		Number of reads		Number of mapped reads		
	Sample	Original	Trimmed	unique paired end	multiple paired end	sum of mapped reads
DMSO	D1	8,168,404	8,152,495	5,254,278 (64.45%)	1,920,049 (23.55%)	7,632,045 (93.62%)
DMSO	D2	9,957,843	9,939,550	6,796,802 (68.38%)	2,066,186 (20.79%)	9,412,813 (94.70%)
DMSO	D3	10,786,595	10,768,858	7,323,091 (68.00%)	2,314,672 (21.49%)	10,233,729 (95.03%)
DMSO	D4	11,799,398	11,777,525	8,055,836 (68.40%)	2,496,782 (21.20%)	11,191,119 (95.02%)
DMSO	D5	14,501,290	14,457,774	9,602,482 (66.42%)	3,175,671 (21.97%)	13,578,935 (93.92%)
PCB 0.3nM	P0.3-9	13,635,225	13,612,736	8,937,657 (65.66%)	3,112,057 (22.86%)	12,757,987 (93.72%)
PCB 0.3nM	P0.3-11	13,674,864	13,640,724	9,334,595 (68.43%)	2,863,975 (21.00%)	12,977,478 (95.14%)
PCB 0.3nM	P0.3-12	11,830,503	11,806,900	8,038,530 (68.08%)	2,579,996 (21.85%)	11,269,978 (95.45%)
PCB 0.3nM	P0.3-13	13,968,933	13,939,004	9,631,230 (69.10%)	2,831,576 (20.31%)	13,279,248 (95.27%)
PCB 0.3nM	P0.3-14	12,536,363	12,507,078	8,589,984 (68.68%)	2,552,200 (20.41%)	11,890,055 (95.07%)
PCB 0.3nM	P0.3-16	12,210,513	12,189,670	8,214,909 (67.39%)	2,711,764 (22.25%)	11,558,361 (94.82%)
PCB 10nM	P10-17	11,824,432	11,802,545	8,155,563 (69.10%)	2,421,282 (20.51%)	11,224,550 (95.10%)
PCB 10nM	P10-18	9,949,807	9,923,567	6,819,261 (68.72%)	2,022,738 (20.38%)	9,424,059 (94.97%)
PCB 10nM	P10-21	10,983,355	10,948,290	7,531,889 (68.80%)	2,226,969 (20.34%)	10,403,421 (95.02%)
PCB 10nM	P10-23	12,880,370	12,854,923	8,907,689 (69.29%)	2,591,122 (20.16%)	12,242,973 (95.24%)
PCB 10nM	P10-24	11,786,444	11,759,002	8,215,663 (69.87%)	2,306,159 (19.61%)	11,224,951 (95.46%)
Average		11,905,896	11,880,040	8,088,091 (68.08%)	2,512,075 (21.15%)	11,268,856 (94.86%)

Table 1: #reads - one read consists of two pairs or fragments; unique/multiple paired end means both fragments are mapped either unique or multiple; "sum of mapped reads" includes additional options, e.g. only one fragment is unique/multiple but not the other; the percentages in "Number of mapped reads" are in relation to the number of trimmed reads;

1 Results

1.1 DNA methylation profiling

General statistics, see Table 1 and Table 2.

1.2 PCB126-induced Changes in DNA Methylation in Testis

1.3 Very short methods

BAT_toolkit with metilene for DMR detection. GREAT for finding corresponding genes and GO analysis.

1.4 PCB126 0.3 nM treatment

There is a total of 37 DMRs with 10 hypermethylated and 27 hypomethylated regions. (Two of the hypomethylated ones are on KN* scaffolds and are therefore not covered by the annotations used later.) None of them showed a percent methylation difference of larger than 40%. Three hypomethylated DMRs had a methylation difference larger than 30%. The highest concentration of DMRs is on chromosome 4 with 9 DMRs, 24% of the total amount. 7 of the 9 DMRs are hypomethylated.

Hypermethylated DMRs are significantly enriched for 17 GO molecular function terms. The three most significant ones beeing GO:0008502 melatonin receptor activity, GO:0005242 inward rectifier potassium channel activity and GO:0005249 voltage-gated potassium channel activity. For biological processes there are 46 terms enriched, e.g. GO:0051876 pigment granule dispersal, GO:0051877 pigment granule aggregation in cell center, GO:0051905 establishment of pigment granule localization.

Hypomethylated DMRs are enriched in 42 molecular function terms. The most significant ones beeing GO:0003774 motor activity, GO:0008528 G-protein coupled peptide receptor activity, GO:0001653 peptide receptor activity. For biological processes GO:0006397 mRNA processing, GO:0042310 vasoconstriction and GO:0009132 nucleoside diphosphate metabolic process are the most significant out of 29 terms.

