



DNA Methylation and Demethylation Are Regulated by Functional DNA Methyltransferases and DnTET Enzymes in *Diuraphis noxia*

Pieter H. du Preez, Kelly Breeds, N. Francois V. Burger, Hendrik W. Swiegers, J. Christoff Truter and Anna-Maria Botha*

Department of Genetics, Stellenbosch University, Stellenbosch, South Africa

OPEN ACCESS

Edited by:

Wei Guo,
Institute of Zoology (CAS), China

Reviewed by:

Bing Chen,
Hebei University, China
Owain Rhys Edwards,
CSIRO Land and Water, Australia

*Correspondence:

Anna-Maria Botha
ambo@sun.ac.za

Specialty section:

This article was submitted to
Epigenomics and Epigenetics,
a section of the journal
Frontiers in Genetics

Received: 14 January 2020

Accepted: 14 April 2020

Published: 23 June 2020

Citation:

du Preez PH, Breeds K, Burger NFV,
Swiegers HW, Truter JC and
Botha A-M (2020) DNA Methylation
and Demethylation Are Regulated by
Functional DNA Methyltransferases
and DnTET Enzymes in *Diuraphis*
noxia. *Front. Genet.* 11:452.
doi: 10.3389/fgene.2020.00452

Aphids are economically important insect pests of crops worldwide. Despite resistant varieties being available, resistance is continuously challenged and eventually broken down, posing a threat to food security. In the current study, the epigenome of two related Russian wheat aphid (*Diuraphis noxia*, Kurdjumov) biotypes (i.e., SA1 and SAM) that differ in virulence was investigated to elucidate its role in virulence in this species. Whole genome bisulfite sequencing covered a total of 6,846,597,083 cytosine bases for SA1 and 7,397,965,699 cytosine bases for SAM, respectively, of which a total of 70,861,462 bases (SA1) and 74,073,939 bases (SAM) were methylated, representing $1.126 \pm 0.321\%$ (SA1) and $1.105 \pm 0.295\%$ (SAM) methylation in their genomes. The sequence reads were analyzed for contexts of DNA methylation and the results revealed that RWA has methylation in all contexts (CpG, CHG and CHH), with the majority of methylation within the CpG context ($\pm 5.19\%$), while the other contexts show much lower levels of methylation (CHG – $\pm 0.27\%$; CHH – $\pm 0.34\%$). The top strand was slightly (0.02%) more methylated than the bottom strand. Of the 35,493 genes that mapped, we also analyzed the contexts of methylation of each of these and found that the CpG methylation was much higher in genic regions than in intergenic regions. The CHG and CHH levels did not differ between genic and intergenic regions. The exonic regions of genes were more methylated ($\pm 0.56\%$) than the intronic regions. We also measured the 5mC and 5hmC levels between the aphid biotypes, and found little difference in 5mC levels between the biotypes, but much higher levels of 5hmC in the virulent SAM. RWA had two homologs of each of the DNA methyltransferases 1 (*DNMT1a* and *DNMT1b*) and *DNMT3s* (*DNMT3a* and *DNMT3b*), but only a single *DNMT2*, with only the expression of *DNMT3* that differed significantly between the two RWA biotypes. RWA has a single ortholog of Ten eleven translocase (*DnTET*) in the genome. Feeding studies show that the more virulent RWA biotype SAM upregulate *DnDNMT3* and *DnTET* in response to wheat expressing antibiosis and antixenosis.

Keywords: whole genome bisulfite sequencing, DNMT and TET, global methylation (5mC), global hydroxymethylation (5hmC), Russian wheat aphid

INTRODUCTION

Diuraphis noxia (Kurdjumov, Hemiptera: Aphididae—or Russian wheat aphid, RWA) biotypes are morphologically similar, yet display vast differences in their capacity to damage wheat cultivars upon feeding (i.e., their virulence) (Botha, 2013). In South Africa, the virulence of the four wild type and the mutant RWA biotypes is as follows in order from least to most virulent: SA1 < SA2 < SA3 < SA4 < SAM (Swanevelder et al., 2010; Jankielsohn, 2016). Despite the emergence of new RWA biotypes in South Africa (Tolmay et al., 2007; Jankielsohn, 2011, 2016), and other parts of the world, including the United States of America (USA) (Haley et al., 2004; Burd et al., 2006; Randolph et al., 2009) and Argentina (Clua et al., 2004), the molecular mechanism(s) underlying the development of new biotypes is currently unknown (Shufran and Payton, 2009; Botha et al., 2014a). The known genealogy of SA1 and SAM (Swanevelder et al., 2010), their genetic similarity (Burger and Botha, 2017) and their position on either end of the virulence spectrum, renders them particularly useful in the present study, to improve the understanding of the process of biotypification. The possibility of a link between RWA methylation and biotype virulence has previously been suggested (Gong et al., 2012; Breeds et al., 2018). In 2012, Gong et al. investigated the methylation of four genes encoding salivary gland proteins (putative effector genes) in RWA biotypes US1 and US2, and found these genes to be differentially methylated in the different biotypes. In the initial investigation of South African RWA methylation (Breeds et al., 2018), the different biotypes exhibited different banding patterns (after restriction of their DNA with methylation-sensitive enzymes), methylation levels and methylation trends, all of which support a role for methylation in biotypification.

The epigenetic modification of DNA methylation involves the covalent addition of a methyl group to the 5' position of cytosine (Glastad et al., 2011; Lyko and Maleszka, 2011). In insects, methylation occurs predominantly within genes (Walsh et al., 2010; Zemach et al., 2010; Glastad et al., 2011; Lyko and Maleszka, 2011), where to date it is reported to perform two major functions. Firstly, intragenic methylation affects alternative splicing by recruiting or interfering with different DNA binding factors (Hunt et al., 2013b; Glastad et al., 2014; Yan et al., 2015), and secondly, it prevents the initiation of spurious transcription at cryptic binding sites within genes (Hunt et al., 2010, 2013a,b). Three classes of DNA methyltransferase (DNMT) proteins are involved in methylation of DNA and these perform different functions, with DNMT3 and DNMT1 establishing and maintaining methylation patterns, respectively, but with a less clear function for DNMT2. This class is known to show strong conservation in sequence and is suggested to be an ancient DNA methyltransferase that changed its substrate specificity from DNA to tRNA (Sunita et al., 2008; Iyer et al., 2011; Jurkowski and Jeltsch, 2011; Raddatz et al., 2013). Insects have a variety of combinations of the DNMT genes, with some lineages having lost one (e.g., *Bombyx mori* and *Tribolium castaneum*) or two (e.g., *Drosophila melanogaster* and *Anopheles gambiae*) classes of DNMTs, and others having multiple homologs (e.g., *Apis mellifera*, *Nasonia vitripennis*, and *Acyrtosiphon pisum*) within

a certain DNMT class (Kunert et al., 2003; Marhold et al., 2004; Walsh et al., 2010; Xiang et al., 2010; Glastad et al., 2011; Feliciello et al., 2013). Despite their important function in DNA methylation, knowledge of RWA DNMTs is still lacking.

DNA methylation is removed through the process of demethylation, which can occur both passively and actively, with 5-hydroxymethylcytosine (5hmC) being a measurable intermediate of one of the active demethylation pathways (Branco et al., 2012; Kohli and Zhang, 2013). Hydroxymethylcytosine is formed through the oxidation of 5mC by ten-eleven translocation enzymes (TETs) (Tahiliani et al., 2009; Shen et al., 2014). The presence of 5hmC has only been reported in a few insects including *A. mellifera*, *T. castaneum*, *N. vitripennis* and *D. melanogaster* (Cingolani et al., 2013; Feliciello et al., 2013; Wojciechowski et al., 2014; Delatte et al., 2016; Pegoraro et al., 2016; Rasmussen et al., 2016). To determine the presence and extent of 5hmC in the RWA, an antibody specific to 5hmC was used, providing the first insight into RWA demethylation.

The objective in this study was firstly to sequence and compare the epigenome of RWA biotypes SA1 and SAM, and determine the level, location (e.g., intergenic or genic, exonic or intronic), and contexts of DNA methylation (i.e., CpG, CHH, CHG) within the genomes of these RWA biotypes with differential virulence. Secondly, to quantify global methylation (5mC) and demethylation (5hmC) in the South African biotypes with shared genealogy; and thirdly, to characterize the DNA methyltransferases (DNMTs) and ten-eleven translocase enzyme-like (*TET*) genes and expression in these aphids, to relate these observations to the reported difference in virulence levels of the South African RWA biotypes SA1 and SAM.

MATERIALS AND METHODS

Aphid Rearing

For whole genome bisulfite sequencing and measurement of global methylation (5mC) and hydroxymethylation (5hmC) levels, colonies of apterous parthenogenetic female aphids of South African RWA biotypes SA1 and SAM were separately established in BugDorm cages (MegaView Science Education Services Co. Ltd., Taiwan) in an insectary with the following conditions: $22.5 \pm 2.5^\circ\text{C}$, 40% relative humidity, and continuous artificial lighting from high pressure sodium lamps as previously described (Breeds et al., 2018). In all instances triplicate colonies of each biotype were established. Stock colonies of RWA biotype SA1 were maintained on the wheat line Tugela (susceptible) and biotype SAM on the near isogenic wheat line Tugela-*Dn1* (resistant). To avoid any environmental effects due to feeding on different wheat plants, aphid biotypes were transferred to the susceptible cultivar “SST356” 1 month prior to DNA extraction for the whole genome bisulfite sequencing. In all instances, treatments were conducted using separate BugDorm cages in triplicate ($n = 3 \times 2$).

For *DnDNMT* expression analysis (0h), RWA biotype SA1 was maintained on the “SST 356” wheat cultivar (susceptible), while the highly virulent SAM biotype was maintained on “SST 398” (RWA resistant), and then transferred to the susceptible

“SST 356” wheat cultivar 1 month prior to RNA extraction and cDNA synthesis.

For the RNaseq analysis, colonies of apterous parthenogenetic female aphids of South African RWA biotypes SA1 and SAM were separately established in BugDorm cages (MegaView Science Education Services Co. Ltd., Taiwan) in an insectary with the following conditions: $22.5 \pm 2.5^\circ\text{C}$, 40% relative humidity, and continuous artificial lighting from high pressure sodium lamps as previously described (Breeds et al., 2018). RWA biotype SA1 was maintained on the wheat line Tugela (susceptible) and biotype SAM on the near isogenic wheat line Tugela-*Dn5* (resistant). Multiple individual replicates, consisting of 50 aphids of various life stages, were collected for each biotype. Collected aphids were flash frozen in liquid nitrogen and RNA was extracted following the protocol of Qiagen’s RNeasy RNA extraction kit performing the optional on column Qiagen DNase treatment. Extracted RNA was assessed for quality through both bleach-gel electrophoresis (Aranda et al., 2012) and with an Agilent 2100 Bioanalyser using the RNA Nano 6000 kit (Babu and Gassmann, 2016). Three RNA samples, from each biotype, representing three biological repeats, with the highest RIN values (at least above 7) were used in subsequent analysis.

For the *DnDNMT* and *DnTET* host-shifts, RWA biotypes SA1 and SAM were maintained on the susceptible “SST 356” wheat cultivar, and then transferred to either near isogenic wheat lines Tugela (susceptible), or Tugela-*Dn1* (resistant), or Tugela-*Dn5* (resistant) prior to RNA extraction and cDNA synthesis. Aphids were sampled at 0, 6, and 48 h post host-shifting. In all instances, treatments were conducted using separate BugDorm cages using separate plants in triplicate ($n = 3 \times 2$).

For the quantitation of DNMT proteins, both biotypes were maintained on the “SST 356” wheat cultivar before protein extraction. In all instances, treatments were conducted using separate BugDorm cages in triplicate ($n = 3 \times 2$). All SST cultivars were obtained from SENSAGO (Pty) Ltd., (South Africa).

Identification, Cloning and Sequencing of RWA DNMTs and Ten Eleven Translocation-Like (TET-Like) Genes

DNMT and TET sequences of the pea aphid (*Acyrtosiphon pisum*) were obtained through GenBank and used as BLAST (Altschul et al., 1997) queries against the NCBI’s non-redundant (nr) database to obtain homologs from the Class Insecta. The obtained sequences were then aligned using MAFFT v.7.4 (Katoh and Standley, 2013) and through use of maximum parsimony the obtained sequences were phylogenetically rendered through use of PAUP v4.0a136.

Primers were designed (Table S1) to amplify the transcripts of identified RWA DNMTs and TET-like genes using Primer3 (Rozen and Skaletsky, 2000). The primers were then used in a primer BLAST analysis against the RWA SAM biotype reference genome (GCA_001465515.1) to ensure they only matched genes of interest. RNA extractions and cDNA synthesis were performed for both RWA biotypes SA1 and SAM as previously described (Burger et al., 2017).

PCR reactions for sequencing were performed using Phusion High-Fidelity DNA Polymerase (NEB) and following the manufacturer’s protocol. PCR products were then ligated into the pTZ57R/T vector (InsTAclone PCR cloning kit, Thermo Scientific) overnight at 4°C . For PCR reactions showing non-specific amplification, gel fragments containing bands of the expected product size were excised and subjected to five freeze-thaw cycles (liquid nitrogen/ 60°C oven) in 20 μl of distilled water and the obtained DNA was quantified through spectrophotometry (NanoDrop 2000, Thermo). Based on these results, differing amounts of freeze-thawed DNA were used, in accordance with the kit’s recommendations on the optimal quantity of PCR product for ligation.

Transformation of DH5 α competent cells (Thermo Scientific) was performed through heat shock following the manufacturers’ protocol and recombinant colonies were cultured and screened as previously performed (Burger et al., 2017). Plasmid minipreps (derived from at least one colony per PCR product) were submitted to the Central Analytical Facility (CAF) of Stellenbosch University for bi-directional Sanger sequencing using universal M13 forward and reverse primers (Table S1).

After Sanger sequencing, raw sequences were imported into Geneious v.9.1.8 and trimmed on either end to remove poor quality or ambiguous base calls. A VecScreen BLAST (<http://www.ncbi.nlm.nih.gov/tools/vecsreen/>) was then performed using the trimmed sequences to remove any vector DNA. The sequences for both SA1 and SAM biotypes (at least one forward and one reverse per PCR product) were aligned with the respective gene from which primers were designed using Primer 3 (Sievers et al., 2011).

