



## Research paper

# High-throughput analyses and Bayesian network modeling highlight novel epigenetic Adverse Outcome Pathway networks of DNA methyltransferase inhibitor mediated transgenerational effects

You Song<sup>a,\*</sup>, Jorke H. Kamstra<sup>b</sup>, Yang Cao<sup>c</sup>, Jana Asselman<sup>d</sup>, Marc Anglès d'Auriac<sup>a</sup>, Nikolai Friberg<sup>a,e,f</sup>

<sup>a</sup> Norwegian Institute for Water Research (NIVA), Gaustadalléen 21, NO-0349 Oslo, Norway

<sup>b</sup> Institute for Risk Assessment Sciences (IRAS), Utrecht University, PO Box 80177, NL-3508 TD Utrecht, The Netherlands

<sup>c</sup> Clinical Epidemiology and Biostatistics, School of Medical Sciences, Örebro University, Fakultetsgatan 1, 70182 Örebro, Sweden

<sup>d</sup> Blue Growth Research Lab, Ghent University, Bluebridge building, Ostend Science Park 1, 8400 Ostend, Belgium

<sup>e</sup> University of Copenhagen, Freshwater Biological Section, Universitetsparken 4, 3rd floor, 2100 Copenhagen, Denmark

<sup>f</sup> University of Leeds, water@leeds, School of Geography, Leeds LS2 9JT UK



## ARTICLE INFO

Editor: Dr. S Nan

## Keywords:

AOP  
Weight of evidence  
DNA methylation  
5-Azacytidine  
*Daphnia*  
Quantitative response-response relationships

## ABSTRACT

A number of epigenetic modulating chemicals are known to affect multiple generations of a population from a single ancestral exposure, thus posing transgenerational hazards. The present study aimed to establish a high-throughput (HT) analytical workflow for cost-efficient concentration-response analysis of epigenetic and phenotypic effects, and to support the development of novel Adverse Outcome Pathway (AOP) networks for DNA methyltransferase (DNMT) inhibitor-mediated transgenerational effects on aquatic organisms. The model DNMT inhibitor 5-azacytidine (5AC) and the model freshwater crustacean *Daphnia magna* were used to generate new experimental data and served as prototypes to construct AOPs for aquatic organisms. Targeted HT bioassays (DNMT ELISA, MS-HRM and qPCR) in combination with multigenerational ecotoxicity tests revealed concentration-dependent transgenerational (F0-F3) effects of 5AC on total DNMT activity, DNA promoter methylation, gene body methylation, gene transcription and reproduction. Top sensitive toxicity pathways related to 5AC exposure, such as apoptosis and DNA damage responses were identified in both F0 and F3 using Gaussian Bayesian network modeling. Two novel epigenetic AOP networks on DNMT inhibitor mediated one-generational and transgenerational effects were developed for aquatic organisms and assessed for the weight of evidence. The new HT analytical workflow and AOPs can facilitate future ecological hazard assessment of epigenetic modulating chemicals.

## 1. Introduction

Heritable epigenetic marks, such as DNA methylation and histone modifications can in many cases reflect the life-time exposure history of an organism to environmental stressors (Mirbahai and Chipman, 2014). Among the epigenetic marks, DNA methylation is considered a master regulator of gene expression in eukaryotes (Law and Jacobsen, 2010) and has been frequently used as an (eco)toxicological biomarker to indicate effects of epigenetic modulators (Vandegheuchte and Janssen, 2011; Kamstra et al., 2015; Meehan et al., 2018). While the rapid development of OMICS techniques allows measurements of genome-wide DNA methylation and gene expression profiles, the high

costs for such analyses still limit our ability to fully understand the epigenetic changes across doses/concentrations of a stressor, exposure durations and multiple generations in a population. In addition to high-content (HC) OMICS tools, targeted high-throughput (HT) bioassays are also needed, allowing the inclusion of more life stages of an organism and exposure conditions of a stressor in the analysis to yield comparative dose/concentration-response data on a temporal scale. Such data can greatly facilitate the development of predictive (eco) toxicological approaches, such as Adverse Outcome Pathways (AOPs) for more efficient chemical safety assessment (Villeneuve et al., 2019).

Adverse Outcome Pathway (AOP) was introduced as a framework to assemble, integrate, evaluate and visualize toxicological data and knowledge relevant for regulatory-relevant adverse effects (Ankley

\* Corresponding author.

E-mail address: [you.song@niva.no](mailto:you.song@niva.no) (Y. Song).

<https://doi.org/10.1016/j.jhazmat.2020.124490>

Received 10 July 2020; Received in revised form 26 October 2020; Accepted 3 November 2020

Available online 5 November 2020

0304-3894/© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Abbreviations	
18 s	18S ribosomal RNA
5AC	5-azacytidine
Acin1	Apoptotic chromatin condensation inducer in the nucleus
Actin	Actin, alpha skeletal muscle
Aifm1	Apoptosis-inducing factor 1, mitochondrial
AO	Adverse outcome
AOP	Adverse outcome pathway
Atm	Serine/threonine-protein kinase ATM
Atp2c1	Calcium-transporting ATPase type 2 C member
ATR	Ataxia telangiectasia and Rad3-related protein
Bhmt	Betaine homocysteine methyltransferase
Casp1	Caspase-1
Casp2	Caspase-2
Casp8	Caspase-8
Cdk	Cyclin-dependent kinase
Dapk	Death associated protein kinase
Dmrta1	Doublesex- and mab-3-related transcription factor A1
DNMT	DNA methyltransferase
Dnmt1	DNA (cytosine-5)-methyltransferase 1
Dnmt3a1	DNA (cytosine-5)-methyltransferase 3A1
Dnmt3a2	DNA (cytosine-5)-methyltransferase 3A2
DSB	Double-strand break
ELISA	Enzyme-linked immunosorbent assay
Galt	Galactose-1-phosphate uridylyltransferase
GBN	Gaussian Bayesian network
Gnmt	Glycine N-methyltransferase
Gsta4	Glutathione S-transferase A4
HC	High-content
Hcfc1	Host cell factor 1
HDI	Highest density interval
HT	High-throughput
KE	Key event
KER	Key event relationship
Ku70	ATP-dependent DNA helicase 2 subunit
Limch1	LIM and calponin domains-containing protein
Mat	S-adenosylmethionine synthase
Mdm2	E3 ubiquitin-protein ligase Mdm2
Met	Methoprene tolerant
Metk	S-adenosylmethionine synthase
MIE	Molecular initiating event
Ms	Methionine synthase, EC:2.1.1.13/sw
MSHRM	Methylation sensitive high resolution melt
Mthfr	Methylenetetrahydrofolate reductase
Mtrr	Methionine synthase reductase
Mycbp	c-Myc-binding protein
Nrf2	Nuclear factor erythroid 2-related factor 2
NTS	Net temperature shift
OCM	One-carbon metabolism
OECD	Organisation for Economic Co-operation and Development
P53	Cellular tumor antigen p53
qPCR	Quantitative real-time reverse transcription polymerase chain reaction
Rad50	DNA repair protein RAD50
Rev1	DNA repair protein REV1
Sahh	Adenosylhomocysteinase adohcyaseec hydrolase
SAM	S-adenosylmethionine
SeqAPASS	Sequence Alignment to Predict Across Species Susceptibility
Tbp	TATA-box-binding protein
Tet2	Methylcytosine dioxygenase TET2
Tnfaip8	Tumor necrosis factor alpha-induced protein 8 protein 3
Vtg1	Vitellogenin-1
Vtg2	Vitellogenin-2
WoE	Weight of evidence

et al., 2010). An AOP causally links the molecular initiating event (MIE) of a chemical with its biological target, a cascade of key downstream events (KEs) at increasing levels of biological organization, and an adverse outcome (AO) into a pathway or pathway network using available information from relevant (forecaster) species and prototypical stressors. The essentiality of the KEs and the strength of the causal linkages (key event relationships, KERs) can be evaluated by Weight of Evidence (WoE) considerations (Becker et al., 2015). The applicability domains of an AOP can be further expanded to other species (taxonomic applicability) based on their phylogenetic similarities, and to other chemicals (chemical applicability) based on their structure-activity properties (Fay et al., 2017). Effects of epigenetic modulators are starting to be considered in the AOP framework in recent years (Angrish et al., 2018; Willett, 2018), however, the initial AOPs proposed are human-centric. Efforts to develop epigenetic AOPs for non-human but ecologically important species may greatly facilitate mechanistic understanding of epigenetic transgenerational effects induced by different environmental stressors and assist ecological hazard assessment of epigenetic modulating chemicals.

Among the epigenetic modulators, DNA methyltransferase (DNMT) inhibitors are a group of cytosine analogues with a common mechanism of action of hypomethylating cytosines in DNA. DNMT inhibitors, such as azacytidine and decitabine suppress DNMT activity by covalently trapping DNMT to the DNA incorporated cytosine analogues followed by degradation and subsequent DNA hypomethylation in organisms (Gnyszka et al., 2013). On the basis of this mechanism of action, DNMT inhibitors have been developed as medicines to treat oncological diseases by re-activating silenced anti-cancer genes such as tumor protein TP53 (Gnyszka et al., 2013). These chemicals, however, have also been

reported to have negative impacts on organisms. It has been realized that incorporation of DNMT inhibitors such as azacytidine into the cellular DNA replication machinery can lead to sequestration of DNMTs through formation of covalent bond between the carbon-6 atom of the cytosine ring and cysteine thiolate of DNMTs (Santi et al., 1984; Chen et al., 1991). The formation of DNA-enzyme adducts is irreversible, as the beta-elimination of this bond via the carbon-5 atom is blocked by azacytosine (Stresemann and Lyko, 2008). This subsequently jeopardizes the normal functions of DNA and triggers DNA damage responses (Kiziltepe et al., 2007). The documented effects of DNMT inhibitors include induction of DNA strand breaks (Covey et al., 1986; Kiziltepe et al., 2007; Palii et al., 2008), activation of excessive apoptosis (Murakami et al., 1995; Kiziltepe et al., 2007; Khan et al., 2008; Ghanim et al., 2012), teratogenic effects, sex shifts and multigenerational growth retardation (Kamstra et al., 2017; Ribas et al., 2017). Although a number of studies have been conducted to understand the effects of DNMT inhibitors on human cells and terrestrial mammals, whether these chemicals are hazardous to human in a long-term perspective remains inconclusive, and relatively little is known about their ecological hazards to aquatic organisms.

