenotype without changing the DNA sequence in response to stress. Stress-associated DMRs compared with DMRs of unstressed linages help identify regions and genes that are transcriptionally silenced by DNA methylation. It is possible that new methylation variants within species only exist for brief periods during evolution in species.

Evolutionary patterns of genomic changes such as insertions, deletion and sequence variations like SNPs make it possible to investigate the potential effects of variants within the genome on phenotypic changes and adaptation. The extended plant lineage in this study takes into account possible evolutionary changes within a species population. The direct transmission of

environmental effects into subsequent generations reflects an advantage which promotes the selection of adapted genotypes in the improvement of crop species.

Practical applications of significance may include the use of induced methylation variants which are stable over time for breeding novel traits by utilizing DMRs as genetic markers to identify quantitative trait loci for certain phenotype. Traits in plants can be improved using molecular techniques that use engineered DNA-binding domains (like ZF, TALE and dCas9) to introduce particular chromatin marks to target modifications on selected responsive loci related to stress tolerance. Identified DNA methylation variants linked to heat stress or correlated with drought-response would allow the formation of stress-tolerant plants for crop breeding and management.

FUTURE DIRECTIONS

Based on our results, further research is required to determine heat-induced response genes. To do this, expression analysis using quantitative real time PCR can be performed on identified stress associated DMRs to highlight specific genomic sequences associated with a tolerance to stress. Also, stochastic variations and association with treatment in the DNA methylome can be investigated further using bioinformatic statistic tools like pairwise comparisons of all individual progeny plants across lineages and rank sum testing of tile- based DMRs respectively. The identification of SNPs in our findings can be explored further with bioinformatics and statistical tools (such as Wilcoxon Test) to establish association between identified SNPs and a phenotype. Other marks of epigenetic modification like Histone acetylation can be tested with histone modification assays such as native chromatin immunoprecipitation (NChIP) and the genomic regions with differential chromatin explored with Illumina deep sequencing.

APPENDIX 1: HISTOGRAM OF %METHYLATION FREQUENCIES OF ALL SAMPLE

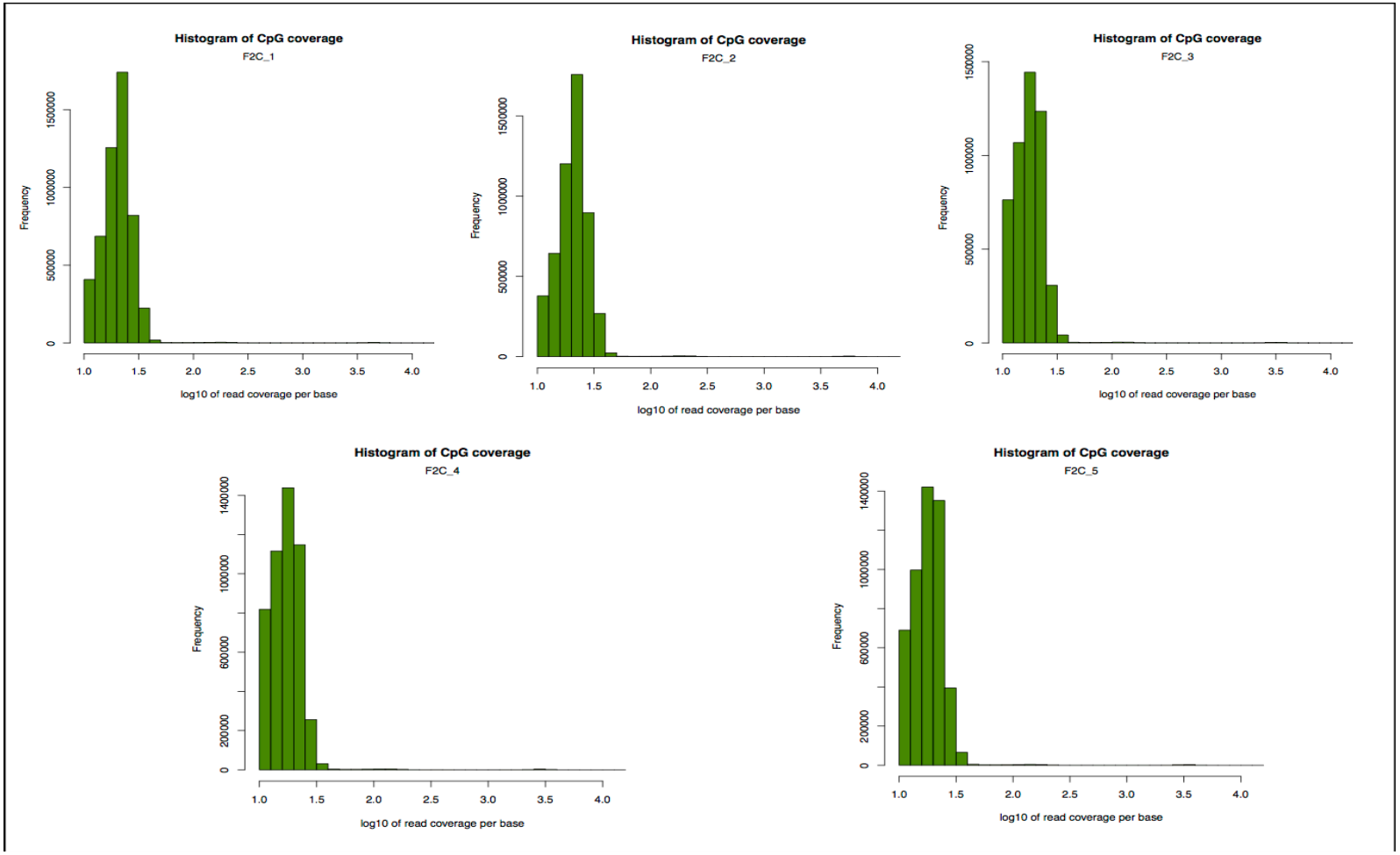

Figure 59: Histogram of CG read coverage for F2C showing %methylation frequency. Each histogram represents a biological replicate where n=5 for F2C.

