Archives of Toxicology
https://doi.org/10.1007/s00204-019-02560-0

ORGAN TOXICITY AND MECHANISMS

Comparative toxicoproteogenomics of mouse and rat liver identifies TCDD-resistance genes

Stephenie D. Prokopec[1] · Aileen Lu[1,2] · Sandy Che-Eun S. Lee[1] · Cindy Q. Yao[1] · Ren X. Sun[1] · John D. Watson[1] · Rabah Soliymani[3] · Richard de Borja[1] · Ada Wong[4] · Michelle Sam[4] · Philip Zuzarte[4] · John D. McPherson[4] · Allan B. Okey[2] · Raimo Pohjanvirta[5,6] · Paul C. Boutros[1,2,7,8,9,10,11]

Received: 26 June 2019 / Accepted: 2 September 2019

Abstract

The aryl hydrocarbon receptor (AHR) mediates many toxic effects of 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD). However, the AHR alone does not explain the widely different outcomes among organisms. To identify the other factors involved, we evaluated three transgenic mouse lines, each expressing a different rat AHR isoform (rWT, DEL, and INS) providing widely different resistance to TCDD toxicity, as well as C57BL/6 and DBA/2 mice which exhibit a ~ tenfold divergence in TCDD sensitivity (exposures of 5-1000 µg/kg TCDD). We supplement these with whole-genome sequencing, together with transcriptomic and proteomic analyses of the corresponding rat models, Long–Evans (L–E) and Han/Wistar (H/W) rats (having a ~ 1000-fold difference in their TCDD sensitivities; 100 µg/kg TCDD), to identify genes associated with TCDD-response phenotypes. Overall, we identified up to 50% of genes with altered mRNA abundance following TCDD exposure are associated with a single AHR isoform (33.8%, 11.7%, 5.2% and 0.3% of 3076 genes altered unique to rWT, DEL, C57BL/6 and INS respectively following 1000 µg/kg TCDD). Hepatic *Pxdc1* was significantly repressed in all three TCDD-sensitive animal models (C57BL/6 and rWT mice, and L–E rat) after TCDD exposure. Three genes, including *Cxxc5*, *Sugp1* and *Hgfac*, demonstrated different AHRE-1 (full) motif occurrences within their promoter regions between rat strains, as well as different patterns of mRNA abundance. Several hepatic proteins showed parallel up- or downward alterations with their RNAs, with three genes (SNRK, IGTP and IMPA2) showing consistent, strain-dependent changes. These data show the value of integrating genomic, transcriptomic and proteomic evidence across multi-species models in toxicologic studies.

Keywords Model organisms · Whole-genome sequencing · Transcriptomics · Proteomics · TCDD · AhR

Introduction

TCDD (2,3,7,8-tetrachlorodibenzo-*p*-dioxin) is a member of the dioxin class of environmental pollutants. It is a persistent, highly lipophilic compound that can be created as a by-product during production of some herbicides and through the incineration of chlorine-containing compounds (Von Burg 1988). TCDD toxicity impacts almost all organ systems in mammals, with effects ranging from chloracne (particularly in humans) to immunosuppression, wasting syndrome, hepatotoxicity and acute lethality (Kransler et al. 2007). There are large differences in the lethality of TCDD, both between and within a given species. Two rodent species, hamster and guinea pig, show roughly 5000-fold difference in sensitivity to TCDD toxicity, with LD_{50} values of 1157–5051 µg/kg and 0.6–2 µg/kg, respectively (Kransler et al. 2007). The DBA/2 mouse

Stephenie D. Prokopec, Aileen Lu, and Sandy Che-Eun S. Lee are contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00204-019-02560-0) contains supplementary material, which is available to authorized users.

✉ Allan B. Okey
allan.okey@utoronto.ca

✉ Raimo Pohjanvirta
raimo.pohjanvirta@helsinki.fi

✉ Paul C. Boutros
pboutros@mednet.ucla.edu

Extended author information available on the last page of the article

Published online: 11 September 2019

strain exhibits a ~ tenfold lower TCDD responsiveness compared with the C57BL/6 strain to a wide variety of biochemical and toxic impacts of TCDD (reviewed in Pohjanvirta 2011). For example, the LD_{50} values for male DBA/2 and C57BL/6 mice are 2570 and 180–305 µg/kg respectively (Chapman and Schiller 1985; Pohjanvirta et al. 2012). Perhaps the most dramatic example of intraspecies differences in TCDD-susceptibility exists between the TCDD-resistant Han/Wistar (*Kuopio*; H/W) rat, which has an LD_{50} of > 9600 µg/kg TCDD and the TCDD-sensitive Long–Evans (*Turku/AB*; L–E) rat, whose LD_{50} values are 9.8 µg/kg for females and 17.7 µg/kg for males (Pohjanvirta et al. 1993).

Structural features of the aryl hydrocarbon receptor (AHR) play a major role in the diversity of TCDD-induced toxicities across species and strains. The AHR is a ligand-dependent transcription factor in the PER-ARNT-SIM (PAS) superfamily, which is evolutionarily conserved across fish, birds and mammals (Hahn et al. 1997). Normally bound to chaperone proteins including hsp90 and XAP2 in the cytosol, AHR can be activated by binding of ligands to the PAS-B domain, leading to nuclear localization (Rowlands and Gustafsson 1997). Once in the nucleus and free of chaperone proteins, AHR dimerizes with the AHR Nuclear Translocator (ARNT) protein and binds to AHR response elements (AHREs) in the genome, altering transcription of specific target genes (Shen and Whitlock 1992). The AHR also has other, 'non-genomic' actions (reviewed in Matsumura 2009). For example, ligand activation of the AHR leads to increased intracellular Ca^{2+}, kinase activation and induction of Cox-2 transcription to promote a rapid inflammatory response (Dong et al. 2010).

Evidence for involvement of the AHR in TCDD toxicity comes from numerous studies, including species with structural AHR variants and *Ahr* knockout models. Mice lacking the *Ahr* gene are phenotypically (Fernandez-Salguero et al. 1996; Schmidt et al. 1996; Mimura et al. 1997) and biochemically (Tijet et al. 2006; Boutros et al. 2009) unresponsive to TCDD, as are *Ahr*-knockout rats (Harrill et al. 2016). Additionally, differences in the structure of the AHR protein result in a wide range of susceptibilities to dioxin toxicity. For example, the H/W rat is TCDD-resistant primarily due to a point mutation in the transactivation domain of the *Ahr* gene. This creates a cryptic splice site, leading to two distinct protein products [termed the deletion (DEL) and insertion (INS) isoforms] which are shorter than the wild type rat AHR (present in the L–E strain) (Pohjanvirta et al. 1998). Of these, expression of the INS isoform is predominant; however both are expressed in a number of tissues (Moffat et al. 2007). Numerous studies have sought to exploit these genetic differences among strains and species to decipher the mechanisms of dioxin toxicity (Boverhof et al. 2006; Boutros et al. 2008, 2011; Yao et al. 2012).

The AHR is not the only mediator of dioxin toxicities: evaluation of rat lines generated through breeding of H/W and L–E rats suggests involvement of a second gene (termed gene "B") in the extreme resistance of H/W rats to TCDD-induced toxicities (Pohjanvirta 1990; Tuomisto et al. 1999). Line A (Ln-A) rats contain the H/W *Ahr* and demonstrate similar resistance to TCDD-induced lethality as H/W rats; however, they appear to harbour the wild type (L–E) form of an as-yet unidentified gene "B" that is proposed to contribute to the phenotypic response to TCDD. Alternatively, Line B (Ln-B) rats express the wild type *Ahr* along with the H/W form of gene "B", and demonstrate an intermediate LD_{50} of 830 µg/kg TCDD. Finally, Line C (Ln-C) rats do not fall far from L–E rats in sensitivity to TCDD (LD_{50} 20–40 µg/kg), and express wild type forms of both the *Ahr* and gene "B" (Tuomisto et al. 1999). It is thus unclear which responses are due solely to the various AHR and gene "B" isoforms and which are artifacts of the genetic heterogeneity at non-AHR loci among species and strains.

To isolate the molecular and phenotypic effects of different AHR genetic variants, we exploit a transgenic mouse model, termed "AHR-ratonized mice", in which the endogenous AHR of C57BL/6 mice is ablated and either the rat wild-type (rWT), the DEL or INS variant is inserted into the mouse genome (Pohjanvirta 2009), as well as the TCDD-resistant DBA/2 mouse strain. The DEL isoform confers only moderate resistance to TCDD toxicity, relative to the INS isoform (Pohjanvirta 2009). We compare hepatic transcriptomic responses of these transgenic mice to one another and to their corresponding TCDD-sensitive and resistant strains/lines of rat. Results are then supplemented with whole-genome sequencing (WGS) of L–E and H/W rats to further isolate the specific genes responsible (i.e. gene "B") for differential toxicities and to further characterize the mechanism by which TCDD activation of the AHR causes toxicity. Finally, the hepatic responses of these two rat strains to TCDD are contrasted by a proteomics approach. These combined proteomic, transcriptomic and genomic resources provide novel, complementary data on the significance of AHR structural properties as determinants of TCDD sensitivity and on candidate target genes for TCDD toxicity which can be utilized also in future studies.

Results

Experimental design

Our experimental strategy is outlined in Fig. 1 and Supplementary Table 1. We examined TCDD-mediated transcriptional changes associated with various AHR isoforms within animals that have different backgrounds (Experiment #1, EXP1) or identical genetic backgrounds (Experiment #2,

