



## Research Article

# Prenatal Developmental Toxicity Testing of Petroleum Substances Using the Zebrafish Embryotoxicity Test

Lenny Kamelia<sup>1</sup>, Sylvia Brugman<sup>2</sup>, Laura de Haan<sup>1</sup>, Hans B. Ketelslegers<sup>3</sup>, Ivonne M. C. M. Rietjens<sup>1</sup> and Peter J. Boogaard<sup>1,4</sup>

<sup>1</sup>Division of Toxicology, Wageningen University and Research, Wageningen, The Netherlands; <sup>2</sup>Animal Sciences Group, Cell Biology and Immunology, Wageningen University and Research, Wageningen, The Netherlands; <sup>3</sup>European Petroleum Refiners Association, Concawe Division, Brussels, Belgium; <sup>4</sup>Shell Health, Shell International B.V., The Hague, The Netherlands

### Abstract

The present study evaluates the applicability of the zebrafish embryotoxicity test (ZET) to assess prenatal developmental toxicity (PDT) potency of the DMSO-extracts of 9 petroleum substances (PS), with variable polycyclic aromatic hydrocarbon (PAH) content, and 2 gas-to-liquid (GTL) products, without any PAHs but otherwise similar properties to PS. All PS extracts induced concentration-dependent PDT as quantified in the ZET and this potency is associated with their 3-5 ring PAH content. In contrast and as expected, GTL products did not induce any effect in the ZET. The potencies obtained in the ZET correlated with those previously reported for the embryonic stem cell test (EST) ( $R^2 = 0.61$ ), while the correlation with potencies reported in *in vivo* studies was higher for the EST ( $R^2 = 0.85$ ) than the ZET ( $R^2 = 0.69$ ). Combining the results of the ZET with those previously reported for the EST (Kamelia et al., 2017), the aryl hydrocarbon (AhR) CALUX assay (Kamelia et al., 2018), and the PAH content, ranked and clustered the test compounds in line with their *in vivo* potencies and chemical characteristics. Our findings indicate that the ZET does not outperform the EST as a stand-alone assay for testing PDT of PS, but confirm the hypothesis that PAHs are the major inducers of PDT by some PS, while they also indicate that the ZET is a useful addition to a battery of *in vitro* tests able to predict the *in vivo* PDT of PS.

## 1 Introduction

Petroleum substances (PS) are UVCBs (substances of Unknown or Variable composition, Complex reaction products and Biological materials) and regulated as such under the European Union (EU) REACH legislation (Registration, Evaluation, Authorisation, and Restriction of Chemicals). REACH has been implemented in the EU since 2007 to improve the protection of human health and the environment from the risks of chemical exposure. REACH legislation requires substances, including PS, that are manufactured or marketed in the EU at a volume of > 100 tons/year to be tested for prenatal developmental toxicity (PDT) (ECHA, 2009). This is a new requirement that was not present in the previous legislation, hence PDT testing would be required for all existing substances produced over a volume of > 100 tons/year. One of the consequences is that most PS (186 currently active registered EINECS numbers, all with a volume of > 1000 tons/year) will need to be tested for their potential adverse effect on

prenatal development according to the Organisation for Economic Co-operation and Development (OECD) 414 testing guideline. Theoretically, this would require substantial numbers of experimental animals (> 2500 animals/test/compound; OECD, 2001) and involve a considerable amount of resources, both monetary and in terms of capacity (van der Jagt et al., 2004; Rovida et al., 2011). Therefore, the use of a battery of *in vitro* alternative assays, such as the zebrafish embryotoxicity test (ZET; Sipes et al., 2011; Strähle et al., 2012) and/or the embryonic stem cell test (EST; Spielmann, 2009), to group PS into a limited number of categories and/or to facilitate read-across from PS for which *in vivo* PDT data are already available, may reduce the number of animals and resources needed to study PDT potencies of PS.

The embryonic stem cell test (EST; Genschow et al., 2004), the whole embryo culture (WEC; Piersma et al., 2004), the limb bud micromass (Spielmann et al., 2004), and the zebrafish embryotoxicity test (ZET; OECD 2011a,b; Busquet et al., 2014), are the four *in vitro* alternative methods that have been validated for PDT testing.

Received August 12, 2018; Accepted December 3, 2018;  
Epub December 7, 2018, 2018; © The Authors, 2018.

ALTEX 36(2), 245-260. doi:10.14573/altex.1808121

Correspondence: Lenny Kamelia, PhD, Division of Toxicology, Wageningen University and Research, PO BOX 8000, 6700 EA Wageningen, The Netherlands (lenny.kamelia@wur.nl)

This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International license (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is appropriately cited.



In the past decades, the ZET and EST have been the most used *in vitro* alternative assays to assess PDT of chemicals compared to the other two alternative methods (Incardona et al., 2004; de Jong et al., 2009; Selderslaghs et al., 2009; Louisse et al., 2011; Hawliczek et al., 2012; Kuske et al., 2012; Strikwold et al., 2012; Goodale et al., 2013; Li et al., 2015; Dimopoulou et al., 2018). The EST does not require harvesting of cells from animals since a permanent stem cell line is used. The validated EST consists of three different assays, including a cytotoxicity assay on differentiated T3 fibroblasts, a cytotoxicity assay on undifferentiated ES-D3 cells, and a differentiation assay of ES-D3 cells into beating cardiomyocytes (Rohwedel et al., 2001; Genschow et al., 2002). The differentiation assay of the EST can be used to rank the potency of chemicals and define concentration-response curves for effects on cell differentiation and development (Seiler and Spielmann, 2011).

The ZET is another emerging alternative method for PDT testing. It makes use of a developing vertebrate organism to investigate the embryotoxicity potential of substances based on the notion that the development of the zebrafish embryo is very similar to the embryogenesis in higher vertebrates, including humans (Sipes et al., 2011). According to EU Directive 2010/63/EU, the ZET is considered an alternative method since the test does not exceed 5 days post-fertilization (dpf) and zebrafish embryos are considered free-living larvae during this time period (EU, 2010). The ZET makes use of newly fertilized eggs and a general morphology scoring (GMS) system, which includes the evaluation of mortality, and the development of tail, somite, heartbeat and eyes to assess the effects of chemicals on the development of zebrafish embryos (Hermesen et al., 2011; Jomaa et al., 2014; Beekhuijzen et al., 2015). The ZET offers several advantages over the EST and the other *in vitro* alternative assays for PDT testing. First, the ZET uses a complete embryo for the test and the chorion of a zebrafish embryo is transparent or is absent from around 2 dpf, facilitating detection of multiple developmental effects on multiple organs during testing (Glaberman et al., 2017). Good predictability, possibility for large-scale high-throughput hazard screening, short duration, ease to perform, and low cost make the ZET one of the most used alternative assays in *in vitro* PDT testing (Scholz et al., 2008; Kari et al., 2007; Brannen et al., 2010; Hill et al., 2011).

PS are complex materials, comprising hundreds to millions of different hydrocarbon constituents, including polycyclic aromatic hydrocarbons (PAHs). PAHs that are present in PS are generally known as petrogenic-origin or petroleum-associated hydrocarbons and differ substantially from PAHs of pyrogenic origin (Pampanin and Sydes, 2003). The level and type of PAH constituents may vary depending on the source of the crude oil and the processing conditions used to manufacture the raw material (Speight, 2006). For instance, heavy PS, e.g., heavy fuel oil (HFO), untreated lubricating oils, and distillate aromatic extract (DAE), contain high amounts of high-molecular-weight (HMW) PAHs, i.e., 4- to 7-ring PAHs, where light PS, such as gas oil (GO), mainly contain low molecular weight (LMW) PAHs, i.e., 2-3 ring

PAHs. In contrast to PS, which generally contain PAHs as one of their constituents, gas-to-liquid (GTL) products are modern synthetic analogues of PS that typically consist of only saturated hydrocarbons and are virtually devoid of unsaturated and aromatic compounds (Boogaard et al., 2017).

Some PAH-containing PS have been reported to induce PDT *in vivo* and show effects in *in vitro* systems designed to test for PDT. Moreover, their potency has been shown to be associated with the presence of 3- to 7-ring PAHs in these products (Feuston et al., 1997; Mackerer et al., 2003; Gray et al., 2013; Kamelia et al., 2017). Animal studies have provided data on developmental toxicity effects induced by some PS, including increased incidence of resorptions (prenatal loss), decreased number of live fetuses/litter, and decreased fetal body weight (Mobil, 1989; ARCO, 1993; Hoberman et al., 1995; Feuston et al., 1996). On the other hand, GTL products, which contain no PAHs but have similar other properties to PS, were negative when tested in PDT studies as well as in two-generation reproductive toxicity studies (Dunster, 2014; Senn, 2014; Boogaard et al., 2017). Altogether, these data led to the hypothesis that the PDT as observed with certain PS is induced by particular 3-7 ring PAH constituents present in these products. In other words, heavier PS containing high concentrations of HMW PAHs may induce PDT where products with no or low concentrations of PAHs will not induce PDT (Tsitou et al., 2015; Kamelia et al., 2017).

This hypothesis is supported by our recent findings, where the EST and a panel of CALUX reporter gene assays were used to evaluate the PDT potency and possible underlying mechanism of PDT of the DMSO-extracts of PS and GTL products (Kamelia et al., 2017, 2018). In the EST, DMSO-extracts of 9 PS, varying in PAH level and content, were able to inhibit the differentiation of ES-D3 cells into beating cardiomyocytes, where GTL products that contain no PAHs showed no effects (Kamelia et al., 2017). Moreover, a pronounced aryl hydrocarbon (AhR)-mediated activity was found upon exposure to increasing concentrations of DMSO-extracts of the same PS, as tested in the AhR CALUX assay. This AhR-mediated activity correlated well with the *in vitro* PDT potency in the EST, suggesting an important role of the AhR in mediating this effect (Kamelia et al., 2018). These earlier results corroborated the notion that PAHs are the primary inducers of PDT in some PS and that the AhR may play an important role in the underlying mode of action. The present study aims to evaluate the PDT potency of the same substances, DMSO-extracts of 9 PS and 2 GTL products, in the ZET in order to investigate the suitability of the ZET to predict relative *in vivo* PDT of these PS extracts.

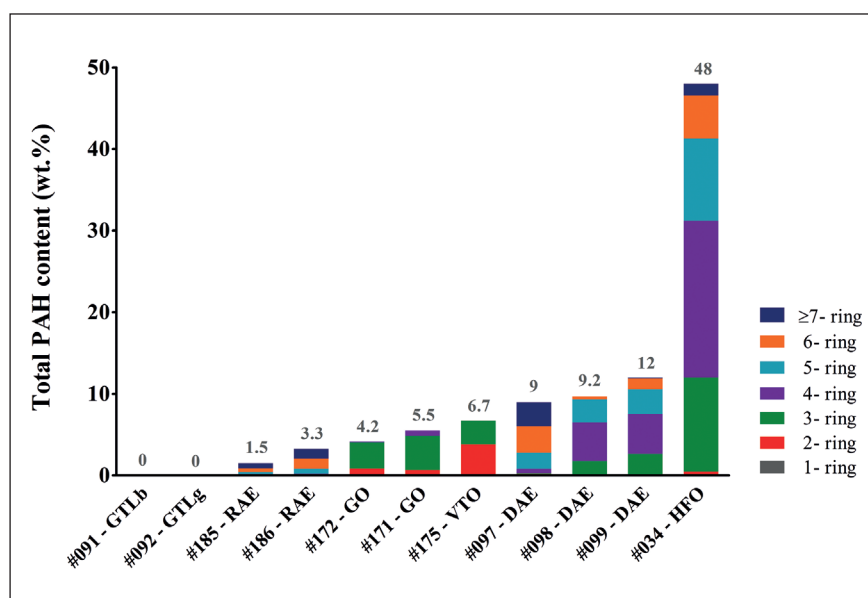
## 2 Materials and methods

### 2.1 Test compounds

DMSO-extracts of 9 PS, varying in PAH content and concentration, and 2 GTL products, which contain no PAHs but have