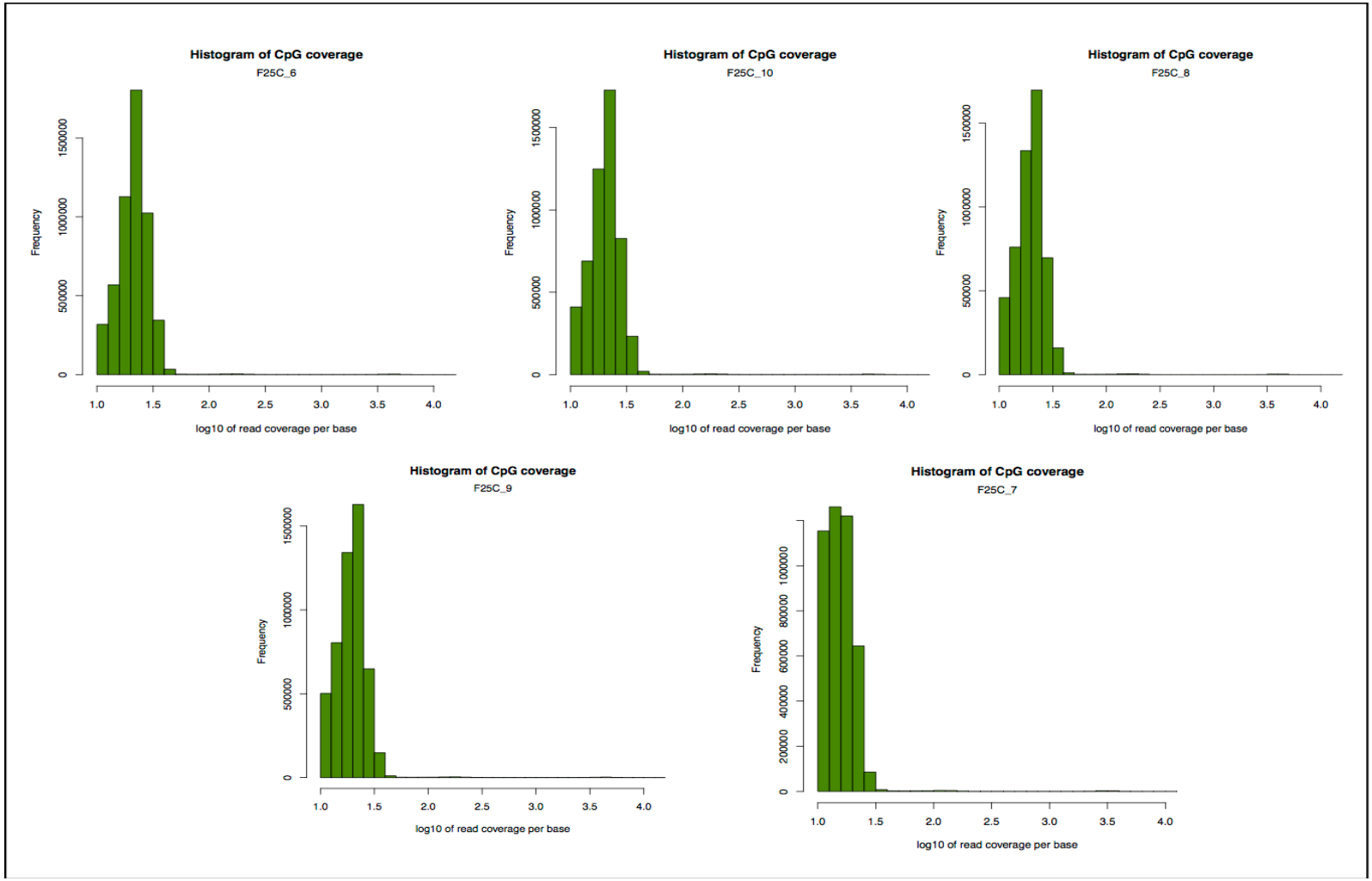

Figure 60: Histogram of CG read coverage for F25C showing %methylation frequency. Each histogram represents a biological replicate where n=5 for F25C.