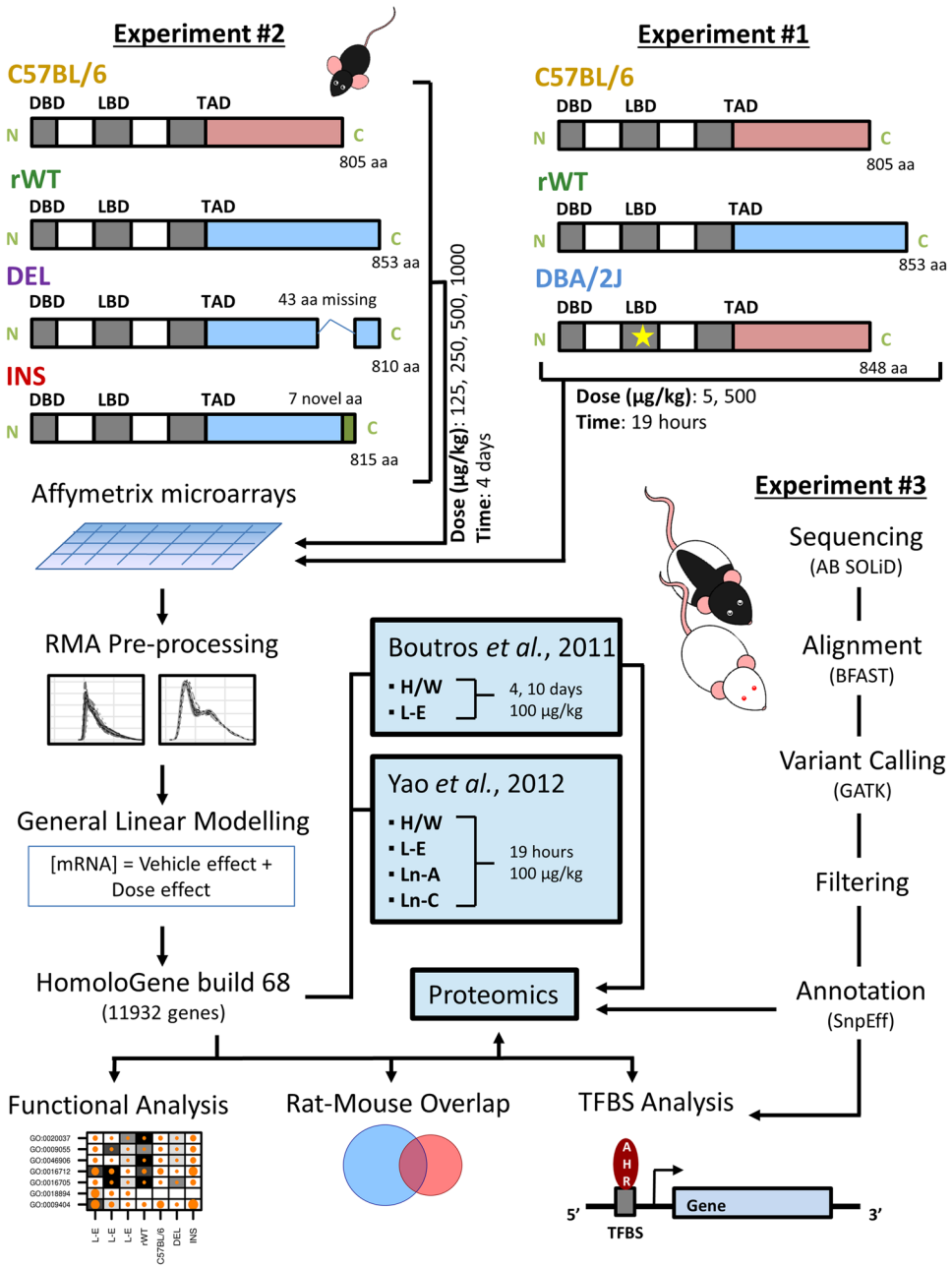

Fig. 1 This study evaluated hepatic transcriptomic profiles of 80 male mice that carry variant AHR isoforms (ratonized mice)—C57BL/6, rWT, DEL and INS—and treated with doses of 0, 125, 250, 500, or 1000 µg/kg TCDD in corn oil vehicle. Livers were excised at 4 days post-exposure and RNA abundance was profiled using microarrays. An additional 36 male mice ($n = 12$ each C57BL/6, rWT and DBA/2 (Ala375Val)) were treated with doses of 0, 5, or 500 µg/kg TCDD in corn oil vehicle and liver tissue collected 19 h post-exposure. Data from two earlier studies that analyzed two strains of rats, Han/Wistar (H/W—INS and DEL) and Long–Evans (rWT), at various times post-treatment were included in the analysis, after filtering for orthologous genes using HomoloGene. Differentially abundant genes were subjected to pathway analysis using GOMiner, transcription factor binding site (TFBS) analysis and overlap visualization of significantly altered genes. Sequencing of genomic DNA isolated from liver of untreated L–E and H/W rats ($n = 2$ each) was performed and processed as shown. SNVs were identified and used to detect novel and/or lost TFBS within each rat strain. Genes containing such sites were further examined for changes to mRNA abundance using the above mentioned microarray data

EXP2). Here we focus on liver as it shows large phenotypic differences between TCDD-resistant and TCDD-sensitive animals following exposure. Moreover, liver exhibits high expression of AHR (Pohjanvirta 2009), thereby making it an appropriate model for determining the effects of AHR isoforms on responses to TCDD. For EXP1, C57BL/6 and DBA/2 mice, along with rWT mice, were treated with a single dose of 0, 5 or 500 µg/kg TCDD in corn oil vehicle, with tissue collected 19 h after. For EXP2, AHR-ratonized (INS/DEL/rWT) and C57BL/6 mice were treated with a single dose of 0, 125, 250, 500 or 1000 µg/kg TCDD in corn oil vehicle. These doses were selected so as to discriminate among the various strains/lines with regard to their overt toxicity responses. The approximate LD_{50} value for male C57BL/6 mice in our laboratory is 305 µg/kg (Pohjanvirta et al. 2012), while male DBA/2 mice are reported to have a LD_{50} of 2570 µg/kg TCDD (Chapman and Schiller 1985). In the case of male rWT, DEL and INS mice, the dose of 500 µg/kg was lethal to 4/6, 2/6 and 0/6 animals respectively (Pohjanvirta 2009). Thus, 5 µg/kg was definitely, and 125 µg/kg probably, sub-lethal to all animals in the present study, whereas the other doses would have been variably fatal over time. The first time point (19 h) should reveal early (and thus most primary) changes in gene expression levels, while

the second (4 days) is the time when histological alterations in the liver are first discernible in rats (Pohjanvirta et al. 1989). Hepatic tissue was collected 4 days following exposure and transcriptional profiling performed. Arrays were pre-processed independently for each group (Supplementary Figures 1–3). Transcriptomic data from 85 animals across 4 rat strains/lines were used for comparison (Boutros et al. 2011; Yao et al. 2012). Genes which demonstrated altered mRNA abundance as well as pathways showing significant enrichment for these genes only in the TCDD-sensitive (C57BL/6, rWT, L–E) or only in TCDD-resistant cohorts (INS, H/W) were identified. Finally for EXP3, WGS was performed on hepatic tissue gDNA extracted from untreated H/W and L–E rats. An average coverage of 85× and 15× was achieved for H/W and L–E respectively (Table 1). Single nucleotide variants (SNVs) were annotated with predicted impact and compared between rat strains (see "Materials and methods"). These datasets were further supplemented with proteomic measurements from H/W rat liver following treatment with 100 µg/kg TCDD or corn oil for 4 days as well as L–E rat liver following treatment with 100 µg/kg TCDD, corn oil vehicle or feed-restriction (Linden et al. 2014), also for 4 days ($n = 5$ animals per group).

Transcriptomic responses to TCDD

TCDD treatment triggered changes in hepatic mRNA abundance in each of the mouse cohorts studied, but with substantial differences in the magnitude, direction and identity of genes affected. Animals treated with corn oil vehicle alone displayed similar transcriptomic profiles regardless of AHR isoform [adjusted Rand index (ARI) = 0.69 for TCDD, control] whereas TCDD-treated animals cluster more closely to animals with the same AHR isoform, regardless of dose (Supplementary Figure 4a, b). Following linear modeling, a p value sensitivity analysis was performed to compare different significance thresholds (Supplementary Figure 4c–i). For downstream analyses, a dual threshold of effect-size ($\log_2$|fold change| > 1) and significance ($p_{adj} < 0.05$) was used to define transcripts with a statistically significantly difference in abundance following TCDD treatment.

Early transcriptomic responses to TCDD differ by Ahr genotype

Previous studies observed transcriptomic changes as early as 6 h in C57BL/6 mouse liver (Prokopec et al. 2015) and 19 h in rat liver (Boutros et al. 2008; Moffat et al. 2010; Yao et al. 2012) following exposure to TCDD. To further study the role that the *Ahr* has in 'early onset' changes, we identified genes with significant differential mRNA abundance following a dose of 5 or 500 µg/kg TCDD in sensitive mouse strains (C57BL/6 or rWT) or resistant (DBA/2) mouse strains. We observed clear trends in response, with the higher dose resulting in an increased number of differentially abundant transcripts; this difference between the two doses was especially pronounced in DBA/2 mice in which only a handful of transcripts changed at 5 µg/kg TCDD (Fig. 2a). The absolute number of changes in transcript abundance was further considerably different among the groups at each dose tested. Intriguingly, liver from the ratonized rWT mouse demonstrated a heightened response relative to both the C57BL/6 and DBA/2 mice, even at 5 µg/kg TCDD. This could suggest differences in binding affinity for either TCDD and/or AHREs of the rWT *Ahr* relative to the mouse *Ahr*. As assessed by a modified sucrose gradient assay and Woolf plot, the apparent binding affinity of AHR for TCDD is quite similar in C57BL/6 mice and L–E rats harboring the WT receptor (Kd 1.8 vs. 2.2 nM respectively), while it is notably lower (16 nM) in the DBA/2 mouse strain (Okey et al. 1989; Pohjanvirta et al. 1999). However, after a lethal dose of TCDD, rWT mice tend to die much more rapidly compared with C57BL/6 mice (Pohjanvirta 2009), which probably bears on the difference in transcript abundance. In further support of this, at 500 µg/kg TCDD a similar number of differentially abundant transcripts was detected in C57BL/6 and DBA/2 mice, but in rWT mice the number was twice as high. We examined the overlap of genes amongst these three groups (Fig. 2b) and identified 34 genes with differential mRNA abundance in all three groups. Five of these are part of the 'AHR-core' gene battery—a set of 11 well-documented TCDD-responsive genes with transcription previously shown to be mediated by the AHR

Table 1 Summary of L–E and H/W rat sequencing. Genomic DNA from hepatic tissue of H/W and L–E rats was sequenced using the AB SOLiD platform

Strain	Average coverage	Number of Variants (pre-filter)	Number of Variants (post-filter)	Number of homozygous SNVs	Number of homozygous Indels	Number of "High Impact" Variants	Number of unique "High Impact" Variants
H/W	85.34	3235191	1138926	176578	127870	756	35 SNVs, 608 indels
L–E	14.73	2004271	1187745	161620	19520	220	116 SNVs, 14 indels

Reads were aligned to the reference (rn6) using BFAST, followed by variant calling using GATK's HaplotypeCaller. Variants were filtered to obtain only novel and unique high-quality variants for each strain, followed by annotation using SnpEff. Final numbers indicate total number of unique [H/W or L–E only, after removal of known variants (Atanur et al. 2013)], homozygous, high impact variants

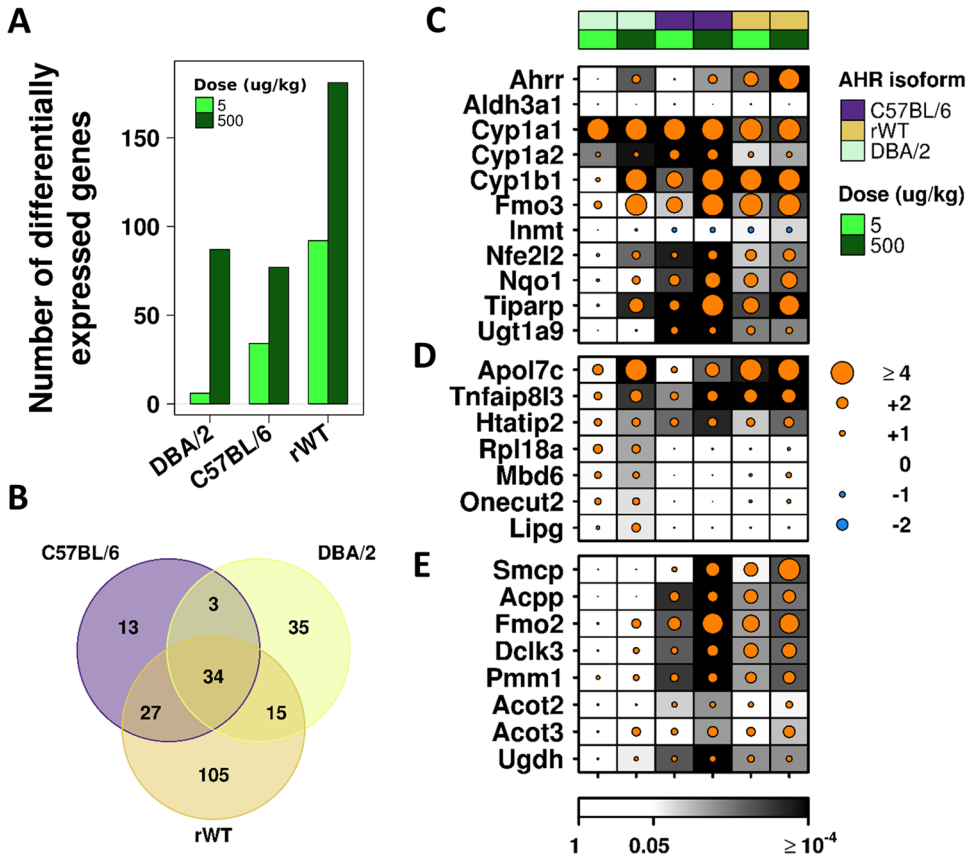

Fig. 2 Differential transcriptomic profiles emerge early following exposure to TCDD. **a** Using a dual threshold of $|\log_2$ fold change$| > 1$, $p_{adj} < 0.05$, genes with differential mRNA abundance were identified. As expected, the TCDD-resistant DBA/2 mouse liver showed a transcriptional response following only the high dose of TCDD, while the sensitive C57BL/6 and rWT strains demonstrated considerable changes following low exposure that increased with dose. **b** Overlap of these genes in each cohort, following exposure to 500 µg/kg TCDD for 19 h. Log_2 fold change of **c** "AHR-core" genes, **d** genes with significantly altered mRNA abundance in resistant mouse liver or **e** sensitive strains. Dot size indicates magnitude of change following exposure to TCDD relative to controls, while colour indicates direction of change (orange = increased abundance, blue = decreased abundance); background shading indicates FDR-adjusted *p* value