The water flea *Daphnia* have been widely used as aquatic invertebrate models for environmental epigenetic research (Harris et al., 2012; Jeremias et al., 2020) and AOP development (Song et al., 2017, 2020a, 2020b) due to their high ecological relevance, rapid reproductive cycles with genetically identical clones via parthenogenesis, ease to maintain under laboratory conditions and worldwide use as a standard OECD regulatory toxicity testing species (OECD, 2012). The present study, therefore used *Daphnia magna* as a prototypical species to investigate the transgenerational effects of a model DNMT inhibitor, 5-azacytidine

(5AC), on aquatic organisms and to develop AOPs. High-throughput bioassays in combination with Gaussian Bayesian network (GBN) analysis were employed to identify key toxicity pathways associated with 5AC toxicity, and to support development and WoE assessment of epigenetic AOP networks. The general hypothesis of the study was that exposure to 5AC could induce transgenerational effects on DNA methylation, gene transcription and reproduction in *D. magna* along defined AOPs. The main objectives of the study were to: 1) understand the relationships between chemical-mediated DNMT inhibition, DNA promoter methylation, gene body methylation, gene expression and multigenerational reproductive quality in *D. magna*; 2) develop and evaluate novel epigenetic AOP networks for DNMT inhibitor-mediated transgenerational effects on aquatic organisms.

## 2. Materials and methods

### 2.1. *Daphnia* culture

*Daphnia magna* DHI strain (DHI Water & Environment, Hørsholm, Denmark) were cultured in 1 L glass beakers (20 daphnids/800 mL M7 medium) under favorable conditions (temperature  $20 \pm 1$  °C, light:dark 16 h:8 h, pH  $7.8 \pm 0.2$ , dissolved oxygen  $>7$  mg/L, density 1 daphnid per 40 mL medium), according to the principles of the OECD Test Guideline 211 (OECD, 2012). The culture media were renewed every two days and offspring were removed to avoid potential crowding effect. The daphnids were daily fed with concentrated green algae *Raphidocelis subcapitata*, corresponding to 0.1 mg carbon/daphnid/day (OECD, 2012).

### 2.2. Test chemical

The model DNA methyltransferase (DNMT) inhibitor 5-azacytidine (5AC, CAS 320-67-2, purity  $\geq 98\%$ ) was purchased from Sigma-Aldrich (St. Louis, USA) and stored at  $-20$  °C. The test solutions of 5AC were freshly prepared shortly before the exposure/medium renewal due to the short half-life of this chemical. A 0.05 mol/L (M) stock solution was prepared first by dissolving 5AC in M7 medium. The final test solutions were made by diluting different volumes of the stock solution in M7 medium. The nominal exposure concentrations of 5AC were 5, 10, 20, 40 and 80  $\mu$ M. This concentration range was chosen, as they were not expected to cause mortality in adult *D. magna* during 7 days exposure, and as low as 10 mg/L (approx. 40  $\mu$ M) was reported to significantly reduce the fecundity in *D. magna* (Lindeman et al., 2019). In addition, a control containing pure medium was also included in the tests.

### 2.3. Exposure and sampling

The laboratory conditions for exposure studies were identical to that used for culturing, except that small (100 mL) glass beakers were used. A multigenerational test was set up to investigate the transgenerational effects of 5AC on different endpoints (Appendix, Fig. A1). Age synchronized (14–15d) adult female *D. magna* was used for exposure studies. Two daphnids were placed in the same test beaker and exposed in 80 mL of test solutions for 7 days (Appendix, Fig. A1), according to a previously described short-term screening (STS) protocol (Abe et al., 2015). A total of 20 beakers per treatment group ( $n = 20$ ) were included at the start of the test (total number of beakers for 6 treatment groups: 120). Each beaker was considered an independent biological replicate and all beakers were randomized every two days. After 7 days, *D. magna* from 14 test beakers were immediately sampled. One individual was snap-frozen in liquid nitrogen and stored at  $-80$  °C until DNA analysis (1 daphnid per replicate,  $n = 6$ ). The other one was sampled in 100  $\mu$ L of RNAlater (Qiagen, Hilden, Germany), stored at 4 °C overnight and further stored at  $-80$  °C until RNA analysis (1 daphnid per replicate,  $n = 6$ ). Eight beakers were pooled into 4 replicates to have sufficient materials, snap-frozen in liquid nitrogen and stored at  $-80$  °C until total DNMT activity analysis (4 daphnids per replicate,  $n = 4$ ). Animals in the remaining six replicates were briefly

rinsed and transferred to new beakers containing clean media (2 daphnids/80 mL medium/beaker) for a 7-day recovery test (Appendix, Fig. A1). All replicates were sampled for DNA methylation (1 daphnid/replicate) and transcriptional (1 daphnid/replicate) analyses immediately after the recovery period. F1 neonates ( $<24$  h old) were collected from the second brood of F0 (i.e., from the 6 replicates immediately sampled after exposure) during the exposure, whereas F2 and F3 offspring were collected from the last broods before termination of the F1 and F2 test, respectively (Appendix, Fig. A1). From F1, each generation ( $n = 6$ , 1 daphnid/40 mL medium/beaker) was tested for 21 days (Appendix, Fig. A1), according to the principles of the OECD *D. magna* Reproduction Test (OECD, 2012). Immediately after the F3 test, *D. magna* were sampled for DNA methylation (1 daphnid/replicate) analysis, but not for transcriptional analysis due to a lack of sufficient material as a result of reduced fecundity after exposure to 5AC. Cumulative fecundity (total number of viable offspring) was recorded for all generations. The test solutions were renewed every two days and *D. magna* were fed daily with concentrated green algae *R. subcapitata*, corresponding to 0.1 mg carbon/daphnid/day (OECD, 2012). All test beakers were randomized to avoid potential bias.

### 2.4. HT-MSHRM

A high-throughput (384-well microplate format) methylation sensitive high-resolution melt (HT-MSHRM) assay was downscaled for low input materials (i.e., DNA from single *D. magna*) and used to determine locus-specific promoter and gene body DNA methylation of 16 genes in F0 after exposure and recovery, and in F3 *D. magna*. The test genes were well-known biomarker genes or key regulators involved in major biological pathways, including one carbon metabolism (*Dnmt3a2*, *Gnmt*, *Tet2*, see Abbreviations for gene full names), apoptotic signaling (*Aifm1*, *Casp2*, *Dapk*, *Mycbp*), DNA damage responses (*Atm*), cell cycle regulation (*Cdk*, *Hcfc1*), oxidative stress responses (*Nrf2*), cell migration (*Limch1*), juvenile hormone signaling (*Met*, *Dmrta4*), calcium signaling (*Atp2c1*) and small sugar metabolism (*Galt*). The gene body sequences were obtained from the WFleaBase (<http://wfleabase.org/>) using the available *D. magna* genome and associated gene models (Orsini et al., 2016), whereas the promoter sequences were considered as 1000 base pairs (bp) upstream of transcription start site (TSS), as previously described (Lindeman et al., 2019). The promoters and the genes with genomic locations and associated wfleabase IDs are listed in Appendix (Tables A1 & A2). Primers for promoter (Appendix, Table A1) and gene body (Appendix, Table A2) methylation analyses were designed using MethPrimer v1.0 (<https://www.urogene.org/cgi-bin/methprimer/methprimer.cgi>) (Li and Dahiya, 2002). Primers were designed for amplification of both methylated and unmethylated DNA. An Agilent Bioanalyzer and High Sensitivity DNA Kit (Agilent technologies, Santa Clara, USA) was used to assess the specificity of PCR products and primer dimers. Unqualified primers were excluded and replaced with new pairs. All PCR products were specific and of the expected size (data not shown).

Genomic DNA (gDNA) from single *D. magna* was isolated using a Quick-DNA™ Tissue/Insect Microprep Kit (Zymo Research, Irvine, USA). The DNA yield ( $>100$  ng) and purity ( $260/280 > 1.8$ ) were determined using a Nanodrop® ND-1000 spectrophotometer (Nanodrop Technologies, Wilmington, USA). The purified gDNA was diluted and bisulfite converted using an EZ DNA Methylation-Lightning™ Kit (Zymo). The conversion efficiency was higher than 99.5% and the DNA recovery rate was higher than 80%, according to the manufacturer's protocol. The human HCT116 DKO methylated (100%) DNA (Zymo) was used as a quality control for bisulfite conversion. An unmethylated (0%) DNA standard was made by whole genome amplification (WGA) of pooled *D. magna* gDNA using a REPLI-g Mini/Midi Kit (Qiagen), following the producer's instructions.

The HT-MSHRM assay ( $n = 6$ ) was performed using the Bio-Rad CFX384 platform (Bio-Rad Laboratories, Hercules, USA). A Biomek 3000 Laboratory Automation Workstation (PerkinElmer, Waltham,

USA) was employed for high-throughput liquid handling. The PCR amplification reaction (10  $\mu$ L) contained 2.5  $\mu$ L (0.4 ng/ $\mu$ L) of bisulfite converted DNA samples or unmethylated DNA standard, 5  $\mu$ L of ZymoTaq™ qPCR PreMix (Zymo) and 2.5  $\mu$ L of 200 nM forward and reverse primer mixture. The PCR cycles were set to: 1) hot start at 95 °C for 15 min; 2) 45 cycles of: denature at 94 °C for 0.5 min, annealing at the primer-specific optimal temperature ( $T_a$ ) for 0.5 min, extension at 63 °C for 1 min; 3) final extension at 63 °C for 10 min. A high-resolution melt analysis was performed between 63 and 95 °C with a temperature increment of 0.1 °C.

The raw HT-MSHRM data was analyzed using the CFX Manager v3.0 and Precision Melt Analysis software v1.2 (Bio-Rad). Net Temperature Shift (NTS) was calculated by comparing the melt curve of a target sample with that of a non-methylated standard, based on the principles of a previous protocol (Kamstra et al., 2014). The NTS in the exposed groups were further normalized to the control to calculate fold changes prior to statistical analysis.

## 2.5. HT-qPCR

A high-throughput (384-well microplate format) quantitative real-time reverse-transcription polymerase chain reaction (HT-qPCR) was employed to measure transcriptional changes of 34 selected genes (Appendix, Table A3). The test genes are well-known biomarker genes or key regulators involved in major biological pathways, including one carbon metabolism (*Dnmt1*, *Dnmt3a1*, *Dnmt3a2*, *Gnmt*, *Mat*, *Ms*, *Mthfr*, *Mtrr*, *Sahh*, *Tet2*), apoptotic signaling (*Acin1*, *Aifm1*, *Casp1*, *Casp2*, *Casp8*, *Dapk*, *Mycbp*, *Tnfrsf8*), DNA damage responses (*Atm*, *Ku70*, *Mdm2*, *P53*, *Rad50*, *Rev1*), cell cycle regulation (*Cdk*, *Hcfc1*), oxidative stress responses (*Nrf2*, *Gsta4*), cell migration (*Limch1*), juvenile hormone signaling (*Met*, *Dmrt4*), embryo development (*Vtg1*), calcium signaling (*Atp2c1*) and small sugar metabolism (*Galt*). Primers (Appendix, Table A3) were designed using Primer3 v4.0.0 (<http://primer3.ut.ee/>) and ordered from Thermo Fisher Scientific (Waltham, USA). The optimal annealing temperatures of the primers were obtained prior to sample analysis (Appendix, Table A3).