#### Abbreviations

AhR, aryl hydrocarbon; BMC, benchmark concentration; BMD, benchmark dose; BMR, benchmark response; DAE, distillate aromatic extract; dpf, days post-fertilization; EST, embryonic stem cell test; GMS, general morphology scoring; GO, gas oil; GTL, gas-to-liquid; HFO, heavy fuel oil; HMW, high-molecular-weight; LMW, low molecular weight; PAH, polycyclic aromatic hydrocarbon; PDT, prenatal developmental toxicity; PS, petroleum substances; WEC, whole embryo culture; ZET, zebrafish embryotoxicity test



**Fig. 1: Aromatic ring class profiles of PS and GTL products tested in the present study**

Aromatic ring class (ARC) profiles of 2 GTL products (contain no aromatics) and 9 PS (vary in PAH content, starting from 1.5% to 48% of total weight PAHs) tested in the present study. ARC profiles represent the weight percent of the DMSO-soluble 1- to  $\geq 7$  aromatic-ring compounds present in each PS and GTL product sample, from the starting raw material of 4.0 g, as determined using the Method II chemical characterization procedure described in detail by Roy et al. (1988). HFO, heavy fuel oil; DAE, distillate aromatic extract; GO, gas oil; VTO, vacuum tower overhead oil; RAE, residual aromatic extract; GTLb, gas-to-liquid base oil; GTLg, gas-to-liquid gas oil

similar other properties to PS, were the model substances tested. The DMSO-extracts of the PS and GTL products were generated using the extraction procedure described by Roy et al. (1988), explained in some more detail in Section 2.2. The raw material of all PS and GTL products that were used for the DMSO extraction were provided by Concawe (Brussels, Belgium) and Shell International by (The Hague, The Netherlands), respectively. These raw materials were: 1 heavy fuel oil (HFO; CAS no. 64741-62-4), 3 distillate aromatic extracts (DAE; all 3 possessing the same CAS no. 64742-04-7), 2 residual aromatic extracts (RAE; CAS no. 64742-10-5 and 91995-70-9), 2 gas oils (GO; CAS no. 68915-96-8 and 64741-43-1), 1 vacuum tower overhead oil (VTO; CAS no. 64741-49-7), 1 GTL base oil (GTLb; CAS no. 848301-69-9), and 1 GTL gas oil (GTLg; CAS no. 848301-67-7). An overview of the PAHs present in each PS and GTL product of the present study, grouped by the number of aromatic rings, is provided in Figure 1 (see also Appendix D<sup>1</sup>).

## 2.2 PAH extraction and analysis

Extraction and analysis of the PS and GTL products was performed essentially as described before by Roy et al. (1988), and was carried out at Port Royal Research Laboratory (Hilton Head, South Carolina, USA). This extraction and analysis procedure is commonly used to obtain and quantify the PAH fraction from the raw material of PS (also known as Method II chemical characterization procedure), and has been widely used and validated also for mutagenicity and carcinogenicity testing of PS (Blackburn et al., 1986; Concawe, 1994; Clonfero et al., 1996; Mackerer et al., 2003).

In brief, 4.0 g of PS or GTL product sample was dissolved in 10 ml of cyclohexane and extracted twice with 10 ml DMSO.

The DMSO extracts were combined and diluted with twice the volume of 4% NaCl solution. The diluted DMSO fraction was back extracted with 20 and 10 ml of cyclohexane. The cyclohexane fractions were combined, washed twice with 5 ml of distilled water, and filtered through anhydrous sodium sulphate. The cyclohexane was then evaporated to dryness at 40°C followed by further evaporation at 80°C for 30 min. This residue was then re-dissolved in cyclohexane (~50 mg/ml). The extracts were analyzed by gas chromatography with mass spectrometry (GC/MS). Naphthalene, phenanthrene, pyrene, benzo(a)pyrene, benzo[ghi]perylene, and coronene were used as standards to define the boundaries of retention times for 2-7 ring PAHs. The results are presented as the aromatic ring class (ARC) profile, the weight percent of the DMSO-soluble 1- to  $\geq 7$ -ring aromatic compounds present in each sample, from the starting raw material of 4.0 g (Fig. 1).

## 2.3 Zebrafish embryotoxicity test (ZET)

Wild-type adult zebrafish (*Danio rerio*) AB line were obtained from the research facility Carus, Wageningen University and Research, The Netherlands, and maintained in a flow-through aquarium system at 27°C with 14 h light/10 h dark cycle. The maintenance of adult zebrafish was done in accordance with the protocols of the Zebrafish Handbook<sup>2</sup>. Zebrafish eggs were produced via spawning groups by placing male and female fish (at a ratio of 1:2) in an individual spawning tank equipped with spawn traps several hours prior to the onset of darkness on the day before the test. A minimum of three parallel spawning tanks was set for the egg production for each independent experiment. Mating, spawning, and fertilization take place within 30 min to 1 h after the onset of the light cycle.

<sup>1</sup> doi:10.14573/altex.1808121s

<sup>2</sup> http://zfin.org



Spawned eggs were collected, rinsed a few times with egg water (1.5 ml salt stock solution in 1 l distilled water), and incubated at 26°C until further use for egg selection. The salt stock solution was prepared by dissolving 40 g “Instant Ocean” sea salt (Blacksburg, Virginia, USA) in 1 liter distilled water. The egg water was aerated for several hours, adjusted for pH (range of pH 7-8), and kept at 26-28°C prior to use. Egg water was used throughout the ZET as the assay medium. Egg selection was done by sorting the zebrafish embryo at the 8- to 32-cell stage using a stereomicroscope and disposable pipette, choosing embryos that followed a normal development, which were pooled in Petri dishes for further use. Zebrafish embryos with obvious anomalies, such as coagulated embryos, and unfertilized eggs, were discarded. The ZET was initiated at 4-5 h post fertilization (hpf) and terminated at 96 hpf, as this covers the entire organogenesis in a zebrafish embryo. The exposure was performed in 24-well plates (Greiner Bio-one, Frickenhausen, Germany) in combination with a self-adhesive film cover (Sigma-Aldrich, Zwijndrecht, The Netherlands). 4-5 hpf was chosen as the start time of exposure as this covers the same development stages as in the developmental toxicity studies performed according to the OECD 414 guideline in rat and rabbit (Beekhuijzen et al., 2015). Hence, this allows comparison of results between the ZET and *in vivo* PDT studies.

Twenty wells of the 24-well plate were used for exposure to one concentration of test compound and the other four wells were used for the internal plate control. Every well of the internal plate control, in the exposed-plate, contained a zebrafish embryo in egg water (negative control), and if more than 1 dead embryo, out of 4 embryos, was observed in these internal plate control wells, the plate was rejected and the results from the respective plate were not used. The exposure medium was prepared by mixing the 400-times concentrated stock solutions of the test compounds (dissolved in DMSO) with egg water. The exposure medium was then transferred into 20 wells of the 24-well plate, at 2 ml exposure medium/well, and for the internal plate control, 2 ml egg water was added into each of the 4 remaining wells. The plate was sealed with self-adhesive film cover to prevent evaporation of test compound throughout the exposure period (up to 96 hpf). All samples were tested at a range of concentrations up to 250 µg raw material/ml, except for sample #034-HFO, which could be dosed up to only 15 µg raw material/ml due to solubility limitations in egg water. Solvent controls (0.25% v/v DMSO; Merck, Darmstadt, Germany), positive controls (4 µg/ml 3,4-dichloroaniline; Sigma-Aldrich), and negative controls (egg water only) were included in each independent experiment. Plates were incubated at 26°C with a photo period of 14 h light: 10 h dark. Embryos were scored daily (every 24 h) for developmental abnormalities and cumulative mortality using the inverted microscope until 96 hpf (the daily assessment time points correspond to 28, 52, 76, and 100 hpf), based on the extended general morphological scoring (GMS) system described by Beekhuijzen et al. (2015). Deviation from normal developmental stages, for example incomplete detachment of tail, incomplete development of eyes, fin, and mouth, unhatched embryos, results in a lower total GMS value corresponding to a certain extent of developmental retardation. A comparison of

results from the daily morphological assessment using the GMS reveals that the most significant results upon exposure to test compounds were achieved with an exposure time window of 0-96 hpf (Appendix E<sup>1</sup>). The GMS used for the exposure time window of 0-96 hpf is based on the 96 hpf endpoints, as listed and described in detail by Beekhuijzen et al. (2015). Exposure time windows of 0-24/0-48/0-72 hpf did not show notable differences between solvent control and test compounds, hence the GMS data of 0-96 hpf were used for further data analysis and comparison, including comparison with published *in vivo* data or PAH content present in each PS sample (Section 2.7). The ZET results of each test compound were presented as fraction of GMS at 96 hpf compared to that of the solvent control (0.25% DMSO). The ZET was considered valid if the following was observed: ≤ 1 dead embryo (out of 4) in the internal plate control of every exposed-plate; ≤ 3 dead embryos (out of 24) in the negative control plate (at least 87.5% survival rate); ≤ 2 dead embryos (out of 20) in the solvent control plate (0.25% DMSO); ≤ 14 live embryos (out of 20) in the positive control plate (4 µg/ml 3,4-dichloroaniline; exposure to positive control should result in a minimum of 30% mortality by 96 hpf). At least 4 independent ZET experiments were performed for each test compound in which the same stock dilutions were used to prepare the concentration ranges tested. A graphical illustration of the ZET performed in the present study and the layout of 24-well plates for the negative, positive, and solvent controls are shown in Appendix H<sup>1</sup>.

## 2.4 Embryonic stem cell test (EST)

The *in vitro* developmental toxicity data of the DMSO-extracts of 9 PS and 2 GTL products, as tested in the EST, were taken from our previous study (Kamelia et al., 2017). In short, the effects on cell viability and cell differentiation upon exposure to increasing concentrations of test compounds were evaluated in the ES-D3 cell viability and ES-D3 cell differentiation assays of the EST.

## 2.5 Aryl hydrocarbon (AhR) CALUX reporter gene assay

Data for the AhR-mediated activity of the DMSO-extracts of 9 PS and 2 GTL products, as tested in the AhR CALUX assay, were taken from our recently published study (Kamelia et al., 2018). Briefly, H4IIE.luc cells were exposed to increasing concentrations of the DMSO-extracts of the aforementioned substances for 6 h. The luciferase induction activity upon exposure was then measured, and the amount of produced luminescence was used to quantify the AhR-mediated activity induced by the respective test compound.

## 2.6 Data analysis

Concentration-response curves upon exposure to the DMSO-extracts of PS and GTL products in the ZET were made using GraphPad Prism 5.0 (California, US). Here, data were fitted to a sigmoid concentration-response curve with three parameters. These curves were used only for graphical illustration of the obtained results, and not for the determination of benchmark concentration (BMC) values. Results obtained from the ZET were



**Tab. 1: Overview of *in vivo* and *in vitro* developmental toxicity data upon exposure to PS and GTL products**

Compounds	CAS number	<i>In vivo</i> data <sup>a</sup>			<i>In vitro</i> data		
		BMD10 increased resorptions (mg/kg bw/day) <sup>b</sup>	BMD10 number of live fetuses/litter (mg/kg bw/day) <sup>c</sup>	BMD10 fetal body weight (mg/kg bw/day) <sup>d</sup>	BMC50 ZET (μg/ml)	BMCd50 EST (μg/ml) <sup>g</sup>	EC50 AhR CALUX (μg/ml) <sup>h</sup>
#034 - HFO	64741-62-4	0.32	0.59	20.64	1.32	0.47	0.0037
#097 - DAE	64742-04-7	13.89	44.38	28.53	76.71	3.74	0.033
#098 - DAE	64742-04-7	13.89	44.38	28.53	21.82	5.05	0.012
#099 - DAE	64742-04-7	13.89	44.38	28.53	21.23	7.42	0.012
#171 - GO	68915-96-8	n.a	n.a	n.a	54.53	171	0.084
#172 - GO	64741-43-1	270.05	174.04	245.21	73.13	142	0.35
#175 - VTO	64741-49-7	151.16	144.03	332.64	64.77	67.2	0.40
#185 - RAE	64742-10-5	n.a	n.a	n.a	> 250 <sup>f</sup>	89.5	0.40
#186 - RAE	91995-70-9	n.a	n.a	n.a	> 250 <sup>f</sup>	95.2	0.13
#091 - GTLb	848301-69-9	(-) <sup>e</sup>	(-) <sup>e</sup>	(-) <sup>e</sup>	(-)	(-)	(-)
#092 - GTLg	848301-67-7	(-) <sup>e</sup>	(-) <sup>e</sup>	(-) <sup>e</sup>	(-)	(-)	(-)