(Nebert et al. 1993; Yeager et al. 2009; Deb and Bandiera 2010; Watson et al. 2014) (Fig. 2c). An additional four of these genes had differential RNA abundance in at least two groups: *Cyp1a2* in C57BL/6 and DBA/2 (borderline significant ($\log_2$ fold change = 0.99) in rWT) and *Fmo3*, *Nqo1* and *Ugt1a9* only in the TCDD-sensitive cohorts. *Inmt* shows significant repression in only rWT at this early time point, while *Aldh3a1* shows no response in mice [consistent with previous studies (Lee et al. 2015; Prokopec et al. 2015, 2017)]. We next examined sets of genes which demonstrated altered RNA abundance in the TCDD-resistant DBA/2 mouse liver or TCDD-sensitive cohorts (Fig. 2d–e). Genes with altered mRNA abundance in DBA/2 mice included *Apol7c*, *Tnfaip8l3* and *Htatip2* (both low and high dose exposure). These genes were similarly altered in the sensitive strains. *Rpl18a* and *Mbd6* had altered mRNA abundance exclusively in the resistant group (high and low dose), while two additional genes, *Onecut2* and *Lipg* had altered mRNA abundance exclusively in the resistant mouse, high dose group (Fig. 2d). Alternatively, Fig. 2e highlights differentially abundant transcripts that appear exclusively in sensitive animals, regardless of dose (including *Acpp*, *Dclk3*, *Fmo2*, *Pmm1* and *Ugdh*) and genes that respond in only sensitive strains but at a higher dose (such as *Acot2*, *Acot3* and *Smcp*). Pathway analysis suggested that genes with altered RNA abundance in DBA/2 mice are involved in lipase activity, while the sensitive strains both demonstrate an enrichment of genes involved in xenobiotic and flavonoid metabolic processes as well as myristoyl- and palmitoyl-CoA hydrolase and oxidoreductase activities.

Late transcriptomic responses to TCDD in ratonized mice

We next sought to identify changes in transcriptomic patterns that occur late following exposure to TCDD. Four days after exposure to TCDD, a considerable difference in the hepatic transcriptomic profiles of H/W and L–E rats has been observed (Boutros et al. 2011). Therefore, we evaluated the responses of ratonized transgenic mice at this same time point and utilizing a dose–response experiment. As above, we identified a large number of genes with differential mRNA abundance in the TCDD-sensitive rWT mouse, as well as the more resistant DEL mouse, and a muted response in highly resistant INS mouse (Fig. 3a). As a clear dose–response regarding number of differentially abundant RNAs was not apparent, we focused downstream analyses to the 500 µg/kg TCDD group, for consistency with EXP1. As explained above, this dose is above the LD_{50} for the TCDD-sensitive mice (C57BL/6 and rWT) but below it in TCDD-resistant cohorts [DEL and INS; (Pohjanvirta 2009)].

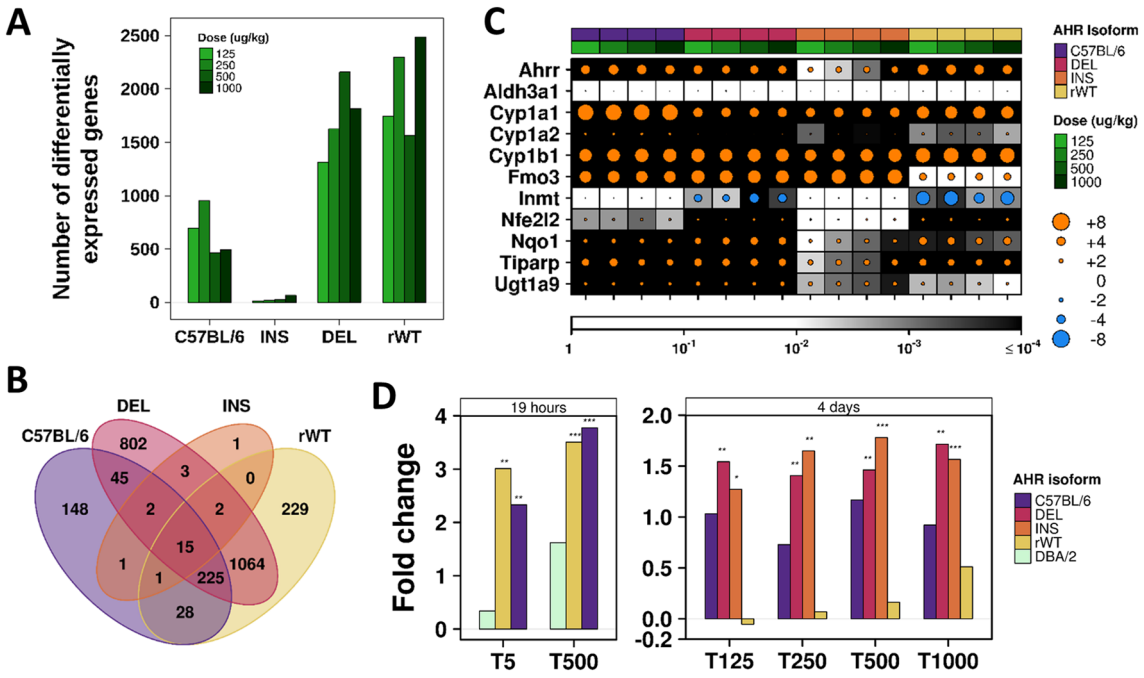

Fig. 3 Late transcriptomic changes in AHR ratonized mouse liver. **a** Using a dual threshold of $|\log_2$ fold change$| > 1$, $p_{adj} < 0.05$, genes with differential mRNA abundance were identified for each AHR isoform. A clear dose–response pattern was not observed; however, a considerable increase in the number of genes demonstrating altered mRNA abundance was detected among the highly TCDD-resistant INS isoform, TCDD-sensitive rWT and C57BL/6 isoforms, and moderately TCDD-resistant DEL isoform. **b** Overlap of these genes in each cohort, following exposure to 500 µg/kg TCDD for 96 h. **c** Log_2 fold change of "AHR-core" genes ordered by AHR isoform and increasing exposure. Dot size indicates magnitude of change following exposure to TCDD relative to controls ($\log_2$ fold change), while colour indicates direction of change (orange = increased abundance, blue = decreased abundance); background shading indicates FDR-adjusted *p* value. **d** *Fmo2* demonstrated significant changes in mRNA abundance earlier among TCDD-sensitive strains that appear later in TCDD-resistant ones. Bar height shows magnitude of change ($\log_2$ fold change); $*p < 0.05$, $**p < 0.01$, $***p < 0.001$

Interestingly, the largest overlap occurred between the DEL and rWT isoforms, consistent across all doses (Fig. 3b, Supplementary Figure 5b–d) and consistent with a previous study demonstrating reduced protection against TCDD toxicity by this DEL isoform in these models (Pohjanvirta 2009). A total of 15 genes were identified in all four groups, including six of the eleven "AHR-core" genes (Fig. 3c). As expected, responses to TCDD of *Cyp1a1* and *Cyp1b1* were consistent across all AHR isoforms and all doses (Pohjanvirta 2009). Similarly, *Ahrr*, *Nqo1*, *Tiparp* and *Ugt1a9* showed consistent changes at the 500 µg/kg TCDD dose across all strains. Interestingly, *Inmt* did not show changes to mRNA abundance in the C57BL/6 or INS but did show dramatic repression in the DEL and rWT, indicating a response specific to these AHR-genotypes. Exclusive to the TCDD-sensitive cohorts, 28 genes were identified with altered transcriptomic response to TCDD, with only three genes exclusive to the TCDD-resistant cohorts. Interestingly, this included *Fmo2* (Fig. 3d, right panel) which was altered exclusively in the TCDD-sensitive cohorts at the early time point used in EXP1 (Fig. 3d, left panel) suggesting a short-lived induction in sensitive strains/lines that is delayed in resistant ones. It is noteworthy, though, that the same resistant models were not used in both time-points, and thus *Fmo2* induction may have had a decreasing trend also in DEL and INS mice over time.

Differences between rat and mouse hepatic response to TCDD

We next expanded the study by contrasting our findings with a rat-transcriptomic dataset previously generated under similar experimental conditions (Moffat et al. 2010; Boutros et al. 2011; Yao et al. 2012). We evaluated 11,932 orthologous genes, obtained from EXP1 and EXP2 (500 µg/kg TCDD), L–E and H/W rats (1/4/10 days, 100 µg/kg TCDD) and Ln-A and Ln-C rats (19 h, 100 µg/kg TCDD). Using the same dual threshold of $|\log_2$ fold change$| > 1$, $p_{adj} < 0.05$, there was little overlap between species (with a higher degree of overlap amongst different strains/lines of the same species independent of TCDD response phenotype; Fig. 4). Thus, in the transgenic mice, the host species was a more important determinant of the resultant responsiveness than the AHR isoform. Four 'AHR-core' genes showed altered mRNA abundance in all cohorts (*Cyp1a1*, *Cyp1b1*, *Nqo1* and *Tiparp*). *Nfe2l2* showed altered transcript abundance in