	Sample	#Cytosines	#mCpG	#un-mCpG	Global CpG methyl.level
DMSO	D1	26,250,848	1,531,846	174,245	73.28%
DMSO	D2	29,817,798	1,785,444	229,466	71.57%
DMSO	D3	29,616,929	1,877,577	247,569	71.41%
DMSO	D4	29,830,337	1,966,073	275,366	70.68%
DMSO	D5	31,389,808	2,251,946	316,096	71.20%
PCB 0.3nM	P0.3-9	31,312,106	2,141,677	279,279	71.28%
PCB 0.3nM	P0.3-11	30,940,958	2,129,553	297,305	71.17%
PCB 0.3nM	P0.3-12	41,028,708	1,937,522	228,690	67.79%
PCB 0.3nM	P0.3-13	31,374,908	2,146,969	299,559	71.14%
PCB 0.3nM	P0.3-14	31,118,080	2,080,073	282,497	71.08%
PCB 0.3nM	P0.3-16	30,420,319	1,940,661	291,059	69.71%
PCB 10nM	P10-17	32,027,515	1,906,940	261,658	69.45%
PCB 10nM	P10-18	30,103,549	1,826,707	237,354	70.88%
PCB 10nM	P10-21	31,013,858	1,953,848	253,455	71.06%
PCB 10nM	P10-23	32,226,790	2,043,756	276,857	70.03%
PCB 10nM	P10-24	31,731,766	1,902,211	261,037	70.25%
Average		31,262,767	1,963,925	263,218	70,75%

Table 2: Bisulfite statistics

1.5 PCB126 10 nM treatment

There is a total of 92 DMRs with 80 hypomethylated and 12 hypermethylated regions. (Six of the hypomethylated ones are on KN* scaffolds and are therefore not covered by the annotations used later.) 13% hypomethylated and 17% hypermethylated show a percent methylation difference larger than 40%. The highest concentration of DMRs is on chromosome 4 with 34 DMRs, 37% of the total amount. 31 of the 34 DMRs are hypomethylated.

Hypermethylated DMRs in the testis showed significant enrichment of a number of GO molecular function terms. There are 42 in total, the three most significant ones beeing GO:0070851 growth factor receptor binding, GO:0008094 DNA-dependent ATPase activity, GO:0005343 organic acid:sodium symporter activity. For biological processes there are 79 terms, e.g. GO:0006753 nucleoside phosphate metabolic process, GO:0007166 cell surface receptor signaling pathway, GO:0044700 single organism signaling.

Hypomethylated DMRs are enriched in 36 molecular function terms. The most significant ones are GO:0043167 ion binding; GO:0003676 nucleic acid binding and GO:0046872 metal ion binding. For biological processes GO:0043631 RNA polyadenylation; GO:0016226 iron-sulfur cluster assembly; GO:0016338 calcium-independent cell-cell adhesion are the most significant of 34 different GO terms.

1.6 PCB126-induced Transcriptional Changes in Testis

1.6.1 Very short methods

DESeq2 with standard parameters. DAVID (https://david.ncifcrf.gov/) for GO analysis.

1.6.2 Results

We obtained an average of 26.1 million reads mapping to ENSEMBL genes in the DMSO-treated control sample. The libraries of individuals treated with PCB126 0.3 nM or 10 nM resulted in 26.6 and 23.1 million reads.

The gene Cyp1a is upregulated in the PCB126 0.3 nM and 10 nM treatment by a log2 fold change of 7.6 and 8.9 respectively. In both cases the adjusted p-value is below 0.05.

There were a total of 767 and 4,708 DEGs in the 0.3 nM and 10 nM treatment with an adjusted p-value < 0.5. Among the 767 DEGs in the 0.3 nM treatment, 458 were upregulated and 309 were downregulated. Among the upregulated genes 214 (46.7%) and in the downregulated genes 144 (46.6%), as well, had a fold change of more than 2.

Biological Process - Upregulated	
Term	Adj. p-value
GO:0009605 response to external stimulus	0.0002
GO:0042330 taxis	0.0028
GO:0042221 response to chemical	0.0024
GO:0006950 response to stress	0.0155
GO:0006955 immune response	0.0367
KEGG - Downregulated	
Term	Adj. p-value
dre04512:ECM-receptor interaction	0.0085
dre04350:TGF-beta signaling pathway	0.0050

Table 3: GO terms of the PCB126 0.3 nM treatment. GOTERM_BP_2, GOTERM_MF_2 and KEGG pathway. Only the five best significant ones (adj. p-value <0.05) are shown. MF and KEGG for upregulated as well as BP and MF for downregulated genes had no significant enrichments.

Biological Process - Upregulated	
Term	Adj. p-value
GO:0006955 immune response	< 0.0001
GO:0009605 response to external stimulus	< 0.0001
GO:0050900 leukocyte migration	< 0.0001
GO:0042330 taxis	< 0.0001
GO:0042221 response to chemical	< 0.0001
Molecular Function - Upregulated	
Term	Adj. p-value
GO:0016491 oxidoreductase activity	< 0.0001
GO:0004129 cytochrome-c oxidase activity	0.0015
GO:0005515 protein binding	0.0052
GO:0030246 carbohydrate binding	0.0469
GO:0030234 enzyme regulator activity	0.0488
KEGG - Downregulated	
Term	Adj. p-value
dre00190:Oxidative phosphorylation	< 0.0001
dre04060:Cytokine-cytokine receptor interaction	< 0.0001
dre04630:Jak-STAT signaling pathway	0.0008

Table 4: GO terms of the PCB126 10 nM treatment. GOTERM_BP_2, GOTERM_MF_2 and KEGG pathway. Only the five best significant (adj. p-value <0.5) ones are shown. Downregulated genes had no significant enrichments.