Total RNA was isolated using a ZR Tissue & Insect RNA MicroPrep™ kit (Zymo), as previously described (Song et al., 2016). The yield (>500 ng) and purity (260/280 > 1.8) of the RNA samples were measured using a Nanodrop® ND-1000 spectrophotometer (Nanodrop). The integrity of the RNA was checked using an Agilent Bioanalyzer and RNA 6000 Nano chips (Song et al., 2016). Qualified samples with clear RNA peaks and flat baseline were stored at -80 °C until HT-qPCR analysis.

The HT-qPCR assay (n = 6) was performed using the Bio-Rad CFX384 platform (Bio-Rad), as previously described (Song et al., 2016). A Biomek 3000 Laboratory Automation Workstation (PerkinElmer) was employed for high-throughput liquid handling. In brief, 600 ng total RNA was reversely transcribed into cDNA using qScript™ cDNA SuperMix (Quanta BioSciences, Gaithersburg, USA). The qPCR reaction (10  $\mu$ L) consisted of 2.5  $\mu$ L diluted cDNA template (1 ng/ $\mu$ L, total mass 2.5 ng), 5  $\mu$ L of PerfeCTa® SYBR® Green FastMix® (Quanta BioSciences) and 2.5  $\mu$ L of forward and reverse primer mixture (400 nM). A series of diluted cDNA (12.5, 6.25, 31.13, 1.56, 0.78, 0.39 ng) were made from pooled *D. magna* RNA and were included to generate standard curves for determination of amplification efficiencies (E) and correlation coefficient ( $R^2$ ). A no-reverse-transcriptase control (NRT) and a non-template control (NTC) were included as additional quality controls. The thermo cycles for qPCR amplification were set to: 1) 95 °C for 3 min; 2) 40 cycles of 95 °C for 20 s, primer-specific  $T_a$  for 20 s, 72 °C for 30 s; 3) a melt curve analysis for 30 s between 65 and 90 °C with an increment of 0.5 °C. Tests with amplification efficiency between 90% and 105% and correlation coefficient ( $R^2$ ) between 0.98 and 1 were considered valid. The relative expression of the target gene was calculated based on a combination of efficiency and the threshold cycle (Cq) value according to the Pfaffl Method (Pfaffl, 2001), and normalized to the geometric mean expression of the reference genes Actin alpha skeletal muscle (*Actin*), 18 S ribosomal RNA (18 s) and

TATA-box-binding protein (*Tbp*), based on the  $\Delta\Delta Cq$  method (Vandesompele et al., 2002).

## 2.6. DNMT ELISA

A high-throughput (96-well microplate format) enzyme-linked immunosorbent assay (ELISA) was employed to determine the total activity of DNMTs in F0 *D. magna* after the exposure. Nuclei were extracted from pooled *D. magna* (4 individuals) using an EpiQuik™ Nuclear Extraction Kit I (Epigentek, New York, USA). The total nuclear protein yield was determined using a Coomassie Plus (Bradford) Assay Kit (Thermo-Fisher Scientific, Waltham, USA), based on the Bradford method (Bradford, 1976). The total DNMT activity HT-ELISA (n = 4) was performed using a fluorometric EpiQuik™ DNMT Activity/Inhibition Assay Ultra Kit (Epigentek). All assay procedures were performed according to the manufacturer's protocols. The absolute DNMT activities (RFU/h/mg nuclear protein) were calculated and further normalized to the control to calculate fold changes.

## 2.7. Basic statistical analyses

A ROUT test (Motulsky and Brown, 2006) was performed first to eliminate outliers in the raw data prior to normality and equal variance assessment. One-way analysis of variance (ANOVA, normal distribution and equal variance) followed with Tukey post-hoc test, or Kruskal-Wallis non-parametric test (no normality or equal variance) followed with Dunn's post-hoc test was used to determine statistical differences ( $p < 0.05$ ) between treatment groups. The basic statistical analyses were performed in Graphpad Prism v8 (Graphpad Software Inc., San Diego, USA).

## 2.8. Gaussian Bayesian network analysis

The quantitative relationships between key events in a toxicity pathway were evaluated using Gaussian Bayesian network (GBN). The GBN approach offers algorithms for prediction and inference where all variables are continuous and defined by a Gaussian prior distribution or a Gaussian conditional distribution. The linear additive relationship was assumed between the KEs, which was supported by the scatter plots and has good mathematical properties such as tractability and the availability of closed-form results for the inference procedures (Buja et al., 1989). The average and variation of the GBN coefficients were estimated using the resampling method. First, different data points of F0 DNMT activity, F0 promoter methylation, F0 gene expression, and F0 fecundity values within each of the 6 treatment groups were randomly sampled with replacement. A total of 180 observations were obtained to build a GBN and estimate the corresponding linear coefficients for the arcs of the GBN. Second, the procedures were repeated 10,000 times to construct 10,000 GBNs which were used to estimate the coefficients of the GBNs. Third, mean, median, and 95% highest density interval (HDI) (Turkkan and Pham-Gia, 1993) were calculated based on the 10,000 estimates. The 95% HDIs that did not include 0 were considered statistically significant ( $p < 0.05$ ) (Appendix, Table A4). Multiplicative coefficient was calculated for each hypothetical toxicity pathway to indicate the overall sensitivity of the pathway subject to changes in 5AC concentrations (Appendix, Table A5). The toxicity pathways were ranked by their absolute multiplicative coefficients. The higher absolute value indicates higher sensitivity of the pathway. The parameters of the second toxicity pathway network for transgenerational effects (i.e. F0-F3) were also estimated using the same method (Appendix, Table A6 & A7). All the analyses were performed in R software 4.01 (R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>).

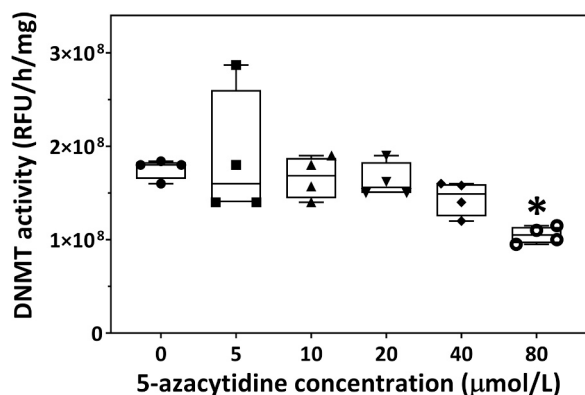


Fig. 1. Total DNA methyltransferase (DNMT) activity ( $n = 4$ ) in F0 *Daphnia magna* after 7 days exposure to 5-azacytidine. \* denotes significant difference from the corresponding control.

## 2.9. AOP assembly and weight of evidence assessment

Conceptual AOPs were assembled following OECD's handbook for AOP development and assessment (OECD, 2018). The WoE of the AOPs was assessed based on the Bradford Hill Considerations (Becker et al., 2015). The confidence levels of the KEs and KERs were scored as "High", "Moderate" or "Low", according to OECD's AOP handbook (OECD, 2018). To expand the taxonomic applicability domain of the AOPs, the Sequence Alignment to Predict Across Species Susceptibility (SeqAPASS) tool (LaLone et al., 2016) was employed to identify susceptible species groups based on protein sequence similarities. Two *D. magna* proteins, *Dnmt3* (GenBank acc. KZS08978.1) and *Casp1* (GenBank acc. KZS14075.1) were used as query sequences for level 1 (primary amino acid sequence alignment) analysis in SeqAPASS, as these were the only *D. magna* proteins found in the database that were directly relevant for the AOPs. *Dnmt3* was used to represent the MIE of the AOPs, whereas *Casp1* was used to represent KE3 and KE9 of the AOPs. Furthermore,

based on the level 1 analysis, the conserved domains of *Dnmt3* (GenBank acc. cd11725) and *Casp1* (GenBank acc. cd00032) were identified and used for level 2 (functional domain alignment) analysis. The chemical applicability domain of the AOPs was defined by data mining in the literature and public databases to collect a list of previously reported DNMT inhibiting chemicals.

## 3. Results and discussion

### 3.1. Molecular initiating event

It has been widely accepted that DNMTs are the primary macromolecular targets of 5AC (Stresemann and Lyko, 2008). Although inhibition of DNMTs by 5AC has been well documented for vertebrates (Stresemann and Lyko, 2008), the potency of this chemical to affect invertebrate DNMTs has been largely unknown. A multigenerational decrease in global DNA cytosine methylation has been previously observed following 5AC exposure (Vandegheuchte et al., 2010; Lindeman et al., 2019), but as of yet no direct activity of DNMTs has been measured in *D. magna*. The present study showed that the total activity of DNMTs in F0 *D. magna* decreased in a concentration-dependent manner after 7 days exposure to 5AC, with a significant reduction of approx. 40% at the highest concentration (Fig. 1). This is to date the first direct evidence to support 5AC mediated DNMT inhibition as a potential MIE in aquatic invertebrates.

### 3.2. Key events

#### 3.2.1. DNA methylation

Hypomethylation of genes as a general consequence of 5AC exposure in both vertebrates (Stresemann and Lyko, 2008) and *D. magna* (Lindeman et al., 2019), was also observed for the majority of the test genes in F0 after 7 days exposure to 40 and 80  $\mu\text{M}$  5AC in the present study (Fig. 2 & see SI-1 Fig. S2 for concentration-response curves). Several genes tested in the current analysis have also been previously reported to have reduced DNA methylation in *D. magna* after chronic exposure