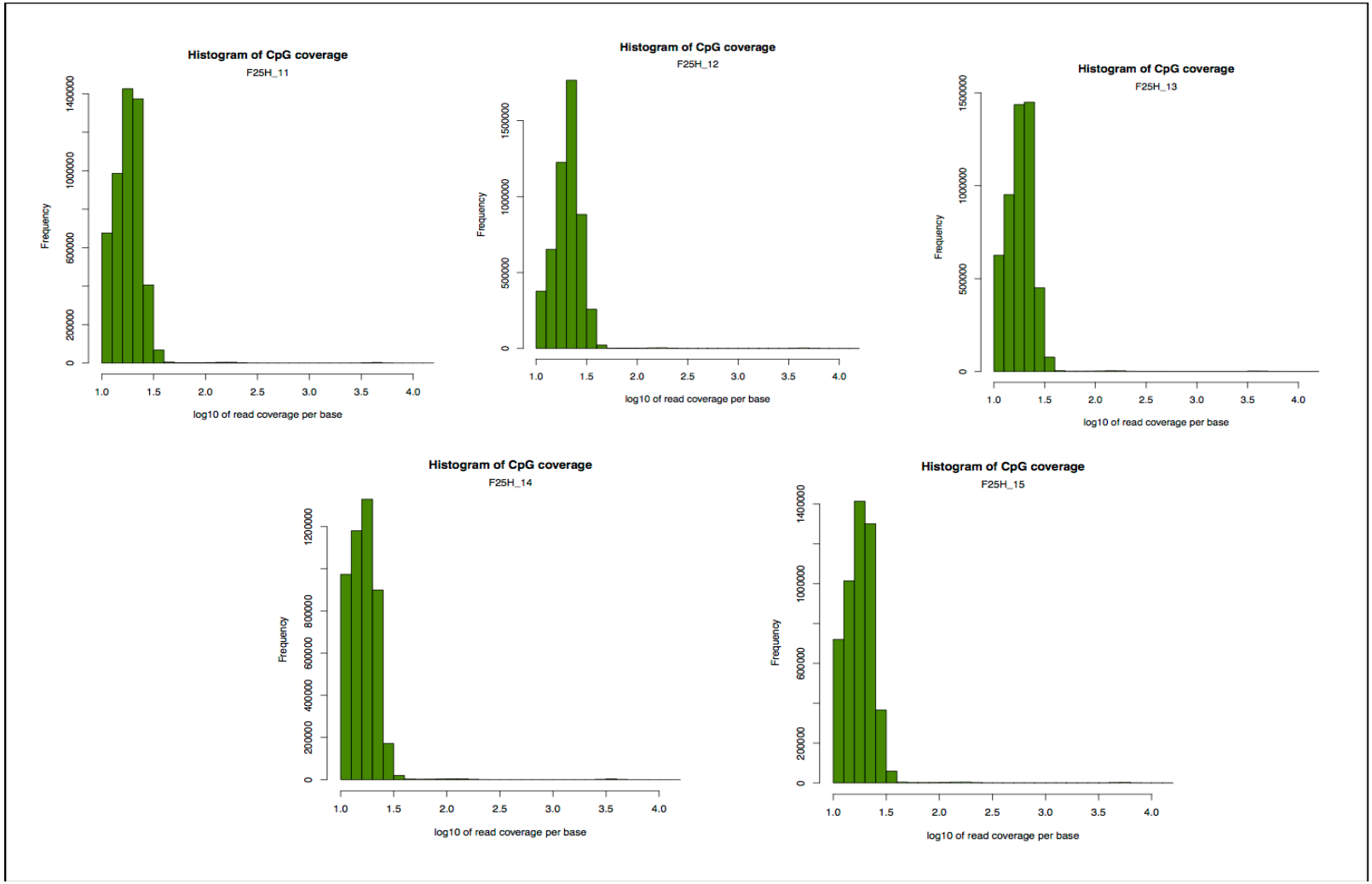

Figure 61: Histogram of CG read coverage for F25H showing %methylation frequency. Each histogram represents a biological replicate where n=5 for F25H.

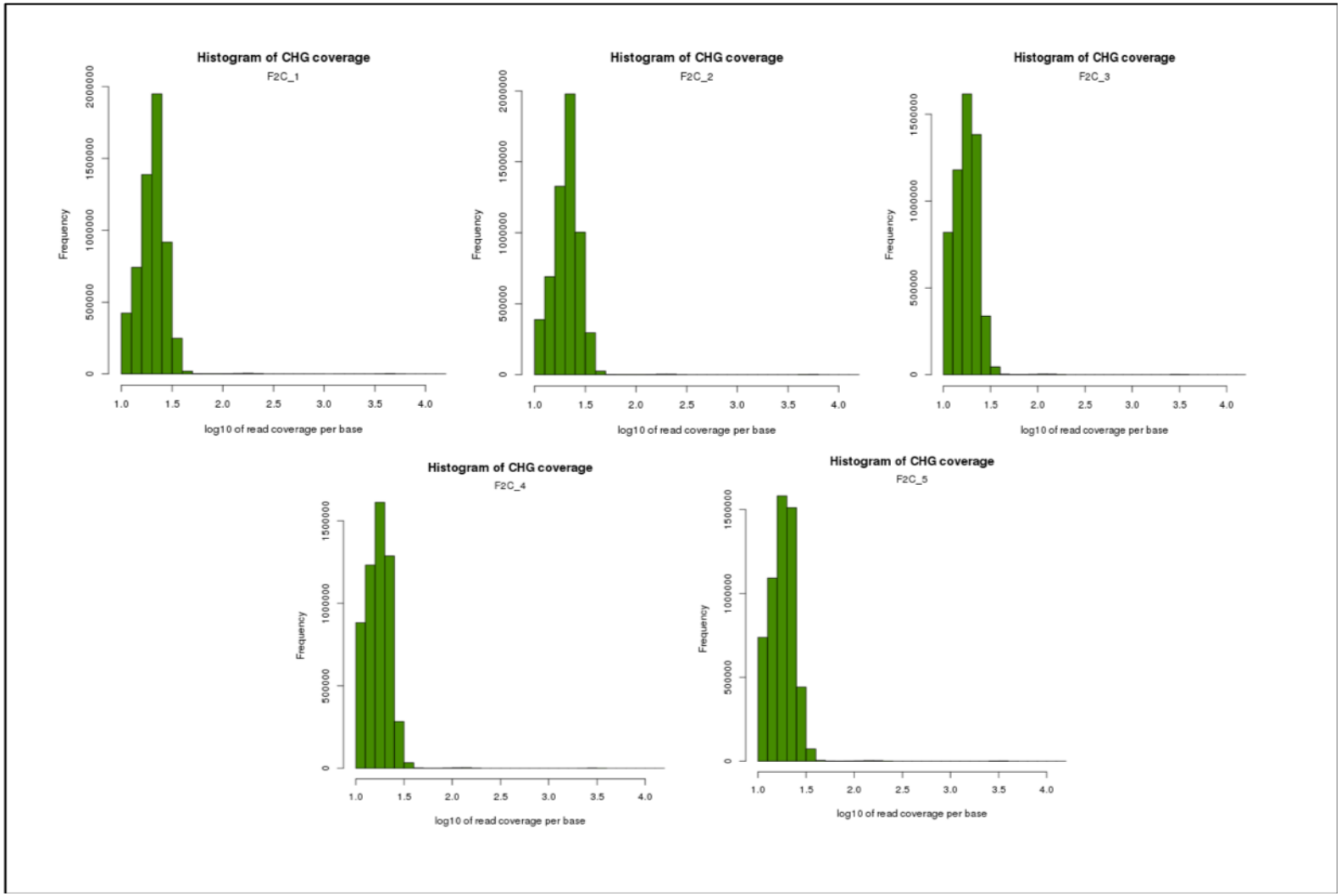

Figure 62: Histogram of CHG read coverage for F2C showing %methylation frequency. Each histogram represents a biological replicate where n=5 for F2C.