Sequencing, Transcriptome Assembly, and Quality Control

RNA samples were sent for sequencing at Macrogen Inc., South Korea where six libraries were prepared using the TruSeq Stranded mRNA Sample Preparation Guide, Part #15031047 Rev. E. Paired-end library construction was performed using the Illumina TruSeq stranded mRNA kit and the subsequent libraries were sequenced on the Illumina NovaSeq 6000 system to obtain 100 bp paired-end reads for three replicates of both the *Diuraphis noxia* SAM and SA1 biotypes.

Raw reads obtained from the NovaSeq 6000 system were analyzed through use of FASTQC (Andrews, 2010) and trimmed of all poor quality reads and sequencing adaptors through use of Trimmomatic (Bolger et al., 2014). The trimmed reads were then used to perform a strand specific *de novo* assembly through use of the Trinity software suite (Haas et al., 2013). The assembled transcriptome’s quality was assessed through mapping the reads back to the assembled transcripts using Bowtie2 (Langmead and Salzberg, 2012), and to assess the percentage of reads utilized to construct the transcriptome. A BUSCO v4.2 (Simão et al., 2015) analysis was also performed using the Insecta homolog set (accessed on 2020/02, https://buscos.ezlab.org/datasets/prerelease/viridiplantae_odb10.tar.gz) to establish the number of represented essential orthologs. To assess if sequencing depth was adequate to generate a high quality *de*

de novo assembly, successively increasing sub-samplings of total sequencing data for individual samples were performed and assembled separately. These were then compared through the use of BLASTx to the NCBI's nr (protein) database, and the SwissProt uniprotKB and TrEMBL databases to assess the number of full-length BLAST matches obtained for the assembled transcripts from the differently sized assemblies.

Basic statistics such as the number of transcripts, transcript average length, transcript average %GC content, transcript N50 and transcript Ex90N50 were also calculated. All transcripts were analyzed through use of OmicsBox v1.2 (OmicsBox—Bioinformatics Made Easy, BioBam Bioinformatics, March 3, 2019, <https://www.biobam.com/omicsbox>) by performing BLASTx and BLASTn searches, respectively to the NCBI's nr and nt databases (accessed on 2019/08/21). Blast2GO (Götz et al., 2008) was then used to assign gene ontologies (GO) and KOG terms to all transcripts.

To compare the differential gene expression between the least and most virulent biotypes, transcript abundance quantification was performed using RSEM (Li and Dewey, 2011) for each sample using the obtained *de novo* transcripts. Average expression and the coefficient of variation was calculated per gene for the two biotypes SA1 and SAM separately. For this purpose FPKM (fragments per kilobase of transcript per million) values were used but also estimated by RSEM. We also identified differentially expressed (DE) genes between biotypes SA1 and SAM using edgeR (Robinson et al., 2010) based on gene-level expected counts estimated by RSEM. Only genes with greater than two counts-per-million in at least three samples were retained for DE analysis and we considered genes DE if they had a fold-change (FC) ≥ 1.5 and $p < 0.05$ after adjusting for multiple testing using the Benjamini–Hochberg (BH) procedure (Benjamini and Hochberg, 1995).

Augustus v3.3.3 (Stanke et al., 2006) was utilized to predict protein coding genes from the assembled transcripts using the *Acyrtosiphon pisum* (pea aphid) training set. Through use of a Trinity provided script, the GATK v3.8 pipeline for variant calling (Van der Auwera et al., 2013) was applied between the transcripts from the biotypes. Variants were accepted as true if they possessed an FS score above 30 (Phred-scaled p -value using Fisher's exact test to detect strand bias) and a QD score < 2 (Variant Confidence/Quality by Depth). Variants were also required to be present in all 3 biological replicates of one biotype and absent in all 3 biological replicates of the other biotype.

Analysis of DNMT and TET Expression

For DNMT gene expression analyses, 20 apterous aphids were collected in triplicate for each biotype ($3 \times n = 60$) and their heads were removed with a liquid nitrogen-cooled scalpel by cutting carefully posterior to the prothorax (Figure S1) and RNA was extracted as previously described (Burger et al., 2017). cDNA synthesis was performed using the iScript™ cDNA Synthesis kit (BioRad) in accordance with the provided protocol, applying 350 to 400 ng of total RNA as template per 20 μ l reaction.

For the host-shift experiment, RNA was isolated from apterous aphid whole-body homogenates prepared using a micro-pestle in liquid nitrogen cooled Eppendorf tubes.

Each treatment was represented by three biological replicates consisting of 30 aphids each ($n = 90$). The frozen aphids were ground with micro-pestles and RNA was extracted using RNeasy Mini Kit (Qiagen), following the manufacturers recommended protocol for insect material. cDNA synthesis was performed using SensiFAST cDNA Synthesis Kit (Bioline), with 200 ng of input RNA.

Primer pairs for RT-qPCR (Table S2) were designed using Primer3 from the CDS regions of the RWA sequenced DNMTs and TET to yield products of between 100 bp and 200 bp in size. Primers were used in a primerBLAST analysis against the assembled RWA SAM biotype reference genome (GCA_001465515.1) to ensure they only matched the DNMT and TET genes from which they were designed. The relative expression of DNMT1, DNMT2, and DNMT3 (in sampled aphid heads of the RWA biotypes SA1 and SAM), as well as the relative expression of DnTET (whole aphids of RWA biotypes SA1 and SAM that underwent host-shifts) was quantified as previously described (Burger et al., 2017). All samples and standards were quantified in triplicate along with a no template control as a measure of contamination. A five point, two times serial dilution of a zero-hour SA1 sample was used to generate quantification standards. The relative expression of DnDNMT3 and TET were calculated using Pfaffl's mathematical model (Pfaffl, 2001) for each time point (0, 6, and 48 h). A CFX96 Real-Time System (Bio-Rad) was used to perform the real-time PCR analysis. Each reaction started with a denaturation step at 95°C for 3 min, followed 40 cycles of amplification, consisting of a denaturation step at 95°C for 10 s, an annealing step at the relevant temperature for each primer set (Table S2) for 30 s, and an extension step at 72°C for 30 s. A melt curve analysis was also performed for each reaction, to verify the absence of non-specific amplification: The incubation temperature was increased in 5 s intervals, 0.5°C at a time, from 65 to 95°C. The ribosomal genes L27 and L32 were used as reference genes as they have previously been shown to be constitutively expressed, respectively, in RWA and the pea aphid (Shakesby et al., 2009; Sinha and Smith, 2014).

Measuring DNMT Protein Activity

For the extraction of aphid protein, three replicates of 150 apterous aphids ($n = 450$) of biotypes SA1 and SAM were collected, flash-frozen and stored at -80°C until use. A micropestle was used to grind aphids into a fine powder, to which 100 μ l phosphate buffered saline (50 mM Na_2HPO_4 , 50 mM Na_2HPO_4 and 150 mM NaCl, pH 7.5), 10 μ l phenylmethylsulphonyl fluoride (1 mM) and 10 μ l dithiothreitol (1 mM) were added. Homogenized mixtures were centrifuged at 15 000 rpm (4°C) for 10 min to pellet the cell debris and the resulting supernatant was transferred to a clean Eppendorf tube. Protein concentrations were quantified using the Bradford protein assay (Bradford, 1976) with Bovine Serum Albumin as standard (BioRad, USA), and the Glomax®-Multi Detection plate reader (Promega, USA) as described by Rylatt and Parish (1982).

DNA methyltransferase protein activity was quantified following the guidelines provided with Abcam's colourimetric DNMT Activity Quantification kit (Abcam, UK), and using

the maximum recommended amount of nuclear extract, 5 μ l (ranging from 7.69 to 10.96 μ g, standardized using the formula below) of each of the three biological replicates per biotype ($n = 3$). DNA methyltransferase activity in OD/h/ μ g (optical density/hour/microgram) was calculated using the formula below.

$$\text{Protein activity} = \frac{(\text{Sample OD} - \text{Blank OD})}{[\text{Protein amount (ug) x hour}]} \times 1000$$

An ANOVA was performed to test for significant differences between the sample means, with the level of significance set at $p \leq 0.05$.

Quantifying Levels of Global Methylation (5mC) and Hydroxymethylation (5hmC)

Global levels of methylation were determined utilizing a colourimetric Methylated DNA Quantification kit (Abcam) using 150 ng DNA of the three biological repeats per biotype ($n = 3$). A slight modification of the protocol was followed in the “methylation capture” section, whereby incubation of DNA and diluted capture antibody was performed for 15 h at room temperature in the dark to allow for optimal antibody binding, as opposed to 1 h at room temperature. The final plate incubation, after addition of the developer solution, was carried out for the maximum recommended time of 10 min. Absorbance at 450 nm was read in triplicate ($n = 9$) within five min of adding the stop solution, using the Glomax[®]-Multi Detection System. Relative methylation levels were calculated for each sample using the following formula:

$$\text{Relative 5mC \%} = \frac{(\text{Sample OD} - \text{Negative control OD}) / S}{(\text{Positive control OD} - \text{Negative control OD}) \times 2/P} \times 100$$

where 5mC is 5-methylcytosine, OD is optical density, S is the amount of sample DNA in ng and P is the amount of positive control in ng. An ANOVA was performed to test for significant differences between the sample means, with the level of significance set at $p \leq 0.05$.

Global hydroxymethylation levels were quantified using a colourimetric Hydroxymethylated DNA Quantification kit (Abcam), in accordance with the provided protocol. Freshly extracted DNA sample from each biotype were loaded in triplicate ($n = 3$), and standardized using the formula below (refer to S, the amount of sample DNA). The final plate incubation was carried out for 10 min, where after absorbance at 450 nm was read using the Glomax[®]-Multi Detection System. Relative hydroxymethylation levels were calculated for each sample using the following formula:

$$\text{Relative 5hmC \%} = \frac{(\text{Sample OD} - \text{Negative control II OD}) / S}{(\text{Positive control OD} - \text{Negative control II OD}) \times 5/P} \times 100$$

where 5hmC is 5-hydroxymethylcytosine, OD is optical density, S is the amount of sample DNA in ng and P is the amount of positive control in ng. An ANOVA was performed to test for significant differences between the sample means, with the level of significance set at $p \leq 0.05$.

Statistical Analysis

Microsoft Excel (2010)/XLSTAT Premium (Addinsoft Inc. USA) were used for the statistical analysis, and SigmaPlot (2001) was used to plot graphs showing the average readings and standard deviation. An ANOVA was performed to test for significant differences between the sample means, with the level of significance set at $p \leq 0.05$. The model assumptions of ANOVA (i.e., homoscedasticity and normality of the residuals), were tested for using Levene's test and the Shapiro-Wilk test, respectively (significance set at $p \leq 0.05$ for both tests). If the ANOVA null hypothesis—that the means of the treatment groups are equal—was rejected, a Fisher's LSD test was then performed.

Whole Genome Bisulfite Sequencing (WGBS) and Analysis

A total of 100 apterous female aphids of South African RWA biotypes SA1 and SAM were used for DNA extraction performed as described previously (Burger and Botha, 2017) (GenBank ID GCA_001465515.1; BioProject PRJNA297165). Three independent biological repeats of each biotype were conducted of each biotype ($n = 3$). Samples consisting of 2 μ g DNA, of both RWA biotypes SA1 and SAM were submitted to Macrogen Inc., South Korea for bisulfite treatment, library preparation and sequencing.

DNA samples were treated with the EZ DNA Methylation Lightning kit (Zymo Research) and used to construct the sequencing library utilizing the TruSeq DNA Methylation Library Kit[™] (Illumina) ($n = 1$) or Accel-NGS[®] Methyl-Seq (Swift Biosciences) for Illumina ($n = 2$), 5' tags were generated through random priming, followed by selective 3' tagging. Illumina P7 and P5 adapters were ligated through amplification to the 5' and 3' ends, respectively. The Illumina HiSeq X platform was used to sequence the bisulfite treated samples.

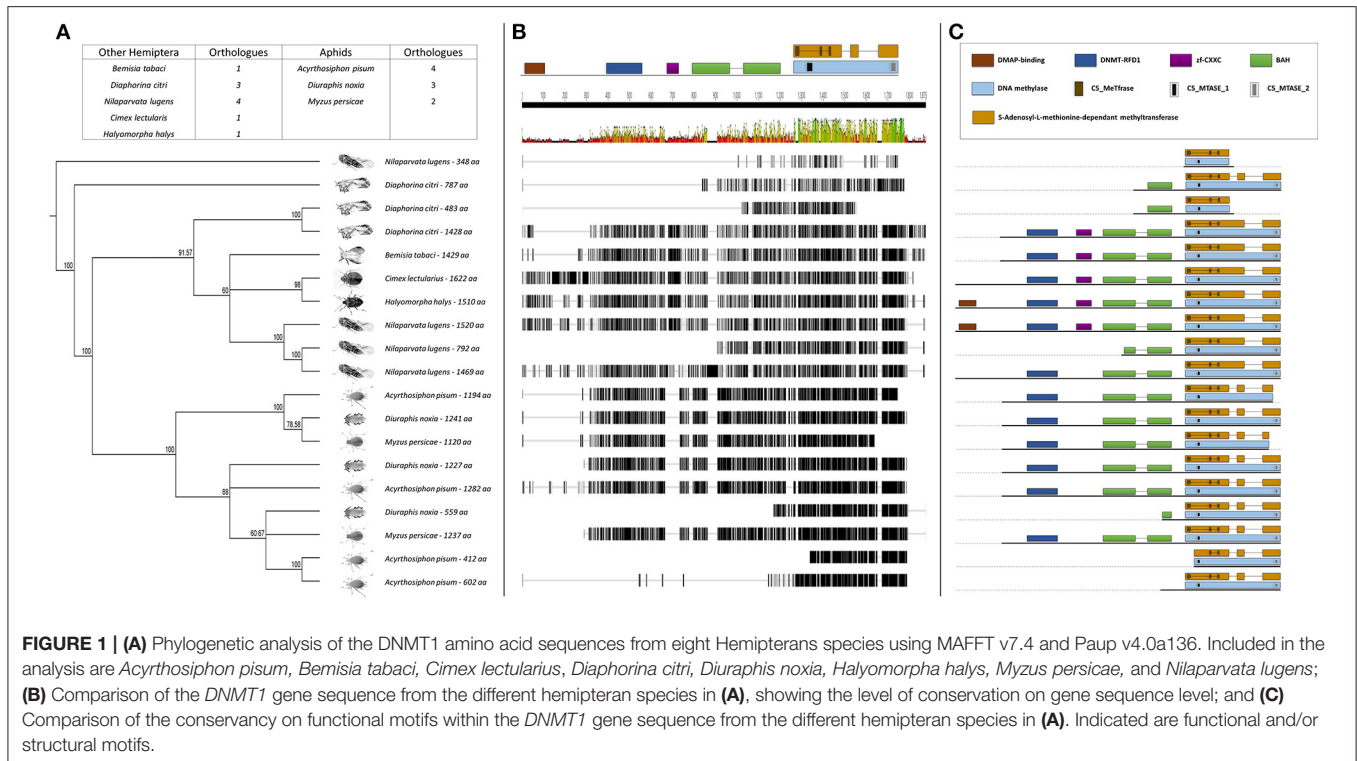
The obtained sequencing data was analyzed for quality using FastQC (Andrews, 2010). After inspecting the adapter content, per base sequence content, and per base sequence quality, Trimmomatic (Bolger et al., 2014) was used to remove adapter sequences and trim the paired-end reads for quality. A sliding window over 15 bp was used to trim for a quality score of 20, along with a headcrop of 10. The Illuminaclip parameter was used to search for and remove adapter sequences from the reads. After trimming, all reads were filtered for a minimum read length of 40 bp.