<sup>a</sup>*In vivo* data for determination of the BMD10 values from different developmental toxicity endpoints were taken from Hoberman et al. (1995) for the HFO; Feuston et al. (1996) for the DAE; ARCO (1993) for the GO; and Mobil (1989) for the VTO. <sup>b</sup>BMD10 increased resorptions: dose (mg/kg bw/day) giving 10% increase in incidence of resorptions. BMD10 values were calculated from a dose-response curve using a dichotomous model (BMD software US-EPA) and data were taken from cited studies. <sup>c</sup>BMD10 number of live fetuses/litter: dose (mg/kg bw/day) giving 10% decrease in live fetuses/litter. BMD10 values were calculated from a dose-response curve using a continuous model (BMD software US-EPA) and data were taken from cited studies. <sup>d</sup>BMD10 fetal body weight: dose (mg/kg bw/day) giving 10% decrease in fetal body weight. BMD10 values were calculated from a dose-response curve using a continuous model (BMD software US-EPA) and data were taken from cited studies. <sup>e</sup>Negative for prenatal developmental toxicity studies; data were taken from Senn (2014) for the GTLb and Dunster (2014) for the GTLg. <sup>f</sup>The calculated BMC50 values are above the highest tested concentration of 250 μg/ml. <sup>g</sup>The BMC<sub>d</sub>50s-EST is taken from Kamelia et al., 2017. <sup>h</sup>The EC50s-AhR CALUX is taken from Kamelia et al. (2018). HFO, heavy fuel oil; DAE, distillate aromatic extract; GO, gas oil; VTO, vacuum tower overhead; RAE, residual aromatic extract; GTLb, gas-to-liquid base oil; GTLg, gas-to-liquid gas oil; n.a, data not available; (-), negative; AhR, aryl hydrocarbon receptor; CALUX, chemical activated luciferase gene expression

expressed as fraction of total GMS score at 96 hpf compared to the solvent control (0.25% DMSO) and are presented as mean ± standard error of the mean (SEM). The DMSO concentration was kept at 0.25% v/v throughout the ZET experiments.

#### Benchmark concentration (BMC) derivation of *in vitro* developmental toxicity data

Results from the ZET were analyzed using the Benchmark Dose software (BMD software US-EPA version 2.6.1) to obtain the benchmark concentration at a benchmark response (BMR) of 50% effect (BMC50). For this purpose, concentration-response curves from the ZET were fitted to all dichotomous concentration-response models (gamma, logistic, log-logistic, probit, log-probit, weibull, multistage-cancer, multistage, and the quantal-linear models; Appendix A<sup>1</sup>) available in the BMD software US-EPA version 2.6.1. The BMR was set to 50%, representing 50% developmental retardation (BMC50) in the ZET upon exposure to the DMSO-extracts of PS and GTL products. The performance of each fitted model was evaluated based on the goodness-of-fit, the scaled residuals, and the visual inspection of model fitting. To allow comparison to data obtained previously in the EST (Kamelia et al., 2017), the respective concentration-re-

sponse curves of the EST were analyzed in a similar manner. The BMC50 values thus obtained (Tab. 1), which account for 50% inhibition of cell differentiation (EST; BMCd50-EST) or 50% developmental retardation (ZET; BMC50-ZET), were selected from all accepted models based on the lowest Akaike's Information Criterion (AIC) (Kamelia et al., 2017; Haber et al., 2018) and were further used for the *in vitro-in vitro* (ZET to EST; ZET to AhR CALUX assay) and the *in vitro-in vivo* comparisons.

#### Benchmark dose (BMD) modeling of *in vivo* developmental toxicity data

*In vivo* data, derived from published PDT studies available for some of the PS of the present study, were used to determine the *in vivo* benchmark dose (BMD) values of the corresponding PS. *In vivo* PDT data of GO (ARCO, 1993) and VTO (Mobil, 1989) were kindly provided by Concawe, and the *in vivo* data of the HFO and DAE were obtained from Feuston et al. (1996) and Hoberman et al. (1995), respectively. Different developmental toxicity endpoints were reported in these studies, such as increased resorptions, number of live fetuses/litter, decreased fetal body weight, and skeletal variations. Increased resorptions, number of live fetuses/litter, decreased fetal body weight, and not



skeletal variations, were selected for the BMD analysis, because all studies contained data on incidence of resorptions, number of live fetuses/litter, and fetal body weight, whereas the studies of HFO and DAE did not provide data on skeletal variations. To determine BMD values, *in vivo* data for fetal body weight and number of live fetuses/litter were fitted to the continuous models (hill, exponential, linear, polynomial, and power models) and the data from the incidence of resorptions were fitted to the dichotomous models (gamma, logistic, log-logistic, probit, log-probit, weibull, multistage-cancer, multistage, and the quantal-linear models), of the BMD software US-EPA version 2.6.1. For these BMD analyses, the BMR was set to 10%, which represents a 10% decrease in fetal body weight (BMD10-fetal body weight), 10% decrease in number of live fetuses/litter (BMD10-live fetuses/litter; Appendix B<sup>1</sup>), and a 10% additional incidence of resorptions (BMD10-increased resorptions). The performance of each fitted model was evaluated based on the goodness-of-fit, the scaled residuals, and the visual inspection of model fitting. The BMD10 values (Tab. 1) derived from the models with the best fit and lowest AIC were selected and further used for the *in vitro-in vivo* comparison.

## 2.7 Correlation analysis

*Correlation of in vitro potencies obtained from three distinct in vitro assays: ZET, EST, and AhR CALUX assay*

Using a linear regression approach (in GraphPad Prism 5.0), the correlation between *in vitro* potencies obtained from three different assays that evaluate the PDT potency and underlying mechanism of PDT of PS and GTL products, i.e., the ZET, EST, and AhR CALUX assay, were determined. Data for the EST and AhR-mediated activity were taken from our previous studies (Kamelia et al., 2017; 2018; Section 2.4 and 2.5, respectively). The correlation analyses were conducted between *in vitro* potencies (BMC50s/EC50s) of the test compounds between each duo of the three different *in vitro* assays, including ZET and EST, ZET and AhR CALUX assay, and EST and AhR CALUX assay. The given R-squared ( $R^2$ ) value reflects the goodness-of-fit of data to the fitted regression line and was considered statistically significant if the p-value was lower than 0.05.

*Correlation between in vitro developmental toxicity potency in the ZET and the in vivo developmental toxicity data or the PAH content*

*In vitro* potencies obtained in the ZET were compared to potencies derived from *in vivo* studies to see whether any correlation exists between *in vitro* and *in vivo* PDT potencies. To this end, the BMC50-ZET values were plotted against the BMD10 values, obtained as described in Section 2.6. No relevant report/literature on an *in vivo* developmental toxicity study was available for the RAE samples, and for that reason they were not included in the *in vitro-in vivo* comparison. Furthermore, correlations between *in vitro* potencies in the ZET and specific PAH constituents present in each PS sample were also investigated. All these correlation analyses were done by performing a linear regression analysis in GraphPad Prism 5.0. The given  $R^2$  value reflects the goodness-of-fit of data to the fitted regression line and was considered statistically significant if the p-value was lower than 0.05.

## 2.8 Data integration and visualization in ToxPi GUI 2.0

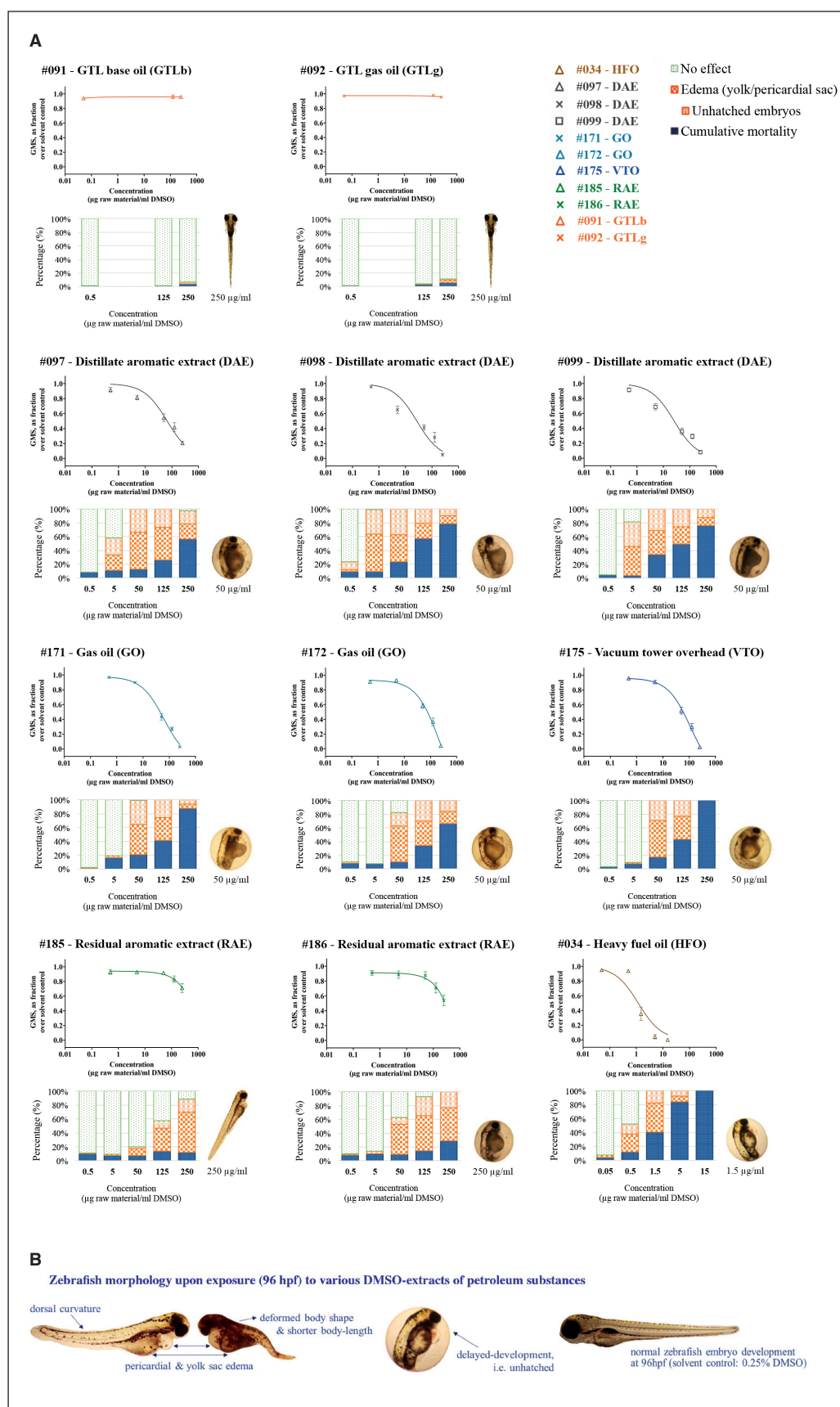
ToxPi Graphical User Interface 2.0 (ToxPi GUI 2.0; Reif et al., 2010; Marvel et al., 2018) was used for data integration, visualization, and comparison of effect-signatures from different *in vitro* assays and the chemical analysis of the 9 PS and 2 GTL products. The *in vitro* potencies (BMC50s/EC50s) in the three different assays (EST, ZET, and AhR CALUX assay), and total PAH content (mg/ml) present in each test compound, were used as point of departure (POD) for ToxPi data input. In cases where no effects were observed, for example the GTL products showed negative results in all of the selected assays, the highest tested concentration was used as the POD value for the respective test compound. Briefly, the POD values (BMC50s/EC50s/total PAH content) of each test compound were listed and inversely normalized on a 0-1 scale. The value of 0 represents the lowest bioactivity in the corresponding *in vitro* assay (the least potent compound of all in the corresponding *in vitro* assay). The value of 1 represents the highest bioactivity or the most potent test compound of all in a given data set. For the chemical analysis data (PAH profiles), the value of 0 represents the test compound that contains the lowest PAH content, i.e., GTL products (0% PAHs), and the value of 1 represents the test compound with the highest PAH content, which is sample #034-HFO (48% PAHs). These normalized POD values were then used as quantitative inputs for bioactivity profiling in ToxPi (Grimm et al., 2016). Further, the ToxPi score was calculated based on the equation provided below, and from this the bioactivity profiling (ToxPi pie-charts), hierarchical clustering, and chemical rank were generated. The terms POD min and POD max that were used to calculate the ToxPi score represent the lowest and highest POD, respectively, observed within one corresponding assay of the present study, i.e., within the ZET. The average method was used for the development of the hierarchical clustering in ToxPi.