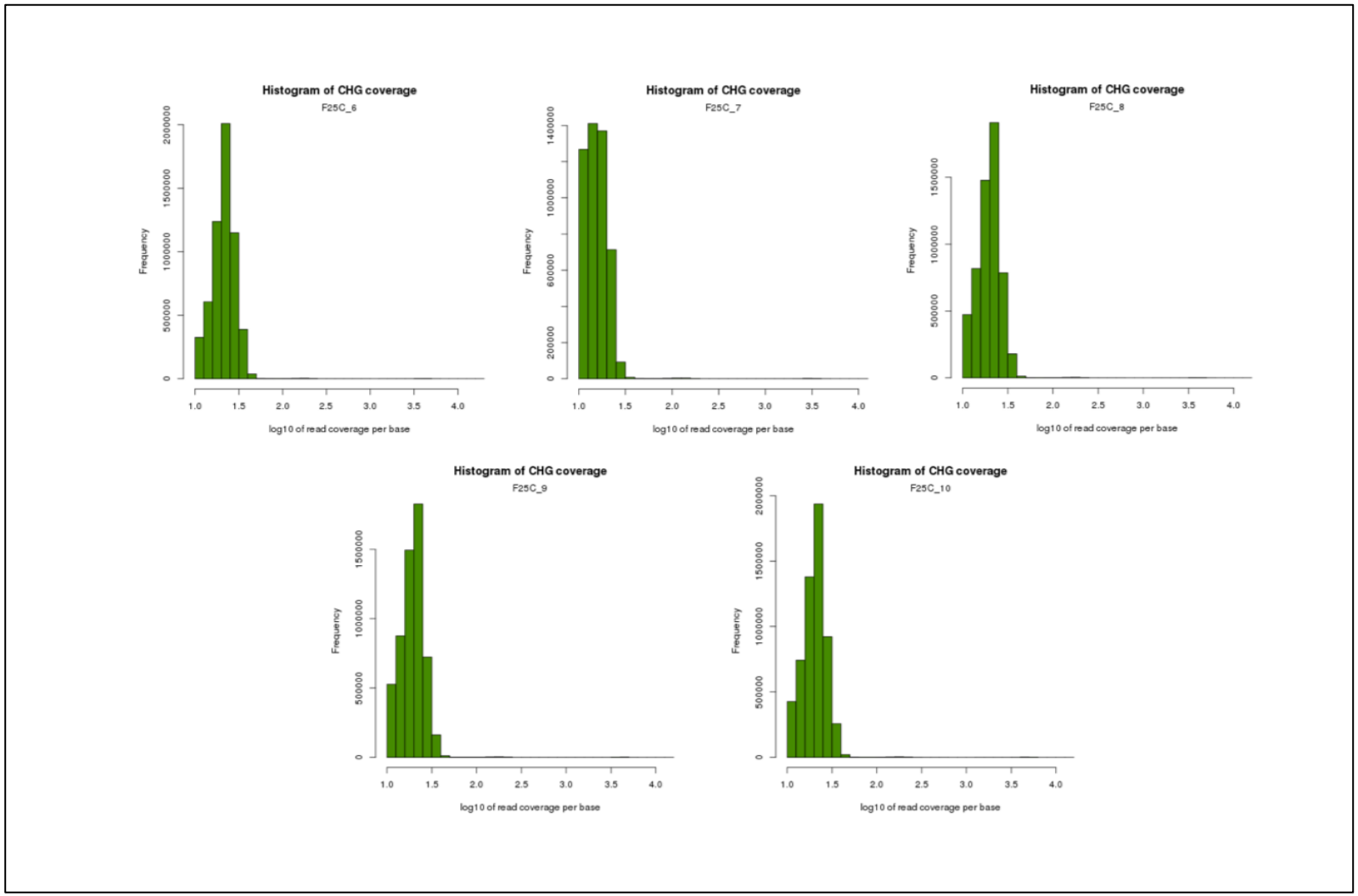

Figure 63: Histogram of CHG read coverage for F25C showing %methylation frequency. Each histogram represents a biological replicate where n=5 for F25C.

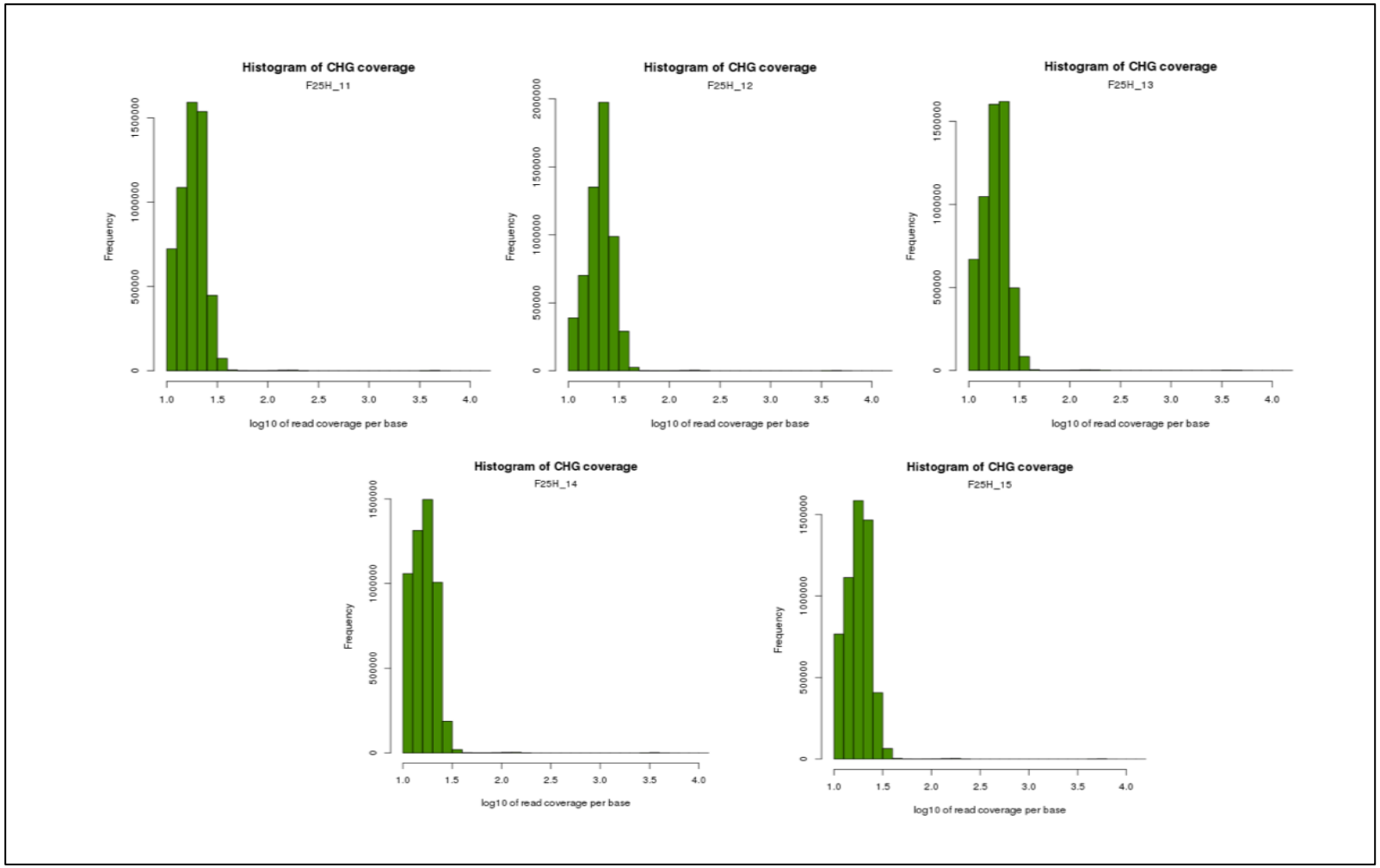

Figure 64: Histogram of CHG read coverage for F25H showing %methylation frequency. Each histogram represents a biological replicate where n=5 for F25H.