The Bismark software program (Krueger and Andrews, 2011) was used to analyze the methylation status of the trimmed and filtered sequence reads. Using the RWA SAM biotype reference genome (GenBank ID GCA_001465515.1; BioProject PRJNA297165), the observed over expected number of cytosine bases for each methylation context was calculated as follows:

$$C_p G_{\frac{O}{E}} = \frac{F_{C_p G}}{F_{C \cdot F_G}}$$

$$CHG_{\frac{O}{E}} = \frac{F_{CAG} + F_{CTG} + F_{CCG}}{3(F_{C \cdot F_1 - G} \cdot F_G)}$$

$$CHH_{\frac{O}{E}} = \frac{F_{CAA} + F_{CAT} + F_{CAC} + F_{CTA} + F_{CTT} + F_{CTG} + F_{CGA} + F_{CGT} + F_{CCG}}{9(F_{C \cdot 2F_1 - G})}$$



Where *F* represents the frequency of the subscripted nucleotide/dinucleotide/trinucleotide, in the reference genome. As a reference genome is not available for biotype SA1, the calculations were only performed for biotype SAM.

The R-package DSS-single (Wu et al., 2015) was used to calculate which genes are significantly differentially methylated (*p*-value < 0.05) between SA1 and SAM from the WGBS data. For the analysis, only genic CpG loci, with at least a ten times coverage in both biotypes, across all three repeats were considered. This amounted to 613,730 CpG sites. A Wald test (Wald, 1943) was conducted for differentially methylated loci with the DMLTest function. The optional “smoothing” algorithm of this function, which uses methylation data of nearby loci to generate “pseudo replicates” is only recommended for datasets where methylation loci are dense (Feng et al., 2014). Due to the high AT content and low methylation levels in RWA, methylation loci are sparse and the “smoothing” algorithm was not employed. The CallDMR function was then used to identify differentially methylated regions using information from the differentially methylated loci, such as the number of CpG sites in a region and the percentage of sites in a region scored as significant. This information is used to calculate a combined score for each region, referred to as an area statistic, which can be used to sort regions based on the degree of differentiation in CpG methylation (Wu et al., 2015). The Blast2GO suite (Conesa et al., 2005; Conesa and Götz, 2008) was used to search for the gene ontologies (GO) and KOG terms (Burger and Botha, 2017) of genes containing a differentially methylated region.

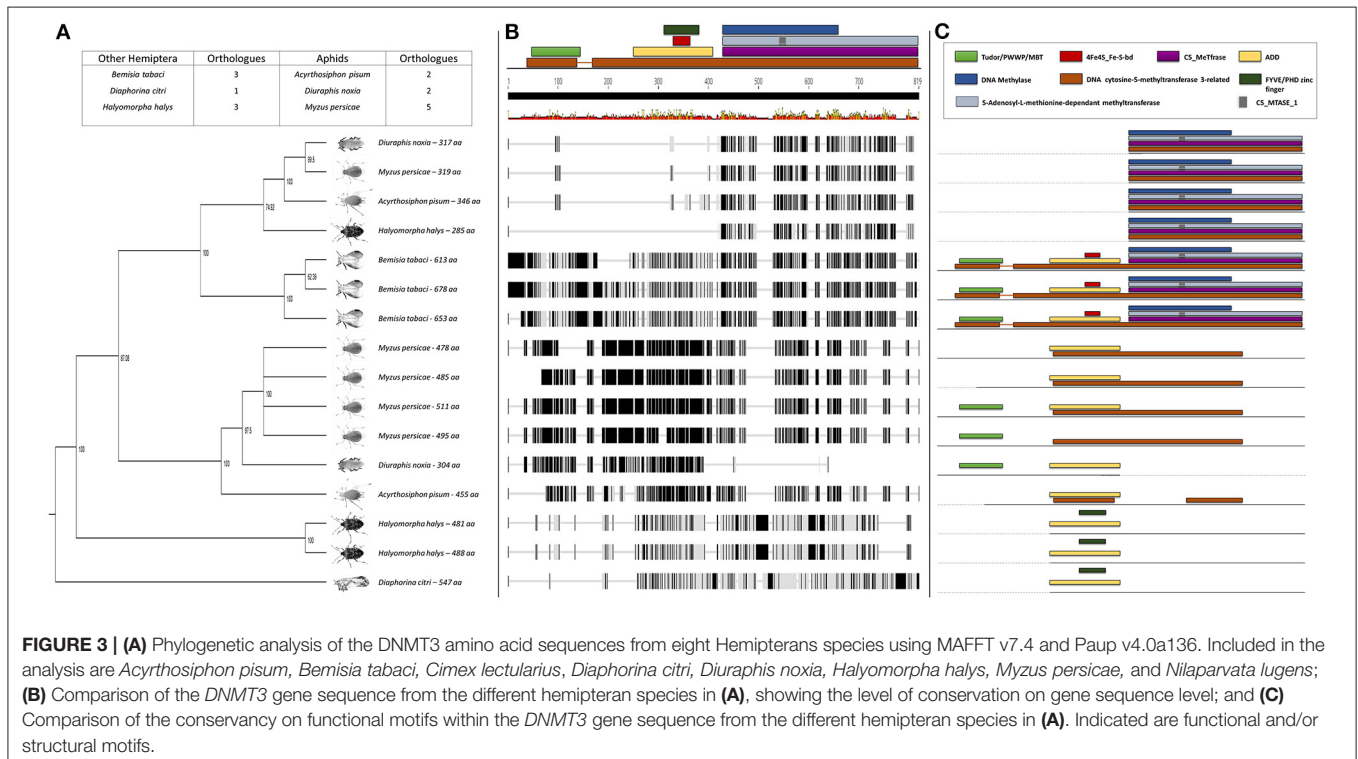
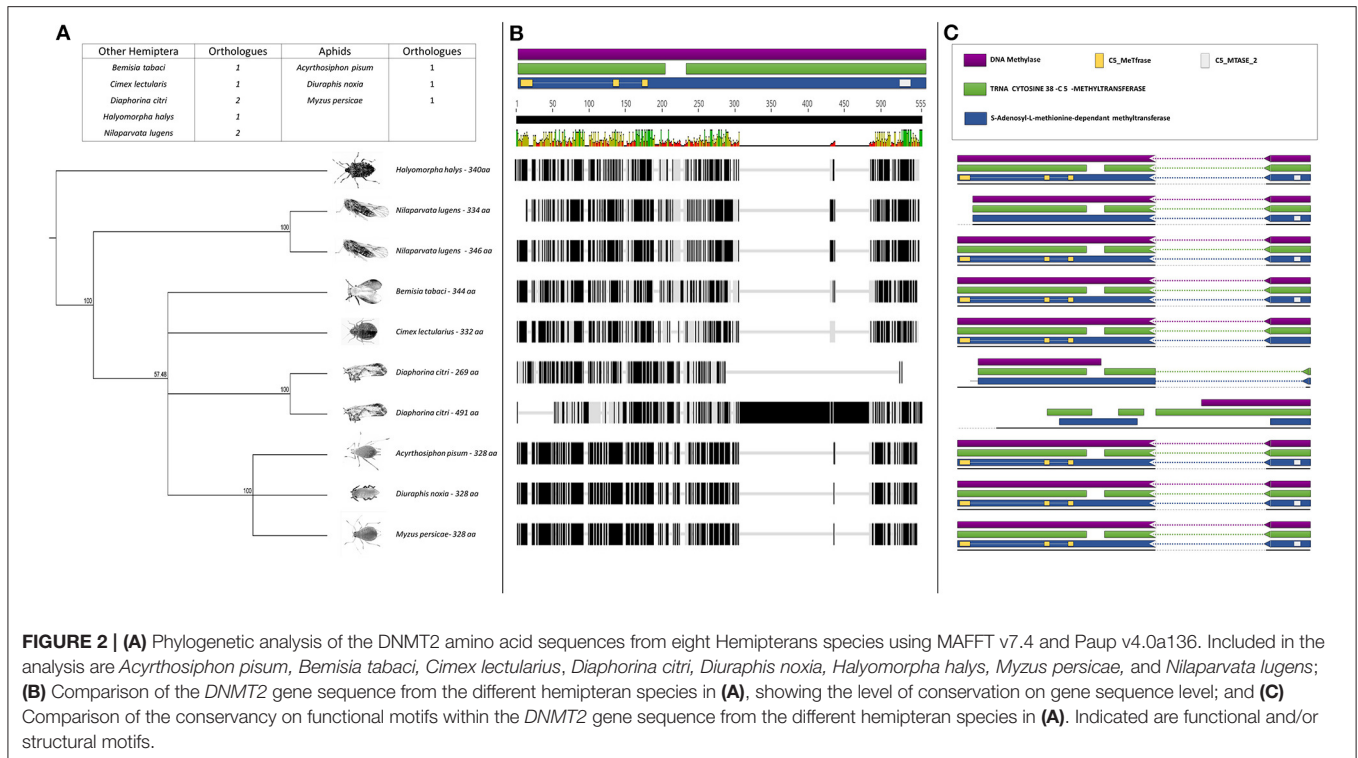
RESULTS

Classes of DNA Methyltransferases in RWA

The BLASTp analysis performed using the insect DNMTs against the RWA proteins, revealed three DNMT subfamilies. Comparison of the DNMTs of RWA with other aphids (i.e., *A. pisum* and *Myzus persicae*), as well as with other distant hemipteran species (i.e., *Bemisia tabaci*, *Cimex lectularius*, *Diaphorina citri*, *Halyomorpha halys*, and *Nilaparvata lugen*) confirmed that as with other hemipterans, RWA have three DNMT subfamilies of genes, i.e., *DNMT1*, *DNMT2*, *DNMT3* (Figures 1–3). With the *DNMT1* and *DNMT3* sequences, those from RWA were mostly similar to that of *M. persicae* and then *A. pisum*, than to the other hemipterans included in the study. Whereas, those from *N. lugens* (the brown planthopper) and *D. citri* (Asian citrus psyllid) the most distant from the plant aphids. In the case of *DNMT2*, the separation was less distinct. These observations were strongly supported by bootstrap values.

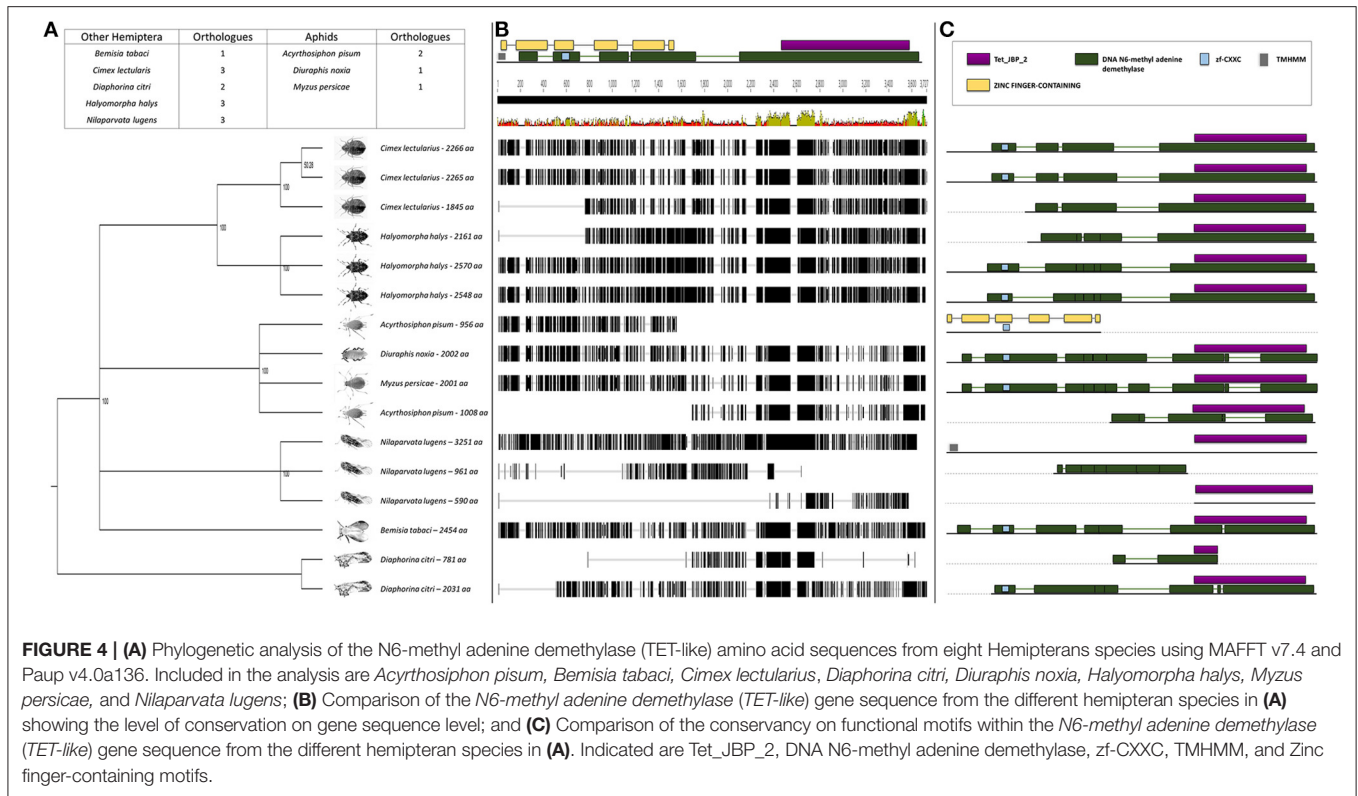
Further sequence analysis revealed that the RWA had two homologs of each of the *DNMT1*s (*DNMT1a* and *DNMT1b*) and *DNMT3*s (*DNMT3a* and *DNMT3b*) (Figures 1A, 3A), but only a single *DNMT2* (Figure 2B) protein. The RWA DNMTs contained several functional motifs that were recognizable and contributing to the ascribed function, all shared between the plant aphids (Figures 1C, 2C, 3C).