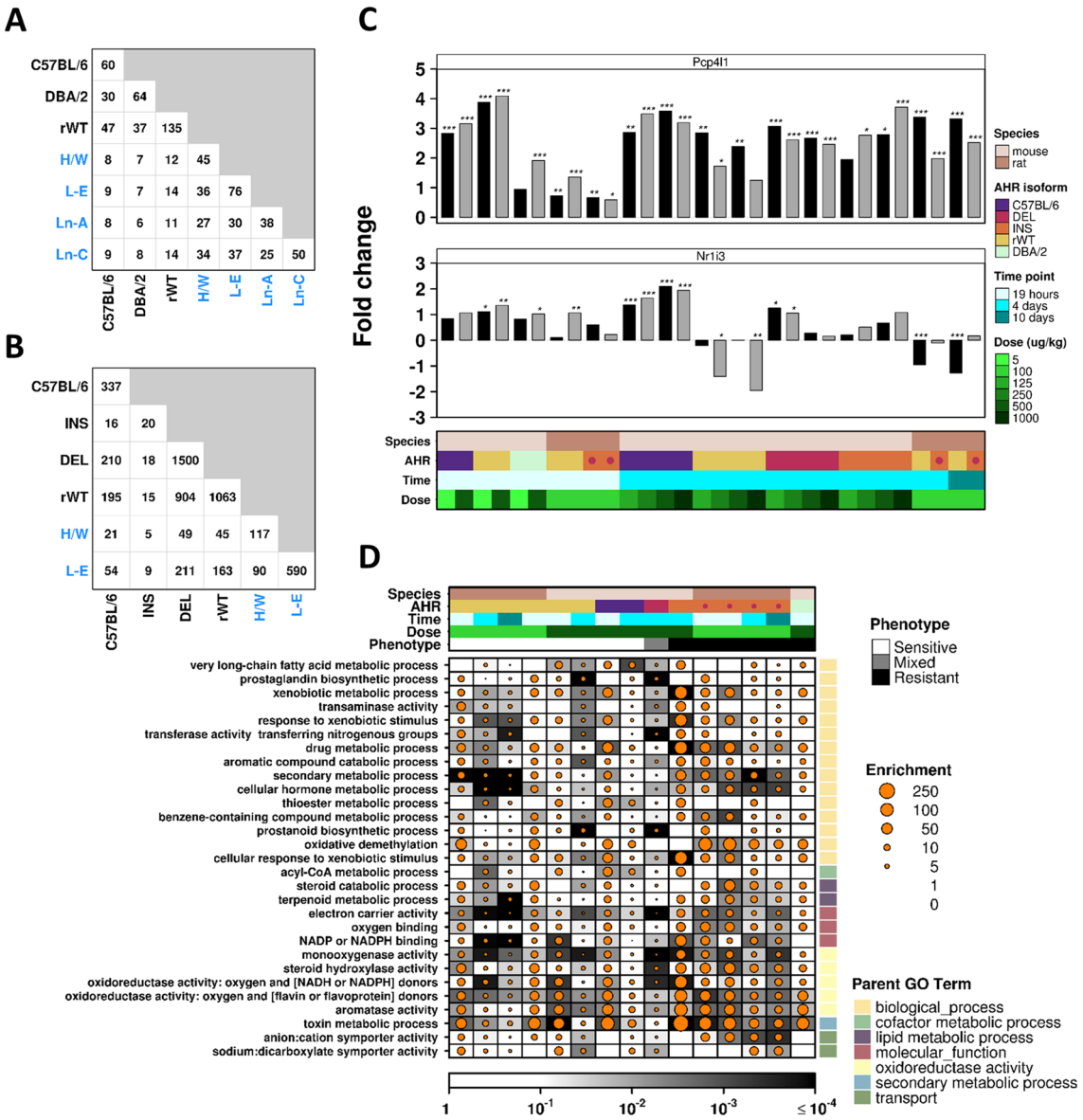

Fig. 4 Comparison of transcriptomic changes between species and AHR isoform. Intersection of the number of genes with significantly altered mRNA abundance ($|\log_2$ fold change$|>1$, $p_{adj}<0.05$) in each examined cohort. Responses following treatment with 500 µg/kg (mouse strains) or 100 µg/kg (rat strains, blue text) TCDD for **a** 19 h or **b** 4 days. Only orthologous genes were examined. **c** Pcp4l1 demonstrated significant changes in mRNA abundance among most cohorts. Bar height shows magnitude of change ($\log_2$ fold change, all with increased abundance relative to controls); $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$. **d** Pathway analysis of significantly differentially abundant orthologous genes in mouse and rat cohorts was performed using GoMiner. Significantly enriched biological pathways ($p_{adj}\leq 0.01$, enrichment>15) were identified within each group and status is shown across all groups. Dot size indicates enrichment score while background shading represents significance level. Empty cells indicate 0 genes within that pathway were differentially abundant

all cohorts except INS from EXP2 and *Inmt* was repressed more consistently in rats (H/W and L–E liver at 19 h, 4 and 10 days, and Ln-A and Ln-C at 19 h) than mice (near significant in rWT at 19 h, DEL/rWT at 4 days, all doses).

Of the 28 genes identified above in the sensitive cohorts (EXP2), 18 had homologs present in the rat dataset. Of these, *Igfbp3* demonstrated differential mRNA abundance ($|\log_2$ fold change$|>1$, $p_{adj}<0.05$) in both the TCDD-sensitive L–E rats and TCDD-resistant H/W rats following a 4-day exposure while *Pxdc1* was altered only in L–E rat liver and *Cyp1a2* was altered only in H/W rat liver (though this showed near significant induction in all animals tested). PX domain-containing protein 1 (*Pxdc1*) is significantly repressed in TCDD-sensitive cohorts (4 days following exposure, regardless of dose in C57BL/6, rWT, as well as in L–E rats; near significant in DEL mice). This gene is poorly characterized and differential

transcript abundance could not be directly attributed to the AHR because this gene demonstrated variable presence of Transcription Factor Binding Sites (TFBSs) among the species and strains/lines used (Supplementary Table 2). Additionally, no genomic differences were detected between the TCDD-sensitive L–E and TCDD-resistant H/W rats (Supplementary Table 3), suggesting that *Pxdc1* is a poor candidate for the proposed gene "B".

Pcp4l1 demonstrated induced RNA abundance in most cohorts (Fig. 4c, top). Genomic analyses indicated the presence of multiple AHRE-1 (core, extended) and ARE motifs within the promoter region of this gene in both species; however, with more occurrences in mice ($n = 8$ AHRE-1 (core) and 2 ARE motifs) than rats ($n = 7$ AHRE-1 (core) and 1 ARE motifs). *Pcp4l1* encodes Purkinje cell protein 4-like 1 and is typically expressed in neuronal tissue; it has been hypothesized to be a calmodulin inhibitor (Morgan and Morgan 2012). Interestingly, this gene is adjacent to and in the reverse orientation of, *Nr1i3*—a gene encoding a transcription factor previously associated with enhanced TCDD sensitivity (Prokopec et al. 2015), showing higher mRNA abundance in TCDD-sensitive male mice than TCDD-resistant female mice. *Nr1i3* demonstrates altered transcriptional response following TCDD exposure across species (Fig. 4c, bottom): hepatic mRNA abundance was increased after TCDD exposure in C57BL/6 mice (4 days, all doses) and earlier in rWT mice and L–E rats (19 h). This response was followed by a significantly reduced mRNA abundance in rWT mice (4 days) and L–E rats (4 and 10 days). This gene also shows different presence of AHREs in its promoter region between species ($n = 2$ AHRE-1 (core) motif in both mice and rats; 4 ARE motifs in mice; 1 AHRE-2 and 1 ARE motif in rats). No differences in AHREs were observed between H/W and L–E rats for either *Pcp4l1* or *Nr1i3*. This makes these genes interesting candidates for involvement in TCDD-induced toxicity.

Finally, functional pathways affected by TCDD were compared across these datasets, using only orthologous genes (Supplementary Table 4). Unsurprisingly, sensitive strains exhibit a larger number of significantly enriched pathways than do resistant animals, and many of these pathways are altered in multiple strains/lines (Fig. 4d). Specifically, TCDD-sensitive animals display a more significant enrichment of altered transcripts among metabolic and oxidoreductase activity pathways than resistant animals whereas resistant animals exhibit responses mostly in metabolism-related processes and transport activities.

Identifying candidates for gene "B" using genomic variants

The H/W rat strain has astonishing resistance to TCDD toxicities, predominantly conferred by a point mutation in the transactivation domain of the *Ahr* gene (Pohjanvirta et al. 1998). Part of this resistance has also been contributed to a second hypothesized gene, termed gene "B" (Tuomisto et al. 1999). Of the 642 nuclear "high impact" homozygous variants unique to the H/W strain (Supplementary Table 3), 209 mapped to genes evaluated in our microarray cohorts. Twenty of these did not show altered transcript abundance in any of our cohorts. Five genes demonstrated altered mRNA abundance exclusively in H/W rat (Supplementary Fig. 6), and 125 showed TCDD-mediated mRNA abundance changes in at least one experimental group, excluding H/W rat liver (28 of these showed significantly altered mRNA abundance in at least 8 non-H/W groups). None of these gene-sets contained more variant genes with altered mRNA abundance than expected by chance alone (hypergeometric test, $p > 0.05$). Additionally, a single "high impact" homozygous mitochondrial SNP was identified in H/W that results in a lost stop codon in mitochondrial gene NADH dehydrogenase 6 (*MT-nd6*). This was not found in other rat strains.

Since the above list does not include variants located within intergenic regions, an alternative analysis was performed to identify genes demonstrating altered transcript abundance associated with modified regulatory regions. Specifically, we searched the H/W genome for novel or lost transcription factor binding sites (TFBSs) specific to the AHR with the hypothesis that a gain of a TFBS may allow AHR-mediated transcription of a TCDD-resistance gene whereas a loss of a TFBS would prevent transcription of a TCDD-susceptibility gene. In total, 13.4% of genes (1446 of 10,772 with available TFBS data) had either a gain or loss of at least one of the AHREs examined (Supplementary Table 2). Of these, 32 exhibited changes to the number of AHRE-1 (full) motifs within their promoter regions as compared to the TCDD-sensitive L–E rat (Table 2), with 17 showing a gain and 15 a loss. Perhaps more interestingly, 15 genes harboured this motif in H/W (with none identified in the same region for L–E), while nine genes demonstrated a complete loss of this motif from the promoter region in H/W rats. TCDD-responsive genes were not enriched in those demonstrating a novel AHRE-1 (full) motif in H/W and TCDD-responsiveness in H/W liver (hypergeometric test; Supplementary Table 5).

We next evaluated differential transcriptional patterns for genes demonstrating differences among these TFBSs between H/W and L–E rats. For example, *Cxxc5* shows a loss of the AHRE1 (full) motif within its promoter region in H/W (one found in H/W and two in L–E/rn6) and exhibits significantly reduced transcript abundance in liver from both L–E (4 and 10 days) and the TCDD-sensitive Ln-C rat (19 h) following exposure to TCDD (Table 2). Similarly, *Sugp1* demonstrated a complete loss of this motif from the H/W strain relative to L–E ($n = 0$ and 1 respectively), with mice (mm9) similarly lacking the AHRE-1 (full) motif.

Table 2 Genes demonstrating altered transcription factor binding sites in rats

Gene ID	Symbol	AHRE-I (full) counts		H/W Coefficient			L–E Coefficient			Other
		H/W	L–E	1 day	4 day	10 day	1 day	4 day	10 day	
300089	Pmm1	3	2	0.58*	0.76*	0.98*	0.82*	1.65*	2.26*	mdrac
338475	Nrep	2	3	− 2.27*	− 1.69*	− 1.81*	− 1.21	− 3.13*	− 2.89*	drac
83589	Apba1	2	3	− 0.07	− 0.01	0.09	− 0.03	0.03	− 0.03	
170582	Fgf19	2	1	0.06	− 0.06	− 0.01	− 0.06	0.03	− 0.09	d
100529260	Ankhd1	1	2	0.26	0.09	− 0.02	0.19	− 0.05	− 0.04	
291,670	**Cxxc5**	1	2	0.11	0.01	0.04	− 0.16	− 0.31*	− 1.00*	c
313878	Galnt14	1	2	NA	0.02	0.16	NA	0.02	− 0.21	
54264	**Mafb**	1	2	0.29	− 0.71	− 0.66*	− 0.49	− 1.08*	− 1.47*	
114214	Dffa	1	0	0.04	0.09	0.02	0.04	0.05	− 0.01	
116565	Lrpap1	1	0	0.08	0.10	0.44*	0.20	0.52*	0.65*	d
192215	**Slc9a5**	1	0	− 0.12	0.01	− 0.05	0.07	− 0.07	− 0.28*	
29146	Jag1	1	0	− 0.07	− 0.22	− 0.07	− 0.02	− 0.20	− 0.09	
291964	Fhod1	1	0	NA	0.08	− 0.03	NA	− 0.15	0.12	
294287	Phf1	1	0	0.06	0.09	− 0.08	0.03	− 0.20	− 0.17	dr
307350	**Afg3l2**	1	0	0.11	− 0.08	− 0.03	0.14	0.00	0.41*	
362456	Arhgdib	1	0	− 0.08	0.01	0.05	0.09	0.45*	0.48*	mdr
363266	Agfg1	1	0	NA	0.14	0.07	NA	0.24	0.16	dr
364403	Blk	1	0	NA	− 0.08	0.21	NA	− 0.15	0.13	
500636	**Rnf144a**	1	0	NA	− 0.21	− 0.05	NA	0.36*	0.11	
56822	Cd86	1	0	− 0.04	0.02	0.02	− 0.09	0.02	0.32*	d
58947	**Hgfac**	1	0	NA	0.57*	0.81*	NA	1.31*	1.67*	
65158	Rab2a	1	0	0.16	− 0.02	0.00	0.16	0.08	0.24*	dc
691504	Zfpm1	1	0	0.17	− 0.25	− 0.13	0.16	0.00	− 0.47*	mdr
25638	Pde4a	0	1	− 0.06	− 0.03	0.13	0.08	0.01	− 0.21	
290666	**Sugp1**	0	1	0.00	0.10	− 0.11	0.08	0.33*	0.14	
290668	Mau2	0	1	0.03	0.00	− 0.02	0.08	0.03	0.05	
293652	**Ndufs8**	0	1	0.04	− 0.13	− 0.29*	− 0.05	0.09	− 0.15*	
302495	**Rap2c**	0	1	NA	0.12	0.15	NA	0.17	0.37*	
361653	**Armc5**	0	1	− 0.11	0.00	− 0.15	0.11	0.29*	0.21*	
361809	**Zfp523**	0	1	0.07	0.12	0.17*	0.09	0.11	− 0.01	
362339	**Creb3l2**	0	1	NA	0.09	0.23*	NA	0.14	− 0.06	
64565	Tmprss11d	0	1	− 0.15	0.06	− 0.19	0.09	0.16	− 0.01	