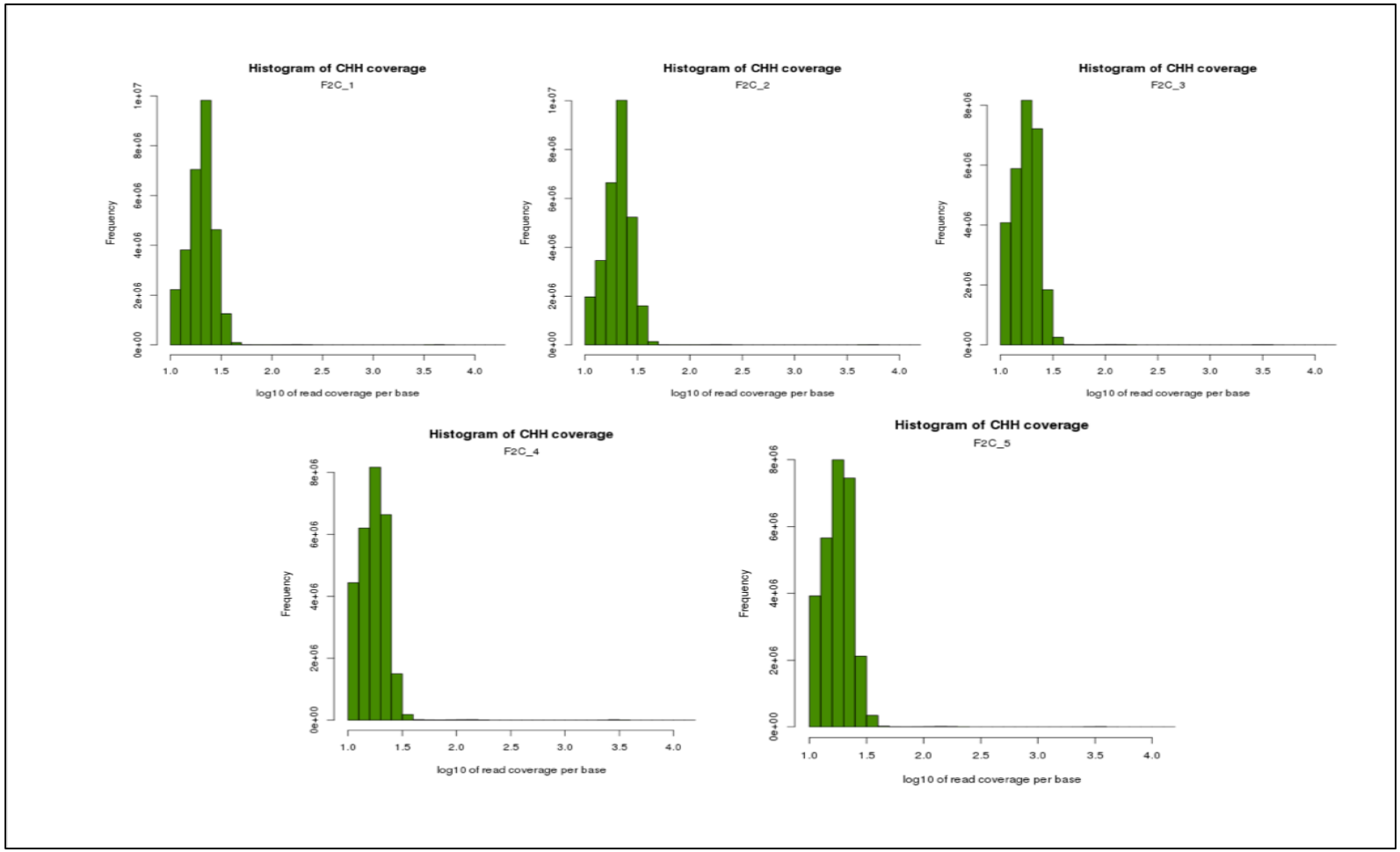

Figure 65: Histogram of CHH read coverage for F2C showing %methylation frequency. Each histogram represents a biological replicate where n=5 for F2C.