To assess whether the DNMTs differ between RWA biotypes, an alignment of the DNMT sequences obtained between the SA1 and SAM biotypes was conducted which revealed 36 SNPs between the biotypes (Figures S2–S7).



We used the transcriptome data to investigate the expression pattern of known methylation genes in RWA biotypes SA1 and SAM (Figures 5, 6). We again found the full complement of DNA methyltransferases which was expressed in different transcript

levels, with minimal differences in *DnDNMT1* and *DnDNMT2* between the RWA biotypes, with fewer *DnDNMT3* transcripts in the more virulent SAM (9.51 ± 6.9) than in the less virulent SA1 (31.45 ± 2.4) ($p = 0.006$).



To measure whether the observed SNPs had any bearing on gene expression between less and more virulent RWA biotypes, the expression of aphid head *DNMTs* among the RWA biotypes was also investigated. Biotype SAM's *DNMT1* expression was higher than that measured in biotype SA1, but the difference in expression was not significant (p -value = 0.416 for *L27* and 0.362 for *L32*), as was the expression of *DNMT2* (Figure 5A). The expression of *DNMT3* however showed the most inter-biotype variation of the three *DNMT* subfamilies, and revealed that the *DNMT3* expression levels of SAM (most virulent aphid biotype) were significantly lower than that measured in SA1 (Fisher's LSD test; p -value of ≤ 0.1).

To establish whether the difference in the *DNMT* gene expression equates into measurable differences in total *DNMT* protein activity, the *DNMT* protein activity was determined (Figure 5B). The concentration of *DNMT* protein activity within the biotypes ranged from 44.80 to 53.54 OD/h/ μ g, with biotype SAM exhibiting the lower *DNMT* protein activity of the biotypes. However, the *DNMT* protein activity levels did not differ significantly between the biotypes ($p \leq 0.05$).

Sequence Analysis and Expression of *DnTET* in RWA Biotypes

To shed light on the observed difference in global demethylation but not methylation levels, the *TET* (*N6-methyl adenine demethylase*) genes responsible for oxidation within the methylation pathway were studied (Wojciechowski et al., 2014). We were able to isolate and sequence the *DnTET* ortholog from RWA. Comparison of the *DnTET* (*N6-methyl*

adenine demethylase) protein sequences in RWA with other aphids (i.e., *A. pisum*, *M. persicae*) and with other distant hemipteran species (i.e., *B. tabaci*, *C. lectularius*, *D. citri*, *H. halys*, and *N. lugens*) confirmed that all these species have recognizable *TET*-like sequences in their genomes (Figure 4). Clustering of the sequences group the aphids closer to each other than to the other hemipterans, with strong bootstrap support (Figure 4A). Further analysis revealed that unlike *A. pisum* and some of the other hemipteran species, RWA has only a single form of *TET* (Figure 4B), which contain several functional motifs including DNA *N6-methyl adenine demethylase* (Figure 4C).

We also used the transcriptome data to investigate the expression pattern of the *TET* (*N6-methyl adenine demethylase*) genes responsible for oxidation within the methylation pathway in RWA biotypes SA1 and SAM (Figure 6C). We found more *DnTET* transcripts in the more virulent SAM (5.81 ± 0.38) than in the less virulent SA1 (2.59 ± 2.3) ($p = 0.014$).

Expression of DNA Methylation Genes During Feeding Studies

To assess further whether the sequenced *DnDNMT* and *DnTET* genes expressed in RWA biotypes SA1 and SAM, differ when challenged with different host plants, the RWA biotypes were reared on susceptible host plants ($n = 3$) (Table 1) and then moved to two resistant wheat cultivars (*Tugela-Dn1* and *Tugela-Dn5*). The lines were selected based on the fact that they are near isogenic and differ only with regard to the resistance gene present (i.e., *Dn1* expressing antibiosis—harmful

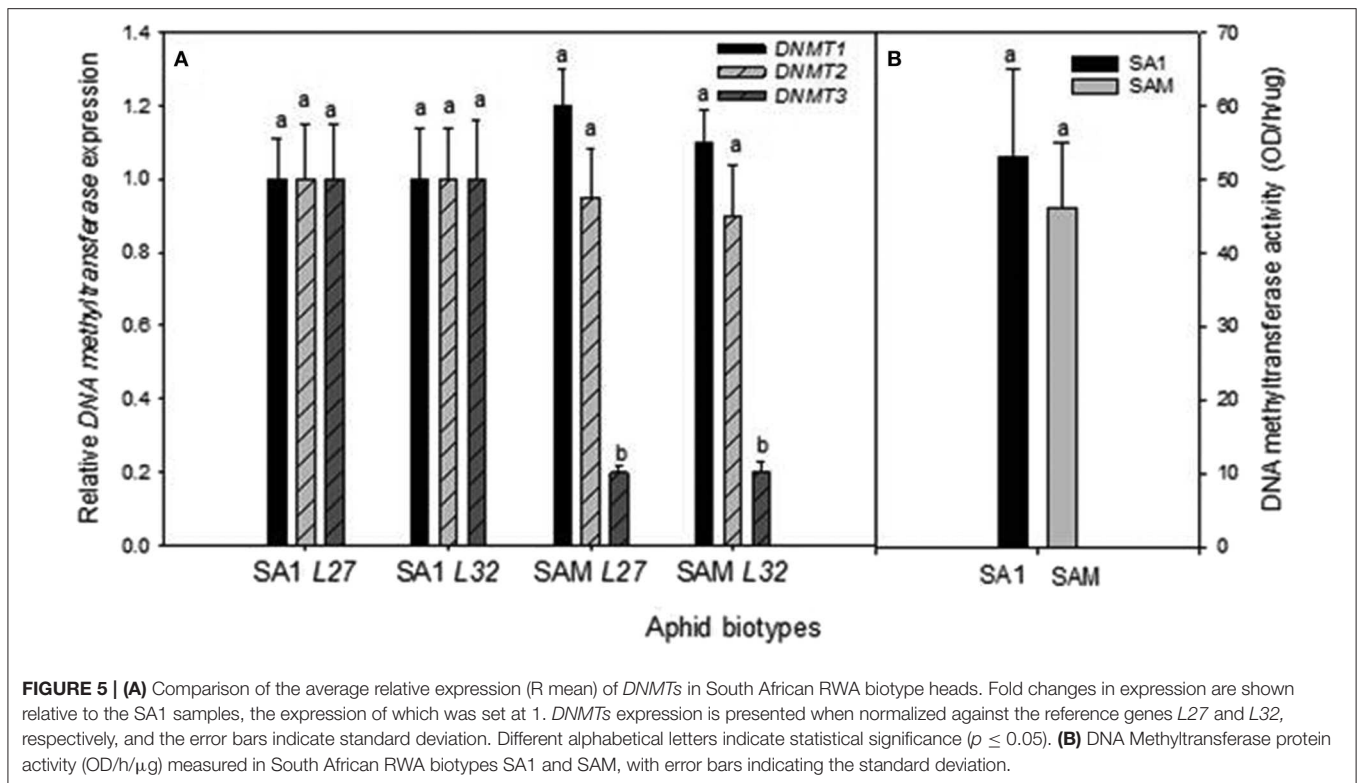


TABLE 1 | Relative expression of *DNMT3* and *TET* in RWA biotypes SA1 and SAM when feeding on a susceptible cultivar SST.

Relative gene expression	Aphid biotype			
	SA1		SAM	
Reference genes	<i>Lr27</i>	<i>Lr32</i>	<i>Lr27</i>	<i>Lr32</i>
DNMT	0.852 ± 0.126	0.734 ± 0.084	1.308 ± 0.531	0.749 ± 0.221
TET	0.758 ± 0.114	0.652 ± 0.072	1.222 ± 0.304	0.721 ± 0.223

Lr27 and *Lr32* was used as reference genes.

to the aphid and *Dn5* expressing antibiosis and antixenosis—non-palatable) (Figure 7). The measured *DnDNMT* expression increased significantly in both aphid biotypes when challenged with a new feeding environment (i.e., *Dn5* expressing both antibiotic and antixenotic) ($p \leq 0.05$). The relative expression almost doubled in the less virulent biotype SA1, but even more in the more virulent biotype SAM (\pm six-fold *Lr27*; \pm 11-fold *Lr32*) within 6 h after host-shifting.

To assess whether the sequenced *DnTET* gene was also differentially expressed in RWA, biotypes SA1 and SAM during feeding after host-shifting, the expression of *DnTET* was also measured (Figure 8). The measured *DnTET* expression increased slightly but not significantly after 6 h in the less virulent SA1 biotypes when challenged with the *Dn5* wheat line (antibiotic and antixenotic) ($p > 0.05$), but significantly in SA1 after 48 h when

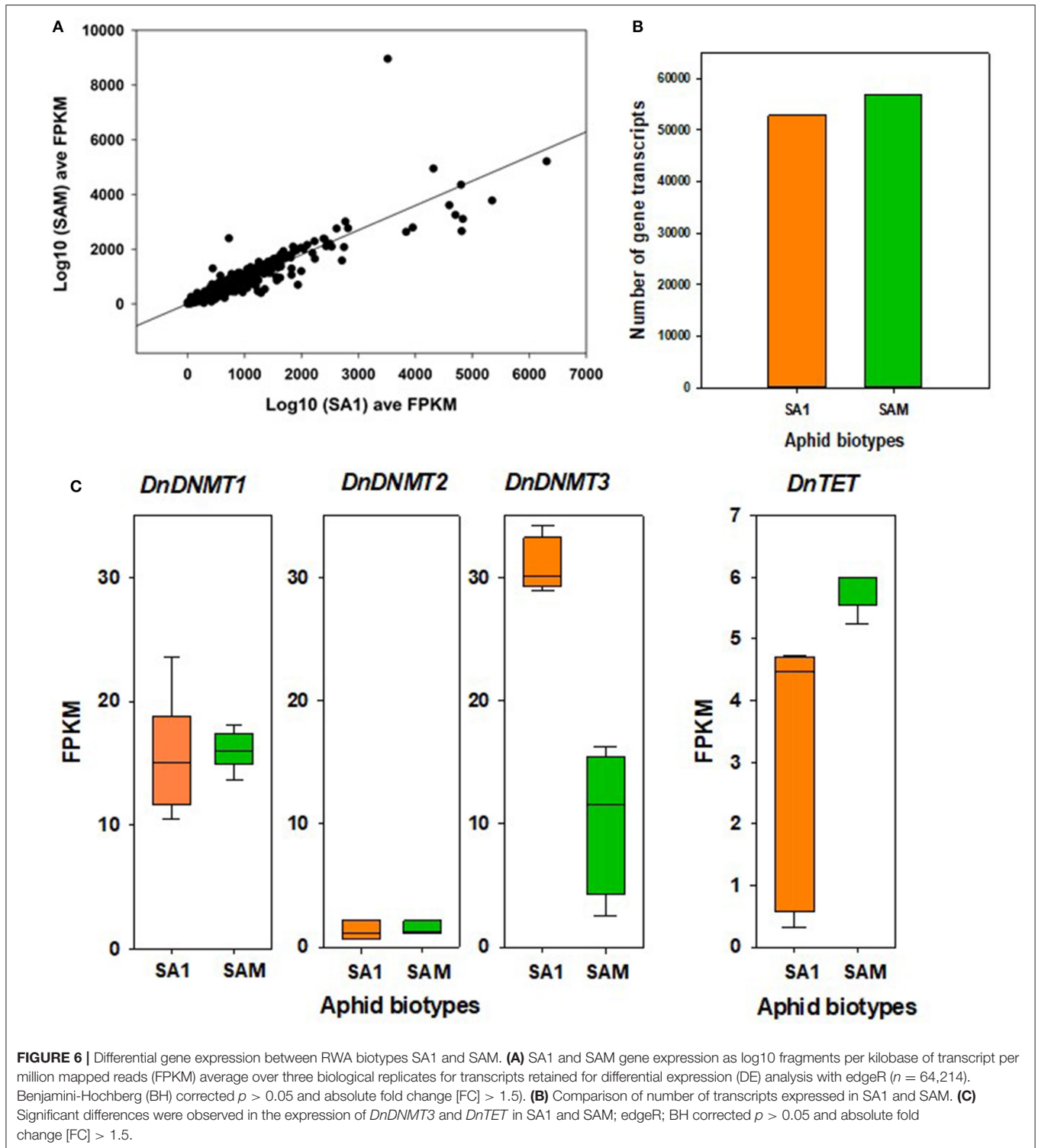
challenged with the antibiotic wheat line Tugela-*Dn1* (*Lr27*) ($p \leq 0.05$). In contrast, the *DnTET* expression of virulent biotype SAM remained the same when challenged with the antibiotic wheat line Tugela-*Dn1*, but increased significantly within the first 6 h after host-shifting from the susceptible Tugela to the wheat line containing the *Dn5* resistance gene ($p \leq 0.05$).

Global Methylation and Hydroxymethylation Quantification

To quantify the global levels of methylation (5mC) and dehydroxymethylation (5hmC), antibodies specific to these (i.e., 5mC and 5hmC) were used. The use of the 5mC antibody revealed similar levels of global methylation between the less virulent SA1 and more virulent SAM biotype, with the measured levels, ranging between 0.14 and 0.16% (Figure 9). The hydroxymethylation levels, however differed significantly ranging from 0.12 to 0.46% (Figure 9), with biotype SA1 displaying the lowest, and biotype SAM displaying the highest 5hmC levels, respectively ($p \leq 0.05$).