The number of occurrences for the AHRE-I (full) motif within a ± 3 kbp region around the transcription start site for each gene was determined for each H/W and L–E rat. Genes which demonstrate either a gain or loss of this motif in H/W relative to L–E may represent gene "B". Of the 11,392 rat/mouse orthologous genes examined, 32 were revealed to have either a gain or loss of this motif in H/W rats. Of these, 19 showed altered mRNA abundance following treatment with TCDD in at least one rat strain at one or more time points evaluated. *significantly altered mRNA abundance in TCDD treated group relative to controls ($p_{adj} < 0.05$). Column labeled 'Other' indicates significantly altered mRNA abundance in: a = Ln-A; c = Ln-C (100 µg/kg TCDD, $p_{adj} < 0.05$); m = C57BL/6 mouse; d = DEL, i = INS, r = rWT AHR-ratonized mice (500 µg/kg TCDD, $p_{adj} < 0.05$)

Gene symbols highlighted with bold font indicate key gene "B" candidates ($n = 12$)

This gene shows significantly increased mRNA abundance in only the TCDD-sensitive L–E rat liver (4 day, $\log_2$ fold change = 0.33, $p_{adj} = 0.0012$). Alternatively, *Hgfac* demonstrates a novel AHRE-1 (full) motif in the H/W rat ($n = 1$ in H/W and 0 in L–E/rn6) and its' mRNA abundance is significantly increased in both strains, 4 and 10 days following treatment, but is much higher in L–E rat liver (Table 2). No changes were detected in the AHR-ratonized mice, and this gene was not included on the arrays used for 1 day exposures, preventing assessment in the Ln-A and Ln-C rats. This list supplies a robust set of 12 gene "B" candidates, in particular *Cxxc5*, with evidence of altered TFBSs and altered transcriptional patterns in only rats that may prove suitable for further mechanistic investigation.

Validation by whole-proteome mass spectrometry

LC–MS was used to assess hepatic protein abundance from L–E and H/W rats, treated with either 100 µg/kg TCDD or control for 4 days ($n=5$ animals per group). As wasting syndrome is a well-documented outcome in L–E rats (Linden et al. 2014), an additional control group of feed restricted, vehicle treated L–E rats (FRC) were used to remove non-TCDD mediated changes to protein abundance. LC–MS provided abundance estimates for a total of 1719 unique proteins. We identified 49 proteins with significantly different abundance ($|\log_2$ fold change$|>1$, $p_{adj}<0.05$) between H/W and L–E treated with vehicle alone (Fig. 5a, basal). A further 22 proteins exhibited differential abundance between feed-restricted and control L–E rat liver (Fig. 5a, FRC). Furthermore, 26 and 113 proteins demonstrated statistically significant differences in abundance in H/W (TCDD vs. vehicle) and L–E (TCDD *vs*. FRC) respectively, of which 20 were similarly altered in both groups (Fig. 5a, H/W and L–E). This overlap included multiple members of the "AHR-core" gene battery: CYP1A1, CYP1A2, CYP1B1, CYP2A1, NQO1 and ALDH3A1. No single gene was changed in all four groups, but abundance of CYP2C70 was reduced by TCDD in both strains ($\log_2$ fold change $=-2.5$, $p_{adj}=6.78\times10^{-17}$ in L–E and $\log_2$ fold change $=-0.99$, $p_{adj}=3.52\times10^{-4}$ in H/W), showed significantly higher basal abundance in H/W ($\log_2$ fold change $=1.08$, $p_{adj}=7.14\times10^{-5}$) but was not affected by feed restriction ($\log_2$ fold change $=0.7$, $p_{adj}=1.37\times10^{-3}$). We compared differential abundance of RNA and protein ($n=1230$ genes with both data types) for each rat strain and found these were poorly correlated overall (Fig. 5b, c) but were well-correlated when focusing on those genes with significant TCDD-associated RNA changes in either strain (Pearson's $r=0.47$, $p=3.8\times10^{-3}$ for H/W; $r=0.24$, $p=2.6\times10^{-3}$ for L–E).

We next examined abundance patterns for those genes with significant differences in the TCDD-sensitive L–E rat liver at both the RNA and protein level, 4 days following exposure to TCDD ($n=21$, ($|\log_2$ fold change$|>1$, $p_{adj}<0.05$). We found increased abundance of Igtp (interferon gamma induced GTPase) and Impa2 (inositol monophosphatase 2) for both RNA and protein that was not observed in H/W. Interestingly, *Igtp* demonstrated multiple high-impact homozygous variants within the H/W genome that may contribute to this lack of change at the RNA and protein levels. Seven genes demonstrated significantly reduced RNA and protein abundance following TCDD exposure in L–E rats but not H/W rats. Five of these (*Kmo*, *Slc27a5*, *Slco1a1*, *Cyp2e1* and *Cyp2j3*) with mouse orthologues showed significant ($|\log_2$ fold change$|>1$, $p_{adj}<0.05$) or near significant ($|\log_2$ fold change $>0.5|$, $p_{adj}<0.05$) reduced transcript abundance in both DEL and rWT ratonized mice (500 µg/kg TCDD, 4 days).

Four genes exhibited significantly reduced protein abundance in H/W rat liver alone. These included SNRK, whose expression was significantly decreased at the protein ($\log_2$ fold change $=-13.6$, $p_{adj}=3.8\times10^{-5}$) but not RNA ($\log_2$ fold change $=-0.03$) levels following TCDD exposure. Fasn demonstrated reduced abundance of both RNA ($\log_2$ fold change $=-1.13$, $p_{adj}=0.036$) and protein ($\log_2$ fold change $=-0.97$, $p_{adj}=1.0\times10^{-8}$) in H/W but it also showed reduced transcript abundance in many TCDD-sensitive mouse models (DEL, rWT, C57BL/6). When looking at proteins with increased abundance unique to H/W ($n=4$), three had measurements for transcript abundance but none of these were statistically significant and they often showed

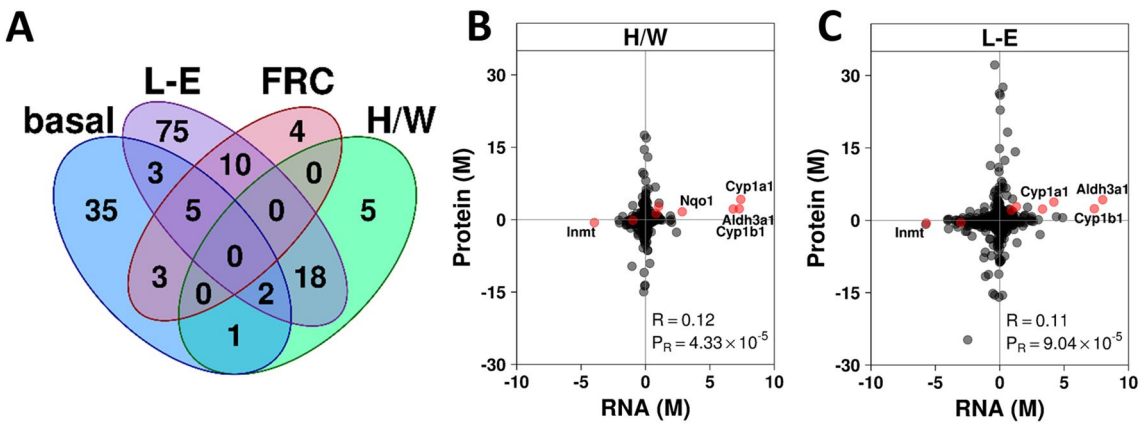

Fig. 5 Integrated comparison of − omics data in rat liver. **a** Overlap of proteins (unique gene IDs) with statistically significantly different protein abundance: basal = H/W (vehicle treated) vs. L–E (vehicle treated); FRC = L–E (feed-restricted, vehicle treated) vs. L–E (vehicle treated); H/W = TCDD vs. vehicle; L–E = TCDD vs. feed-restricted, vehicle treated. Correlation of RNA and protein differential abundance (*M* TCDD relative to control) in liver from **b** H/W and **c** L–E rats; red points show position of "AHR-core" genes, with select genes labeled. Pearson's correlations (*R*) and *p* values are shown

changes in the opposite direction. This pattern of non-significant, yet often reduced, abundance was also observed in our mouse models.

Finally, we sought to examine protein abundance for those genes having phenotype-associated transcriptomic changes (as described in above sections), however none were confidently detected via mass-spectrometry. Differential protein abundance was also not observed amongst genes with high impact sequence variants. Amongst genes with high impact sequence variants, differential abundance of CYP1B1 was not strain specific, while IGTP showed altered (increased) abundance of both RNA and protein in only L–E rats and RYR1 showed increased protein abundance, also in L–E rats. Interestingly, *Igtp* also showed a loss of the AHRE-2 binding motif in L–E rats, with no changes detected in H/W rats. Having altered AHRE-1 (full) binding sites did not result in any detectable differences in protein abundance between strains, however there was an additional AHRE-1 (extended) motif adjacent to *Snrk* in H/W that may contribute to its decreased abundance.

Discussion

There is abundant evidence that the AHR structure is a primary determinant of susceptibility to TCDD toxicity (Fernandez-Salguero et al. 1996; Schmidt et al. 1996; Mimura et al. 1997; Tijet et al. 2006; Moffat et al. 2007; Boutros et al. 2009; reviewed in Pohjanvirta et al. 2011). Two rodent species, mice and rats, show inter-strain differences in response to TCDD-insult depending on AHR genotype (Chang et al. 1993; Boutros et al. 2011; Yao et al. 2012). Furthermore, previous studies have shown little overlap in the transcriptomic response following TCDD exposure in TCDD-sensitive strains of mouse vs. rat (Boverhof et al. 2006; Boutros et al. 2008). These differences may in part be due to the intrinsic genetic variation between model organisms. In order to remove this confounding factor, we examined transgenic mouse models harbouring different rat AHR isoforms within an identical genetic background. To further support this analysis, we performed whole-genome sequencing of two common rat models with highly different responses to TCDD exposure, due to bearing the variant AHR isoforms utilized in our transgenic models. By contrasting the profiles of homozygous SNVs in each strain with the transcriptional landscape of various rat and mouse models, we aim to identify phenotype-pertinent genes.

Previous studies contrasting basal transcriptomic profiles of mice and rats with different AHR genotypes (Tijet et al. 2006; Sun et al. 2014) identified large differences in mRNA abundance of many genes in the absence of xenobiotic AHR ligands. Contrary to this, there was little difference among basal transcriptomic profiles among our transgenic mice, suggesting that these variations in AHR structure have little effect on basal gene expression. This may be partly explained by differences in the mRNA abundance of Ahr itself—in liver tissue, basal expression of TCDD-sensitive Ahr isoforms (rWT, DEL) is considerably higher than for the TCDD-resistant INS isoform in these transgenic mice (Pohjanvirta 2009). This represents an intriguing pattern and will require further study to better understand the inherent differences in activity of these isoforms.