$$\text{ToxPi score} = 1 - \frac{\log_{10} \text{POD} - \log_{10} \text{POD min.}}{\log_{10} \text{POD max.} - \log_{10} \text{POD min.}}$$

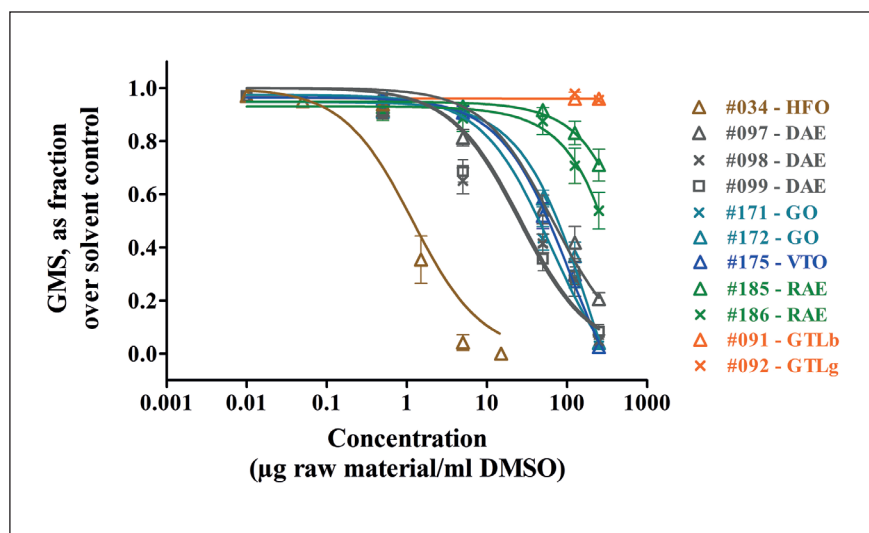
## 3 Results

### 3.1 Zebrafish embryotoxicity test (ZET)

As shown in Figure 2A, exposure to increasing concentrations of the DMSO-extracts of PS induced concentration-dependent developmental retardation in zebrafish embryos (scored at 96 hpf), while both GTL products showed no effects (data are provided in Appendix C<sup>1</sup>). The effects induced by PS in the ZET at this time point include the absence of circulation and movement, delayed development (i.e., unhatched embryos), deformed body shape (dorsal curvature), the reduction of body and tail length, and malformations of the heart and yolk sac (Fig. 2B). Figure 2A presents the 100% stacked-bar graph illustrating the percentage of particular developmental anomalies at 96 hpf, including pericardial and yolk sac edemas (in both hatched and unhatched embryos) and cumulative mortality, below the concentration-response curves of the GMS. It should be noted that the deformed body shape and the reduction of body and tail length can only be observed when the embryo has hatched at 96 hpf.







**Fig. 3: Comparison of the concentration-dependent inhibition of zebrafish embryo development in the ZET (at 96 hpf) upon exposure to DMSO-extracts of PS and GTL products**

Results represent data from at least four independent experiments ( $n \geq 4$ ) and are presented as mean  $\pm$  standard error of the mean (SEM). HFO, heavy fuel oil; DAE, distillate aromatic extract; GO, gas oil; VTO, vacuum tower overhead oil; RAE, residual aromatic extract; GTLb, gas-to-liquid base oil; GTLg, gas-to-liquid gas oil; hpf, hours post fertilization

As a consequence, only the occurrence of pericardial and yolk sac edema, unhatched embryos, and cumulative mortality were noted and presented here in the stack-bar graph of Figure 2A. From the stacked-bar graph, it is clear that the percentage of cumulative mortality increases with the concentration of the test compounds, except for the RAE samples. The main manifestations upon exposure to increasing concentrations of RAE samples were pericardial and yolk-sac edemas. It is worth mentioning that the high incidence of pericardial and yolk sac edemas induced by heavy PS like HFO and DAEs happened at a lower concentration, ranging from 0.5 (HFO) to 5 µg/ml (DAE), than for the light PS, such as GO and VTO, for which the effect occurred only from 50 µg/ml onwards.

Using the concentration-response curves of the GMS, the BMC50-ZET value was determined and the results were listed in Table 1 and presented in Figure 3. Among all samples tested, sample #034-HFO appeared to be the most potent substance in inducing the developmental retardation of zebrafish embryos, followed by two of the DAE samples (#098 and #099), then the GO, VTO, and sample #097-DAE, and lastly the RAE samples (Fig. 3). For RAE samples, a decline in the GMS occurred only at the highest tested concentrations (Fig. 2A and 3), thus, the obtained BMC50 values of these two samples were above the highest tested concentration of 250 µg/ml (Tab. 1). The BMC50 values for both of the GTL products could not be determined as they did not induce any effects in the ZET.

### 3.2 Correlation between *in vitro* developmental toxicity in the ZET and EST, and between the *in vitro* PDT in the ZET or EST and the observed effects in the AhR-mediated gene expression assay testing the same substances

The results obtained in the ZET were compared to results obtained in our previous studies (Kamelia et al., 2017, 2018) for the same DMSO-extracts of PS and GTL products tested in the EST and AhR CALUX assay. Figure 4 presents the correlation between the results obtained in these *in vitro* assays. The *in vitro* potencies in the ZET had a moderate correlation with either the

EST ( $R^2 = 0.61$ ; Fig. 4A) or the AhR CALUX assay ( $R^2 = 0.66$ ; Fig. 4B). Further, a good correlation also exists between the *in vitro* potencies in the EST and the AhR CALUX assay, showing an  $R^2$  value of 0.80 (Kamelia et al., 2018; data and graph are provided in Appendix E<sup>1</sup>).

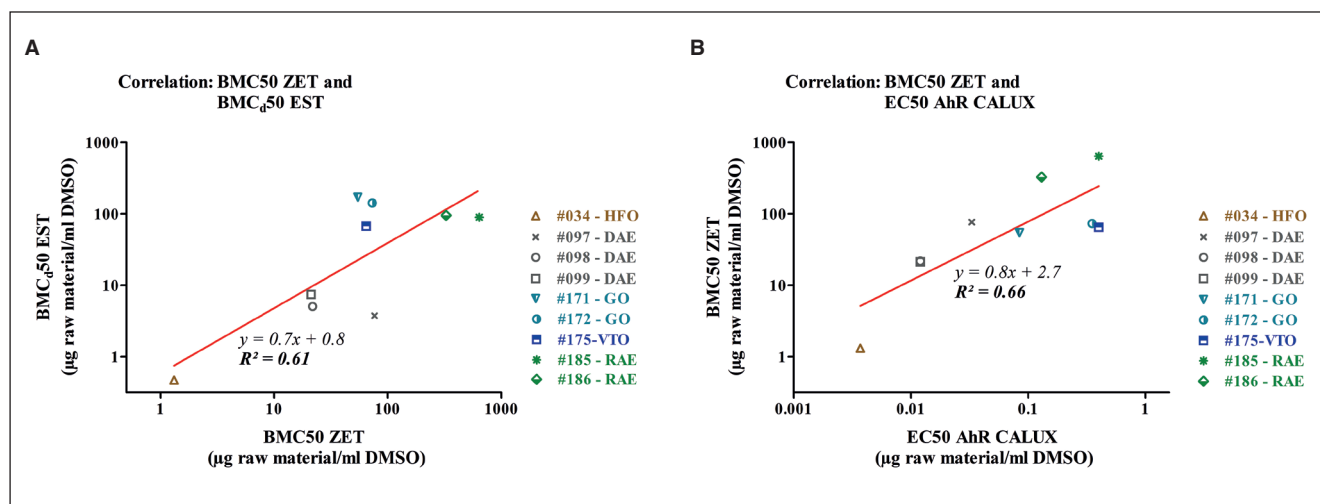
### 3.3 Correlation between *in vitro* developmental toxicity potency in the ZET and potencies observed *in vivo*

Linear regression analysis was performed to see whether a correlation exists between the obtained *in vitro* potencies in the ZET and published *in vivo* developmental toxicity data of the selected PS samples. For this purpose, the BMC50s-ZET were plotted against the BMD10 values for increased resorptions, number of live fetuses/litter, and fetal body weight endpoints. As depicted in Figure 5A, the *in vitro* potencies in the ZET correlate best with the developmental effect reflected by the number of live fetuses/litter ( $R^2 = 0.91$ ), and also correlate well with the increased resorptions endpoint ( $R^2 = 0.80$ ). However, a very poor correlation was obtained when plotting the BMC50s-ZET with the BMD10s of the fetal body weight endpoint, resulting in an  $R^2$  of 0.35 (Fig. 5A). To sum up, good correlations exist between *in vitro* potencies in the ZET and potencies observed *in vivo*, giving an average  $R^2$  for *in vitro-in vivo* correlation (from three different PDT endpoints) of 0.69.

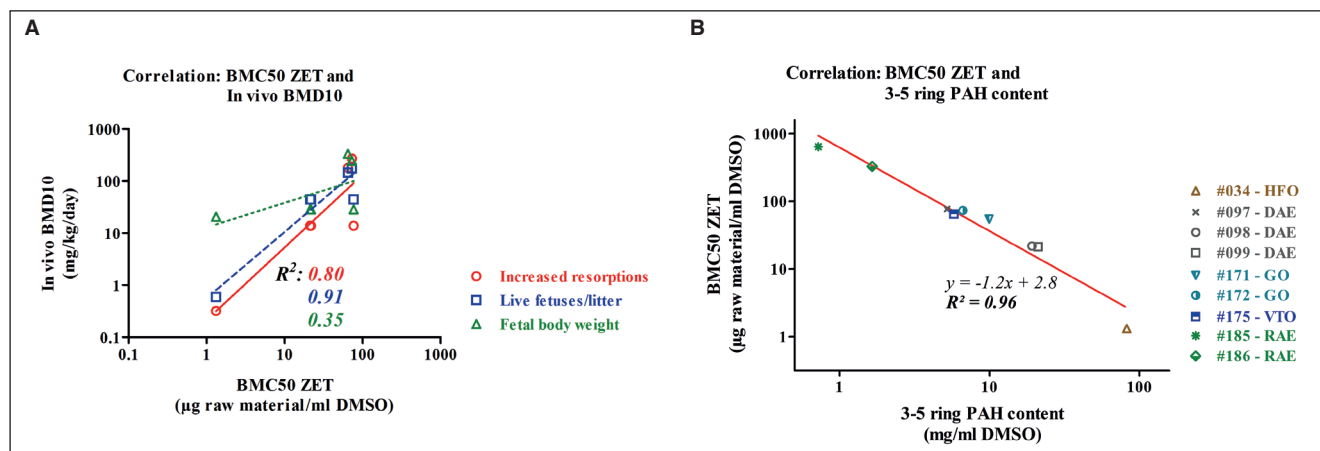
### 3.4 Relation of *in vitro* developmental toxicity potency in the ZET to specific PAH content present in each PS sample

The *in vitro* potencies obtained from the ZET, expressed as BMC50s, were also compared to the amount and type of PAHs in each of the PS samples to see if there is any relation between the observed *in vitro* effects and their PAH content. Data correlations were made with the amount of 2- to 7-ring PAHs present in PS samples, in total giving 15 combinations/assay for which the correlation between PAH content and BMC50s were determined. The results and details for all of these correlation analyses are provided in Appendix F<sup>1</sup>. The goodness-of-fit of correlation analysis is expressed in the  $R^2$  value and the most relevant correlation with





**Fig. 4: Correlation between (A) developmental toxicity potency of PS in the ZET and the EST; (B) developmental toxicity potency of PS in the ZET, expressed as BMC50s, and agonist activity of the same substances in the AhR CALUX assay, expressed as EC50s** Data of the PS in the EST and AhR CALUX assay were taken from our previous published studies Kamelia et al. 2017 and 2018, respectively. HFO, heavy fuel oil; DAE, distillate aromatic extract; GO, gas oil; VTO, vacuum tower overhead oil; RAE, residual aromatic extract; GTLb, gas-to-liquid base oil; GTLg, gas-to-liquid gas oil