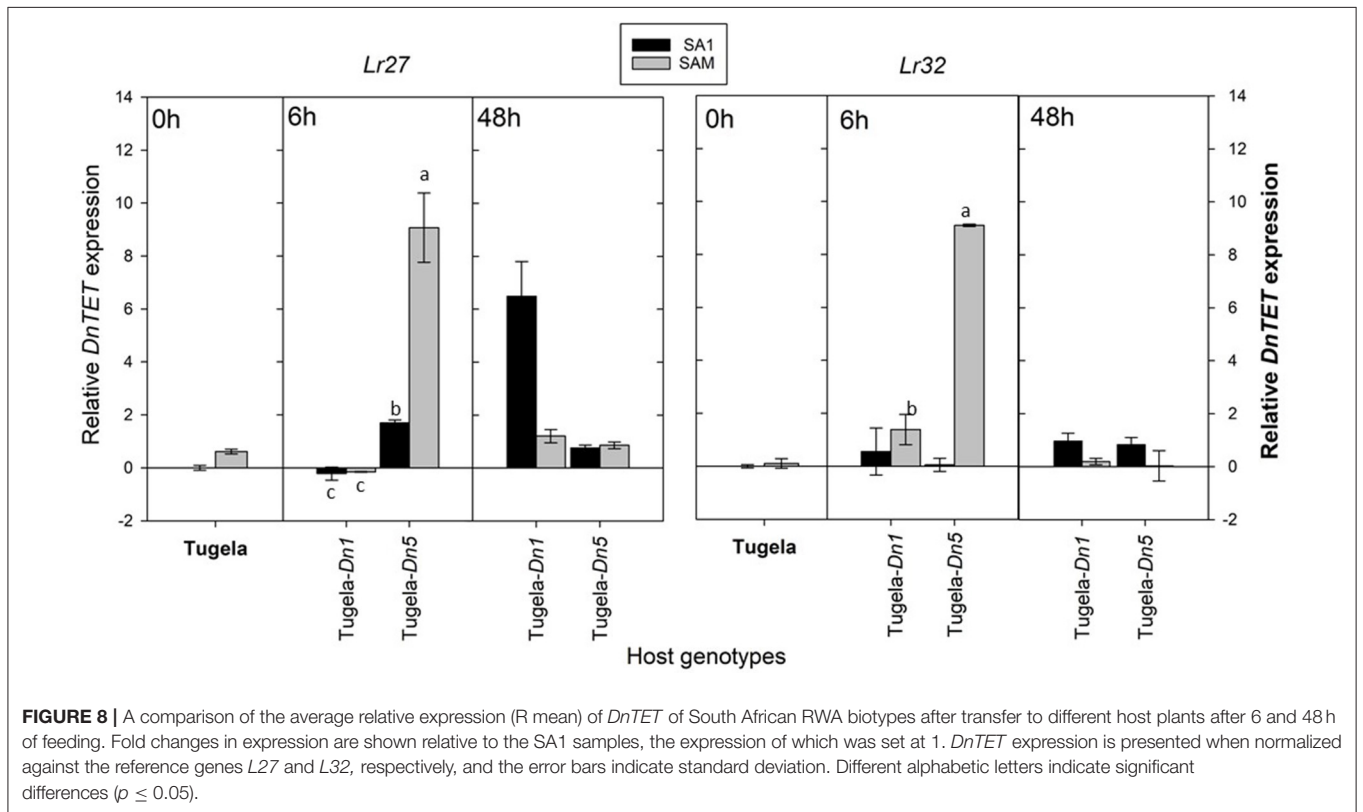
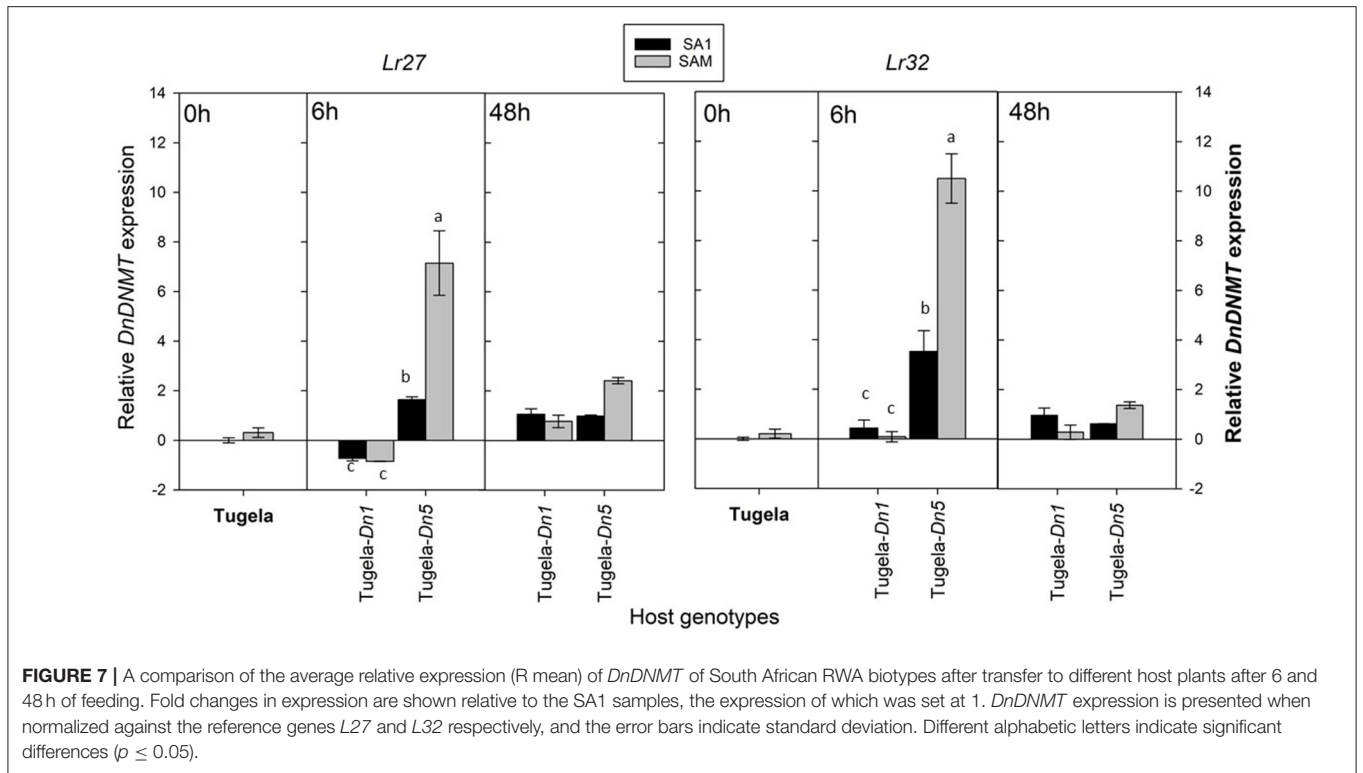
Whole Genome Bisulfite Sequencing

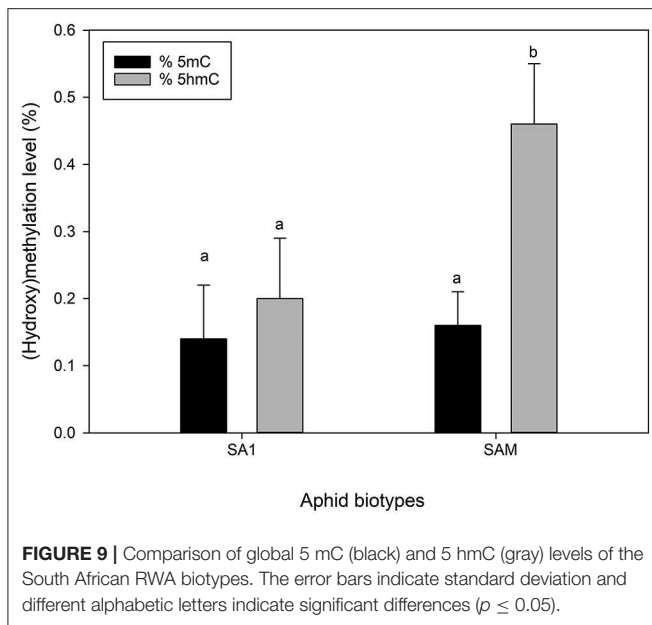
The whole genome bisulfite sequencing produced a total of 6,846,597,083 raw reads for SA1 and 7,397,965,699 raw reads for SAM, respectively, of which a total of 70,861,462 bases (SA1) and 74,073,939 bases (SAM) were methylated, which represents 1.126 ± 0.321% (SA1) and 1.105 ± 0.295% (SAM) methylation in the genome (Tables S3, S4). The sequence reads were analyzed for contexts of DNA methylation within the genome (Table S5) and the results revealed that RWA has methylation in all contexts



(CpG, CHG, and CHH), with the majority of methylation within the CpG context ($\pm 5.19\%$), while the other contexts show much lower levels of methylation (CHG— $\pm 0.27\%$; CHH— $\pm 0.34\%$). The reads were then subjected to quality analysis, aligned and mapped to the RWA biotype SAM reference genome

(GenBank ID GCA_001465515.1; BioProject PRJNA297165). Of the methylated reads, most of the methylation was located in genic regions ($\pm 1.58\%$), but intergenic methylation was also present ($\pm 0.808\%$) (Table S6). Methylation was evenly distributed between both strands, with the top strand only





containing 0.02% more methylated calls than the bottom strand (Table S7). Within genes exonic regions were found to be overall more methylated ($\pm 0.56\%$) than the intronic regions (Table S8), and the most represented context of methylation (i.e., CpG, CHG, or CHH) was in the CpG context, followed by the CHG context, with the least in the CHH context (Table S9).

Using the SAM biotype reference genome, the observed over expected number of cytosine bases were also calculated (Figure 10). This is commonly seen for the CpG context, denoted as $CpG_{O/E}$ (Hunt et al., 2010), with the data on all three contexts of cytosine methylation available from the Bismark pipeline, the other, often overlooked $CHG_{O/E}$ and $CHH_{O/E}$ ratios were also included in this study. After calculating these ratios, it was revealed that the observed CpG context, unlike with the other contexts was lower than expected (Figure 10). This is particularly so for the observed CpG context in exonic regions (Figure 10A).

After analysis, we identified 40 differentially methylated genes (DMEs) when we compared the genes that were differentially methylated between the less and more virulent biotypes (Figure 11; Figures S8A1-3, B1-9, C1-13). Even though based on broad functional categories, the differences between the least virulent SA1 and most virulent SAM seems minimal (Figure 11), further analyses of their involvement into biochemical pathways revealed that these DMEs had distinctly different predicted functions (even when involved within the same pathway, Figures S8A1-3). Interesting examples include the selective methylation of genes in more virulent SAMs, but not SA1, which include include DMEs involved in the irinotecan metabolism (Figure S8A3); metabolism of cytotoxicity by cytochrome P450 (Figures S8C10, C11); steroid hormone biosynthesis (Figure S8C2) and wax biosynthesis (Figure S8C).

DISCUSSION

Integrated pest management programs against RWA depend heavily on the breeding of wheat cultivars that provide resistance (Tolmay et al., 1997; Smith and Clement, 2012; Botha, 2013; Sinha and Smith, 2014). The effectiveness of these cultivars, however, is often short-lived as aphids overcome the resistance they impart (Botha et al., 2005, 2010; Tagu et al., 2008; Sinha and Smith, 2014). Understanding how new aphid biotypes develop, as well as the mechanisms they employ to exert their virulence enabling them to breakdown plant resistance, are of utmost importance if resistant cultivars are to be used to their full potential (Botha et al., 2014a). The availability of the highly virulent mutant RWA biotype (SAM) (Swanevelder et al., 2010), alongside South Africa's naturally occurring biotypes (SA1, SA2, SA3, and SA4) (Walters et al., 1980; Tolmay et al., 2007; Jankielsohn, 2011, 2016) presents a unique opportunity for the study of biotypification. Despite having developed from SA1, which only renders *dn3*-containing cultivars susceptible (Jankielsohn, 2011), SAM has the remarkable ability to overcome the resistance of all the *Dn* genes that have been introduced and/or documented (Botha, 2013; Botha et al., 2014a). SAM thus serves as a model to resolve aphid biotypification.

In the present study, we investigated the DNMT protein family, as they catalyze the covalent addition of a methyl group to the 5' position of cytosine in the methylation pathway. *DNMT2* seemed the most conserved with only a single form of the protein responsible for stabilizing tRNA and the regulation of protein synthesis in response to environmental cues (Becker and Weigel, 2015). In contrast, *DNMT1* and *DNMT3* with two forms each, are responsible for maintaining and establishing methylation patterns, respectively (Goll and Bestor, 2005; Goll et al., 2006; Jeltsch et al., 2006). Owing to the important role of these proteins in changing methylation patterns, it is not surprising that variations in these genes occur. RWA is also not the only insect showing multiple homologs within a specific DNMT class. Some insect lineages were shown to lack one (e.g., *B. mori* and *T. castaneum*) or two (e.g., *D. melanogaster* and *A. gambiae*) classes of DNMTs, while others have multiple homologs (e.g., *A. mellifera*, *N. vitripennis* and *A. pisum*) within a certain DNMT class (Kunert et al., 2003; Marhold et al., 2004; Walsh et al., 2010; Xiang et al., 2010; Glastad et al., 2011; Feliciello et al., 2013).