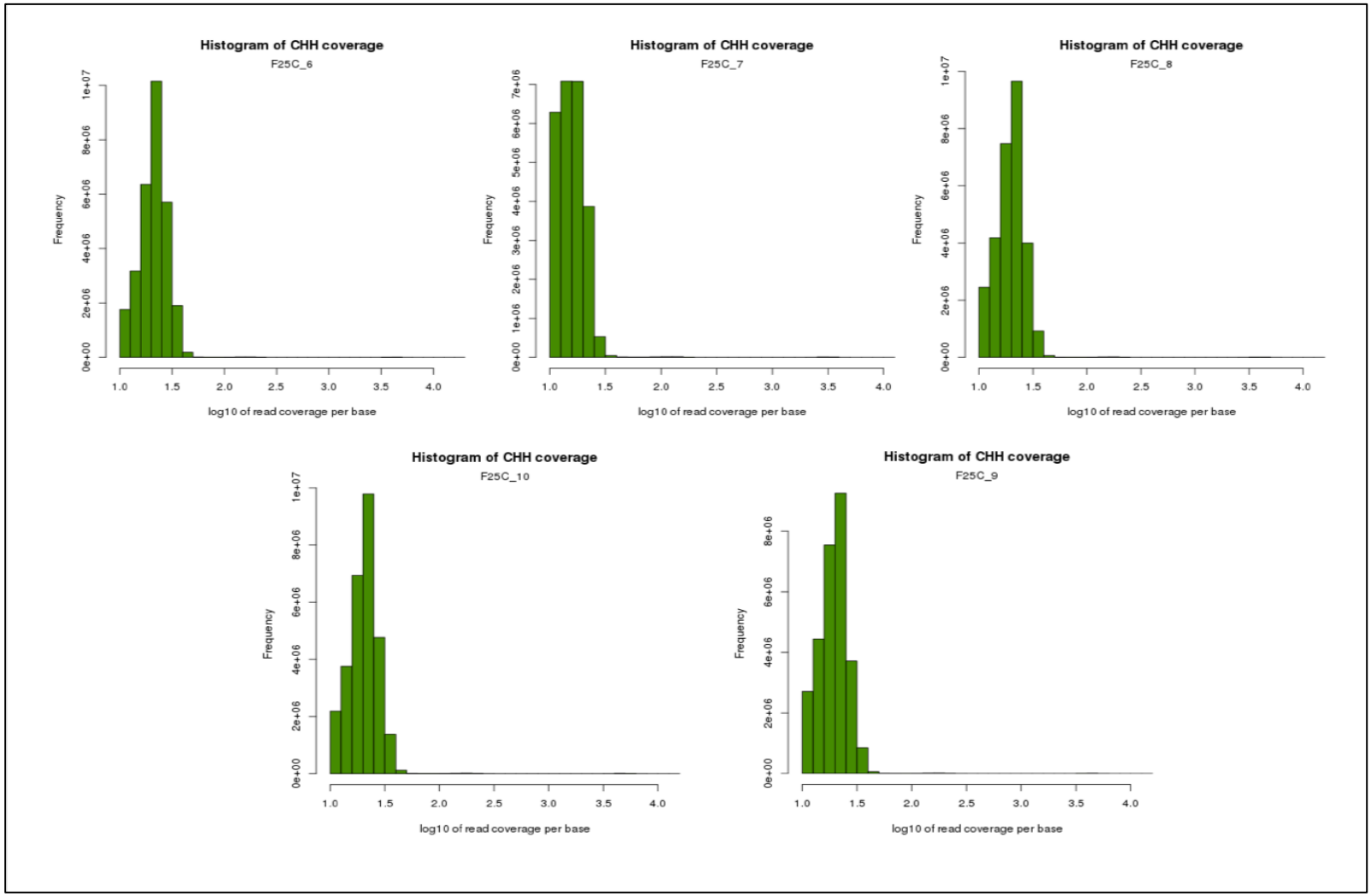

Figure 66: Histogram of CHH read coverage for F25C showing %methylation frequency. Each histogram represents a biological replicate where n=5 for F25C.

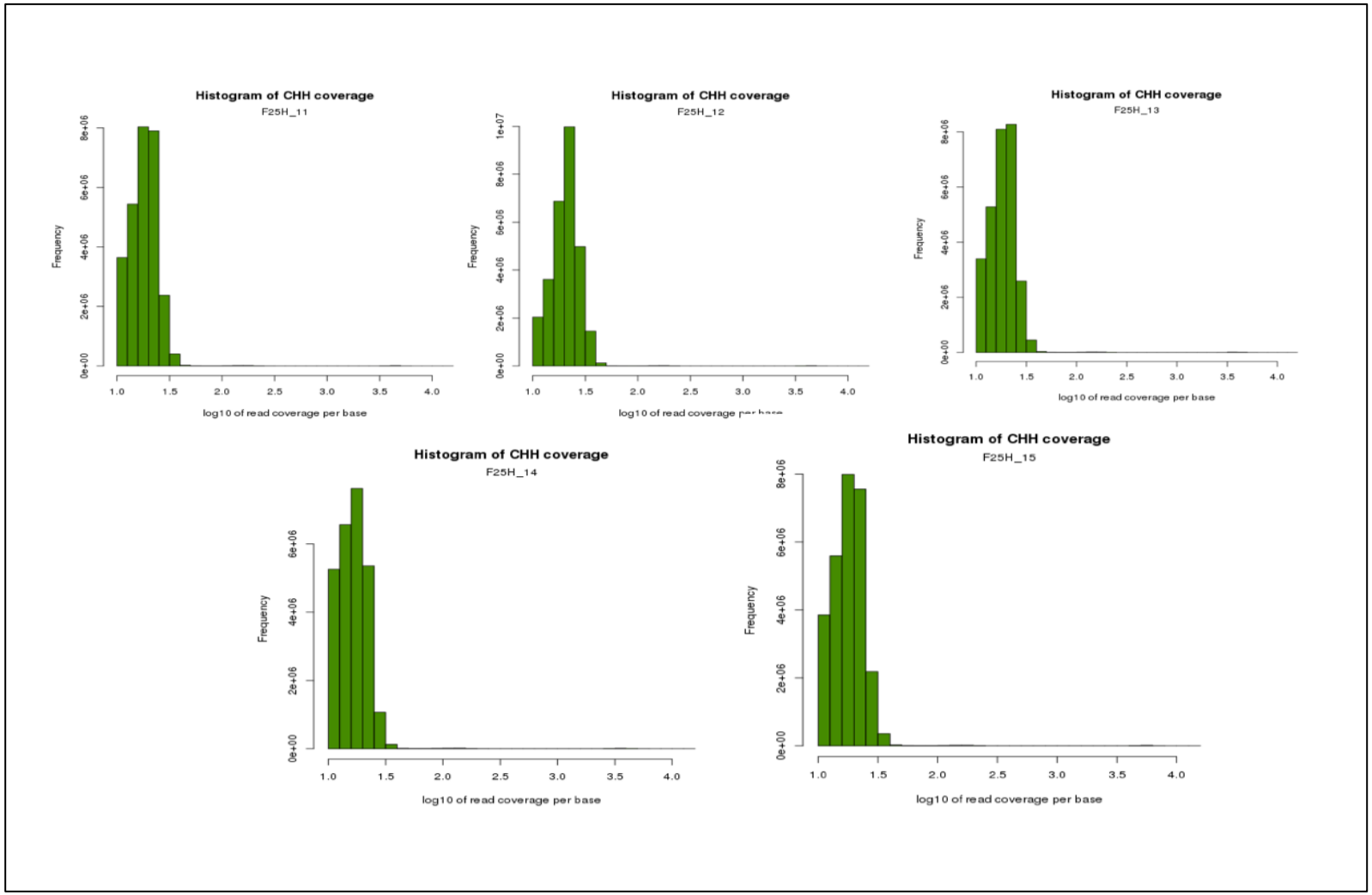

Figure 67: Histogram of CHH read coverage for F25H showing %methylation frequency. Each histogram represents a biological replicate where n=5 for F25H.