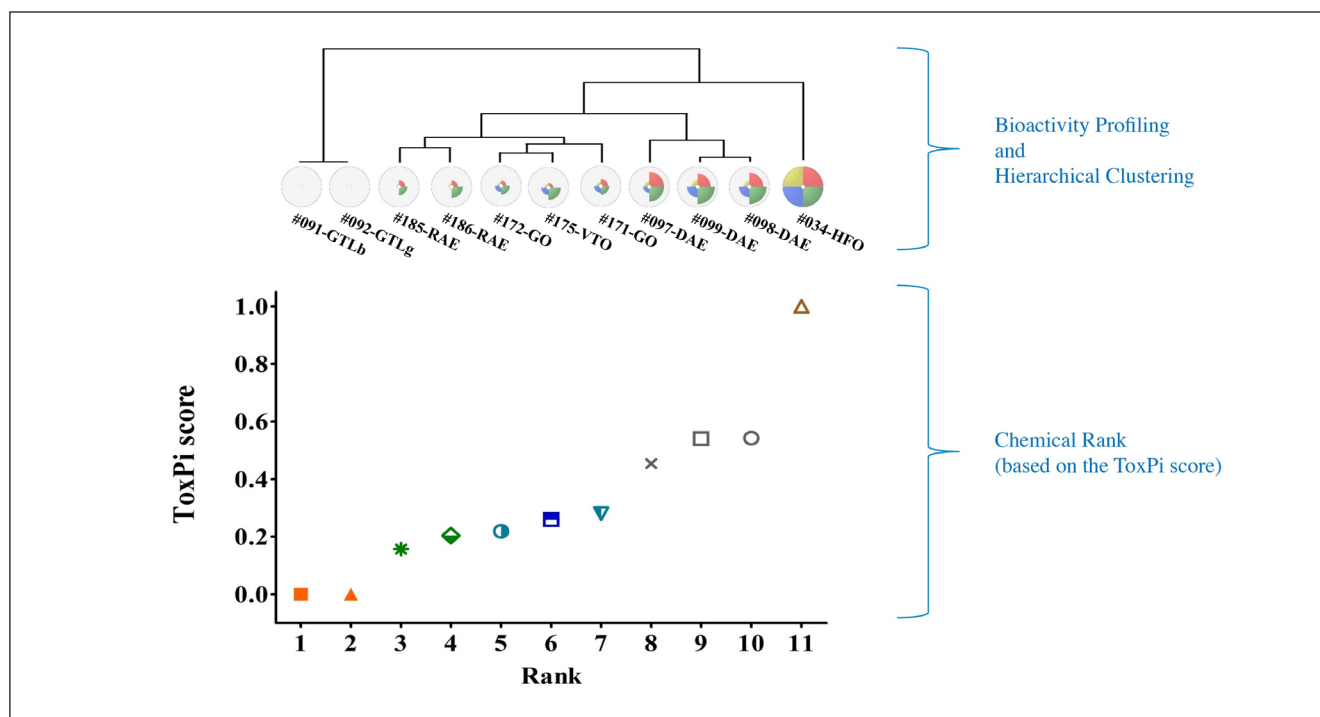
**Fig. 5: (A) Correlation between BMC50 values, obtained from the ZET, and *in vivo* BMD10 values based on three different developmental toxicity endpoints: increased incidence of resorptions, number of live fetuses/litter, and fetal body weight; (B) Correlation between PDT potencies in the ZET (BMC50s) and 3-5 ring PAH content present in each PS sample** BMD10: the benchmark dose for 10% decrease in fetal body weight (BMD10-fetal body weight), 10% decrease in number of live fetuses/litter (BMD10-live fetuses/litter), and a 10% additional incidence of resorptions (BMD10-increased resorptions), as determined using the BMD software (US-EPA); BMC50: the benchmark concentration at which the fraction of GMS is reduced by 50% in the ZET or EST. HFO, heavy fuel oil; DAE, distillate aromatic extract; GO, gas oil; VTO, vacuum tower overhead oil; RAE, residual aromatic extract; GTLb, gas-to-liquid base oil; GTLg, gas-to-liquid gas oil; GMS, general morphology scoring; hpf, hours post fertilization; BMD, benchmark dose; BMC, benchmark concentration

the highest  $R^2$  is presented in Figure 5B. The *in vitro* potencies in the ZET were best correlated to the amount of 3- to 5-ring PAHs present in the corresponding PS sample ( $R^2 = 0.96$ ; Fig. 5B).

### 3.5 Bioactivity profiling and grouping of the DMSO-extracts of PS and GTL products

Figure 6 shows the bioactivity profiling (ToxPi pie-charts) and grouping of the DMSO-extracts of 9 PS and 2 GTL products

based on the data obtained from the three different *in vitro* assays: ZET, EST, and AhR CALUX assay, and the chemical analysis (PAH content) of the respective test compound. In the resulting ToxPi pie-charts (Fig. 6; top), each slice is associated with a specific variable (bioassay/chemical analysis data), and the area covered by the slice is proportional to the relative activity (defined by ToxPi score equation, see Section 2.8) of the test compound within this data set. The red slice in the pie-chart represents data



**Fig. 6: Bioactivity profiling, hierarchical clustering, and chemical ranking using ToxPi GUI 2.0**

Bioactivity profiling, illustrated by the pie-charts, represents data from combinatorial integration of *in vitro* potencies of the DMSO-extracts of 9 PS and 2 GTL products (expressed either as BMC50s or EC50s) from three different *in vitro* assays: ZET (blue quadrant), EST (red quadrant), and AhR CALUX assay (green quadrant); and data from the chemical analysis for PAH content (yellow quadrant) present in each of these substances. The same data were used as data input for hierarchical clustering and global chemical rank in ToxPi. HFO, heavy fuel oil; DAE, distillate aromatic extract; GO, gas oil; VTO, vacuum tower overhead oil; RAE, residual aromatic extract; GTLb, gas-to-liquid base oil; GTLg, gas-to-liquid gas oil; AhR, aryl hydrocarbon; CALUX, chemical activated luciferase gene expression

from the EST (BMC50 EST; Kamelia et al., 2017), the green slice represents data from the AhR CALUX assay (EC50 AhR CALUX assay; Kamelia et al., 2018), the blue slice reflects data from the ZET (BMC50 ZET), and the yellow slice portrays the level of PAHs (mg/ml) present in each test compound. It is worth mentioning that the pie-chart was developed by giving the same weight to each dataset: EST, AhR CALUX assay, ZET, and PAH content, in each slice of the pie-chart. This means, each dataset has 25% weight in the pie-chart to make a total of 100%. Looking closely at the bioactivity profiles illustrated by the ToxPi pie-charts, similarities in bioactivity profiles of test compounds that belong to the same class of substances were also observed (Fig. 6; top).

ToxPi also derives a global chemical rank based on the cumulative ToxPi score (Fig. 6; bottom). The cumulative ToxPi score was calculated based on the sum and average of all 4 slice areas (see Section 2.8 for the ToxPi score equation and determination), in which each area represents the ToxPi score from one input-variable (EST/AhR CALUX assay/ZET/PAH content). A high ToxPi score reflects a higher activity (a more potent) and a low ToxPi score reflects a lower activity (a less potent) of the corresponding test compound in all of the assays used for the present study. As shown in Figure 6 (bottom), the global chemical rank showed that sample #034-HFO (ToxPi score = 1) was the most active/potent compound, where both of the RAEs (Tox-

Pi score = 0.16-0.20) were the least active test compounds of the present study. Furthermore, the DAEs, GOs, and VTO were ranked in between the HFO and RAE, with DAEs (ToxPi score = 0.45-0.54) being more active/potent compared to the GOs and VTO (ToxPi score = 0.22-0.28). Both of the GTL products showed no effects in any of the assays: EST, AhR CALUX assay, and ZET, while they also contain no PAH (0% PAHs), hence their cumulative ToxPi score was 0. As a consequence, they occupy the two lowest ranks in the global chemical rank. Altogether, in the ToxPi score, albeit based on four endpoints, test compounds were ranked proportionally according to their PAH level from the highest, HFO, to the lowest PAH content, RAEs. The list of cumulative ToxPi scores presented in Figure 6 is provided in Appendix G<sup>1</sup>.

In addition to the bioactivity profiling (ToxPi pie-charts) and global chemical rank also a hierarchical clustering was obtained from ToxPi GUI 2.0. The dendrogram of the hierarchical clustering (Fig. 6; top) reveals that two main clusters of test compounds can be distinguished on the basis of their potencies in the three different *in vitro* assays and their PAH content. The GTL extracts are grouped in one cluster and all PS extracts are grouped in another cluster (Fig. 6; top). Within the cluster of PS extracts, test compounds that belong to the same class of substances are grouped together, for example DAEs and RAEs. VTO appears to

be in the same cluster with the GOs, as a result of having similar PAH profiles (Fig. 1) and of inducing similar effects in the ZET, EST, and AhR CALUX assay.

#### 4 Discussion and conclusions

Published *in vivo* studies show that some PS are able to cause PDT in pregnant rats, and it is suspected that this adverse effect is induced by specific constituents present in these substances, mainly 3- to 7-ring PAHs (Feuston et al., 1994; Murray et al., 2013). The present study evaluates the applicability of the ZET to assess PDT potency of the DMSO-extracts of 9 PS and 2 GTL products. Our results indicate that all PS extracts, varying in PAH level and content, were able to inhibit the development of zebrafish embryos in a concentration-dependent manner and this potency is associated with their 3-5 ring PAH content. On the contrary, DMSO-extracts of both GTL products, with no aromatics, showed no effects in the ZET. This points to a major role of PAHs present in the PS extracts used in the study in causing the PDT effects observed in the ZET.

Prominent developmental aberrations, specifically pericardial and yolk sac edemas, dorsal curvature (only in the hatched embryo at 96 hpf), and cumulative mortality, were observed in the zebrafish embryo exposed to PS extracts in addition to the effects scored using the extended GMS. This particular observation is believed to be compound-specific, or only induced by PAHs and PAH-containing PS. These findings are in line with what has been reported in the literature, where exposure to individual PAHs and mixtures of PAHs caused similar developmental effects (Goodale et al., 2013; Gu et al., 2010; Huang et al., 2012; Zhang et al., 2012; Wincent et al., 2015; Geier et al., 2018). For example, studies by Incardona et al. (2004, 2006, 2011) showed that some 3-ring (fluorene, phenanthrene, dibenzothiophene), 4-ring (pyrene, benz[a]anthracene), and 5-ring PAHs (benzo[a]pyrene, benzo[k]fluoranthene), and mixtures of PAHs cause severe pericardial and yolk sac edemas, dorsal curvature of the trunk and tail, and growth retardation in the exposed-zebrafish embryo at 48-96 hpf. In addition, these authors also concluded that the relative amount of specific PAH constituents is more important than the total PAH content in a PAH mixture for the observed developmental effects in the exposed zebrafish embryo. In the present study, it seems like both specific and total PAH content play an important role in mediating the PDT induced by some PS.

The ZET, which was performed in this study, is the third assay of the test battery that we have applied to study the PDT potency (and modes-of-action) of a similar series of PS extracts. Hence, it is now possible to compare the results in the ZET with the results obtained from the other *in vitro* assays, i.e., the EST and the AhR CALUX assay. The main goal of this cross-model comparison is to see whether each of these assays could be used either as a stand-alone assay or be combined to predict the PDT potency of PS extracts. The comparison revealed that PDT potencies of the DMSO-extracts of the 9 PS and 2 GTL products in the ZET and EST were moderately correlated ( $R^2 = 0.61$ ). This is in line with results reported by de Jong et al. (2011), who found

a moderate correlation ( $R^2 = 0.57$ ) of potencies in the ZET and EST when assessing the embryotoxicity potency of azoles.