The limited available literature on aphid DNMTs prompted an investigation into the baseline DNMT expression (i.e., expression of aphids not challenged with resistance) of South African RWA. It is widely assumed that the insect DNMTs have the same functions as their mammalian orthologs (Wang et al., 2006; Glastad et al., 2014). *DNA methyltransferase 3 (DNMT3)* was the only gene of which the expression was significantly different between the two RWA biotypes. *DNMT3* has long been known as a *de novo* methyltransferase (Okano et al., 1999; Goll and Bestor, 2005), which establishes new methylation patterns by methylating previously unmethylated sites (Kunert et al., 2003; Schaefer and Lyko, 2007). The *DNMT3* expression of the virulent biotype used in this study, SAM, is down-regulated

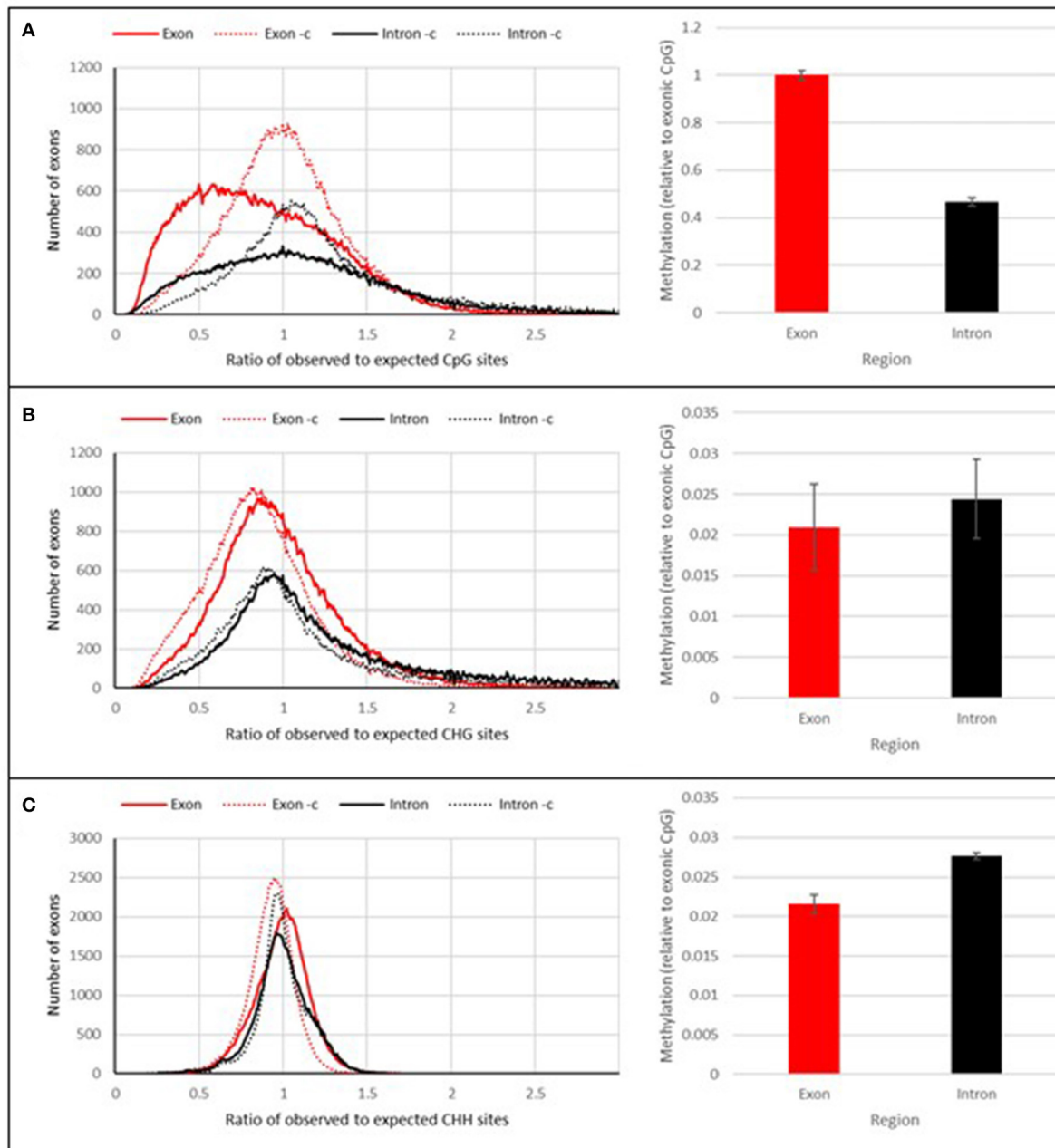
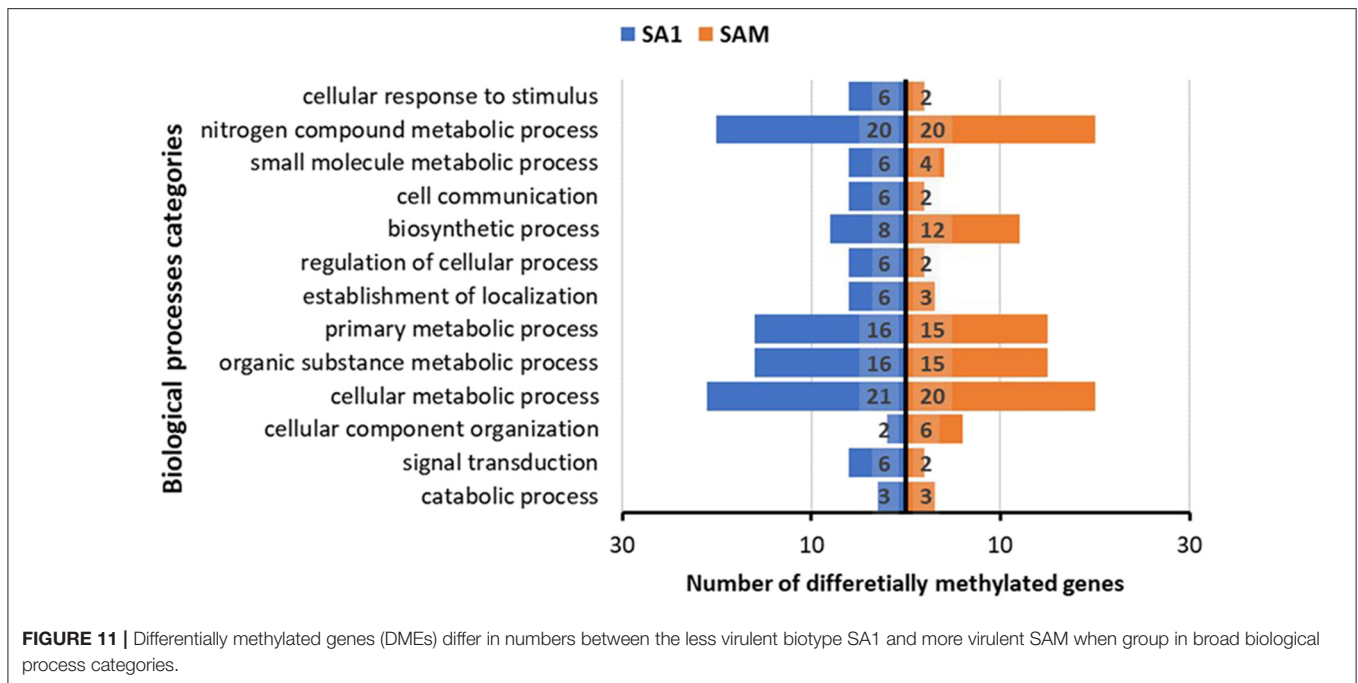


FIGURE 10 | DNA methylation context. The calculated observed over expected number of cytosine bases in the genome of RWA. **(A)**, CpG_(O/E) **(B)**, CHG_(O/E). **(C)** CHH_(O/E).

in comparison to the less virulent biotype, SA1, and this decrease in expression could therefore be advantageous from a virulence perspective.

A role for *DNMT3A* in the facilitation of transcription has also been identified, with *DNMT3A*-dependent methylation of gene bodies promoting transcription by antagonizing polycomb repression (Wu et al., 2010). Although the aphid effector genes are yet to be identified (Botha et al., 2005, 2014b), it is possible that they contain *DNMT3A* binding sites within their gene

bodies, and that their transcription could be facilitated by *DNMT3A* binding and subsequent methylation. In the current study, SA1's *DNMT3A* expression, and therefore *DNMT3A* protein production, is up-regulated in comparison to the more virulent biotype. The fact that SA1 has higher *DNMT3A* expression (and perhaps greater effector protein production) under unchallenged conditions, may provide some insight into why SA1 is the least virulent biotype. Therefore, quantifying the *DNMT3A* expression of aphids challenged by resistance may



yield valuable information on *DNMT3A*'s possible involvement in effector transcription.

Other functions of *DNMT3* include its role in the removal of 5mC and 5hmC (Chen et al., 2012, 2013) and a proposed involvement in the maintenance of methylation, by being able to “methylate sites missed by DNMT1 activity” (Jones and Liang, 2009). However, as the *DNMT3*-mediated removal of 5mC and 5hmC is dependent on certain redox conditions (Chen et al., 2012, 2013), and has only been shown to occur *in vitro* (Chen et al., 2012, 2013), it is difficult to draw conclusions regarding the *DNMT3* expression and its potential demethylating and dehydroxymethylating activities in RWA. The DNA methyltransferase 3 protein is assumed to help maintain methylation in densely methylated areas of mammalian genomes (Jones and Liang, 2009).

The hydroxylation of methylated cytosines by TET enzymes, resulting in the formation of 5hmC, is one of various active demethylation mechanisms (Tahiliani et al., 2009; Branco et al., 2012). The initial functional characterization of TETs was performed in mammals, which have three TET enzymes, namely TET1, TET2, and TET3 (Iyer et al., 2009; Tahiliani et al., 2009). In contrast to this, invertebrates possess only a single TET ortholog (Pastor et al., 2013; Wojciechowski et al., 2014), which has been identified in insects containing hydroxymethylation, including *A. mellifera* (Cingolani et al., 2013; Wojciechowski et al., 2014), *T. castaneum* (Felicciello et al., 2013), *N. vitripennis* (Pegoraro et al., 2016) and *D. melanogaster* (Dunwell et al., 2013). In 2014, Wojciechowski et al. functionally characterized the *A. mellifera* TET ortholog, AmTET, and concluded that, like the mammalian TETs, AmTET is capable of hydroxylating 5mC to form 5hmC. This provided the first evidence that TETs play a similar role in insects, as they do in mammals. The presence of measurable amounts of 5hmC in the RWA biotypes tested,

suggests that at least one active demethylation pathway (i.e., hydroxylation of 5mC by TET) is present in RWA, as confirmed in the current study when we sequenced the *DnTET* ortholog. We then studied the expression of this gene in the RWA biotypes after challenging the aphids through differential feeding, as this was previously shown to be perceived as stressful (Burger et al., 2017). Interestingly, when SA1 feeds on wheat with an antibiotic mode of resistance (e.g., *Tugela-Dn1*), an oxidative burst (elevated H₂O₂) occurs at the feeding sites (Botha et al., 2014b; Burger et al., 2017), and the expression of *DnTET* more than doubles. Whereas, when SA1 feed on wheat expressing both antibiosis and antixenosis (e.g., *Tugela-Dn5*), not only is the aphid challenged by the elevated H₂O₂ but also by volatile substances that make the wheat unpalatable (Botha et al., 2014b), resulting in the tripling of *DnTET* expression. A similar trend is not observed with the expression of *DnTET* in SAM. Biotype SAM feeding, however, is not associated with an oxidative burst or increased peroxidase activity levels, because SAM “avoids” detection by wheat hosts (Botha et al., 2014a).

In the present study, we also studied the epigenome of RWA biotypes SA1 (least virulent SA biotype) and SAM (most virulent SA biotype). The whole-genome bisulfite sequencing indicated that the genomes of these biotypes were globally more methylated (i.e., 1.126 ± 0.321% for SA1; 1.105 ± 0.295% for SAM) than previously reported for insect genomes. For example, the global methylation levels of *A. mellifera* (Lyko et al., 2010), *B. mori* (Xiang et al., 2010), the ants *Camponotus floridanus* and *Harpegnathos saltator* (Bonasio et al., 2012) and *N. vitripennis* (Beeler et al., 2014) are all between 0.1 and 0.2%.

However, when quantified using the antibody-based methods, the global methylation levels (0.14–0.16%) are in line with other reports of insect methylation. Panikar et al. (2015) investigated adult *D. melanogaster* methylation, also through an

antibody-based method, found the adult *D. melanogaster* genome to be ~0.5% methylated. Russian wheat aphids thus have low, but detectable levels of methylation which are ~0.2 to 0.4-fold of that of the model organism *D. melanogaster*, as measured using the same technique, which allows a more direct comparison. Although other authors have reported lower levels of adult *D. melanogaster* methylation using bisulfite sequencing (0% – Lyko et al., 2000), liquid chromatography tandem mass spectrometry (0.034% – Capuano et al., 2014) and thin layer chromatography (0.05–0.1% – Gowher et al., 2000).

The sequence reads were analyzed for contexts of DNA methylation within the genome and the results revealed that RWA has methylation in all contexts (CpG, CHG and CHH), with the majority of methylation within the CpG context ($\pm 5.19\%$), but still notable methylation in the other contexts (CHG— $\pm 0.27\%$; CHH— $\pm 0.34\%$), with most of the methylation located in the genic regions. A similar finding was recently reported by Mathers et al. (2019) in the green peach aphid, *Myzus persicae*. The authors found that exons are highly enriched for methylated CpGs, particularly at the 3' end of genes. Their findings also alludes to sex-biased differential methylation of genes involved in aphid sexual differentiation.

The model organisms, mice (*Mus musculus*) and zebra fish (*Danio rerio*), showed very low levels of CHH and CHG methylation (1% and lower) when compared to the 74.2% and 80.3% CpG methylation observed, respectively. Low levels of CHG (0.26%) and CHH (0.17%) were also reported for the honeybee (*Apis mellifera*) (Feng et al., 2010). In another insect example, the over expression of DNMT2-like protein in fruit flies (*Drosophila melanogaster*) resulted in observable CpT and CpA methylation (Kunert et al., 2003).

We also wanted to assess whether the DNA strands (top vs. bottom) were methylated equally and found that the top strands were slightly more methylated (0.06%) than the bottom ones. However, interestingly the difference between the methylation of the top and bottom strands were significantly more in the more virulent SAM (0.09%), than SA1 (0.02%). This was not an unexpected finding, as biotype SAM was previously found to exhibit the highest level of hemimethylation (at the external cytosine) when its methylation was investigated using the Methylation-Sensitive Amplification Polymorphism (MSAP) technique (Breeds et al., 2018). Hemimethylated DNA arises during DNA replication, as the newly synthesized daughter strand contains unmodified cytosines (Jeltsch, 2002; Goll and Bestor, 2005). When the levels of 5-hmC were measured using the antibody based method, earlier observations (Breeds et al., 2018) were confirmed, as the levels of 5hmC differed significantly between the aphid biotypes, with SAM much higher than that measured in the less virulent SA1 ($p < 0.05$; **Figure 9**).