The upregulated genes were enriched in GO terms such as response to external stimulus and taxis (both biological process). The downregulated genes were only enriched in two KEGG pathways ECM-receptor interaction and TGF-beta signaling pathway, see Table 3 for details.

The PCB126 10 nM exposure resulted in the differential expression of 4,708 genes. Among these 2,822 genes were upregulated and 1,886 genes were downregulated. Among the upregulated genes 1,534 (54.4%) and in the downregulated genes 324 (17.2%) had a fold change of more than 2.

The upregulated genes of the PCB126 10 nM treatment are enriched in similar GO terms compared to the lighter PCB126 treatment, For example immune response and response to external stimulus. But opposite to the 0.3 nM treatment they are also enriched in molecular function GO terms, e.g. oxidoreductase activity and cytochrome-c oxidase activity but not in KEGG pathways. Downregulated genes of the 10 nM treatment were only enriched in three KEGG pathways Oxidative phosphorylation, Cytokine-cytokine receptor interaction and Jak-STAT signaling pathway, see Table 4 for details.

DMR ID	Gene ID	Difference of mean methylation rates per group	log2 fold change of Gene Expression
DMR_3	ENSDARG00000028661	0.14	0.78
DMR_14	ENSDARG00000103318	-0.34	0.39
DMR_15	ENSDARG00000103318	-0.22	0.39
DMR_2	ENSDARG00000052037	-0.20	-3.56

Table 5: DMRs with DEGs and methylation difference as well as expression log2 fold change.

DMR ID	Gene ID	Difference of mean methylation rates per group	log2 fold change of Gene Expression
DMR_2	ENSDARG00000030289	0.54	-0.45
DMR_12	ENSDARG00000005482	0.22	-0.47
DMR_6	ENSDARG00000005185	-0.24	3.10
DMR_34	ENSDARG00000070845	-0.18	1.89
DMR_70	ENSDARG00000070845	-0.18	1.89
DMR_35	ENSDARG00000069311	-0.20	1.51
DMR_69	ENSDARG00000089382	-0.20	1.21
DMR_10	ENSDARG00000015472	-0.23	1.16
DMR_1	ENSDARG00000069996	-0.40	0.90
DMR_3	ENSDARG00000052361	-0.13	0.87
DMR_11	ENSDARG00000102824	-0.21	0.75
DMR_73	ENSDARG00000036567	-0.21	0.61
DMR_17	ENSDARG00000103318	-0.16	0.41
DMR_18	ENSDARG00000103318	-0.47	0.41
DMR_22	ENSDARG00000020730	-0.20	-0.55
DMR_22	ENSDARG00000044718	-0.20	-0.58
DMR_30	ENSDARG00000013312	-0.20	-0.60

Table 6: DMRs with DEGs and methylation difference as well as expression log2 fold change.

1.7 Relationship between DMRs and Altered Gene Expression

1.8 Very short methods

DMRs defined using metilene, corresponding gene using GREAT. Subsequently, they were searched in differentially expressed genes in the DESeq2 output.

1.9 PCB126 0.3 nM treatment

One hypermethylated DMR corresponds to an upregulated gene while three hypomethylated DMRs correspond to two upd and one down-regulated gene, for details see Table 5 (Please note: the DMR IDs between both treatments are not the same).

1.10 PCB126 10 nM treatment

Two hypermethylated DMRs have an corresponding gene which is significiantly upregulated. In both cases the gene is downregulated. 15 hypomethylated DMRs correspond to an differentially regulated gene. 12 of these genes are upregulated while 3 are downregulated. The fold change ranges from 1.3 up to 8.6, for details see Table 6 (Please note: the DMR IDs between both treatments are not the same).

1.11 Problem with Gene Annotation

For some genes the gene symbol used by GREAT can not be found in our annotation file 'GRCz10.90.genes.gff'. Sometimes they can be identified as "Previous names" in ZFIN but not always.

PCB126 0.3 nM treatment, hypermethylated DMRs: 2 gene names edited, one not found (zgc:110222). Hypomethylated DMRs 2 gene names edited, six not found (zgc:175264, si:dkey-63b1.1, chst13, trim63, cxcr7b, pc).

PCB126 10 nM treatment, hypermethylated DMRs: 1 gene name edited; onenot found (maf). Hypomethylated DMRs 4 gene names edited, 18 genes not found (si:ch211-202h22.10; tsc2; zgc:175264; si:dkey-63b1.1; zgc:112437; si:dkey-4c15.4; si:dkey-51d8.7; si:dkey-78o7.1; si:dkey-77p23.2; zgc:123060; si:dkey-45d16.8; si:dkey-10p5.9; pc; rab35; zgc:112433; rhoad; slc16a1; tns1).

Todo: Should we try other ways to find corresponding gene names?

1.12 Reanalysis of Aluru et al., 2018

Todo: Do mapping of brain and liver again. Still running...