A moderate correlation also exists between potencies in the ZET and the AhR-mediated activity of the same substances ( $R^2 = 0.66$ ), pointing at a possible role of the AhR in mediating the observed PDT in the ZET. It is generally accepted that the PDT induced by PAHs is partly mediated via activation of the AhR (Puga et al., 2005; Goodale et al., 2013). Similar observations of developmental effects, such as pericardial and yolk sac edemas, dorsal curvature, and increased mortality rates, have previously been found to occur upon exposure of zebrafish embryos to PAHs and PAH-containing extracts, which supports these effects to be typical phenotypic effects of AhR-mediated embryotoxicity in zebrafish embryos (Billiard et al., 2006; Incardona, 2004, 2006; Wincent et al., 2015). A lower correlation between potencies in the ZET and AhR-mediated activity ( $R^2 = 0.66$ ), compared to one obtained between the EST and AhR-mediated activity ( $R^2 = 0.80$ ; Kamelia et al., 2018), could be explained by the fact that not all individual PAHs induce PDT via the AhR pathway. For instance, pyrene, a 4-ring PAH, is known to cause developmental defects in zebrafish embryos via ion transportation and homeostasis pathways, not the AhR, while benzo[a]pyrene (5-ring PAHs) induces embryotoxicity via the AhR pathway (Huang et al., 2012; Goodale et al., 2013).

Interestingly, a different potency ranking is seen when comparing the relative potencies obtained in the ZET (results from the current study) with the EST (Kamelia et al., 2017). One of the prominent disparities is that *in vitro* PDT potencies among the classes of PS in the EST differed by at least 1-order of magnitude (10 times), where this is not the case for the ZET. In the ZET, it is evident that all PS extracts were able to induce concentration-dependent developmental effects, however, there is no wide separation of potency among some classes of PS extracts, for example between the GOs and the DAEs. In the ZET, the relative PDT potency of GOs differed only 3-3.5-fold from that of the DAEs, while in the EST, their BMCd50s varied at least 25-fold. Also, the RAEs appear to be the least potent substances to induce PDT in the ZET whilst they were ranked as the second last in the EST. This discrepancy could be attributed to the fact that the observed effects in both assays might be induced by a different group of PAH constituents, involving different molecular pathways for the observed PDT. Furthermore, the disparity of PDT potencies between ZET and EST might also be explained by the difference in complexity between the two assays, the compound kinetics, and/or the experimental set-up of the assay itself. For example, a saltwater-based medium is used for the ZET and a serum-based medium is used for the EST, which may affect the solubility and bioavailability of test compounds in the assay medium (Fischer et al., 2017) throughout the test. In spite of these differences in potency ranking, the overall correlation between the assays was still moderate with an  $R^2 = 0.61$ . Another reason for the differences between the ZET and EST might be related to the fact that zebrafish embryos are metabolically active, due to the cytochrome P450 (CYP)-related activity during organogenesis (Verbueken et al., 2017), and to some extent able to transform the PAHs present in PS extracts into their reactive metabolites. This suggests that the potency obtained in the ZET reflects the PDT effects in the presence of a biotransformation system, where this is not the case for the EST. Such a role of metabolic activation in the PDT



of PAHs and PAH-containing mixtures remains to be established/investigated in further detail.

Comparing the BMC50s, it appears that the BMC50 values obtained for the EST span a wider range than those of the ZET. The PDT potency of heavy PS, i.e., HFO and two of the most similar DAEs, #098 and #099, is higher in the EST (lower BMC50) than in the ZET, while in the ZET the obtained BMC50s of light PS, such as GOs and VTO, are lower than those quantified in the EST (see Tab. 1). This result may be associated with the role of hydrophobicity of particular PAHs present in the PS extracts under study, which eventually affect their bioavailability in the assay medium throughout the test. The hydrophobicity of PAHs rises with the molecular mass (Sverdrup et al., 2002), which means the LMW PAHs are more soluble in aqueous medium than the HMW PAHs. Heavy PS of the present study comprise mainly HMW PAHs (4-7 ring PAHs), while the light PS contain more LMW PAHs (2-3 ring PAHs). Hence, it is suspected that the effects induced by light PS of the present study in the ZET are due to the role of these LMW PAHs. This is supported by the fact that sample #097-DAE does not result in a proportionate potency in the ZET, compared to the other 2 DAEs (#098 and #099), which could partly be explained by the difference in aromatic ring class (ARC) profile and PAH content among these samples, but not by their total PAH content (Fig. 1). Sample #097-DAE mainly consists of 5-7 ring PAHs, where the other two DAEs contain 3-5 ring PAHs. This supports the above-mentioned postulation, where PS extracts that contain a high amount of 3-5 ring PAHs, including samples #034-HFO, #098-DAE, #099-DAE, result in lower BMC50s compared with those containing only 5-7 ring PAHs, like the RAEs. So, the fact that the RAEs are low in total PAH content (1.5-3.3 wt.%) and consist of mainly 5-7 ring PAHs explains why both of them are the least potent substances in the ZET.

The discrepancy between potencies in the ZET and EST may also be related to the presence of a protective envelope, called the chorion, prior to hatching of the zebrafish embryo in the ZET. The chorion of the zebrafish embryo acts as a permeability barrier during development, which could limit the uptake of certain toxicants by the embryo (Henn and Braunbeck, 2011). A recent publication by Geier et al. (2018), using the dechorionated zebrafish embryo model, demonstrated that HMW PAHs are significantly more developmentally toxic than LMW PAHs. Hence, the role of the chorion for assessing the PDT potency of PAH-containing PS and whether that will result in different outcomes remains to be investigated further.

PDT potencies obtained in the ZET were also compared to *in vivo* data. The correlations with potencies reported in *in vivo* studies were lower for the ZET ( $R^2 = 0.69$ ) than previously reported for the EST ( $R^2 = 0.85$ ; Kamelia et al., 2017) or the AhR CALUX ( $R^2 = 0.80$ ) (Kamelia et al., 2018). It may be considered that the EST (using mouse ES cells) and the AhR CALUX assay (using rat H4IIE.luc cells) are ontogenetically closer to rats (*in vivo* data) than to fish, which may be the reason for a better *in vivo-in vitro* correlation. It should be noted that the better correlation between the mammalian-based *in vitro* and *in vivo* assays, as seen here, refers to PAH-containing PS. Whether this observation would also hold for other types of developmental toxicants remains to be established. The ZET uses a complete embryo and covers a wider se-

ries of events during embryogenesis, but, as previously mentioned, the hydrophobicity of a specific group of PAHs, mainly the HMW PAHs, in the egg water may limit the uptake of the PS extracts by the zebrafish embryo throughout the test. As a consequence, the ZET might not quantitatively cover the whole range of PDT potencies among PS extracts, thereby, correlating less with the *in vivo* data, in which the BMD10 values between different classes of PS extracts varied by at least 1-order of magnitude (see Tab. 1). Lastly, the differences between potencies in the ZET and *in vivo* findings may also be explained by the differences in toxicokinetics, developmental stage of the embryos used in the assay, and complexity of the respective assay itself. Taking possible differences in toxicokinetics between *in vivo* and *in vitro* models into account may further improve the *in vitro-in vivo* correlation.

Results of the present study show that PDT potency of the PS extracts in the ZET is well-correlated with their level of 3-5 ring PAHs. Available evidence from the published literature confirms our findings that only some 3- to 5-ring PAHs, including phenanthrene, pyrene, benz[a]anthracene, benzo[a]pyrene, may cause developmental effects in the zebrafish embryo (Incardona et al., 2004, 2006, 2011; Goodale et al., 2013). Hence, it is possible that the observed developmental retardations, as seen in the ZET in the present study, were induced by a specific group of PAH constituents, and not all of them. Another possible explanation why potencies in the ZET are well-correlated with especially the 3-5 ring PAH content may be related to the hydrophobicity feature of these and > 5 ring PAHs influencing their solubility in the egg water medium used for the ZET. It is known that individual PAHs with a  $\log K_{ow} > 5$  (Sverdrup et al., 2002; Lu et al., 2008), including 5- to 7-ring PAHs like BaP, benzo[g,h,i]perylene, benzo[b]fluoranthene, indeno[1,2,3-c,d]pyrene, and coronene have solubility limitations in aqueous based medium. This means, the solubility of these 5- to 7-ring PAH constituents, present in the PS extracts under study, in the egg water medium may be limited, limiting the bioavailability and uptake and thus the developmental toxicity of these PAHs in the ZET.

Combining the results of the ZET with those reported for the EST (Kamelia et al., 2017), AhR CALUX assay (Kamelia et al., 2018), and the PAH content, ranked and clustered the test compounds in line with their *in vitro* potencies and chemical characteristics. Hierarchical clustering grouped PS that belong to the same class of substances together. Moreover, there is a clear separation of cluster between the PS and GTL products, showing an opposite bioactivity of substances with and without PAHs. Altogether, ToxPi analysis successfully integrated and visualized the multivariate data obtained from distinct informative domains, thereby clustering and differentiating between the PS, within and among classes, and also between PS and GTL extracts, in a way that is in line with the differences in their PAH composition and their bioactivity in several *in vitro* assays for PDT testing. This result suggests the applicability of the present assay battery to group similar substances.

The use of the BMD approach to estimate and determine *in vitro* concentrations (BMC) that reflect the potencies obtained from different *in vitro* assays for developmental toxicity testing, including the EST, WEC, and ZET, has been widely accepted (Piersma et al., 2008, 2013; Hermesen et al., 2011; de Jong et al.,



2011). In the present study, we applied the BMD approach to determine both *in vivo* (BMDs) and *in vitro* (BMCs) PDT potencies of the test compounds under study, thereby allowing us to compare both potencies derived using the same approach. Nonetheless, additional analysis of the *in vitro* data using GraphPad Prism software resulted in EC50s in the ZET that were comparable and correlate well with the calculated BMC50s ( $R^2 = 0.99$ ), resulting in similar results and conclusions (data are provided in Appendix I<sup>1</sup>).

Several published studies (Hermsen et al., 2011; Selderslaghs et al., 2012) used the so-called teratogenicity index (TI) for classifying teratogenic compounds using the ZET. The TI, within the ZET, is defined as a ratio between the 50% lethal concentration (LC50; based on the GMS) and the 50% effect concentration (EC50; based on teratogenic endpoints) (Selderslaghs et al., 2012; Beekhuijzen et al., 2015). According to the publications by Beekhuijzen et al. (2015) and Selderslaghs et al. (2009, 2012), 6 teratogenic endpoints/effects are scored from 96 hpf to 144 hpf to define the EC50. PAH-containing PS tested in the present study induced mainly pericardial and yolk sac edemas when scored at 96 hpf. However, given that most of the zebrafish exposed to PS did not hatch at 72-96 hpf, full scoring of these teratogenic endpoints was seriously hampered, although deformed body shape and malformation of the tail (i.e., kinked tail) were observed in several of the zebrafish embryos that hatched at 72-96 hpf. The other 2 endpoints, including malformation of the head and malformation of the otoliths, were not observed upon exposure to PAH-containing PS extracts for up to 96 hpf in the ZET, suggesting that they may not represent effects following exposure to PAH-containing substances. As a result, for the PS samples tested, the EC50 for teratogenic endpoints could not be derived and thus also the TI could not be established. These observed effects are in line with those of previous studies where prominent developmental aberrations upon exposure to PAHs and PAH-mixture are limited to pericardial edema, yolk sac edema, and cumulative mortality (Goodale et al., 2013; Gu et al., 2010; Huang et al., 2012; Zhang et al., 2012; Wincent et al., 2015; Geier et al., 2018). In view of this, one may argue about the suitability of the ZET performed in the present study to evaluate possible PDT potency of PAH-containing PS and also about how the morphological effects relate to mortality upon exposure to these substances. Based on Wilson's Principles of Teratology (Finnel, 1992), embryoletality (i.e., cumulative mortality), growth retardation (i.e., unhatched zebrafish embryos), and functional deficit (i.e., pericardial and yolk sac edemas), as seen in the ZET results of the present study, are all manifestations of developmental toxicity. Further, the *in vivo* developmental effects induced by some PAH-containing PS, especially those tested in the present study, are mainly associated with increased incidence of resorptions (prenatal loss) and decreased fetal body weight, and not so much with malformations of the fetuses (ARCO, 1993; Feuston et al., 1989, 1994, 1996; Feuston and Mackerer, 1996; Hoberman et al., 1995; Mackerer et al., 2003). In this case, given that a good correlation exists between the obtained PDT potency in the ZET and the *in vivo* BMD10 values (particularly for live fetuses/litter and increased resorptions endpoints; Fig. 5A) for developmental toxicity, this validates the ZET for detection of developmental toxicity of the PAH-containing PS under study.