Analysis of the 40 differentially methylated genes (DMEs) revealed that the less and more virulent biotypes had distinctly different DMEs. As previously indicated, DMEs in the more virulent biotype SAM include DMEs involved in the irinotecan metabolism where it seemingly regulates the conversion between SN-38 and SN38G which regulates secretion. SN-38 produced in the body by carboxylesterase is the active metabolite of irinotecan (Fujita et al., 2015), with its mechanism of action

thought to be its interaction with the cleavable complex of DNA and a nuclear protein topoisomerase I. This then results in a blockade of DNA replication (Hsiang and Liu, 1988; Hertzberg et al., 1989) which causes double-strand DNA breakage and cell death. This process may be linked to the metabolism of xenobiotics by cytochrome P450 and ascorbate and aldarate metabolism that may enable SAM to counter its feeding environment better than its parent, the less virulent SA1. In locusts, it was demonstrated that an ascorbate-recycling system in the midgut lumen can act as an effective antioxidant defense in caterpillars that feed on prooxidant-rich foods, emphasizing the importance of a defensive strategy in herbivorous insects based on the maintenance of conditions in the gut lumen that reduce or eliminate the potential prooxidant behavior of ingested phenols (Barbehenn et al., 2001). Interestingly, more virulent SAM also selectively methylates genes associated with steroid hormone biosynthesis and wax biosynthesis, both pathways producing chemicals that affect the physiological processes associated with insect development and insect survival (Niwa and Niwa, 2016). The less virulent SA1 have DMEs associated with the glutathione biosynthesis which may signal stress responses.

Collectively, the results suggest that RWA biotype SAM has a greater capacity to actively methylate/demethylate its DNA than its parent SA1. Thus, it can be concluded with fair confidence that SAM undergoes more methylation/demethylation than its parent biotype SA1. However, the question that remains is whether this ability to actively methylate/demethylate occurs at specific sets of genes depending on the environmental cue/stress RWA is faced with, as opposed to occurring globally (although global, genome-wide demethylation was measured). As gene bodies are the predominant sites of methylation in insects (Zemach et al., 2010; Glastad et al., 2011; Lyko and Maleszka, 2011), it is likely that it is in these regions that methylation will be removed. Also, removal of intragenic methylation of certain genes may alter the transcripts that are produced, by exposing cryptic binding sites or intragenic promoters (Maunakea et al., 2010; Hunt et al., 2013a) and/or affect the splice variants that are produced, through methylation's involvement in alternative splicing (Lyko and Maleszka, 2011; Shukla et al., 2011; Bonasio et al., 2012; Maunakea et al., 2013; Glastad et al., 2014; Yan et al., 2015). As demethylation can occur in a matter of hours (Glastad et al., 2011), the greater capability of SAM to demethylate its genome, may provide SAM with more flexibility to adapt to changing environments, and therefore may underlie SAM's ability to overcome plant resistance. However, this is an aspect that requires further investigation in future.

DATA AVAILABILITY STATEMENT

The whole epigenome sequencing data from SA1 (accession SRX4643785) and SAM (accession SRX4643786) were deposited to the National Center for Biotechnology Information (NCBI) (GEO SUBMISSION GSE119504) (<https://www.ncbi.nlm.nih.gov/gds/?term=diuraphis%20noxia>) with Bioproject number PRJNA489432.

AUTHOR CONTRIBUTIONS

A-MB conceived and designed the experiments as well as drafted the manuscript. Aphid rearing and sampling was performed by KB and NB. DNMT expression and TET expression was performed by KB and JT respectively, while DNMT and TET sequencing was performed by KB and HS, respectively. Alignments and phylogenetic analyses of DNMT and TET sequences were performed by NB while 5mC and 5hmC analyses

were performed by KB, NB, and A-MB. Whole genome bisulfite sequencing and its analysis was performed by PP and NB. All authors edited and reviewed the final manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00452/full#supplementary-material>

REFERENCES

- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. doi: 10.1093/nar/25.17.3389
- Andrews, S. (2010). *FastQC: A Quality Control Tool for High Throughput Sequence Data*. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed October 6, 2011).
- Aranda, P. S., Lajoie, D. M., and Jorcyk, C. L. (2012). Bleach gel: a simple agarose gel for analyzing RNA quality. *Electroph* 33, 366–369. doi: 10.1002/elps.201100335
- Babu, C. V. S., and Gassmann, M. (2016). *Assessing Integrity of Plant RNA With the Agilent 2100 Bioanalyzer System*. Waldbronn: Agilent Technologies.
- Barbehenn, R. V., Bumgarner, S. L. F., Roosen, E., and Martin, M. M. (2001). Antioxidant defenses in caterpillars: role of the ascorbate-recycling system in the midgut lumen. *J. Insect Physiol.* 47, 349–357. doi: 10.1016/S0022-1910(00)00125-6
- Becker, C., Weigel, D. (2015). Epigenetic variation: origin and transgenerational inheritance. *Curr. Opin. Plant Biol.* 15, 562–7. doi: 10.1016/j.pbi.2012.08.004
- Beeler, S. M., Wong, G. T., Zheng, J. M., Bush, E. C., Remnant, E. J., Benjamin, P., et al. (2014). Whole-genome DNA methylation profile of the jewel wasp (*Nasonia vitripennis*). *G3* 4, 383–388. doi: 10.1534/g3.113.008953
- Benjamini, V., and Hochberg, V. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B.* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Bonasio, R., Li, Q., Lian, J., Mutti, N. S., Jin, L., Zhao, H., et al. (2012). Genome-wide and caste-specific DNA methylomes of the ants *Camponotus floridanus* and *Harpegnathos saltator*. *Curr. Biol.* 22, 1755–1764. doi: 10.1016/j.cub.2012.07.042
- Botha, A.-M. (2013). A coevolutionary conundrum: the arms race between *Diuraphis noxia* (Kurdjumov) a specialist pest and its host *Triticum aestivum* (L.). *Arthropod Plant Interact.* 7, 359–372. doi: 10.1007/s11829-013-9262-3
- Botha, A.-M., Burger, N. F. V., and van Eck, L. (2014a). Hypervirulent *Diuraphis noxia* (Hemiptera: Aphididae) biotype SAM avoids triggering defenses in its host (*Triticum aestivum*) (Poales: Poaceae) during feeding. *Environ. Entomol.* 43, 672–681. doi: 10.1603/EN13331
- Botha, A.-M., Li, Y., and Lapitan, N. L. V. (2005). Cereal host interactions with Russian wheat aphid: a review. *J. Plant Interact.* 1, 211–222. doi: 10.1080/17429140601073035
- Botha, A.-M., Swanevelder, Z. H., and Lapitan, N. L. V. (2010). Transcript profiling of wheat genes expressed during feeding by two different biotypes of *Diuraphis noxia*. *Environ. Entomol.* 39, 1206–1231. doi: 10.1603/EN09248
- Botha, A.-M., van Eck, L., Burger, N. F. V., and Swanevelder, Z. H. (2014b). Near-isogenic lines of *Triticum aestivum* with distinct modes of resistance exhibit dissimilar transcriptional regulation during *Diuraphis noxia* feeding. *Biol. Open* 3, 1116–1126. doi: 10.1242/bio.201410280
- Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72, 248–254. doi: 10.1016/0003-2697(76)90527-3
- Branco, M. R., Ficuz, G., and Reik, W. (2012). Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat. Rev. Genet.* 13, 7–13. doi: 10.1038/nrg3080
- Breeds, K., Burger, N. F. V., and Botha, A.-M. (2018). New insights into the methylation status of virulent *Diuraphis noxia* (Hemiptera: Aphididae) biotypes. *J. Econ. Entomol.* 111, 1395–1403. doi: 10.1093/jeet/toy039
- Burd, J. D., Porter, D. R., Puterka, G. J., Haley, S. D., and Peairs, F. B. (2006). Biotypic variation among North American Russian wheat aphid (Homoptera: Aphididae) populations. *J. Econ. Entomol.* 99, 1862–1866. doi: 10.1093/jeet/99.5.1862
- Burger, N. F. V., and Botha, A. M. (2017). Genome of Russian wheat aphid an economically important cereal aphid. *Stand. Genomic Sci.* 12:90. doi: 10.1186/s40793-017-0307-6
- Burger, N. F. V., Venter, E., and Botha, A.-M. (2017). Profiling *Diuraphis noxia* (Hemiptera: Aphididae) transcript expression of the biotypes SA1 and SAM feeding on various *Triticum aestivum* varieties. *J. Econ. Entomol.* 110, 692–701. doi: 10.1093/jeet/tow313
- Capuano, F., Mülleder, M., Kok, R., Blom, H. J., and Ralser, M. (2014). Cytosine DNA methylation is found in *Drosophila melanogaster* but absent in *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, and other yeast species. *Anal. Chem.* 86, 3697–3702. doi: 10.1021/ac500447w
- Chen, C.-C., Wang, Y., and Shen, K. J. (2012). The mammalian *de novo* DNA methyltransferases DNMT3A and DNMT3B are also DNA 5-hydroxymethylcytosine dehydroxymethylases. *J. Biol. Chem.* 287, 33116–33121. doi: 10.1074/jbc.C112.406975
- Chen, C.-C., Wang, Y., and Shen, K. J. (2013). DNA 5-methylcytosine demethylation activities of the mammalian DNA methyltransferases. *J. Biol. Chem.* 288, 9084–9091. doi: 10.1074/jbc.M112.445585
- Cingolani, P., Cao, X., Khetani, R. S., Chen, C.-C., Coon, M., Sammak, A., et al. (2013). Intronic non-CG DNA hydroxymethylation and alternative mRNA splicing in honey bees. *BMC Genomics* 14:666. doi: 10.1186/1471-2164-14-666
- Clua, A., Castro, A. M., Ramos, S., Gimenez, D. O., and Vasicek, A. (2004). The biological characteristics and distribution of the greenbug, *Schizaphis graminum*, and Russian wheat aphid, *Diuraphis noxia* (Hemiptera: Aphididae), in Argentina and Chile. *Eur. J. Entomol.* 101, 193–198. doi: 10.14411/eje.2004.024
- Conesa, A., and Götz, S. (2008). Blast2GO: a comprehensive suite for functional analysis in plant genomics. *Int. J. Plant Genomics.* 2008:619832. doi: 10.1155/2008/619832
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610
- Delatte, B., Wang, F., Ngoc, L. V., Collignon, E., Bonvin, E., Deplus, R., et al. (2016). Transcriptome-wide distribution and function of RNA hydroxymethylcytosine. *Science* 351, 282–285. doi: 10.1126/science.aac5253
- Dunwell, T. L., McGuffin, L. J., Dunwell, J. M., and Pfeifer, G. P. (2013). The mysterious presence of a 5-methylcytosine oxidase in the *Drosophila* genome. Possible explanations. *Cell Cycle* 12, 3357–3365. doi: 10.4161/cc.26540
- Feliciello, I., Parazajder, J., Akrap, I., and Ugarković, D. (2013). First evidence of DNA methylation in insect *Tribolium castaneum*. Environmental regulation of DNA methylation within heterochromatin. *Epigenetics* 8, 534–541. doi: 10.4161/epi.24507