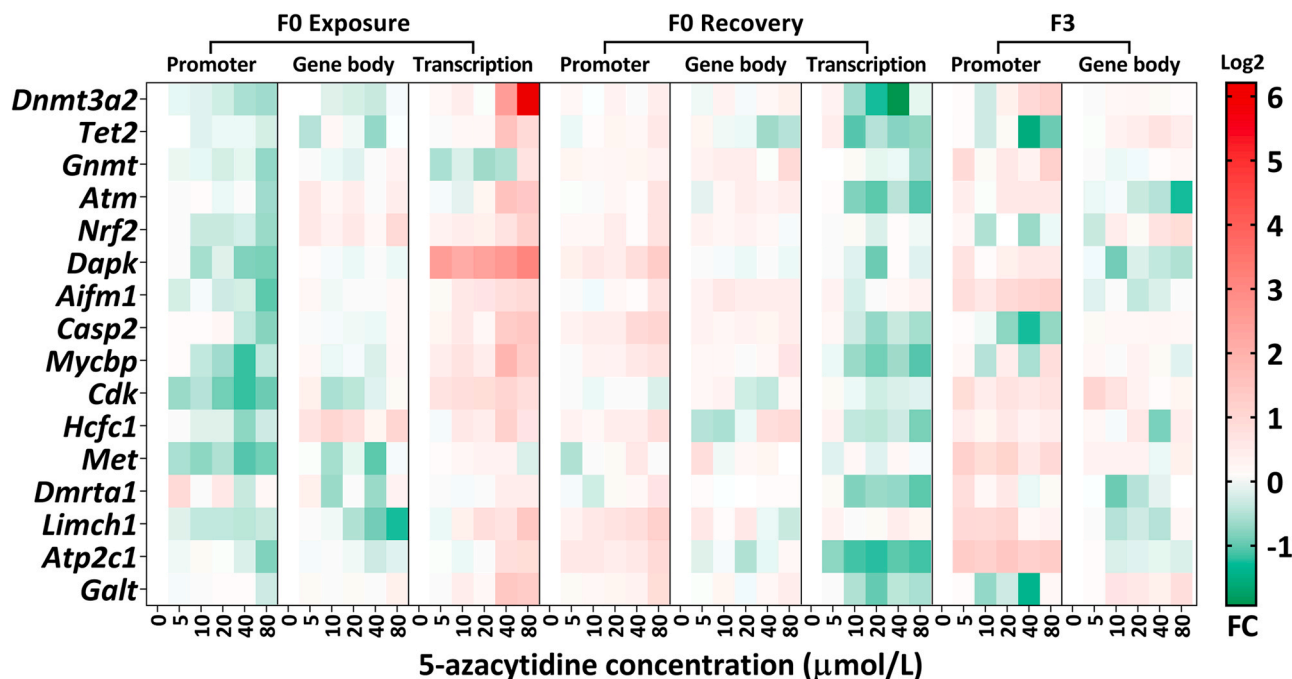


Fig. 2. A heatmap displaying DNA promoter methylation ( $n = 6$ ), gene body methylation ( $n = 6$ ) and gene transcription ( $n = 6$ ) in F0 *Daphnia magna* after 7 days exposure to 5-azacytidine, in F0 after 7 days recovery, and in F3. Each cell represents the log<sub>2</sub> transformed mean fold change compared to the control. Red: increased compared to the control (0); green: decreased compared to the control. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

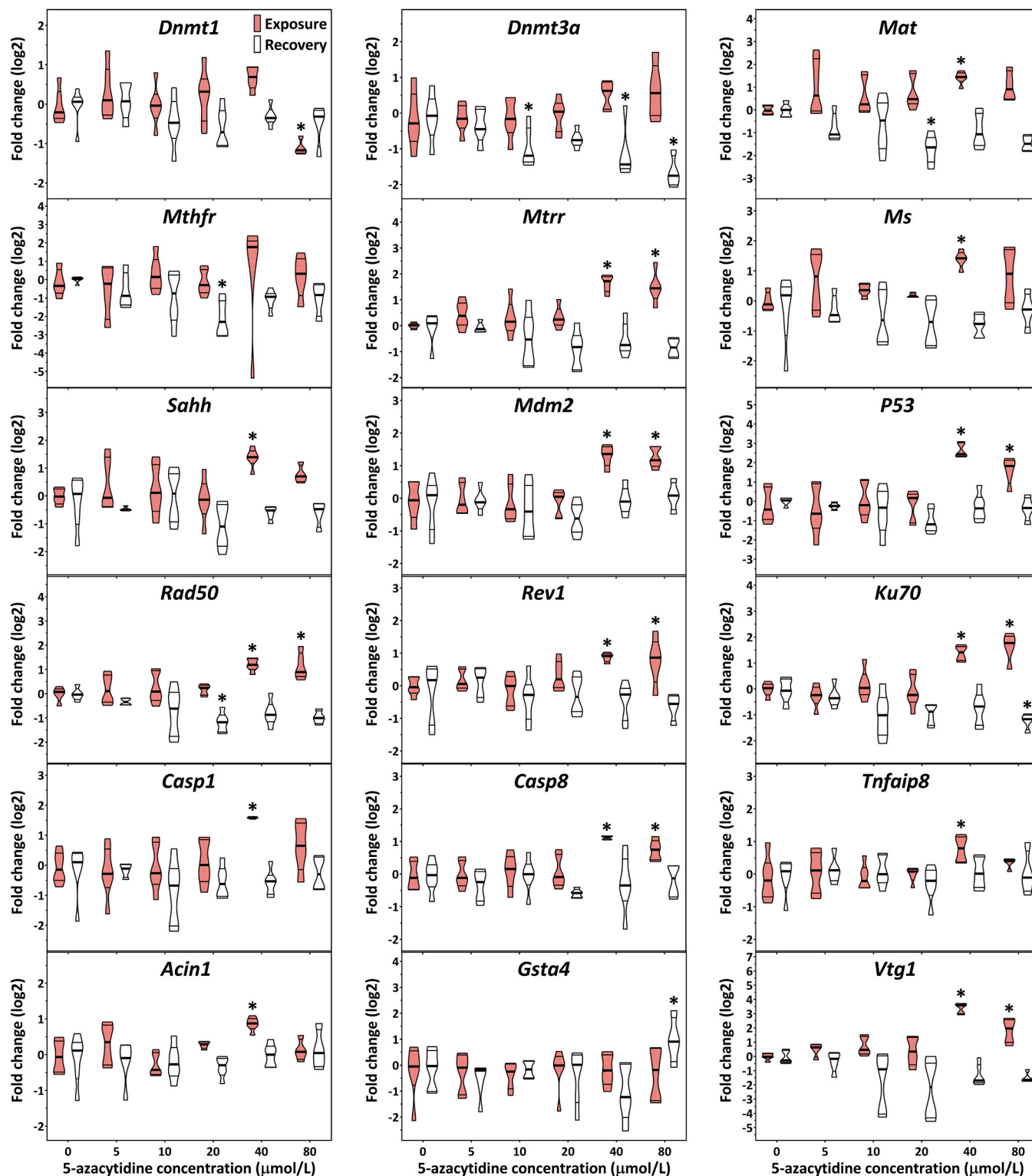


Fig. 3. Violin plots displaying transcriptional responses of biomarker genes in F0 *Daphnia magna* after 7 days exposure to 5-azacytidine (n = 6) and after 7 days recovery (n = 6). \* denotes significant difference from the control (0).

(until releasing of the third brood) to 3.7 mg/L (approx. 15 µM) of 5AC (Athanasio et al., 2018), such as the cell migration regulator *Limch1*, the cell cycle regulators *Hcfc1* and *Cdk*, and the plasma membrane calcium transporter *Atp2c1*. In contrast, no clear patterns of gene body methylation were identified (Fig. 2 and see Appendix, Fig. A2 for detailed concentration-response curves), except for the master regulator of oxidative response *Nrf2* and *Limch1* which showed significantly

increased gene body methylation at the highest concentration and a clear concentration-dependent decrease in gene body methylation, respectively (Fig. 2 and Appendix, Fig. A2).

Interestingly, after 7 days recovery, the majority of the genes displayed significantly increased promoter methylation at 80 µM (Fig. 2 and Appendix, Fig. A2), whereas only the S-adenosylmethionine (SAM) regulator in the OCM *Gnmt* and the mitochondrial initiator of apoptotic

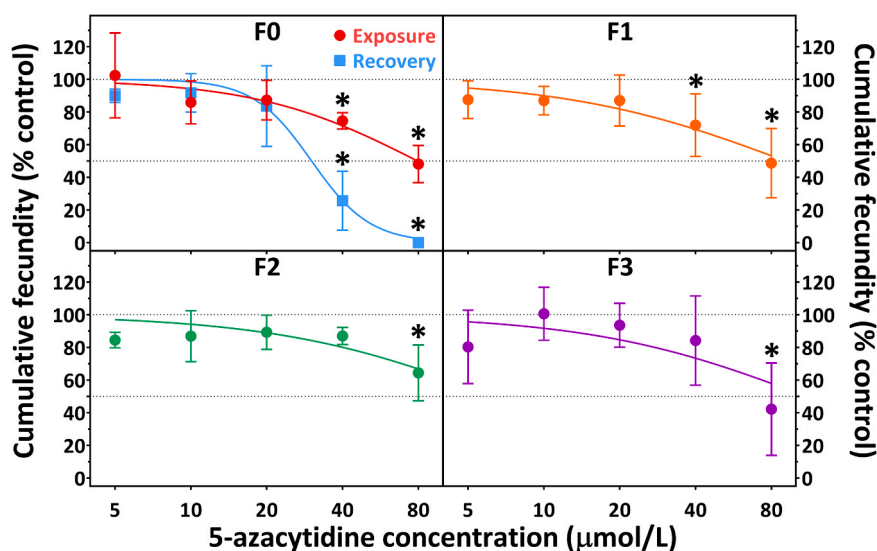


Fig. 4. Concentration-response curves of cumulative fecundity in F0 *Daphnia magna* after 7 days exposure to 5-azacytidine ( $n = 12$ ), cumulative fecundity in F0 after 7 days recovery in clean media ( $n = 6$ ), and 21 days cumulative fecundity in F1 ( $n = 6$ ), F2 ( $n = 6$ ) and F3 ( $n = 6$ ). Data are presented as mean $\pm$ SD. \* denotes significant difference from the corresponding control.

signaling pathway *Afim1* had significant increases in gene body methylation at 80 and 10  $\mu\text{M}$ , respectively (Fig. 2 and Appendix, Fig. A2). These findings clearly indicate that 5AC mediated promoter hypomethylation was reversible. It is plausible that certain plasticity mechanisms exist to reactivate the DNA methylation machinery and compensate for the excessive demethylation during exposure.

In addition, transgenerational effects of 5AC on DNA methylation were also observed in the present study. In F3, most of the genes had increased promoter methylation at 40 and/or 80  $\mu\text{M}$  (Fig. 2 and Appendix, Fig. A2), with exceptions of the DNA demethylation regulator *Tet2*, the apoptosis activator *Casp2* and the regulator of small sugar metabolism *Galt* which showed reduced promoter methylation (Fig. 2 and Appendix, Fig. A2). Significantly increased gene body methylation was found for *Tet2* (20, 40 and 80  $\mu\text{M}$ ), *Casp2* (20 and 80  $\mu\text{M}$ ) and *Cdk* (5  $\mu\text{M}$ ), whereas significantly reduced gene body methylation was identified for the DNA double-strand break (DSB) sensor *Atm* and the apoptosis regulator *Dapk*.