To conclude, here we show the applicability of the ZET as an *in vitro* alternative assay to predict the relative *in vivo* PDT potency of DMSO-extracts of PS and GTL products. Our findings confirm the hypothesis that PAH are the major inducers of PDT in PS; for PS it seems that the EST is a better predictive model for *in vivo* PDT, but it should be taken into account that this may not be the case for other classes of compounds. In addition, the ZET does not outperform the EST as a stand-alone assay for testing PDT of PS, but the ZET is a useful addition to the battery of *in vitro* tests able to predict the *in vivo* PDT of PS. Zebrafish embryos are able to bioactivate some parent compounds thanks to the expression of cytochrome P450 enzymes, allowing evaluation of PDT potency of PS extracts in the presence of a biotransformation system. PAHs and PAH-containing PS require metabolic activation to exert their genotoxicity effect, but it is still unknown whether bioactivation is essential to induce PDT. PS used in the present study are categorized as UVCBs or substances of unknown and variable composition, hence the observed PDT in both *in vivo* and *in vitro* studies may be caused by a wide range of compounds and underlying mechanisms. Having said that, it can be foreseen that it is almost impossible to assess the PDT potency of these substances by using just one single assay. The use of a battery of *in vitro* assays that focuses on the relevant modes-of-action of PDT by some PS will answer the challenge for PDT testing of these substances. Upon addition of the ZET, our testing battery of PDT consists of three *in vitro* assays: the EST, ZET, and AhR CALUX assay. The results of the present study revealed that such a battery enables prediction of the PDT potency and possible underlying mechanisms of PS. Future studies will focus on gene expression studies to see whether these may provide additional value to the current testing battery, and if so to what extent. It would also be interesting to study the PDT potency of individual PAHs present as major constituents in the DMSO-extracts of PS samples to fully understand and unravel the association between PAHs in PS and PDT.

## References

- ARCO (1993). Developmental toxicity (embryo-fetal toxicity and teratogenic potential) study of F-193 administered percutaneously to CrI:CD®BRK VAF/Plus® presumed pregnant rats. Report ATX-92-0011.
- Billiard, S. M., Timme-Laragy, A. R., Wassenberg, D. M. et al. (2006). The role of the aryl hydrocarbon receptor pathway in mediating synergistic developmental toxicity of polycyclic aromatic hydrocarbons to zebrafish. *Toxicol Sci* 92, 526-536. doi:10.1093/toxsci/kfl011
- Beekhuijzen, M., de Koning, C., Flores-Guillén, M-E. et al. (2015). From cutting edge to guideline: A first step in harmonization of the zebrafish embryotoxicity test (ZET) by describing the most optimal test conditions and morphology scoring system. *Reprod Toxicol* 56, 64-76. doi:10.1016/j.reprotox.2015.06.050
- Blackburn, G. R., Deitch, R. A., Schreiner, C. A. and Mackerer, C. R. (1986). Predicting carcinogenicity of petroleum distillation fractions using a modified Salmonella mutagenicity assay. *Cell Biol Toxicol* 2, 63-84. doi:10.1007/BF00117708
- Boogaard, P. J., Carrillo, J. C., Roberts, L. G. and Whale, G. F.



- (2017). Toxicological and ecotoxicological properties of gas-to-liquid (GTL) products. 1. Mammalian toxicology. *Crit Rev Toxicol* 47, 121-144. doi:10.1080/10408444.2016.1214676
- Brannen, K. C., Panzica-Kelly, J. M., Danberry, T. L. and Augustine-Rauch, K. A. (2010). Development of a zebrafish embryo teratogenicity assay and quantitative prediction model. *Birth Defects Res B Dev Reprod Toxicol* 89, 66-77. doi:10.1002/bdrb.20223
- Busquet, F., Strecker, R., Rawlings, J. M. et al. (2014). OECD validation study to assess intra- and inter-laboratory reproducibility of the zebrafish embryo toxicity test for acute aquatic toxicity testing. *Regul Toxicol Pharmacol* 69, 496-511. doi:10.1016/j.yrtph.2014.05.018
- Clonfero, E., Nardinim B., Machioro, M. et al. (1996). Mutagenicity and contents of polycyclic aromatic hydrocarbons in used and recycled motor oils. *Mutat Res* 368, 283-291. doi:10.1016/S0165-1218(96)90070-1
- Concawe (1994). The Use of the Dimethyl Sulphoxide (DMSO) Extract by the IP 346 Method as an Indicator of the Carcinogenicity of Lubricant Base Oils and Distillate Aromatic Extracts. Report no 94/51. <https://www.concawe.eu/wp-content/uploads/2017/01/rpt9451ocr-2005-00417-01-e.pdf>
- de Jong, E., Louisse, J., Verwei, M. et al. (2009). Relative developmental toxicity of glycol ether alkoxy acid metabolites in the embryonic stem cell test as compared with the in vivo potency of their parent compounds. *Toxicol Sci* 110, 117-124. doi:10.1016/j.taap.2011.03.014
- de Jong, E., Barenys, M., Hermesen, S. A. et al. (2011). Comparison of the mouse embryonic stem cell test, the rat whole embryo culture and the zebrafish embryotoxicity test as alternative methods for developmental toxicity testing of six 1,2,4-triazoles. *Toxicol Appl Pharmacol* 253, 103-111. doi:10.1016/j.taap.2011.03.014
- Dimopoulou, M., Verhoef, A., Gomes, C. A. et al. (2018). A comparison of the embryonic stem cell test and whole embryo culture assay combined with the BeWo placental passage model for predicting the embryotoxicity of azoles. *Toxicol Lett* 286, 10-21. doi:10.1016/j.toxlet.2018.01.009
- Dunster, J. (2014). GTL Diesel (Distillates (Fischer-Tropsch), C8-26 Branched and Linear): Oral (Gavage) Pre-natal Development Toxicity Study in the Rat. Report 41301744.
- ECHA (European Chemicals Agency) (2009). Information requirements for repeated dose toxicity and reproductive toxicity-substances over 100 (and 1000) tonnes. [https://echa.europa.eu/documents/10162/13644/reach\\_factsheet\\_testing\\_en.pdf](https://echa.europa.eu/documents/10162/13644/reach_factsheet_testing_en.pdf) (accessed 18.05.2018).
- EU (2010). Directive 2010/63/EU of the European parliament and of the council of 22 September 2010 on the protection of animals used for scientific purposes. *J Eur Comm L* 276, 34-79.
- Feuston, M. H., Kerstetter, S. L., Singer, E. J. and Mehlmán, M. A. (1989). Developmental toxicity of clarified slurry oil applied dermally to rats. *Toxicol Ind Health* 3, 587-599. doi:10.1177/074823378900500313
- Feuston, M. H., Low, L. K., Hamilton, C. E. and Mackerer, C. R. (1994). Correlation of systemic and developmental toxicities with chemical component classes of refinery streams. *Toxicol Sci* 22, 622-630. doi:10.1093/toxsci/22.4.622
- Feuston, M. H. and Mackerer, C. R. (1996). Developmental toxicity of clarified slurry oil, syntower bottoms, and distillate aromatic extract administered as a single oral dose to pregnant rats. *J Toxicol Environ Health* 49, 45-66. doi:10.1080/009841096160989
- Feuston, M. H., Hamilton, C. E. and Mackerer, C. R. (1996). Systemic and developmental toxicity of dermally applied distillate aromatic extract in rats. *Fundam Appl Toxicol* 30, 276-284. doi:10.1006/faat.1996.0065
- Feuston, M. H., Hamilton, C. E., Schreiner, C. A. and Mackerer, C. R. (1997). Developmental toxicity of dermally applied crude oils in rats. *J Toxicol Environ Health* 52, 79-93. doi:10.1080/00984109708984054
- Finnell, R. H. (1992). Teratology: General considerations and principles. *J Allergy Clin Immunol* 103, S337-S342. doi:10.1016/S0091-6749(99)70259-9
- Fischer, F. C., Henneberger, L., König, M. et al. (2017). Modeling exposure in the Tox21 in vitro bioassays. *Chem Res Toxicol* 30, 1197-1208. doi:10.1021/acs.chemrestox.7b00023
- Geier, M. C., Chlebowski, A. C., Truong, L. et al. (2018). Comparative developmental toxicity of a comprehensive suite of polycyclic aromatic hydrocarbons. *Arch Toxicol* 92, 571-586. doi:10.1007/s00204-017-2068-9
- Genschow, E., Spielmann, H., Scholz, G. et al. (2002). The ECVAM international validation study on in vitro embryotoxicity tests: Results of the definitive and evaluation of prediction models. *Altern Lab Anim* 30, 151-176.
- Genschow, E., Spielmann, H., Scholz, G. et al. (2004). Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* 32, 209-244.
- Glaberman, S., Padilla, S. and Barron, M. G. (2017). Evaluating the zebrafish embryo toxicity test for pesticide hazard screening. *Environ Toxicol Chem* 36, 1221-1226. doi:10.1002/etc.3641
- Gray, T. M., Simpson, B. J., Nicolich, M. J. et al. (2013). Assessing the mammalian toxicity of high-boiling petroleum substances under the rubric of the HPV program. *Regul Toxicol Pharmacol* 67, S4-S9. doi:10.1016/j.yrtph.2012.11.014
- Goodale, B. C., Tilton, S. C., Wilson, G. et al. (2013). Structurally distinct polycyclic aromatic hydrocarbons induce differential transcriptional responses in developing zebrafish. *Toxicol Appl Pharmacol* 272, 656-670. doi:10.1016/j.taap.2013.04.024
- Grimm, F. A., Iwata, Y., Sirenko, O. et al. (2016). A chemical-biological similarity-based grouping of complex substances as a prototype approach for evaluating chemical alternatives. *Green Chem* 18, 4407-4419. doi:10.1039/c6gc01147k
- Gu, H., Wang, X., Bai, C. et al. (2010). Phenanthrene and fluorene-mediated early development toxicity in zebrafish. *Fresenius Environ Bull* 19, 57-62.
- Haber, L. T., Dourson, M. L., Allen, B. C. et al. (2018). Benchmark dose (BMD) modeling: Current practice, issues, and challenges. *Crit Rev Toxicol* 48, 387-415. doi:10.1080/10408444.2018.1430121
- Hawliczek, A., Nota, B., Cenijn, P. et al. (2012). Developmental toxicity and endocrine disrupting potency of 4-azapyrene, benzo[b]fluorene and retene in the zebrafish *Danio rerio*. *Reprod Toxicol* 33, 213-223. doi:10.1016/j.reprotox.2011.11.001
- Henn, K. and Braunbeck, T. (2011). Dechoriation as a tool to