The AHR alone is sufficient to explain much of the observed variation in sensitivity to TCDD: at all doses assessed, ~ 50% of genes demonstrating TCDD-mediated changes in transcription were unique to a single transgenic model. For example, at the highest dose studied (1000 µg/kg TCDD), 33.8%, 11.7%, 5.2%, and 0.3% of differentially abundant transcripts were specific to the rWT, DEL, C57BL/6 and INS cohorts respectively (of a total 3076 with altered mRNA abundance). These proportions were fairly consistent across doses, with the proportion of genes with altered mRNA abundance that were unique to C57BL/6 negatively correlated with dose, while INS showed a positive correlation; DEL and rWT were less consistent, with each showing ~ 30% unique changes at 500 or 1000 µg/kg TCDD, respectively. At first glance, it seems unusual that a considerable transcriptomic response was detected in the transgenic model expressing the TCDD-semi-resistant H/W DEL isoform. Both the DEL and INS isoforms are expressed naturally in the H/W rat, in roughly 15–85% proportions; however, the DEL isoform displays a significantly higher intrinsic transactivation activity (Moffat et al. 2007), partially explaining the reduced resistance observed in these models (Pohjanvirta 2009). Since our ratonized mice varied only at the AHR locus, the remaining 50% of differentially transcribed genes may be due to additional transcription factors as either primary or secondary effects, as suggested by the low overlap between early (EXP1) and late (EXP2) studies.

We identified twelve candidates for the hypothesized TCDD-resistance associated gene "B". These genes demonstrate altered TFBS within their promoter regions and altered transcript abundance in rats, but not the corresponding AHR-ratonized mice. Here we focused on the AHRE-1 (full) motif, because it results in the most productive receptor-DNA interaction (Lusska et al. 1993). In particular, three genes were identified that had lost this motif in H/W rats, and demonstrated altered mRNA abundance in only L–E (Sugp1, Rap2c and Armc5). SURP and G patch domain containing 1 (Sugp1) encodes a splicing factor that had increased mRNA abundance in TCDD-exposed L–E rat liver (4 days); however, the magnitude of change was small ($\log_2$ fold change = 0.33). Rap2c is member of RAS oncogene family that showed increased mRNA abundance in TCDD-exposed L–E rat liver (10 days), again however,

the magnitude of change was small ($\log_2$ fold change = 0.37). Armadillo repeat containing 5 (Armc5), showed significantly increased mRNA abundance following treatment in L–E rat liver, as well as in liver of mice expressing the rWT-Ahr, while it was repressed in livers of C57BL/6 mice (despite also containing a proximal AHRE-1 (full) motif). Little is known about this gene; however, it is a putative tumour suppressor gene (Assie et al. 2013).

Genomic sequencing further revealed loss of one of two AHRE-1 (full) occurrences from the promoter region of the CXXC finger protein 5 (Cxxc5) in the H/W rat along with significantly repressed mRNA abundance in livers of TCDD-sensitive L–E (4 days: $\log_2$ fold change = − 0.31, $p_{adj} = 8.8 \times 10^{-3}$ and 10 days: $\log_2$ fold change = − 1, padj = 3.1×10^{-6}) and Ln-C (19 h, $\log_2$ fold change = − 0.45, $p_{adj} = 0.014$) rats. Of interest, the transcription factor CXXC5 has been shown to inhibit expression of cytochrome c oxidase by binding to an oxygen response element in the proximal promoter of Cox4l2 in human lung cells (Aras et al. 2013). The delayed onset of repression of Cxxc5 in TCDD-sensitive cohorts may be a secondary response due to the presence of oxidative stress, and possibly corresponds with altered energy metabolism observed in these animals. Another gene, Mafb (v-maf musculoaponeurotic fibrosarcoma oncogene family, protein B), has a single AHRE-1 (full) motif in H/W rats but two in L–E rats and exhibited significantly repressed mRNA abundance in L–E rats. This may bear on the disparate developmental toxicity outcomes upon TCDD exposure observed in these strains: cleft palate was not seen at any dose in H/W rat progeny but it occurred in 71.4% of offspring in L–E rats treated with 5 µg/kg TCDD (Huuskonen et al. 1994). In humans, Mafb variants have been associated with cleft palate and lip (Zhang et al. 2018). Finally, a novel AHRE-1 (full) motif was detected in the promoter region of Hgfac (HGF activator) in H/W rats that was absent in L–E rats. mRNA abundance for this gene was significantly increased in TCDD-exposed liver from L–E rats (4 and 10 days, $\log_2$ fold change = 1.3 and 1.7, $p_{adj} = 1.16 \times 10^{-12}$ and 1.16×10^{-17}) with a more muted response in H/W rats (4 and 10 days, $\log_2$ fold change = 0.57 and 0.81, $p_{adj} = 3.7 \times 10^{-4}$ and 4.28×10^{-8}). These genes provide interesting candidates for gene "B" that require further studies into its potential involvement in the onset of TCDD-toxicities.

In H/W rats, the proteomics data revealed a high correlation with transcriptomic findings when genes with significant TCDD-associated RNA changes were considered ($r = 0.47$, $p = 3.8 \times 10^{-3}$); in L–E rats the correlation was less striking. Our combined liver proteomic and transcriptomic analyses also brought to light several candidate proteins to underlie the strain sensitivity differences in TCDD toxicity. These included Igtp, Impa2 and Snrk. In all of these cases, parallel changes were recorded at both mRNA and protein levels. Thus, the roles of their genes in TCDD toxicity should be studied further in the future.

The purpose of this study was to ascertain the mechanism of classic TCDD toxicity using various model systems, including transgenic mice to compare various rat Ahr variants in a system with a homogeneous genetic background, and various strains of rat, each with differing phenotypic responses to TCDD. To accomplish this, we generated unique proteomic, transcriptomic and genomic datasets that provide multiple levels of evidence. Using this valuable resource, we identified several genes whose transcription was selectively altered by TCDD in either TCDD-sensitive or TCDD-resistant cohorts, a differential response that can be attributed to the particular AHR isoform expressed in each cohort. Pxdc1 in particular demonstrated differential transcription between TCDD-sensitive and TCDD-resistant models across both mice and rats. However, the transcriptional responsiveness of this gene could not be explained by genomic differences in AHR-binding sites, as the transcription factor binding site analysis revealed highly variant sites between species, and no major difference between strains of rat. Our genomic sequence analysis allowed identification of differences between sensitive and resistant rat strains, which are potential "gene B" candidates. For instance, Cxxc5 harboured fewer occurrences of AHRE-1 (full) TFBSs in H/W relative to L–E, and had reduced RNA abundance in sensitive strains/lines. This is therefore a promising candidate for further study in relation to mechanisms of TCDD toxicity and regulatory roles of the AHR.

Materials and methods

Animal handling

Three separate experiments were performed (Fig. 1). In the first, adult male C57BL/6Kuo, rWT and DBA/2 J mice were evaluated. In the second, adult male C57BL/6 mice carrying 4 different *Ahr* isoforms were examined: C57BL/6Kuo and rWT, DEL, and INS transgenic mice. Transgenic mice were generated as described previously (Pohjanvirta 2009). Briefly, animals were bred from *Ahr*-null mice to avoid interference by the C57BL/6 AHR. Once established, transgenic colonies were bred at the National Public Health Institute, Division of Environmental Health in Kuopio, Finland. For this study, animal ages varied from 12 to 23 weeks. All animals were housed singly in Makrolon cages. The housing environment was maintained at 21 ± 1 °C with relative humidity at 50 ± 10%, and followed a 12-h light cycle. Tap water and R36 pellet feed (Lactamin, Stockholm, Sweden) or Altromin 1314 pellet feed (DBA/2 mice; Altromin Spezialfutter GmbH & Co. KG, Lage, Germany) were available ad libitum.

In the first experiment (EXP1), a total of 36 mice were used ($n = 12$ per genotype), divided into groups of four mice per treatment group. Mice were treated by oral gavage with a single dose of 0, 5, or 500 µg/kg TCDD dissolved in corn oil vehicle (Supplementary Table 1). Animals were euthanized by carbon dioxide, followed immediately by cardiac exsanguination 19 h following treatment. In the second experiment (EXP2), a total of 84 mice were used, divided into groups according to *Ahr* isoform. Mice were treated by oral gavage with a single dose of 0, 125, 250, 500, or 1000 µg/kg TCDD dissolved in corn oil vehicle (Supplementary Table 1). Animals were euthanized by cervical dislocation 4 days following exposure and their livers excised. A single rWT animal from the 1000 µg/kg TCDD group died prematurely and was thus excluded from the study, thereby leaving 83 animals.

Microarray hybridization

Mouse livers were frozen in liquid nitrogen immediately upon excision and stored at − 80 °C. Tissue samples were homogenized and RNA was isolated using an RNeasy Mini Kit (Qiagen, Mississauga, Canada) according to the manufacturer's instructions. Total RNA was quantitated using a NanoDrop UV spectrophotometer (Thermo Scientific, Mississauga, ON) and RNA quality was verified using RNA 6000 Nano kits on an Agilent 2100 Bioanalyzer (Agilent Technologies, Mississauga, ON). RNA abundance levels were assayed using Affymetrix Mouse Gene 1.1 ST arrays (Affymetrix Mouse Gene 2.0 ST arrays were used for the C57BL/6 mice for EXP1) at The Centre for Applied Genomics (TCAG; Toronto, ON) as described previously (Boutros et al. 2011).

Data pre-processing and statistical analysis

Raw data (CEL files) were loaded into the R statistical environment (v3.4.3) using the affy package (v1.48.0) (Gautier et al. 2004) of the BioConductor library (Gentleman et al. 2004) and pre-processed using the RMA algorithm (Irizarry et al. 2003). Each group (experiment/AHR isoform) was pre-processed independently to avoid masking any differences between isoforms. Probes were mapped to EntrezGene IDs using the custom mogene11stmmentrezgcdf and mogene20stmmentrezgcdf packages (v22.0.0) (Dai et al. 2005). Visual inspection of array distribution and homogeneity of the results suggested the presence of outliers. These were further identified using Dixon's Q test (Dixon 1950) as implemented in the outliers package (v0.14) in R, which was used to compare abundance patterns of "AHR-core" genes for each genotype/treatment combination (Supplementary Figure 1). Three TCDD-treated rWT animals were removed from EXP2, along with one control DBA animal and one TCDD-treated rWT animal from EXP1, as mRNA abundance for these AHR-core genes closely resembled the control animals (FDR-adjusted Dixon's Q test, $p < 0.05$ for 2 or more AHR-core genes, requiring at least one of *Cyp1a1*, *Cyp1a2* or *Cyp1b1*; Supplementary Figure 1). Remaining arrays were re-processed without outliers (Supplementary Figures 2 and 3). ComBat was run on the combined dataset, using the sva package (v3.24.4) for R, to adjust RMA normalized values for comparison across batches (Supplementary Fig. 4). General linear modeling was performed to determine which genes were significantly altered following exposure to TCDD. Specifically, for each gene (i) the abundance (Y) was modelled as $Y_i = \beta_0 + \beta_1 X$, where X indicates the dose of TCDD administered. Each experiment/AHR isoform was modelled independently, using RMA normalized values (without ComBat correction as each group was processed as a single batch) with each dose treated as a factor. Genes were modelled as a univariate combination of a basal effect (β_0) (represented by the vehicle control) and TCDD effect (β_1) with contrasts fit to compare each dose–effect to baseline. Standard errors of the coefficients were smoothed using an empirical Bayes method (Smyth 2004) and significance was identified using model-based t-tests, followed by FDR adjustment for multiple testing (Storey and Tibshirani 2003). Sensitivity of results to various p_{adj}-value cut-offs was assessed (Fig. 2b, Supplementary Figure 4c–i). Those genes deemed significantly altered ($p_{adj} < 0.05$ and $|\text{coefficient}| > 1$) following treatment were examined in downstream analyses. All statistical analyses were performed using the limma package for R (v3.32.10) (Smyth 2004). Venn diagrams were created using the VennDiagram package (v1.6.21) for R (Chen and Boutros 2011) to visualize overlap between groups. All other data visualizations were generated using the BPG plotting package (v5.9.8) (P'ng et al. 2017), using the lattice (v0.20-38) and latticeExtra (v0.6-28) packages for R.