### 3.2.2. Transcriptional responses

It has been suggested that promoter methylation is in general associated with transcriptional suppression, whereas gene body methylation is normally correlated with increased gene expression (Moore et al., 2013). In the present study, the majority of the genes in F0 were significantly upregulated after 7 days exposure to 40 and/or 80  $\mu\text{M}$  of 5AC, whereas downregulated after 7 days recovery (Figs. 2 and 3). Apparent inverse relationships between gene expression and promoter methylation were observed, whereas no clear patterns were identified between gene expression and gene body methylation (Fig. 2). Since the promoter is highly dynamic and smaller than a gene body which often spans multiple kilobases in length, differences in gene body methylation are more difficult to assess. In some invertebrates, methylated exons have been associated with alternative splicing sites (Flores et al., 2012). Differential methylation at alternative splicing sites has been previously observed in *Daphnia* exposed to *Microcystis aeruginosa* (Asselman et al., 2017) and to salinity (Jeremias et al., 2018). Given the lack of clear patterns in both gene body methylation and its relation to gene expression, it seems that promoter methylation rather than gene body methylation is driving the transcriptional responses to 5AC. At the functional level, transcriptional responses of the 34 test genes indicate that multiple pathways, such as OCM, apoptotic signaling, DNA damage responses, cell cycle regulation, oxidative stress responses, cell migration, juvenile hormone signaling,

embryo development, calcium signaling and small sugar metabolism (*Galt*) were activated after exposure to 5AC, whereas suppressed after the recovery period (Figs. 2 and 3). Exposure to 5AC is known to affect OCM (Krushkal et al., 2016), induce ATR-mediated DNA DSB responses and cell cycle regulation (Kiziltepe et al., 2007), and trigger both caspase dependent and independent apoptotic pathways (Murakami et al., 1995; Kiziltepe et al., 2007) in mammalian cells. In *D. magna*, exposure to 10 mg/L (approx. 41  $\mu\text{M}$ ) of 5AC for 7 days was reported to alter the expression of genes involved in OCM (*Dnmt1*, *Dnmt3a2*, *Gnmt*, *Metk*, *Sahh*, *Mthfr*), juvenile hormone signaling (*Met*) and embryo development (*Vtg1*, *Vtg2*) in F0 and F1, however, with dissimilar directions of regulation (Lindeman et al., 2019). Another study showed that after chronic exposure (until releasing of the third brood) to 3.7 mg/L (approx. 15  $\mu\text{M}$ ) of 5AC, both genes and metabolites involved in OCM (*Mat*, *Ms*, *Sahh*, *Bhmt*, *Mtr*, *Gnmt*) were affected, albeit the responses were highly time-dependent (Athanasio et al., 2018).

### 3.3. Adverse outcomes

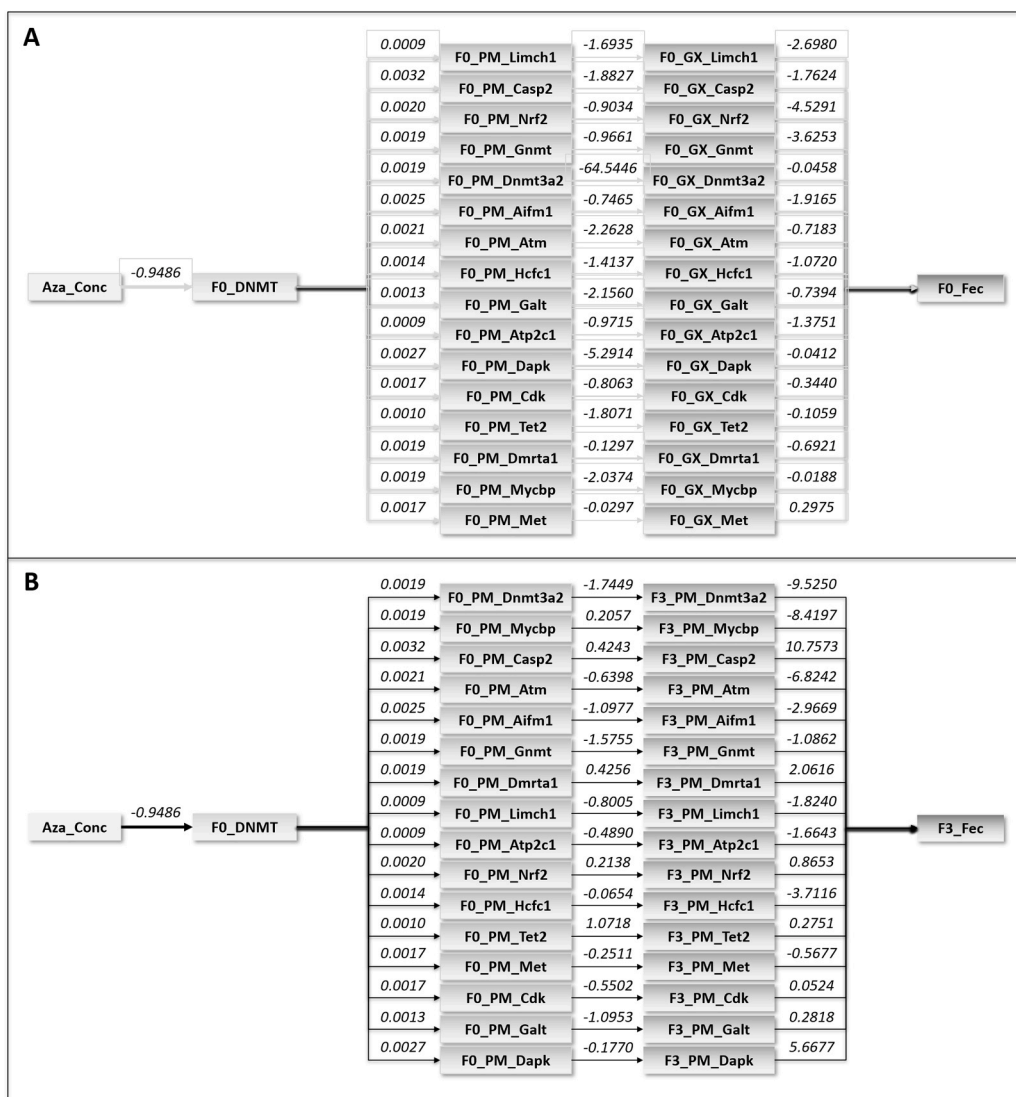
The adverse outcomes of 5AC have been documented for different species including *D. magna*. In the previous study by Vandegheuchte and coworkers, exposure to 5AC for 21 days significantly reduced reproduction in F0 *D. magna*, with a lowest-observed-effect-concentration (LOEC) of 16 mg/L (approx. 65  $\mu\text{M}$ ) corresponding to 54% reduction in cumulative fecundity (Vandegheuchte et al., 2010). A more recent study by Lindeman and colleagues also showed a significant reduction (12%) in cumulative fecundity in adult *D. magna* after 7 days exposure to 10 mg/L (approx. 41  $\mu\text{M}$ ) of 5AC (Lindeman et al., 2019). In line with the previous knowledge, the present study has also demonstrated that after 7 days exposure, the reproduction in F0 *D. magna* decreased in a concentration-dependent manner, with exposure to 40 and 80  $\mu\text{M}$  of 5AC leading to a significant reduction in cumulative fecundity (Fig. 4). Interestingly, the 7 days recovery did not mitigate the effects of 5AC, with 40 and 80  $\mu\text{M}$  of 5AC causing a more severe decline in fecundity (Fig. 4), thus indicating irreversible adverse effects of this chemical at the organismal level.

Through a multigenerational design, the present study further revealed that parental (F0) exposure to 5AC also led to reduced reproduction in the three successive (F1–3) generations in *D. magna*, with the concentration-response curves of F1–F3 in general resembling that obtained for F0 immediately after exposure (Fig. 4). As F3 is considered the

first truly unexposed progeny in this exposure setup, the results clearly indicate transgenerational effects of 5AC on *D. magna* reproduction. Such transgenerational reproductive effects have not been documented for *D. magna* or other aquatic invertebrates, albeit a multigenerational study reported that parental exposure to 2.9 mg/L (approx. 12 µM) of 5AC for 16 days led to significantly reduced reproduction in F1 *D. magna*, whereas a non-significant decrease in reproduction was observed for F2 (Vandegheuchte et al., 2010). As no truly unexposed generation was investigated, this study could not confirm transgenerational effects of the chemical. Although not well studied in crustaceans, the transgenerational effects of 5AC have been reported for other aquatic organisms. For example, exposure of parental zebrafish (*Danio rerio*) to 10 µM of 5AC during the early embryonic period (0–6 day-post-fertilization, dpf) led to a significant reduction in body length in unexposed F2 and a shifted sex ratio to males in F1 (Kamstra et al., 2017). In another study where zebrafish were exposed to 25 and 75 µM of 5-aza-2'-deoxycytidine from 0 to 6 dpf, a shift towards females was found in F0 (Ribas et al., 2017). In addition, ovaria were analyzed for RNA expression with a number of differentially expressed genes related to reproduction (Ribas et al., 2017). However, no direct effect on fecundity was reported in this study (Ribas et al., 2017).

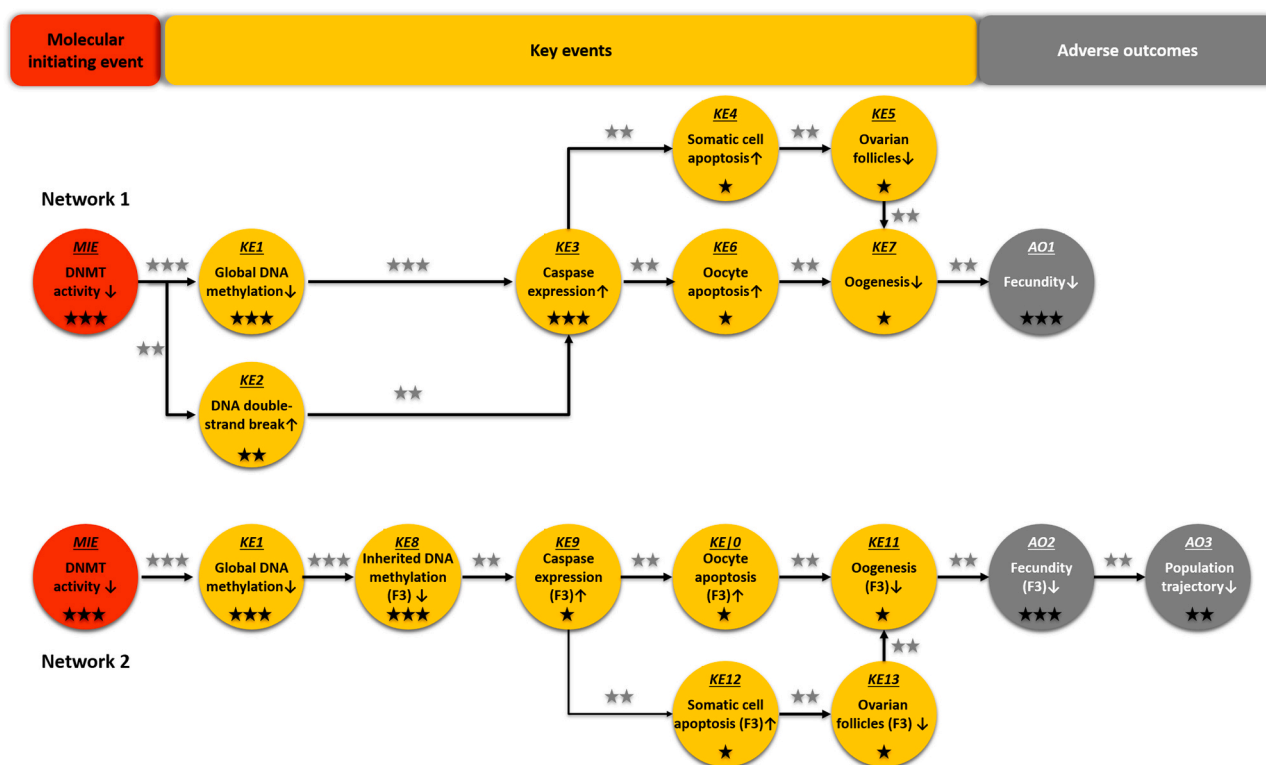
### 3.4. Toxicity pathway characterization