- improve the fish embryo toxicity test (FET) with the zebrafish (*Danio rerio*). *Comp Biochem Physiol C Toxicol Pharmacol* 153, 91-98. doi:10.1016/j.cbpc.2010.09.003
- Hermesen, S. A. B., van den Brandhof, E.-J., van der Ven, L. T. M. and Piersma, A. H. (2011). Relative embryotoxicity of two classes of chemicals in a modified zebrafish embryotoxicity test and comparison with their in vivo potencies. *Toxicol In Vitro* 25, 745-753. doi:10.1016/j.tiv.2011.01.005
- Hill, A., Jones, M., Dodd, A. and Diekmann, H. (2011). A review of developmental toxicity screening using zebrafish larvae. *Toxicol Lett* 205S, S60-S179. doi:10.1016/j.toxlet.2011.05.413
- Hoberman, A. M., Christian, M. S., Lovre, S. et al. (1995). Developmental toxicity study of clarified slurry oil (CSO) in the rat. *Fundam Appl Toxicol* 28, 34-40. doi:10.1006/faat.1995.1143
- Huang, L., Wanga, C., Zhang, Y. et al. (2012). Benzo[a]pyrene exposure influences the cardiac development and the expression of cardiovascular relative genes in zebrafish (*Danio rerio*) embryos. *Chemosphere* 87, 369-375. doi:10.1016/j.chemosphere.2011.12.026
- Incardona, J. P., Collier, T. K. and Scholz, N. L. (2004). Defects in cardiac function precede morphological abnormalities in fish embryos exposed to polycyclic aromatic hydrocarbons. *Toxicol Appl Pharmacol* 196, 191-205. doi:10.1016/j.taap.2003.11.026
- Incardona, J. P., Day, H. L., Collier, T. K. and Scholz, N. L. (2006). Developmental toxicity of 4-ring polycyclic aromatic hydrocarbons in zebrafish is differentially dependent on AH receptor isoforms and hepatic cytochrome P4501A metabolism. *Toxicol Appl Pharmacol* 217, 308-321. doi:10.1016/j.taap.2006.09.018
- Incardona, J. P., Linbo, T. L. and Scholz, N. L. (2011). Cardiac toxicity of 5-ring polycyclic aromatic hydrocarbons is differentially dependent on the aryl hydrocarbon receptor 2 isoform during zebrafish development. *Toxicol Appl Pharmacol* 257, 242-249. doi:10.1016/j.taap.2011.09.010
- Jomaa, B., Hermesen, S. A., Kessels, M. Y. et al. (2014). Developmental toxicity of thyroid-active compounds in a zebrafish embryotoxicity test. *ALTEX* 31, 303-317. doi:10.14573/altex.1402011
- Kamelia, L., Louisse, J., de Haan, L. et al. (2017). Prenatal developmental toxicity testing of petroleum substances: Application of the mouse embryonic stem cell test (EST) to compare in vitro potencies with potencies observed in vivo. *Toxicol In Vitro* 44, 303-312. doi:10.1016/j.tiv.2017.07.018
- Kamelia, L., Louisse, J., de Haan, L. et al. (2018). The role of endocrine and dioxin-like activity of extracts of petroleum substances in developmental toxicity as detected in a panel of CALUX reporter gene assays. *Toxicol Sci* 164, 576-591. doi:10.1093/toxsci/kfy114
- Kari, G., Rodeck, U. and Dicker, A. P. (2007). Zebrafish: An emerging model system for human disease and drug discovery. *Clin Pharmacol Ther* 82, 70-80. doi:10.1038/sj.clpt.6100223
- Kuske, B., Pulyanina, P.Y. and Nieden, N. I. (2012). Embryonic stem cell test: Stem cell use in predicting developmental cardiotoxicity and osteotoxicity. In C. Harris and J. Hansen (eds.), *Developmental Toxicology. Methods in Molecular Biology (Methods and Protocols)*. Volume 889. Totowa, NJ, USA: Humana Press. doi:10.1007/978-1-61779-867-2\_10
- Li, H., Rietjens, I. M., Louisse, J. et al. (2015). Use of the ES-D3 cell differentiation assay, combined with the BeWo transport model, to predict relative in vivo developmental toxicity of anti-fungal compounds. *Toxicol In Vitro* 29, 320-328. doi:10.1016/j.tiv.2014.11.012
- Louisse, J., Gönen, S., Rietjens, I. M. and Verwei, M. (2011). Relative developmental toxicity potencies of retinoids in the embryonic stem cell test compared with their relative potencies in in vivo and two other in vitro assays for developmental toxicity. *Toxicol Lett* 203, 1-8. doi:10.1016/j.toxlet.2011.02.012
- Lu, G. N., Tao, X. Q., Dang, Z. et al. (2008). Estimation of n-octanol/water partition coefficients of polycyclic aromatic hydrocarbons by quantum chemical descriptors. *Cent Eur J Chem* 6, 210-318. doi:10.2478/s11532-008-0010-y
- Mackerer, C. R., Griff, L. C., Grabowski Jr, J. S. and Reitman, F. A. (2003). Petroleum mineral oil refining and evaluation of cancer hazard. *Appl Occup Environ Hyg* 18, 890-901. doi:10.1080/10473220390237467
- Marvel, S. W., To, K., Grimm, F. A. et al. (2018). ToxPi graphical user interface 2.0: Dynamic exploration, visualization, and sharing of integrated data models. *BMC Bioinformatics* 19, 80. doi:10.1186/s12859-018-2089-2
- Mobil (1989). Developmental Toxicity Study in Rats Exposed Dermally to Vacuum Tower Overheads. Mobil Environmental and Health Sciences Laboratory (Report 62328).
- Murray, F. J., Roth, R. N., Nicolich, M. J. et al. (2013). The relationship between developmental toxicity and aromatic-ring class profile of high-boiling petroleum substances. *Regul Toxicol Pharmacol* 67, S46-S59. doi:10.1016/j.yrtph.2013.05.003
- OECD – Organization for Economic Co-operation and Development (2001). OECD Guideline for Testing of Chemicals 414: Prenatal Developmental Toxicity Study. [https://ntp.niehs.nih.gov/iccvam/suppdocs/feddocs/oecd/oecd\\_gl414.pdf](https://ntp.niehs.nih.gov/iccvam/suppdocs/feddocs/oecd/oecd_gl414.pdf) (accessed 18.05.2018).
- OECD (2011a). Validation Report (Phase 1) for the Zebrafish Embryo Toxicity Test. *Series on Testing and Assessment No. 157*. OECD, Paris, France. <http://www.oecd.org/env/ehs/testing/48572244.pdf> (accessed 18.05.2018).
- OECD (2011b). Validation Report (Phase 2) for the Zebrafish Embryo Toxicity Test. *Series on Testing and Assessment No. 179*. OECD, Paris, France. <http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282012%2925&doclanguage=en> (accessed 18.05.2018).
- Pampanin, D. M. and Sydnese, M. O. (2013). Polycyclic aromatic hydrocarbons a constituent of petroleum: Presence and influence in the aquatic environment. *Intech* 5, 83-113.
- Piersma, A. H., Genschow, E., Verhoef, A. et al. (2004). Validation of the post implantation rat whole-embryo culture test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* 32, 275-307.
- Piersma, A. H., Janer, G., Wolterink, G. et al. (2008). Quantitative extrapolation of in vitro whole embryo culture embryotoxicity data to developmental toxicity in vivo using the benchmark dose approach. *Toxicol Sci* 101, 91-100. doi:10.1093/toxsci/kfm253
- Piersma, A. H., Bosgra, S., van Duursen, M. B. et al. (2013). Evaluation of an alternative in vitro test battery for detecting re-





- productive toxicants. *Reprod Toxicol* 38, 53-64. doi:10.1016/j.reprotox.2013.03.002
- Puga, A., Tomlinson, C. R. and Xia, Y. (2005). Ah receptor signals cross-talk with multiple developmental pathways. *Biochem Pharmacol* 69, 199-207. doi:10.1016/j.bcp.2004.06.043
- Reif, D. M., Martin, M. T., Tan, S. W. et al. (2010). Endocrine profiling and prioritization of environmental chemicals using ToxCast data. *Environ Health Perspect* 118, 1714-1720. doi:10.1289/ehp.1002180
- Rohwedel, J., Guan, K., Hegert, C. and Wobus, A. M. (2001). Embryonic stem cells as an in vitro model for mutagenicity, cytotoxicity and embryotoxicity studies: Present state and future prospects. *Toxicol In Vitro* 15, 741-753. doi:10.1016/S0887-2333(01)00074-1
- Rovida, C., Longo, F. and Rabbit, R. R. (2011). How are reproductive toxicity and developmental toxicity addressed in reach dossiers? *ALTEX* 28, 273-294. doi:10.14573/altex.2011.4.273
- Roy, T. A., Johnson, S. W., Blackburn, G. R. and Mackerer, C. R. (1988). Correlation of mutagenic and dermal carcinogenic activities of mineral oils with polycyclic aromatic compound content. *Fundam Appl Toxicol* 10, 466-476. doi:10.1016/0272-0590(88)90293-X
- Scholz, S., Fischer, S., Gündel, U. et al. (2008). The zebrafish embryo model in environmental risk assessment – Applications beyond acute toxicity testing. *Environ Sci Pollut Res* 15, 394-404. doi:10.1007/s11356-008-0018-z
- Seiler, A. E. and Spielmann, H. (2011). The validated embryonic stem cell test to predict embryotoxicity in vitro. *Nat Protoc* 6, 961-978. doi:10.1038/nprot.2011.348
- Selderslaghs, I. W., Van Rompay, A. R., Coen, W. D. and Wittersa, H. E. (2009). Development of a screening assay to identify teratogenic and embryotoxic chemicals using the zebrafish embryo. *Reprod Toxicol* 28, 308-320. doi:10.1016/j.reprotox.2009.05.004
- Selderslaghs, I. W., Blust, R. and Witters, H. E. (2012). Feasibility study of the zebrafish assay as an alternative method to screen for developmental toxicity and embryotoxicity using a training set of 27 compounds. *Reprod Toxicol* 33, 142-154. doi:10.1016/j.reprotox.2011.08.003
- Senn, C. (2014). GTL Base Oil Distillate (Distillates (Fischer-Tropsch), Heavy, C18-50-Branched, Cyclic and Linear): Prenatal Developmental Toxicity Study in the Han Wistar Rat. Report D80072.
- Sipes, N. S., Martin, M. T., Reif, D. M. et al. (2011). Predictive models of prenatal developmental toxicity from ToxCast high-throughput screening data. *Toxicol Sci* 124, 109-127. doi:10.1093/toxsci/kfr220
- Speight, J. G. (2006). *The Chemistry and Technology of Petroleum*. 4<sup>th</sup> edition. Boca Raton, USA: CRC Press, Taylor & Francis Group.
- Spielmann, H., Genschow, E., Brown, N. A. et al. (2004). Validation of the rat limb bud micromass test in the international ECVAM validation study on three in vitro embryotoxicity tests. *Altern Lab Anim* 32, 245-274.
- Spielmann, H. (2009). The way forward in reproductive/developmental toxicity testing. *Altern Lab Anim* 37, 641-656.
- Strähle, U., Scholz, S., Geisler, R. et al. (2012). Zebrafish embryos as an alternative to animal experiments – A commentary on the definition of the onset of protected life stages in animal welfare regulations. *Reprod Toxicol* 33, 128-132. doi:10.1016/j.reprotox.2011.06.121
- Strikwold, M., Woutersen, R., Spenkelink, B. et al. (2012). Relative embryotoxic potency of p-substituted phenols in the embryonic stem cell test (EST) and comparison to their toxic potency in vivo and in the whole embryo culture (WEC) assay. *Toxicol Lett* 213, 235-242. doi:10.1016/j.toxlet.2012.07.005
- Sverdrup, L. E., Nielsen, T. and Krogh, P. H. (2002). Soil ecotoxicity of polycyclic aromatic hydrocarbons in relation to soil sorption, lipophilicity, and water solubility. *Environ Sci Technol* 36, 2429-2435. doi:10.1021/es010180s
- Tsitou, P., Heneweer, M. and Boogaard, P. J. (2015). Toxicogenomics in vitro as an alternative tool for safety evaluation of petroleum substances and PAHs with regard to prenatal developmental toxicity. *Toxicol In Vitro* 29, 299-307. doi:10.1016/j.tiv.2014.11.005
- van der Jagt, K., Munn, S., Tørsløv, J. and de Bruijn, J. (2004). Alternative Approaches Can Reduce the Use of Test Animals under REACH. European Commission Report EUR 21405EN.
- Verbueken, E., Pype, C., Bars, C. et al. (2017). Localization of cytochrome P450 activity in the zebrafish embryo and larva. *Reprod Toxicol* 72, 31-52. doi:10.1016/j.reprotox.2017.06.140
- Wincent, E., Jonsson, M. E., Bottai, M. et al. (2015). Aryl hydrocarbon receptor activation and developmental toxicity in zebrafish in response to soil extracts containing unsubstituted and oxygenated PAHs. *Environ Sci Technol* 49, 3869-3877. doi:10.1021/es505588s
- Zhang, Y., Wang, C., Huang, L. et al. (2012). Low-level pyrene exposure causes cardiac toxicity in zebrafish (*Danio rerio*) embryos. *Aquat Toxicol* 114-115, 119-124. doi:10.1016/j.aquatox.2012.02.022

### Conflict of interest

LK, SB, LH, and IMCMR declare that they have no conflict of interest.

Peter J. Boogaard is employed by Shell International, a member company of Concawe, and chairman of the toxicology group of Concawe. Hans B. Ketelslegers is science executive for health at Concawe. Both Prof. Boogaard and Dr Ketelslegers are totally free (by contract) to freely design and conduct research and express their own scientific opinion without any obligation towards either Shell or Concawe. The current findings are not intended to constitute any product endorsement.

### Acknowledgements

This work was supported by Concawe (Grant number: 201506110).