- Feng, H., Conneely, K. N., and Wu, H. (2014). A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data. *Nucl. Acids Res.* 42:e69. doi: 10.1093/nar/gku154
- Feng, S., Cokus, S. J., Zhang, X., Chen, P.-Y., Bostick, M., Goll, M. G., et al. (2010). Conservation and divergence of methylation patterning in plants and animals. *Proc. Natl. Acad. Sci. U S A.* 107, 8689–8694. doi: 10.1073/pnas.1002720107
- Fujita, K. I., Kubota, Y., Ishida, H., and Sasaki, Y. (2015). Irinotecan, a key chemotherapeutic drug for metastatic colorectal cancer. *World J. Gastroenterol.* 21, 12234–12248. doi: 10.3748/wjg.v21.i43.12234
- Glastad, K. M., Hunt, B. G., and Goodisman, M. A. D. (2014). Evolutionary insights into DNA methylation in insects. *Curr. Opin. Insect Sci.* 1, 25–30. doi: 10.1016/j.cois.2014.04.001
- Glastad, K. M., Hunt, B. G., Yi, S. V., and Goodisman, M. A. D. (2011). DNA methylation in insects: on the brink of the epigenomic era. *Insect Mol. Biol.* 20, 553–565. doi: 10.1111/j.1365-2583.2011.01092.x
- Goll, M. G., and Bestor, T. H. (2005). Eukaryotic cytosine methyltransferases. *Annu. Rev. Biochem.* 74, 481–514. doi: 10.1146/annurev.biochem.74.010904.153721
- Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, L., Zhang, X., et al. (2006). Methylation of tRNA^{Asp} by the DNA methyltransferase homolog Dnmt2. *Science* 311, 395–398. doi: 10.1126/science.1120976
- Gong, L., Cui, F., Sheng, C., Lin, Z., Reeck, G., Xu, J., et al. (2012). Polymorphism and methylation of four genes expressed in salivary glands of Russian wheat aphid (Homoptera: Aphididae). *J. Econ. Entomol.* 105, 232–241. doi: 10.1603/EC11289
- Götz, S., Garcia-Gomez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420–3435. doi: 10.1093/nar/gkn176
- Gowher, H., Leismann, O., and Jeltsch, A. (2000). DNA of *Drosophila melanogaster* contains 5-methylcytosine. *EMBO J.* 19, 6918–6923. doi: 10.1093/emboj/19.24.6918
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. (2013). *De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Prot.* 8:1494. doi: 10.1038/nprot.2013.084
- Haley, S. D., Peairs, F. B., Walker, C. B., Rudolph, J. B., and Randolph, T. L. (2004). Occurrence of a new Russian wheat aphid biotype in Colorado. *Crop Sci.* 44, 1589–1592. doi: 10.2135/cropsci2004.1589
- Hertzberg, R. P., Caranfa, M. J., Holden, K. G., Jakas, D. R., Gallagher, G., Mattern, M. R., et al. (1989). Modification of the hydroxyl lactone ring of camptothecin: inhibition of mammalian topoisomerase I and biological activity. *J. Med. Chem.* 32, 715–720. doi: 10.1021/jm00123a038
- Hsiang, Y. H., and Liu, L. F. (1988). Identification of mammalian DNA topoisomerase I as an intercellular target of the anticancer drug camptothecin. *Cancer Res.* 48, 1722–1726.
- Hunt, B. G., Brisson, J. A., Yi, S. V., and Goodisman, M. A. D. (2010). Functional conservation of DNA methylation in the pea aphid and the honeybee. *Genome Biol. Evol.* 2, 719–728. doi: 10.1093/gbe/evq057
- Hunt, B. G., Glastad, K. M., Yi, S. V., and Goodisman, M. A. D. (2013a). Patterning and regulatory associations of DNA methylation are mirrored by histone modifications in insects. *Genome Biol. Evol.* 5, 591–598. doi: 10.1093/gbe/evt030
- Hunt, B. G., Glastad, K. M., Yi, S. V., and Goodisman, M. A. D. (2013b). The function of intragenic DNA methylation: insights from insect epigenomes. *Integr. Comp. Biol.* 53, 319–328. doi: 10.1093/icb/ict003
- Iyer, L. M., Abhiman, S., and Aravind, L. (2011). Natural history of eukaryotic DNA methylation systems. *Prog. Mol. Biol. Transl. Sci.* 101, 25–104. doi: 10.1016/B978-0-12-387685-0.00002-0
- Iyer, L. M., Tahiliani, M., Rao, A., and Aravind, L. (2009). Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell Cycle* 8, 1698–1710. doi: 10.4161/cc.8.11.8580
- Jankielsohn, A. (2011). Distribution and diversity of Russian wheat aphid (Hemiptera: Aphididae) biotypes in South Africa and Lesotho. *J. Econ. Entomol.* 104, 1736–1741. doi: 10.1603/EC11061
- Jankielsohn, A. (2016). Changes in the Russian wheat aphid (Hemiptera: Aphididae) biotype complex in South Africa. *J. Econ. Entomol.* 109, 907–912. doi: 10.1093/jee/tov408
- Jeltsch, A. (2002). Beyond Watson and crick: DNA methylation and molecular enzymology of DNA methyltransferases. *ChemBiochem* 3, 274–293. doi: 10.1002/1439-7633(20020402)3:4<274::AID-CBIC274>3.0.CO;2-S
- Jeltsch, A., Nellen, W., and Lyko, F. (2006). Two substrates are better than one: dual specificities for Dnmt2 methyltransferases. *Trends Biochem. Sci.* 31, 306–308. doi: 10.1016/j.tibs.2006.04.005
- Jones, P. A., and Liang, G. (2009). Rethinking how DNA methylation patterns are maintained. *Nat. Rev. Genet.* 10, 805–811. doi: 10.1038/nrg2651
- Jurkowski, T. P., and Jeltsch, A. (2011). On the evolutionary origin of eukaryotic DNA methyltransferases and Dnmt2. *PLoS ONE* 6:e28104. doi: 10.1371/journal.pone.0028104
- Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010
- Kohli, R. M., and Zhang, Y. (2013). TET enzymes, TDG and the dynamics of DNA demethylation. *Nature* 502, 472–479. doi: 10.1038/nature12750
- Krueger, F., and Andrews, S. R. (2011). Bismark: a flexible aligner and methylation caller for bisulfite-seq applications. *Bioinformatics* 27, 1571–1572. doi: 10.1093/bioinformatics/btr167
- Kunert, N., Marhold, J., Stanke, J., Stach, D., and Lyko, F. (2003). A Dnmt2-like protein mediates DNA methylation in *Drosophila*. *Development* 130, 5083–5090. doi: 10.1242/dev.00716
- Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9:357. doi: 10.1038/nmeth.1923
- Li, B., and Dewey, C. N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* 12:323. doi: 10.1186/1471-2105-12-323
- Lyko, F., Foret, S., Kucharski, R., Wolf, S., Falckenhayn, C., and Maleszka, R. (2010). The honey bee epigenomes: differential methylation of brain DNA in queens and workers. *PLoS Biol.* 8:e1000506. doi: 10.1371/journal.pbio.1000506
- Lyko, F., and Maleszka, R. (2011). Insects as innovative models for functional studies of DNA methylation. *Trends Genet.* 27, 127–131. doi: 10.1016/j.tig.2011.01.003
- Lyko, F., Ramsahoye, B. H., Jaenisch, R. (2000). DNA methylation in *Drosophila melanogaster*. *Nature* 408, 538–540. Available online at: <https://www.nature.com/articles/35046205>
- Marhold, J., Rothe, N., Pauli, A., Mund, C., Kuehle, K., Brueckner, B., et al. (2004). Conservation of DNA methylation in dipteran insects. *Insect Mol. Biol.* 13, 117–123. doi: 10.1111/j.0962-1075.2004.00466.x
- Mathers, T. C., Mugford, T. S., Percival-Alwyn, L. Y., Chen, K., and Swarbreck, D. (2019). Sex-specific changes in the aphid DNA methylation landscape. *Mol. Ecol.* 28, 4228–4241. doi: 10.1111/mec.15216
- Maunakea, A. K., Chepelev, I., Cui, K., and Zhao, K. (2013). Intragenic DNA methylation modulates alternative splicing by recruiting MeCP2 to promote exon recognition. *Cell Res.* 23, 1256–1269. doi: 10.1038/cr.2013.110
- Maunakea, A. K., Nagarajan, R. P., Bilenyk, M., Ballinger, T. J., D'Souza, C., Fouse, S. D., et al. (2010). Conserved role of intragenic DNA methylation in regulating alternative promoters. *Nature* 466, 253–257. doi: 10.1038/nature09165
- Niwa, Y. S., and Niwa, R. (2016). Transcriptional regulation of insect steroid hormone biosynthesis and its role in controlling timing of molting and metamorphosis. *Dev. Growth Differ.* 58, 94–105. doi: 10.1111/dgd.12248
- Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. *Cell* 99, 247–257. doi: 10.1016/S0092-8674(00)81656-6
- Panikar, C. S., Rajpathak, S. N., Abhyankar, V., Deshmukh, S., and Deobagkar, D. D. (2015). Presence of DNA methyltransferase activity and CpC methylation in *Drosophila melanogaster*. *Mol. Biol. Rep.* 42, 1615–1621. doi: 10.1007/s11033-015-3931-5
- Pastor, W. A., Aravind, L., and Rao, A. (2013). TETonic shift: biological roles of TET proteins in DNA demethylation and transcription. *Nat. Rev. Mol. Cell Biol.* 14, 341–356. doi: 10.1038/nrm3589
- Pegoraro, M., Bafna, A., Davies, N. J., Shuker, D. M., and Tauber, E. (2016). DNA methylation changes induced by long and short photoperiods in *Nasonia*. *Genome Res.* 26, 203–210. doi: 10.1101/gr.196204.115
- Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29:e45. doi: 10.1093/nar/29.9.e45
- Raddatz, G., Guzzardob, P. M., Olovac, N., Fantappied, M. R., Ramppe, M., Schaefera, M., et al. (2013). Dnmt2-dependent methylomes lack defined

- DNA methylation patterns. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8627–8631. doi: 10.1073/pnas.1306723110
- Randolph, T. L., Peairs, F., Weiland, A., Rudolph, J. B., and Puterka, G. J. (2009). Plant responses to seven Russian wheat aphid (Hemiptera: Aphididae) biotypes found in the United States. *J. Econ. Entomol.* 102, 1954–1959. doi: 10.1603/029.102.0528
- Rasmussen, E. M. K., Vågbo, C. B., Münch, D., Krokan, H. E., Klungland, A., Amdam, G. V., et al. (2016). DNA base modifications in honey bee and fruit fly genomes suggest an active demethylation machinery with species- and tissue-specific turnover rates. *Biochem. Biophys. Rep.* 6, 9–15. doi: 10.1016/j.bbrep.2016.02.011
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinform.* 26, 139–140. doi: 10.1093/bioinformatics/btp616
- Rozen, S., and Skaletsky, H. (2000). “Primer3 on the WWW for general users and for biologist programmers,” in *Bioinformatics Methods and Protocols*, edited by S. Misener and S. A. Krawetz (Totowa, NJ: Humana Press), 365–386. doi: 10.1385/1-59259-192-2:365
- Rylatt, D. B. and Parish, C. R. (1982). Protein determination on an automatic spectrophotometer. *Anal. Biochem.* 121, 213–214. doi: 10.1016/0003-2697(82)90578-4
- Schaefer, M., and Lyko, F. (2007). DNA methylation with a sting: an active DNA methylation system in the honeybee. *BioEssays* 29, 208–211. doi: 10.1002/bies.20548
- Shakesby, A. J., Wallace, I. S., Isaacs, H. V., Pritchard, J., Roberts, D. M., and Douglas, A. E. (2009). A water-specific aquaporin involved in aphid osmoregulation. *Insect Biochem. Mol. Biol.* 39, 1–10. doi: 10.1016/j.ibmb.2008.08.008
- Shen, L., Song, C. X., He, C., and Zhang, Y. (2014). Mechanism and function of oxidative reversal of DNA and RNA methylation. *Annu. Rev. Biochem.* 83, 585–614. doi: 10.1146/annurev-biochem-060713-035513
- Shufan, K. A., and Payton, T. L. (2009). Limited genetic variation within and between Russian wheat aphid (Hemiptera: Aphididae) biotypes in the United States. *J. Econ. Entomol.* 102, 440–445. doi: 10.1603/029.102.0157
- Shukla, S., Kavak, E., Gregory, M., Imashimizu, M., Shutinoski, B., Kashlev, M., et al. (2011). CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 479, 74–79. doi: 10.1038/nature10442
- Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75
- Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351
- Sinha, D. K., and Smith, C. M. (2014). Selection of reference genes for expression analysis in *Diuraphis noxia* (Hemiptera: Aphididae) fed on resistant and susceptible wheat plants. *Sci. Rep.* 4:5059. doi: 10.1038/srep05059
- Smith, C. M., and Clement, S. L. (2012). Molecular bases of plant resistance to arthropods. *Annu. Rev. Entomol.* 57, 309–328. doi: 10.1146/annurev-ento-120710-100642
- Stanke, M., Schöffmann, O., Morgenstern, B., and Waack, S. (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinform.* 7:62. doi: 10.1186/1471-2105-7-62
- Sunita, S., Tkaczuk, K. L., Purta, E., Kasprzak, J. M., Douthwaite, S., Bujnicki, J. M., et al. (2008). Crystal structure of the *Escherichia coli* 23S rRNA:m5C methyltransferase RlmI (YccW) reveals evolutionary links between RNA modification enzymes. *J. Mol. Biol.* 383, 652–666. doi: 10.1016/j.jmb.2008.08.062
- Swanevelder, Z. H., SurrIDGE, A. K. J., Venter, E., and A.-Botha, M. (2010). Limited endosymbiont variation in *Diuraphis noxia* (Hemiptera: Aphididae) biotypes from the United States and South Africa. *J. Econ. Entomol.* 103, 887–897. doi: 10.1603/ec09257
- Tagu, D., Klingler, J. P., Moya, A., and Simon, J.-C. (2008). Early progress in aphid genomics and consequences for plant-aphid interactions studies. *Mol. Plant Microbe Interact.* 21, 701–708. doi: 10.1094/MPMI-21-6-0701
- Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935. doi: 10.1126/science.1170116
- Tolmay, V. L., Lindeque, R. C., and Prinsloo, G. J. (2007). Preliminary evidence of a resistance-breaking biotype of the Russian wheat aphid, *Diuraphis noxia* (Homoptera: Aphididae), in South Africa. *Afr. Entomol.* 15, 228–230. doi: 10.4001/1021-3589-15.1.228
- Tolmay, V. L., van Lill, D., and Smith, M. F. (1997). The influence of Demeton-S-Methyl/Parathion and Imidacloprid on the yield and quality of Russian wheat aphid resistant and susceptible wheat cultivars. *S. Afr. J. Plant Soil* 14, 107–111. doi: 10.1080/02571862.1997.10635091
- Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Prot. Bioinform.* 43, 11–10. doi: 10.1002/0471250953.bi1110s43
- Wald, A. (1943). Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Trans. Am. Math. Soc.* 54, 426–482. doi: 10.1090/S0002-9947-1943-0012401-3
- Walsh, T. K., Brisson, J. A., Robertson, K., Gordon, K., Jaubert-Possamai, S., Tagu, D., et al. (2010). A Functional DNA methylation system in the pea aphid, *Acyrtosiphum pisum*. *Insect Mol. Biol.* 19, 215–228. doi: 10.1111/j.1365-2583.2009.00974.x
- Walters, M. C., Penn, F., du Toit, F., Botha, T. C., and Aalbersberg, K. (1980). The Russian wheat aphid. Farming in South Africa, Leaflet series, wheat, G3, 1–6.
- Wang, Y., Jorda, M., Jones, P. L., Maleszka, R., Ling, X., Robertson, H. M., et al. (2006). Functional CpG methylation system in a social insect. *Science* 314, 645–647. doi: 10.1126/science.1135213
- Wojciechowski, M., Rafalski, D., Kucharski, R., Miształ, K., Maleszka, J., Bochtler, M., et al. (2014). Insights into DNA hydroxymethylation in the honeybee from in-depth analyses of TET dioxygenase. *Open Biol.* 4:140110. doi: 10.1098/rsob.140110
- Wu, H., Coskun, V., Tao, J., Xie, W., Ge, W., Yoshikawa, K., et al. (2010). Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* 329, 444–448. doi: 10.1126/science.1190485
- Wu, H., Xu, T., Feng, H., Chen, L., Li, B., Yao, B., et al. (2015). Detection of differentially methylated regions from whole-genome bisulfite sequencing data without replicates. *Nucleic Acids Res.* 43:e141. doi: 10.1093/nar/gkv715
- Xiang, H., Zhu, J., Chen, Q., Dai, F., Li, X., Li, M., et al. (2010). Single base-resolution methylome of the silkworm reveals a sparse epigenomic map. *Nat. Biotechnol.* 28, 516–520. doi: 10.1038/nbt.1626
- Yan, H., Bonasio, R., Simola, D. F., Liebig, J., Berger, S. L., and Reinberg, D. (2015). DNA methylation in social insects: how epigenetics can control behavior and longevity. *Annu. Rev. Entomol.* 60, 435–452. doi: 10.1146/annurev-ento-010814-020803
- Zemach, A., McDaniel, I. E., Silva, P., and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* 328, 916–919. doi: 10.1126/science.1186366

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 du Preez, Breeds, Burger, Swiegers, Truter and Botha. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.