To identify the most sensitive (influential) pathways contributing to the 5 AC toxicity and use these for AOP development, quantitative relationships between total DNMT activity, promoter methylation, gene expression and reproduction were established for the 16 genes (which have complete data) using the GBN approach (Appendix, Tables A4 & A6), and assembled into two hypothetical networks covering both F0 (Fig. 5A) and F0-F3 transmission pathways (Fig. 5B). By calculating the multiplicative coefficient for each hypothetical pathway (Appendix, Tables A6 & A7), the overall relevance of the pathways subject to the 5AC concentrations were ranked (Fig. 5). The results suggest that in the F0 network (Fig. 5A), the most sensitive (top 5) toxicity pathways were associated with cell migration (*Limch1*), apoptosis (*Casp2*), oxidative stress responses (*Nrf2*) and OCM (*Gnmt*, *Dnmt3a2*). In the F0-F3 network (Fig. 5B), the top 5 pathways were related to OCM (*Dnmt3a2*), apoptosis (*Mycbp*, *Casp2*, *Aifm1*) and DNA damage responses (*Atm*). On the basis of the current findings and existing knowledge on the mechanisms of action of 5AC, it is likely that besides direct effects on OCM, 5AC mediated apoptosis and DNA damage were also major contributors to the observed reproductive effects. Interestingly, besides *Dnmt3a2*, the *Casp2* pathway was scored highly relevant for both F0 and F0-F3 networks, potentially highlighting a key role of caspase-mediated apoptosis in 5AC mediated reproductive toxicity. In fact, previous studies have suggested causal



**Fig. 5.** Bayesian network analysis of the F0 toxicity pathways (A) and F0 to F3 transmission toxicity pathways (B) of 5-azacytidine in *Daphnia magna*. Aza\_Conc: 5-azacytidine concentration; F0\_DNMT: total enzymatic activity of DNMTs in F0; F0\_PM: promoter methylation in F0; F0\_GX: gene expression in F0; F3\_PM: promoter methylation in F3; F0\_Fec: fecundity in F0; F3\_Fec: fecundity in F3. The numbers indicate pathway coefficients. The pathways are ranked by their absolute multiplicative coefficients from the top (highest value) to the bottom (lowest value).





**Fig. 6.** Adverse Outcome Pathway (AOP) networks linking DNA methyltransferase (DNMT) inhibition to population decline. Network 1: One-generational AOP network for effects on F0. Network 2: Transgenerational AOP network for transmission of effects from F0 to F3. ↑: Increase; ↓: Decrease. Stars indicate the weight of evidence of key events (KE, black) and key event relationships (KERs, grey): 3 stars = high, 2 stars = moderate, 1 star = low.

relationships between somatic cell apoptosis, ovarian follicle integrity, oocyte apoptosis, oogenesis and fecundity in mammals (Hussein, 2005; Tiwari et al., 2015) and invertebrates (Song et al., 2020a). It is therefore highly plausible that the observed reproductive decline in the present study was a direct consequence of impaired oogenesis, possibly due to oocyte apoptosis and/or somatic cell apoptosis mediated ovarian follicle breakdown. On the basis of this, the observed epigenetic effects occurring at the molecular level in the present study could be connected to the reported tissue/organ effects documented for *D. magna* via the common event of apoptosis. It should also be noted that the cell migration and oxidative stress responses pathways were also identified to be highly relevant for 5AC toxicity by the data-driven approach. These pathways may warrant further investigations.

### 3.5. AOP development and evaluation

#### 3.5.1. AOP assembly

On the basis of relevant toxicity pathways identified in the present study and existing AOPs related to DNA damage, apoptosis and reproductive toxicity in the AOP repository database AOPWiki (<https://aopwiki.org/>, AOP#216), two new AOP networks on inhibition of DNMT activity leading to reduced population trajectory (Fig. 6) were assembled and submitted to AOPWiki (AOP#336–341). In network 1 (Fig. 6), both DNMT inhibition mediated global hypomethylation and DNA DSB can trigger caspase-dependent apoptosis, which affects oogenesis directly by inducing oocyte apoptosis, or indirectly via somatic cell death and subsequent destruction of ovarian follicles. Impaired oogenesis can lead to reduced fecundity (Fig. 6). “Caspase expression” is considered an important KE, as caspases are key regulators of apoptosis (McIlwain et al., 2013) and the caspase-dependent apoptotic pathway has been associated with 5AC toxicity (Kiziltepe et al., 2007). Several caspase genes were upregulated in F0 after exposure, and the *Casp2* promoter methylation was consistently suppressed in F0 and F3 in the

present study. In addition, the *Casp2* pathway was ranked as one of the top influential toxicity pathways in both F0 and F3. The KE term is generalized to be applicable to more species with different caspases. In network 2, DNMT inhibition mediated global hypomethylation in F0 can be transmitted across generations to the unexposed (F3) progeny. Through a similar caspase-dependent apoptotic pathway to impair oogenesis, fecundity in F0 (AO1) and fecundity in F3 (AO2) are considered “High” (Fig. 6 and Appendix, Table A8), as these KEs are supported by direct evidence associated with DNMT inhibitor exposures by both present and previous studies. In contrast, the essentialities of increased “DNA DSB” (KE2) and decreased “population trajectory” (AO3) are considered “Moderate” (Fig. 6 and Appendix, Table A8), as no direct supporting evidence was obtained from the present study, whereas some indirect evidence exists in the literature (Kiziltepe et al., 2007). The rest, especially the KEs associated with 5AC mediated effects at higher levels of biological organization (i.e., tissue/organ) and in F3, are scored as “Low” (Fig. 6 and Appendix, Table A8), as these are hypothetical in the present study and are still lacking direct supporting evidence from the previous studies. Nevertheless, these KEs are highlighted as knowledge gaps and warrant further investigations. Among the 18 KERs, “reduced DNMT activity leading to reduced global DNA methylation”, “reduced global DNA methylation leading to increased caspase expression”, and “reduced global DNA methylation leading to inherited DNA methylation (F3)” were well supported by the current data and existing studies thus scored as “High”, whereas the rest with

#### 3.5.2. WoE assessment of KEs and KERs

The two AOP networks consist of a total of 17 KEs (MIE and AO as special KEs) and 18 KERs (Fig. 6). Among the KEs, the essentialities of reduced “DNMT activity” (MIE), reduced “global DNA methylation” (KE1), increased “caspase expression”, “inherited DNA methylation in F3” (KE8), fecundity in F0 (AO1) and fecundity in F3 (AO2) are considered “High” (Fig. 6 and Appendix, Table A8), as these KEs are supported by direct evidence associated with DNMT inhibitor exposures by both present and previous studies. In contrast, the essentialities of increased “DNA DSB” (KE2) and decreased “population trajectory” (AO3) are considered “Moderate” (Fig. 6 and Appendix, Table A8), as no direct supporting evidence was obtained from the present study, whereas some indirect evidence exists in the literature (Kiziltepe et al., 2007). The rest, especially the KEs associated with 5AC mediated effects at higher levels of biological organization (i.e., tissue/organ) and in F3, are scored as “Low” (Fig. 6 and Appendix, Table A8), as these are hypothetical in the present study and are still lacking direct supporting evidence from the previous studies. Nevertheless, these KEs are highlighted as knowledge gaps and warrant further investigations. Among the 18 KERs, “reduced DNMT activity leading to reduced global DNA methylation”, “reduced global DNA methylation leading to increased caspase expression”, and “reduced global DNA methylation leading to inherited DNA methylation (F3)” were well supported by the current data and existing studies thus scored as “High”, whereas the rest with