Rat-mouse overlap of differentially abundant RNAs

Publicly available data for male H/W and L–E rats, as well as two lines (Ln-A and Ln-C) derived from L–E × H/W crosses, were used for comparison (Moffat et al. 2010; Boutros et al. 2011; Yao et al. 2012; Prokopec et al. 2017). Data consisted of the hepatic transcriptomic responses of rats to a single dose of TCDD (100 µg/kg in corn oil) at three time points (19 h, 4 and 10 days). Notably, in the 4 day exposure group, L–E rats treated with TCDD were contrasted with feed-restricted, vehicle treated controls (FRC). All data were available from the TCDD.Transcriptomics (v2.2.5) package (Prokopec et al. 2017) for R. Due to the use of different microarrays between studies, HomoloGene (build 68) was used to identify orthologous genes. Homologene IDs which mapped to multiple EntrezGeneIDs were discarded.

In total, 11,932 genes were found to be orthologous between species and were used for overlap analyses (Supplementary Table 6).

Pathway analysis

Pathway analysis was conducted in order to elucidate the functions of genes demonstrating significantly altered mRNA abundance for each AHR isoform cohort. Analysis was performed using the High-Throughput GoMiner web interface (application build 469, database build 2011-01, accessed 2019-02) (Zeeberg et al. 2005). A separate run was performed for each mouse AHR isoform and each available rat dataset. We compared significantly altered genes against a randomly drawn sample from all orthologous mouse (or rat) genes in the dataset, using an FDR threshold of 0.1, 1000 randomizations, all mouse (or rat) databases and look-up options, and all GO evidence codes and ontologies (Supplementary Table 4). Results were further filtered using a threshold of $p_{adj} < 0.01$ and an enrichment score > 15 in at least one of the examined groups. This resulted in 29 unique gene ontologies to examine further.

Genome sequencing

Untreated adult male outbred H/W (Kuopio) and inbred L–E (Turku/AB) (Pohjanvirta and Tuomisto 1990) rats were euthanized by decapitation and their livers excised. Tissue was frozen in liquid nitrogen and shipped to the analytic facility on dry ice. Genomic DNA (gDNA) was isolated and whole genome sequencing performed by Genome Technologies at the Ontario Institute for Cancer Research and Applied Biosystems (Burlington, ON) using the AB SOLiD platform using mate-pair and fragment libraries. For mate-pair libraries, 100 µg of gDNA was sheared to 1-2 kb fragments using the GeneMachine HydroShear standard shearing assembly and 1.5 kb fragments isolated using 1% agarose gel size selection. Mate-pair libraries were circularized and constructed according to standard SOLiD Long Mate-Pair library protocols (Applied Biosystems, Burlington, ON). Following library quantitation via TaqMan qPCR (Applied Biosystems, Burlington, ON), libraries underwent emulsion PCR and bead enrichment according to standard SOLiD protocols (Applied Biosystems, Burlington, ON). Enriched libraries were then sequenced 2 × 50 bp using SOLiD 3 sequencing chemistry (Applied Biosystems, Burlington, ON). A similar procedure was used to produce fragment libraries, with the following changes: 1 µg of gDNA was sheared to 70-90 bp fragments using the Covaris S220 (Covaris Inc., Woburn, MA) and 150 bp fragment libraries were constructed according to standard SOLiD Fragment library protocols (Applied Biosystems, Burlington, ON). Libraries were quantified and enriched as described above and sequenced 1x50 bp using SOLiD 3 sequencing chemistry (Applied Biosystems, Burlington, ON).

Sequence alignment and variant calling

Raw reads were split into manageable chunks ($n = 10{,}000{,}000$ reads) and aligned to the rat reference genome (rn6) using BFAST (v0.7.0a) with default parameters. Resulting SAM format files were converted to BAM format and coordinate sorted, followed by mark duplicates, merging of partial files and indexing using Picard (v1.92). Indel realignment/recalibration were performed using GATK (v3.7.0), followed by variant detection using GATK's HaplotypeCaller, with known variants identified using dbSNP (build 149, downloaded from UCSC on 2018-11-19). Resulting variants were filtered such that any variants with a depth < 6 reads and SNPs common to previously sequenced rat strains (Atanur et al. 2013) were excluded [GATKs LiftoverVariants; using rn4 to rn6 chain file from UCSC, and vcftools (v0.1.15)]. Variant annotation was performed using SnpEff (v4.3t) with the rn6 database (Rnor_6.0.86). As these are highly controlled strains, only homozygous variants were carried forward for analysis. This reduced the number of variants by 95% to 176,578 for H/W and by 92% to 161,620 for L–E.

Transcription factor binding site analyses

The modified rat genomes (H/W and L–E) were used to identify differing transcription factor binding sites (TFBS) between strains. Conservation scores, REFLINK and REFFLAT tables for rn6 were downloaded from the UCSC genome browser on August 09, 2018 (Karolchik et al. 2003). For each rat strain examined (H/W and L–E), the unique variants identified above were inserted into the reference genome (rn6) FASTA file. For each rat strain, as well as for mouse (mm9), the genome was searched for the following motifs: AHRE-I (core) GCGTG (Denison et al. 1988), AHRE-I (extended) T/NGCGTG (Swanson et al. 1995), AHRE-I (full) [T|G]NGCGTG[A|C][G|C]A (Lusska et al. 1993) and AHRE-II CATG(N6)C[T|A]TG (Sogawa et al. 2004) and ARE TGAC(N3)GC (Rushmore et al. 1991). Exposed motifs were annotated to specific genes if they occurred within a promoter region (± 3 kbp of the transcription start site) and a PhyloHMM conservation score from 0 (weak conservation) to 1 (strong conservation) was calculated (Supplementary Table 2).

Sample preparation and proteolytic digestion of proteins

Adult male H/W and L–E rats were treated, and tissue collected as described previously (Linden et al. 2014).

Proteomics analyses were carried out at the Meilahti Clinical & Basic Proteomics Core Facility of the University of Helsinki. Liver pieces were homogenized in 7 M urea, 2 M thiourea, 4% CHAPS using Percellys® 24 homogenizer (Bertin Technologies). For each sample, 10 µg of total protein lysate was processed for digestion, using a modified FASP protocol (Wisniewski et al. 2009; Scifo et al. 2015). In brief, the lysate buffer was exchanged by washing it several times with 8 M urea, 0.1 M Tris, pH 8 (Urea buffer, UB). The proteins were reduced by addition of 10 mM DTT in UB and alkylated with 50 mM iodoacetamide in UB, after washing out the DTT-containing solution. 1:50 w/w of Lysine-C endopeptidase (Wako) was added to ~ 4 M urea/0.1 M Tris pH 8 and incubated at room temperature overnight. The peptide digests were collected by centrifugation and trypsin solution was added in a ratio of 1:50 w/w in 50 mM ammonium bicarbonate. As before, the digests were collected and combined. The peptide samples were cleaned using C18-reverse phase ZipTip™ (Millipore), resuspended in 1%TFA and sonicated in water bath for 1 min.

LC-HDMS^e

300 ng of digested proteins/replicate (3 replicate runs per sample) was injected for LC–MS analysis. The peptides were separated by nanoAcquity UPLC system (Waters) equipped with a trapping column 5 µm Symmetry C18 180 µm × 20 mm C18 reverse phase (Waters), followed by an analytical 1.7 µm, 75 µm × 250 mm BEH-130 C18 reversed-phase column (Waters), in a single pump trapping mode. The injected sample analytes were trapped at a flow rate of 15 µl/min in 99.5% of solution A (0.1% formic acid). After trapping, the peptides were separated with a linear gradient of 3–35% of solution B (0.1% formic acid/acetonitrile), for 100 min at a flow rate 0.3 µl/min and stable column temperature of 35 °C. Every sample run was succeeded by two empty runs to wash out any remaining peptides from previous runs. Samples were run in ion mobility assisted data-independent analysis mode (HDMS^E), in a Synapt G2-S mass spectrometer (Waters), by alternating between low collision energy (6 V) and high collision energy ramp in the transfer compartment (20–45 V) and using 1 s cycle time. The separated peptides were detected online with mass spectrometer, operated in positive, resolution mode in the range of *m/z* 50-2000 amu. 150 fmol/µl of human [Glu^1]-fibrinopeptide B (Sigma) in 50% acetonitrile/0.1% formic acid solution at a flow rate of 0.3 µl/min was used for a lock mass correction, applied every 30 s.

Database mining of LC-HDMS^E

Relative quantification between samples using precursor ion intensities was performed with Progenesis QI for Proteomics™ Informatics for Proteomics software (Nonlinear Dynamics, Waters) and ProteinLynx Global Server (PLGS V3.0). MS^e parameters were set as follows: low energy threshold of 135 counts, elevated energy threshold of 30 counts, and intensity threshold of precursor/fragment ion cluster 750 counts. Chromatograms were automatically aligned by the Progenesis QI for Proteomics™ software using the default values by following the wizard, and those that had alignment score ≥ 70% to the reference run were selected for further analysis. To compare the control(s) to other subjects we utilized the between-subject design scheme of Progenesis QI for Proteomics™ software. The ANOVA calculation applied by this scheme assumes that the conditions are independent and applies the statistical test which presumes the means of the conditions are equal.

Database searches were carried out against *Rattus norvegicus* UniProtKB-SwissProt, reviewed, database (release 2019_5, 8062 entries) with Ion Accounting algorithm and using the following parameters: peptide and fragment tolerance: automatic, maximum protein mass: 500 kDa, min fragment ions matches per protein ≥ 7, min fragment ions matches per peptide ≥ 3, min peptide matches per protein ≥ 1, primary digest reagent: trypsin, missed cleavages allowed: 2, fixed modification: carbamidomethylation C, variable modifications: deamidation of NQ residues, oxidation of Methionine and false discovery rate (FDR) < 4%. All identifications were subsequently refined to *Rattus norvegicus* only identifiers and the protein lists simplified by protein grouping. Protein quantitation was performed entirely on non-conflicting proteins identifications, using precursor ion intensity data and standardized expression profiles. Additional filters in Progenesis QI for Proteomics™ applied to final data increased the stringency of accepted protein leads by limiting the ratio between treated as compared to controls (fold change > 2 in either direction of up- or down-regulation, computed from averaged, normalized protein intensities, and $p < 0.05$ by ANOVA for all comparisons).

Statistical analysis of proteomic data

Following above filtering, normalized protein intensities were provided for 1734 unique protein/protein groups. Additional filtering was applied to remove proteins with low confidence score (< 10), leading to removal of 15 proteins. Cases with a normalized intensity of 0 were replaced with the minimum, non-zero intensity present in the dataset. Data were then $\log_2$-transformed and linear modeling applied, with technical replicates treated as a random effect using the limma package for R (v3.32.10) (Smyth 2004). Contrasts were fit to compare each dose–effect to baseline (TCDD vs. vehicle alone for H/W; TCDD vs. vehicle treated, feed-restricted controls for L–E), as well as identify basal differences between strains. Standard errors of the coefficients

were smoothed using an empirical Bayes method (Smyth 2004) and significance was identified using model-based t-tests, followed by FDR adjustment for multiple testing (Storey and Tibshirani 2003). Model results are available in Supplementary Table 7.