APPENDIX 2: COMPARISON OF COMMON BASES BETWEEN ALL SAMPLES

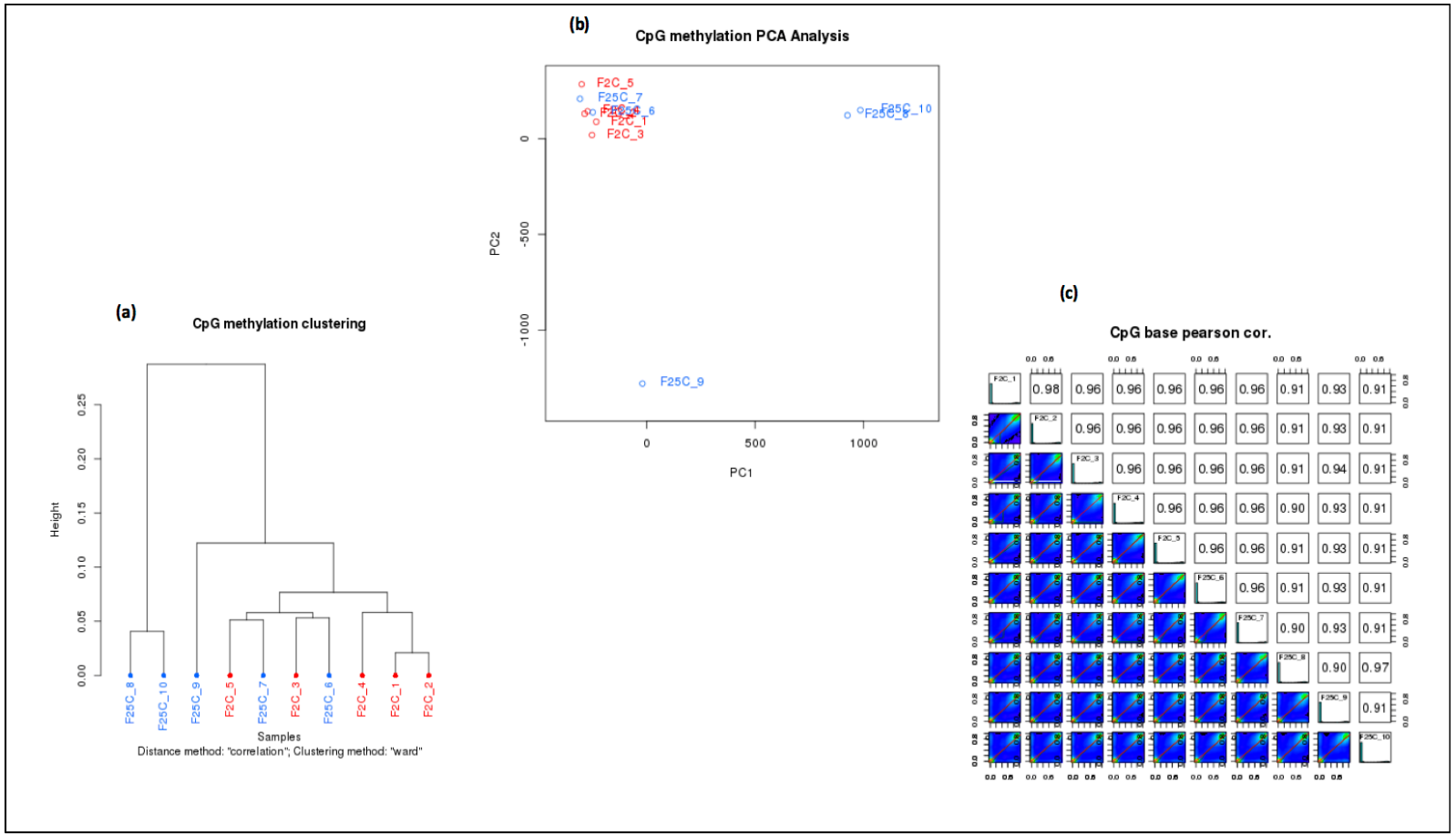

Figure 68: Comparison of common bases between F25C versus F2C. (a) Hierarchical clustering using 1-Pearson's correlation distance. (b) PCA comparing methylation profiles, plot shows principal component 1 and principal component 2 of each sample. Samples closer to each other in principal component space are similar in methylation profiles. (c) Correlation matrix. Scatter plots of %methylation values for samples in F25H, F2C and F25C. Numbers on upper right corner denotes pairwise Pearson's correlation scores.