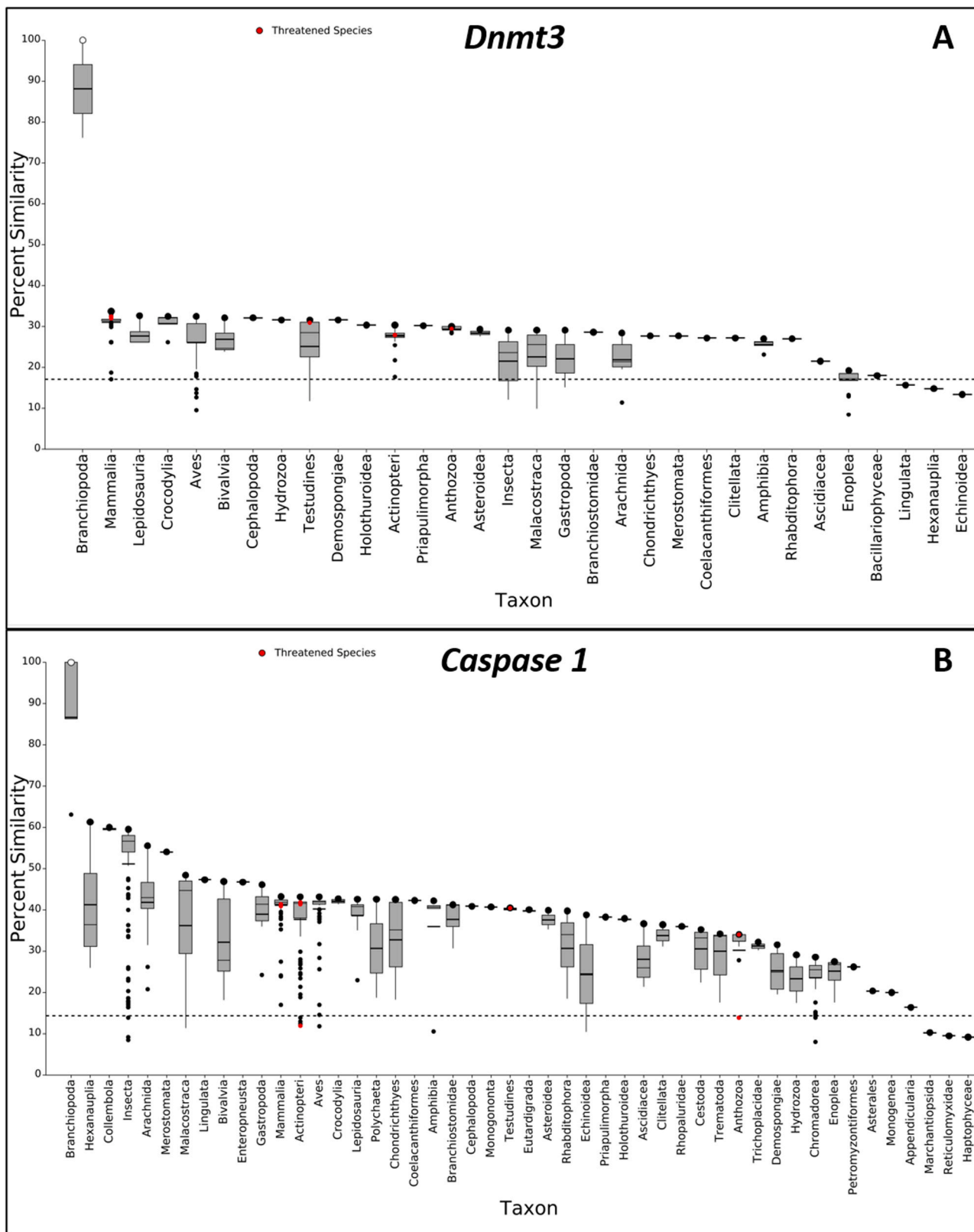


Fig. 7. Multiple protein sequence alignment of the *Daphnia magna* DNA methyltransferase 3 (*Dnmt3*, A) and caspase 1 (*Casp1*, B) conserved domains with that in other taxa to define the taxonomic applicability domain of the Adverse Outcome Pathway (AOP) networks. Dashed lines indicate susceptibility cutoffs, above the line: susceptible; below the line: not susceptible.

partial supporting evidence are scored as “Moderate” (Fig. 6 and Appendix, Table A9).

### 3.5.3. Applicability domains

The AOPs are considered applicable to female animals at both juvenile and adult stages. The SeqAPASS analysis showed that the *D. magna Dnmt3* (accession: KZS08978.1) conserved domain (accession: cd11725) had high protein sequence similarities to that of 32 taxonomic groups (439 species), of which 30 taxa (437 species) were considered susceptible groups to DNMT modulators (Fig. 7A and Appendix, Table A10). The MIE of the AOPs is therefore potentially applicable to a wide range of taxa, except for Lingulata and Hexanauplia. Alignment of the *D. magna* caspase 1 (accession: KZS14075.1) conserved domain (accession: cd00032) also showed high protein sequence similarities to 48 taxa (704 species), of which 46 taxa (703 species) were considered susceptible (Fig. 7B and Appendix, Table A10). The KE of “caspase expression” is therefore potentially applicable to most species except for Reticulomyxidae and Haptophyceae. Chemicals that have previously been classified as DNMT inhibitors in different literatures are collected in a list as the chemical applicability domain of the AOPs (Appendix, Table A11).

### 3.5.4. Potential applications of the AOPs

The AOPs are applicable to a wide range of species/chemicals and can potentially be used to reduce future needs for toxicity testing of DNMT inhibitors. Besides DNMT inhibitors, the proposed AOPs may also serve as a mechanistic basis for understanding epigenetic effects mediated by other environmental stressors which share common pathways. For instance, depletion of *s*-adenosyl methionine (SAM) as a consequence of Arsenic exposure (Reichard and Puga, 2010) may likely lead to similar pathways downstream of DNMT inhibition in the proposed AOP networks. In the context of ecological hazard and risk assessment at the population level, epigenetic marks are considered sentinels of multigenerational population dynamics, and as such, potential early warnings for impact, or identifying possible barriers to population recovery following exposure to stressors. The proposed AOPs may aid such analyses by providing extensive mechanistic knowledge and multiple lines of evidence at different levels of biological organization. As AOPs are living documents (Villeneuve et al., 2014), these generalized models will be continuously refined with accumulating evidence.

## 4. Conclusion

The present study has integrated a high-throughput laboratory analytical workflow, Bayesian network modeling and bioinformatics to understand the transgenerational effects of DNMT inhibitors (5-azacytidine) on aquatic organisms (water flea), and to support the development of novel epigenetic AOP networks for cost-efficient ecological hazard and risk assessment of epigenetic modulators. The targeted laboratory tests have generated substantial concentration-response and multigenerational data to provide in-depth mechanistic and quantitative insights into the hazards of DNMT inhibitors. The high-throughput analytical workflow developed by the present study can be easily adopted by others, as the required instrument is normally available in most laboratories. The Bayesian network assisted toxicity pathway characterization was presented as a novel data-driven approach to effectively identifying key information from high volume and complex (eco)toxicological data. The world's first epigenetic AOP network models for aquatic organisms were developed and have great potentials to inform and improve future risk assessment of DNMT inhibitors and other epigenetic modulators, thus greatly reducing (eco)toxicological testing needs. The complete experimental and data analytical workflow presented by this study also sets a good example for future *de novo* AOP and quantitative AOP (qAOP) development and evaluation based on limited empirical data.

## CRediT authorship contribution statement

**You Song:** Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing - original draft, Writing - review & editing, Project administration. **Jorke H. Kamstra:** Methodology, Investigation, Visualization, Writing - review & editing. **Yang Cao:** Methodology, Formal analysis, Visualization, Writing - review & editing. **Jana Asselman:** Methodology, Investigation, Writing - review & editing. **Marc Angles d'Auriac:** Methodology, Writing - review & editing. **Nikolai Friberg:** Writing - review & editing, Project administration, Funding acquisition.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This project was funded by the NIVA institutional funding scheme “Strategic Institutional Initiatives programme (SIS) for Restoration Ecology” and the Research Council of Norway (RCN) through the project “RiskAOP - Quantitative Adverse Outcome Pathway assisted risk assessment of mitochondrial toxicants” (grant No. 301397, <https://www.niva.no/en/projectweb/riskaop>). Special thanks to the general support from the NIVA Computational Toxicology Program (NCTP, [www.niva.no/nctp](http://www.niva.no/nctp)). During the project period, Jana Asselman was supported by the FWO Science Foundation Flanders.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at [doi:10.1016/j.jhazmat.2020.124490](https://doi.org/10.1016/j.jhazmat.2020.124490).

## References

- Abe, R., Watanabe, H., Yamamuro, M., Iguchi, T., Tatarazako, N., 2015. Establishment of a short-term, in vivo screening method for detecting chemicals with juvenile hormone activity using adult *Daphnia magna*. *J. Appl. Toxicol.* 35, 75–82.
- Angrish, M.M., Allard, P., McCullough, S.D., Druwe, I.L., Helbling Chadwick, L., Hines, E., Chorley, B.N., 2018. Epigenetic applications in Adverse Outcome Pathways and environmental risk evaluation. *Environ. Health Perspect.* 126, 045001.
- Ankley, G.T., Bennett, R.S., Erickson, R.J., Hoff, D.J., Hornung, M.W., Johnson, R.D., Mount, D.R., Nichols, J.W., Russom, C.L., Schmieder, P.K., Serrano, J.A., Tietge, J.E., Villeneuve, D.L., 2010. Adverse outcome pathways: a conceptual framework to support ecotoxicology research and risk assessment. *Environ. Toxicol. Chem.* 29, 730–741.
- Asselman, J., De Coninck, D.I.M., Beert, E., Janssen, C.R., Orsini, L., Pfrender, M.E., Decaestecker, E., De Schampelaere, K.A.C., 2017. Bisulfite sequencing with *daphnia* highlights a role for epigenetics in regulating stress response to microcystin through preferential differential methylation of serine and threonine amino acids. *Environ. Sci. Technol.* 51, 924–931.
- Athanasio, C.G., Sommer, U., Viant, M.R., Chipman, J.K., Mirbahai, L., 2018. Use of 5-azacytidine in a proof-of-concept study to evaluate the impact of pre-natal and post-natal exposures, as well as within generation persistent DNA methylation changes in *Daphnia*. *Ecotoxicology* 27, 556–568.
- Becker, R.A., Ankley, G.T., Edwards, S.W., Kennedy, S.W., Linkov, I., Meek, B., Sachana, M., Segner, H., Van Der Burg, B., Villeneuve, D.L., Watanabe, H., Barton-Maclaren, T.S., 2015. Increasing scientific confidence in Adverse Outcome Pathways: application of tailored Bradford-Hill considerations for evaluating weight of evidence. *Regul. Toxicol. Pharm.* 72, 514–537.
- 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.
- Buja, A., Hastie, T., Tibshirani, R., 1989. Linear smoothers and additive models. *Ann. Stat.* 17, 453–510.
- Chen, L., MacMillan, A.M., Chang, W., Ezaz-Nikpay, K., Lane, W.S., Verdine, G.L., 1991. Direct identification of the active-site nucleophile in a DNA (cytosine-5)-methyltransferase. *Biochemistry* 30, 11018–11025.
- Covey, J.M., D'Incalci, M., Tilchen, E.J., Zaharko, D.S., Kohn, K.W., 1986. Differences in DNA Damage Produced by Incorporation of 5-Aza-2'-deoxycytidine or 5,6-Dihydro-5-azacytidine into DNA of Mammalian Cells. *Cancer Res.* 46, 5511–5517.