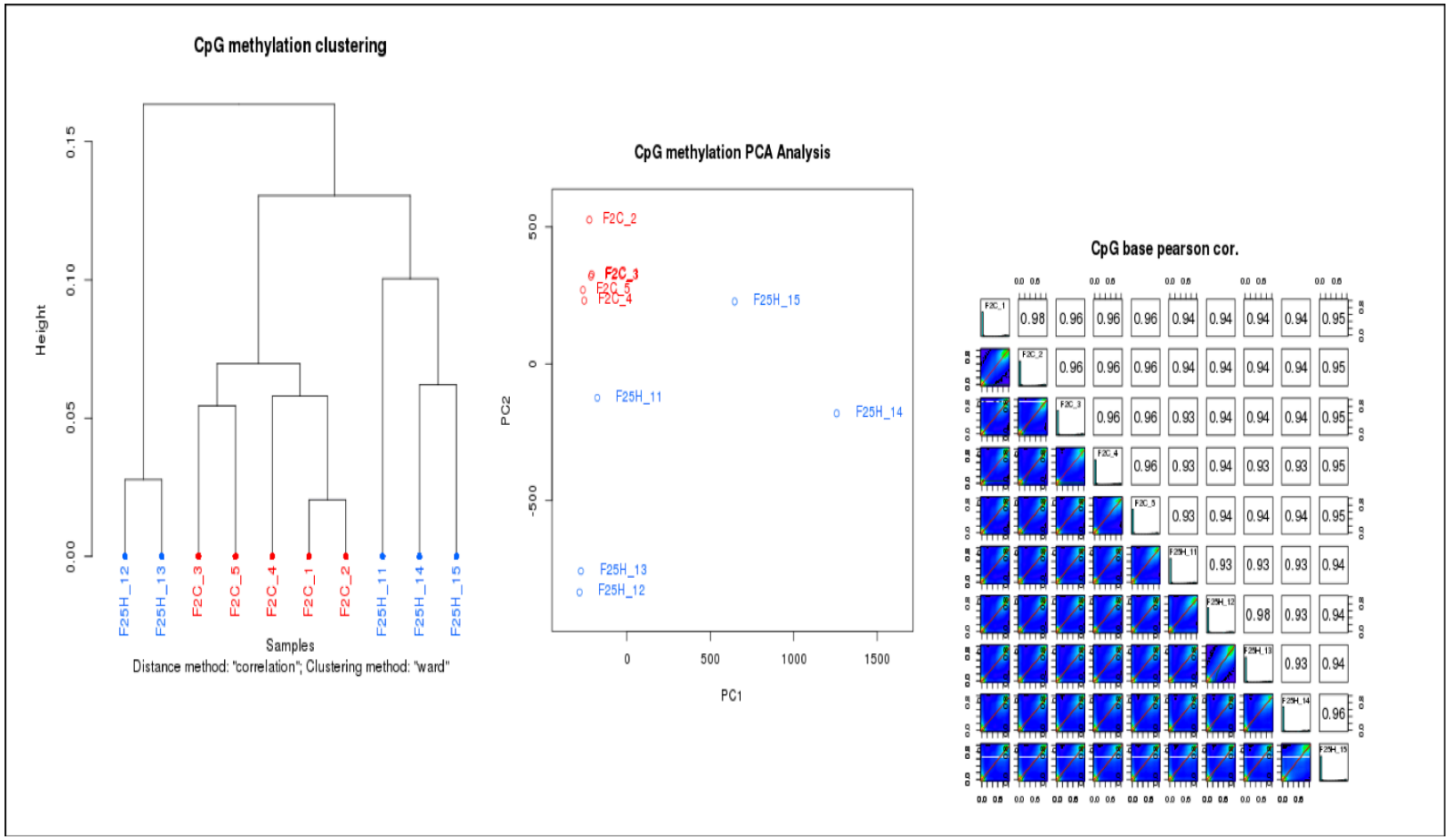

Figure 69: Comparison of common bases between F25H versus F2C. (a) Hierarchical clustering using 1-Pearson's correlation distance. (b) PCA comparing methylation profiles, plot shows principal component 1 and principal component 2 of each sample. Samples closer to each other in principal component space are similar in methylation profiles. (c) Correlation matrix. Scatter plots of %methylation values for samples in F25H, F2C and F25C. Numbers on upper right corner denotes pairwise Pearson's correlation scores.

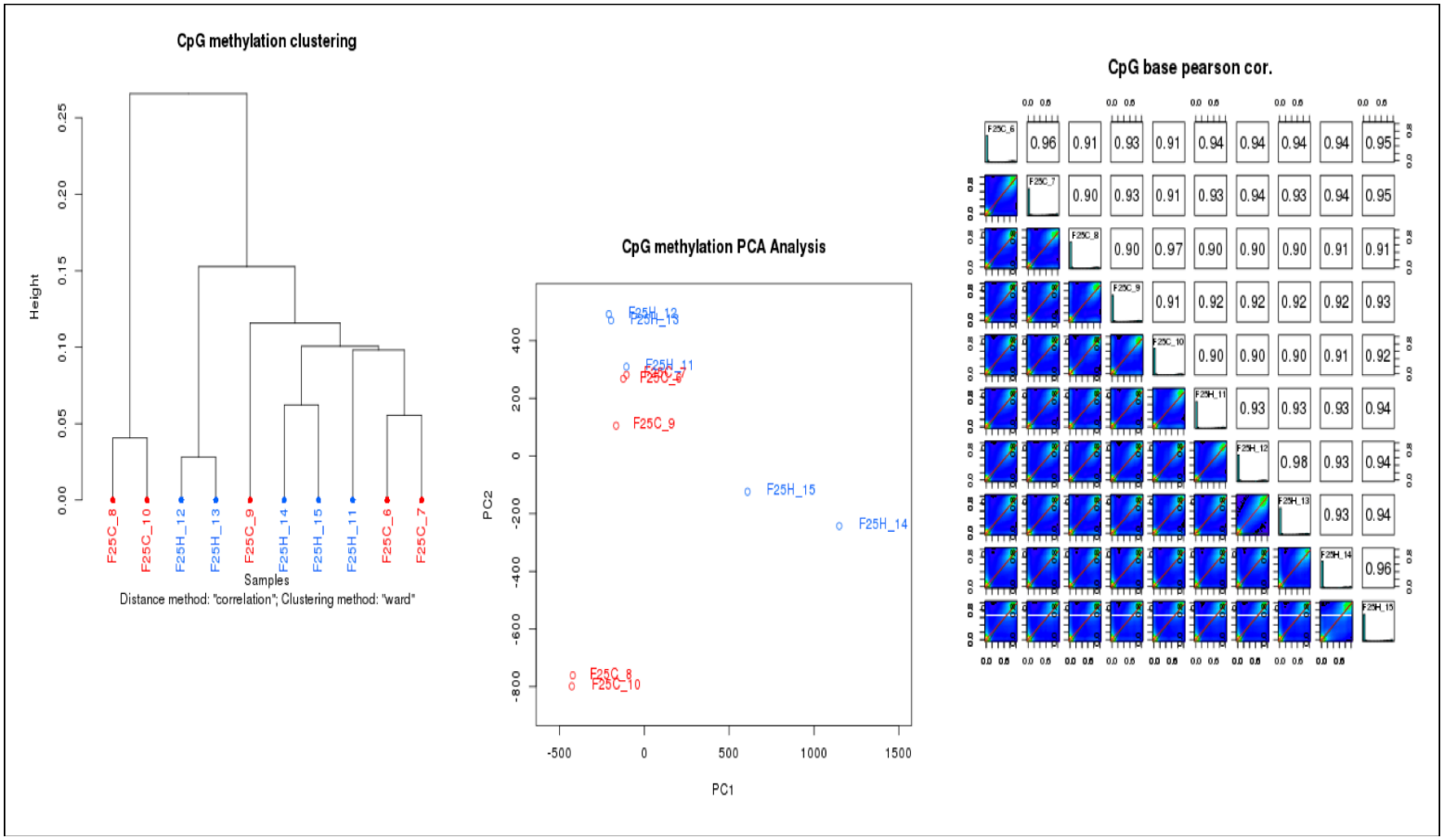

Figure 70: Comparison of common bases between F25H versus F25C. (a) Hierarchical clustering using 1-Pearson's correlation distance. (b) PCA comparing methylation profiles, plot shows principal component 1 and principal component 2 of each sample. Samples closer to each other in principal component space are similar in methylation profiles. (c) Correlation matrix. Scatter plots of %methylation values for samples in F25H, F2C and F25C. Numbers on upper right corner denotes pairwise Pearson's correlation scores.