- Fay, K.A., Villeneuve, D.L., LaLone, C.A., Song, Y., Tollefsen, K.E., Ankley, G.T., 2017. Practical approaches to adverse outcome pathway development and weight-of-evidence evaluation as illustrated by ecotoxicological case studies. *Environ. Toxicol. Chem.* 36, 1429–1449.
- Flores, K., Wolschin, F., Corneveaux, J.J., Allen, A.N., Huentelman, M.J., Amdam, G.V., 2012. Genome-wide association between DNA methylation and alternative splicing in an invertebrate. *BMC Genom.* 13, 480.
- Ghanim, V., Herrmann, H., Heller, G., Peter, B., Hadzijušufovic, E., Blatt, K., Schuch, K., Cerny-Reiterer, S., Mirkina, I., Karlic, H., Pickl, W.F., Zochbauer-Muller, S., Valent, P., 2012. 5-azacytidine and decitabine exert proapoptotic effects on neoplastic mast cells: role of FAS-demethylation and FAS re-expression, and synergism with FAS-ligand. *Blood* 119, 4242–4252.
- Gnyszka, A., Jastrzebski, Z., Flis, S., 2013. DNA methyltransferase inhibitors and their emerging role in epigenetic therapy of cancer. *Anticancer Res.* 33, 2989–2996.
- Harris, K.D., Bartlett, N.J., Lloyd, V.K., 2012. Daphnia as an emerging epigenetic model organism. *Genet. Res. Int.* 2012, 147892.
- Hussein, M.R., 2005. Apoptosis in the ovary: molecular mechanisms. *Hum. Reprod. Update* 11, 162–177.
- Jeremias, G., Barbosa, J., Marques, S.M., De Schamphelaere, K.A.C., Van Nieuwerburgh, F., Deforce, D., Gonçalves, F.J.M., Pereira, J.L., Asselman, J., 2018. Transgenerational inheritance of DNA hypomethylation in daphnia magna in response to salinity stress. *Environ. Sci. Technol.* 52, 10114–10123.
- Jeremias, G., Gonçalves, F.J.M., Pereira, J.L., Asselman, J., 2020. Prospects for incorporation of epigenetic biomarkers in human health and environmental risk assessment of chemicals. *Biol. Rev. Camb. Philos. Soc.* 95, 822–846.
- Kamstra, J.H., Hruba, E., Blumberg, B., Janesick, A., Mandrup, S., Hamers, T., Legler, J., 2014. Transcriptional and epigenetic mechanisms underlying enhanced in vitro adipocyte differentiation by the brominated flame retardant BDE-47. *Environ. Sci. Technol.* 48, 4110–4119.
- Kamstra, J.H., Aleström, P., Kooter, J.M., Legler, J., 2015. Zebrafish as a model to study the role of DNA methylation in environmental toxicology. *Environ. Sci. Pollut. Res. Int.* 22, 16262–16276.
- Kamstra, J.H., Sales, L.B., Aleström, P., Legler, J., 2017. Differential DNA methylation at conserved non-genic elements and evidence for transgenerational inheritance following developmental exposure to mono(2-ethylhexyl) phthalate and 5-azacytidine in zebrafish. *Epigenetics Chromatin* 10, 20.
- Khan, R., Schmidt-Mende, J., Karimi, M., Gogvadze, V., Hassan, M., Ekstrom, T.J., Zhivotovskiy, B., Hellstrom-Lindberg, E., 2008. Hypomethylation and apoptosis in 5-azacytidine-treated myeloid cells. *Exp. Hematol.* 36, 149–157.
- Kiziltepe, T., Hideshima, T., Catley, L., Rajee, N., Yasui, H., Shiraiishi, N., Okawa, Y., Ikeda, H., Vallet, S., Pozzi, S., Ishitsuka, K., Ocio, E.M., Chauhan, D., Anderson, K.C., 2007. 5-Azacytidine, a DNA methyltransferase inhibitor, induces ATR-mediated DNA double-strand break responses, apoptosis, and synergistic cytotoxicity with doxorubicin and bortezomib against multiple myeloma cells. *Mol. Cancer Ther.* 6, 1718–1727.
- Krushkal, J., Zhao, Y., Hose, C., Monks, A., Doroshow, J.H., Simon, R., 2016. Concerted changes in transcriptional regulation of genes involved in DNA methylation, demethylation, and folate-mediated one-carbon metabolism pathways in the NCI-60 cancer cell line panel in response to cancer drug treatment. *Clin. Epigenetics* 8, 73.
- LaLone, C.A., Villeneuve, D.L., Lyons, D., Helgen, H.W., Robinson, S.L., Swintek, J.A., Saari, T.W., Ankley, G.T., 2016. Sequence alignment to predict across species susceptibility (SeqAPASS): a web-based tool for addressing the challenges of cross-species extrapolation of chemical toxicity. *Toxicol. Sci.* 153, 228–245.
- Law, J.A., Jacobsen, S.E., 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat. Rev. Genet.* 11, 204–220.
- Li, L.C., Dahiya, R., 2002. MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 18, 1427–1431.
- Lindeman, L.C., Thaulow, J., Song, Y., Kamstra, J.H., Xie, L., Asselman, J., Alestrom, P., Tollefsen, K.E., 2019. Epigenetic, transcriptional and phenotypic responses in two generations of Daphnia magna exposed to the DNA methylation inhibitor 5-azacytidine. *Environ. Epigenetics* 5, dvz016.
- Mclwain, D.R., Berger, T., Mak, T.W., 2013. Caspase functions in cell death and disease. *Cold Spring Harb. Perspect. Biol.* 5, a008656.
- Meehan, R.R., Thomson, J.P., Lentini, A., Nestor, C.E., Pennings, S., 2018. DNA methylation as a genomic marker of exposure to chemical and environmental agents. *Curr. Opin. Chem. Biol.* 45, 48–56.
- Mirbahai, L., Chipman, J.K., 2014. Epigenetic memory of environmental organisms: a reflection of lifetime stressor exposures. *Mutat. Res. Genet. Toxicol. Environ. Mutagen.* 764–765, 10–17.
- Moore, L.D., Le, T., Fan, G., 2013. DNA methylation and its basic function. *Neuropsychopharmacology* 38, 23–38.
- Motulsky, H.J., Brown, R.E., 2006. Detecting outliers when fitting data with nonlinear regression – a new method based on robust nonlinear regression and the false discovery rate. *BMC Bioinforma.* 7, 123.
- Murakami, T., Li, X., Gong, J., Bhatia, U., Tragano, F., Darzynkiewicz, Z., 1995. Induction of apoptosis by 5-azacytidine: drug concentration-dependent differences in cell cycle specificity. *Cancer Res.* 55, 3093–3098.
- OECD, 2012. Test No. 211: Daphnia magna Reproduction Test.
- OECD, 2018. Users' Handbook Supplement to the Guidance Document for Developing and Assessing Adverse Outcome Pathways, OECD Series on Adverse Outcome Pathways, No. 1. OECD Publishing, Paris.
- Orsini, L., Gilbert, D., Podicheti, R., Jansen, M., Brown, J.B., Solari, O.S., Spanier, K.I., Colbourne, J.K., Rusch, D.B., Decaestecker, E., Asselman, J., De Schamphelaere, K.A., Ebert, D., Haag, C.R., Kvist, J., Laforsch, C., Petrusek, A., Beckerman, A.P., Little, T.J., Chaturvedi, A., Pfrender, M.E., De Meester, L., Frilander, M.J., 2016. Daphnia magna transcriptome by RNA-Seq across 12 environmental stressors. *In: Sci. Data*, 3.
- Palii, S.S., Van Emburgh, B.O., Sankpal, U.T., Brown, K.D., Robertson, K.D., 2008. DNA methylation inhibitor 5-aza-2'-deoxycytidine induces reversible genome-wide DNA damage that is distinctly influenced by DNA methyltransferases 1 and 3B. *Mol. Cell Biol.* 28, 752–771.
- Pfaffl, M.W., 2001. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res.* 29, e45.
- Reichard, J.F., Puga, A., 2010. Effects of arsenic exposure on DNA methylation and epigenetic gene regulation. *Epigenomics* 2, 87–104.
- Ribas, L., Vanezis, K., Imues, M.A., Piferrer, F., 2017. Treatment with a DNA methyltransferase inhibitor feminizes zebrafish and induces long-term expression changes in the gonads. *Epigenetics Chromatin* 10, 59.
- Santi, D.V., Norment, A., Garrett, C.E., 1984. Covalent bond formation between a DNA-Cytosine methyltransferase and DNA containing 5-azacytosine. *Proc. Natl. Acad. Sci. U.S.A.* 81, 6993–6997.
- Song, Y., Rundberget, J.T., Evensen, L.M., Xie, L., Gomes, T., Hogasen, T., Iguchi, T., Tollefsen, K.E., 2016. Whole-organism transcriptomic analysis provides mechanistic insight into the acute toxicity of emamectin benzoate in Daphnia magna. *Environ. Sci. Technol.* 50, 11994–12003.
- Song, Y., Villeneuve, D.L., Toyota, K., Iguchi, T., Tollefsen, K.E., 2017. Ecdysone receptor agonism leading to lethal molting disruption in arthropods: review and Adverse Outcome Pathway development. *Environ. Sci. Technol.* 51, 4142–4157.
- Song, Y., Xie, L., Lee, Y., Brede, D.A., Lyne, F., Kassaye, Y., Thaulow, J., Caldwell, G., Salbu, B., Tollefsen, K.E., 2020a. Integrative assessment of low-dose gamma radiation effects on Daphnia magna reproduction: toxicity pathway assembly and AOP development. *Sci. Total Environ.* 705, 135912.
- Song, Y., Xie, L., Lee, Y., Tollefsen, K.E., 2020b. De novo development of a quantitative adverse outcome pathway (qAOP) network for ultraviolet B (UVB) radiation using targeted laboratory tests and automated data mining. *Environ. Sci. Technol.* 54, 13147–13156.
- Stresemann, C., Lyko, F., 2008. Modes of action of the DNA methyltransferase inhibitors azacytidine and decitabine. *Int. J. Cancer* 123, 8–13.
- Tiwari, M., Prasad, S., Tripathi, A., Pandey, A.N., Ali, I., Singh, A.K., Shrivastav, T.G., Chaube, S.K., 2015. Apoptosis in mammalian oocytes: a review. *Apoptosis* 20, 1019–1025.
- Turkkan, N., Pham-Gia, T., 1993. Computation of the highest posterior density interval in bayesian analysis. *J. Stat. Comput. Simul.* 44, 243–250.
- Vandegheuchte, M.B., Janssen, C.R., 2011. Epigenetics and its implications for ecotoxicology. *Ecotoxicology* 20, 607–624.
- Vandegheuchte, M.B., Lemiere, F., Vanhaecke, L., Vanden Berghe, W., Janssen, C.R., 2010. Direct and transgenerational impact on Daphnia magna of chemicals with a known effect on DNA methylation. *Comp. Biochem. Physiol. C Toxicol. Pharm.* 151, 278–285.
- Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., Speleman, F., 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* 3.
- Villeneuve, D.L., Crump, D., Garcia-Reyero, N., Hecker, M., Hutchinson, T.H., LaLone, C.A., Landesmann, B., Lettieri, T., Munn, S., Nepelska, M., Ottinger, M.A., Vergauwen, L., Whelan, M., 2014. Adverse outcome pathway (AOP) development I: strategies and principles. *Toxicol. Sci.* 142, 312–320.
- Villeneuve, D.L., Coady, K., Escher, B.I., Mihaich, E., Murphy, C.A., Schlegel, T., Garcia-Reyero, N., 2019. High-throughput screening and environmental risk assessment: state of the science and emerging applications. *Environ. Toxicol. Chem.* 38, 12–26.
- Willett, K.L., 2018. Considering epigenetics in Adverse Outcome Pathways. In: Garcia-Reyero, N., Murphy, C.A. (Eds.), *A Systems Biology Approach to Advancing Adverse Outcome Pathways for Risk Assessment*. Springer International Publishing, Cham, pp. 219–234.