Acknowledgements The authors thank Hanbert Chen, Alexander Wu, Ashley Smith, Janne Korkalainen, Arja Moilanen, and Virpi Tiihonen for excellent technical assistance and support. Additional thanks to Marc Baumann and staff at the Meilahti Clinical & Basic Proteomics Core Facility. This work was supported by the Canadian Institutes of Health Research [grant number MOP-57903 to ABO and PCB], the Academy of Finland [grant number 123345 to RP], and with the support of the Ontario Institute for Cancer Research to PCB through funding provided by the Government of Ontario. PCB was supported by a Terry Fox Research Institute New Investigator Award and a Canadian Institutes of Health Research New Investigator Award.

Author contributions JDW and SDP carried out the sample preparation for transcriptomic analyses. AW, MS and PZ were involved in library preparation and genome sequencing. RS performed proteomics work. AL, SDP, SL and RDB performed statistical and bioinformatics analyses. AL and SDP wrote the first draft of the manuscript. AL, CQY, SDP, RXS and RP generated tools and reagents. ABO, RP and PCB initiated the project. JDM, ABO, RP and PCB supervised the research. All authors approved the manuscript.

Data availability The data sets supporting the results of this article are publicly available. All transcriptomic datasets are available from NCBI's Gene Expression Omnibus (GEO; http://www.ncbi.nlm.nih.gov/geo/): C57BL/6 mouse data are available at accession GSE61038 (Lee et al. 2015); other mouse data are available from GSE127217 (EXP1) and GSE72270 (EXP2); rat transcriptomic data are available from accessions GSE31411 (Yao et al. 2012) and GSE13513 (Boutros et al. 2011). Combined transcriptomic data are additionally available from the TCDD. Transcriptomics (v2.2.5) package for the R statistical environment (Prokopec et al. 2017). The sequence data generated in this study have been submitted to the NCBI BioProject database (http://www.ncbi.nlm.nih.gov/bioproject/) under accession number PRJNA480994. Raw and processed LC–MS data are available from MassIVE under accession MSV000083870.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All study plans were approved by the Finnish National Animal Experiment Board (Eläinkoelautakunta, ELLA; permit code: ESLH-2008-07223/Ym-23). All animal handling and reporting comply with ARRIVE guidelines (Kilkenny et al. 2010).

References

Aras S, Pak O, Sommer N, Finley R Jr, Huttemann M, Weissmann N, Grossman LI (2013) Oxygen-dependent expression of cytochrome c oxidase subunit 4-2 gene expression is mediated by transcription factors RBPJ, CXXC5 and CHCHD2. Nucleic Acids Res 41:2255–2266

Assie G, Libe R, Espiard S, Rizk-Rabin M, Guimier A, Luscap W, Barreau O, Lefevre L, Sibony M, Guignat L et al (2013) ARMC5 mutations in macronodular adrenal hyperplasia with Cushing's syndrome. N Engl J Med 369:2105–2114

Atanur SS, Diaz AG, Maratou K, Sarkis A, Rotival M, Game L, Tschannen MR, Kaisaki PJ, Otto GW, Ma MC et al (2013) Genome sequencing reveals loci under artificial selection that underlie disease phenotypes in the laboratory rat. Cell 154:691–703

Boutros PC, Yan R, Moffat ID, Pohjanvirta R, Okey AB (2008) Transcriptomic responses to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in liver: comparison of rat and mouse. BMC Genom 9:419

Boutros PC, Bielefeld KA, Pohjanvirta R, Harper PA (2009) Dioxin-dependent and dioxin-independent gene batteries: comparison of liver and kidney in AHR-null mice. Toxicol Sci 112:245–256

Boutros PC, Yao CQ, Watson JD, Wu AH, Moffat ID, Prokopec SD, Smith AB, Okey AB, Pohjanvirta R (2011) Hepatic transcriptomic responses to TCDD in dioxin-sensitive and dioxin-resistant rats during the onset of toxicity. Toxicol Appl Pharmacol 251:119–129

Boverhof DR, Burgoon LD, Tashiro C, Sharratt B, Chittim B, Harkema JR, Mendrick DL, Zacharewski TR (2006) Comparative toxicogenomic analysis of the hepatotoxic effects of TCDD in Sprague Dawley rats and C57BL/6 mice. Toxicol Sci 94:398–416

Chang C, Smith DR, Prasad VS, Sidman CL, Nebert DW, Puga A (1993) Ten nucleotide differences, five of which cause amino acid changes, are associated with the Ah receptor locus polymorphism of C57BL/6 and DBA/2 mice. Pharmacogenetics 3:312–321

Chapman DE, Schiller CM (1985) Dose-related effects of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in C57BL/6 J and DBA/2 J mice. Toxicol Appl Pharmacol 78:147–157

Chen H, Boutros PC (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. BMC Bioinform 12:35

Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG, Bunney WE, Myers RM, Speed TP, Akil H et al (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. Nucleic Acids Res 33:e175

Deb S, Bandiera SM (2010) Characterization of a new cytochrome P450 enzyme, CYP2S1, in rats: its regulation by aryl hydrocarbon receptor agonists. Toxicology 267:91–98

Denison MS, Fisher JM, Whitlock JP Jr (1988) The DNA recognition site for the dioxin-Ah receptor complex. Nucleotide sequence and functional analysis. J Biol Chem 263:17221–17224

Dixon WJ (1950) Analysis of extreme values. Ann Math Statist 21:488–506

Dong B, Nishimura N, Vogel CF, Tohyama C, Matsumura F (2010) TCDD-induced cyclooxygenase-2 expression is mediated by the nongenomic pathway in mouse MMDD1 macula densa cells and kidneys. Biochem Pharmacol 79:487–497

Fernandez-Salguero PM, Hilbert DM, Rudikoff S, Ward JM, Gonzalez FJ (1996) Aryl-hydrocarbon receptor-deficient mice are resistant to 2,3,7,8-tetrachlorodibenzo-p-dioxin-induced toxicity. Toxicol Appl Pharmacol 140:173–179

Gautier L, Cope L, Bolstad BM, Irizarry RA (2004) affy–analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20:307–315

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J et al (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5:R80

Hahn ME, Karchner SI, Shapiro MA, Perera SA (1997) Molecular evolution of two vertebrate aryl hydrocarbon (dioxin) receptors (AHR1 and AHR2) and the PAS family. Proc Natl Acad Sci USA 94:13743–13748

Harrill JA, Layko D, Nyska A, Hukkanen RR, Manno RA, Grassetti A, Lawson M, Martin G, Budinsky RA, Rowlands JC et al (2016) Aryl hydrocarbon receptor knockout rats are insensitive to the

pathological effects of repeated oral exposure to 2,3,7,8-tetrachlorodibenzo-p-dioxin. J Appl Toxicol 36:802–814
Huuskonen H, Unkila M, Pohjanvirta R, Tuomisto J (1994) Developmental toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in the most TCDD-resistant and -susceptible rat strains. Toxicol Appl Pharmacol 124:174–180
Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP (2003) Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res 31:e15
Karolchik D, Baertsch R, Diekhans M, Furey TS, Hinrichs A, Lu YT, Roskin KM, Schwartz M, Sugnet CW, Thomas DJ et al (2003) The UCSC genome browser database. Nucleic Acids Res 31:51–54
Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG (2010) Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. PLoS Biol 8:e1000412
Kransler KM, McGarrigle BP, Olson JR (2007) Comparative developmental toxicity of 2,3,7,8-tetrachlorodibenzo-p-dioxin in the hamster, rat and guinea pig. Toxicology 229:214–225
Lee J, Prokopec SD, Watson JD, Sun RX, Pohjanvirta R, Boutros PC (2015) Male and female mice show significant differences in hepatic transcriptomic response to 2,3,7,8-tetrachlorodibenzo-p-dioxin. BMC Genom 16:625
Linden J, Lensu S, Pohjanvirta R (2014) Effect of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) on hormones of energy balance in a TCDD-sensitive and a TCDD-resistant rat strain. Int J Mol Sci 15:13938–13966
Lusska A, Shen E, Whitlock JP Jr (1993) Protein-DNA interactions at a dioxin-responsive enhancer. Analysis of six bona fide DNA-binding sites for the liganded Ah receptor. J Biol Chem 268:6575–6580
Matsumura F (2009) The significance of the nongenomic pathway in mediating inflammatory signaling of the dioxin-activated Ah receptor to cause toxic effects. Biochem Pharmacol 77:608–626
Mimura J, Yamashita K, Nakamura K, Morita M, Takagi TN, Nakao K, Ema M, Sogawa K, Yasuda M, Katsuki M et al (1997) Loss of teratogenic response to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) in mice lacking the Ah (dioxin) receptor. Genes Cells 2:645–654
Moffat ID, Roblin S, Harper PA, Okey AB, Pohjanvirta R (2007) Aryl hydrocarbon receptor splice variants in the dioxin-resistant rat: tissue expression and transactivational activity. Mol Pharmacol 72:956–966
Moffat ID, Boutros PC, Chen H, Okey AB, Pohjanvirta R (2010) Aryl hydrocarbon receptor (AHR)-regulated transcriptomic changes in rats sensitive or resistant to major dioxin toxicities. BMC Genom 11:263
Morgan MA, Morgan JI (2012) Pcp4l1 contains an auto-inhibitory element that prevents its IQ motif from binding to calmodulin. J Neurochem 121:843–851
Nebert DW, Puga A, Vasiliou V (1993) Role of the Ah receptor and the dioxin-inducible [Ah] gene battery in toxicity, cancer, and signal transduction. Ann NY Acad Sci 685:624–640
Okey AB, Vella LM, Harper PA (1989) Detection and characterization of a low affinity form of cytosolic Ah receptor in livers of mice nonresponsive to induction of cytochrome P1-450 by 3-methylcholanthrene. Mol Pharmacol 35:823–830
P'ng C, Green J, Chong LC, Waggott D, Prokopec SD, Shamsi M, Nguyen F, Mak DYF, Lam F, Albuquerque MA et al (2017) BPG: seamless, automated and interactive visualization of scientific data. *bioRxiv*
Pohjanvirta R (1990) TCDD resistance is inherited as an autosomal dominant trait in the rat. Toxicol Lett 50:49–56
Pohjanvirta R (2009) Transgenic mouse lines expressing rat AH receptor variants–a new animal model for research on AH receptor function and dioxin toxicity mechanisms. Toxicol Appl Pharmacol 236:166–182
Pohjanvirta R (ed) (2011) The AH receptor in biology and toxicology. Wiley, Hoboken
Pohjanvirta R, Tuomisto J (1990) Mechanism of action of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). Toxicol Appl Pharmacol 105:508–509
Pohjanvirta R, Kulju T, Morselt AF, Tuominen R, Juvonen R, Rozman K, Mannisto P, Collan Y, Sainio EL, Tuomisto J (1989) Target tissue morphology and serum biochemistry following 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) exposure in a TCDD-susceptible and a TCDD-resistant rat strain. Fundam Appl Toxicol 12:698–712
Pohjanvirta R, Unkila M, Tuomisto J (1993) Comparative acute lethality of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), 1,2,3,7,8-pentachlorodibenzo-p-dioxin and 1,2,3,4,7,8-hexachlorodibenzo-p-dioxin in the most TCDD-susceptible and the most TCDD-resistant rat strain. Pharmacol Toxicol 73:52–56
Pohjanvirta R, Wong JM, Li W, Harper PA, Tuomisto J, Okey AB (1998) Point mutation in intron sequence causes altered carboxyl-terminal structure in the aryl hydrocarbon receptor of the most 2,3,7,8-tetrachlorodibenzo-p-dioxin-resistant rat strain. Mol Pharmacol 54:86–93
Pohjanvirta R, Viluksela M, Tuomisto JT, Unkila M, Karasinska J, Franc MA, Holowenko M, Giannone JV, Harper PA, Tuomisto J et al (1999) Physicochemical differences in the AH receptors of the most TCDD-susceptible and the most TCDD-resistant rat strains. Toxicol Appl Pharmacol 155:82–95
Pohjanvirta R, Korkalainen M, Moffat ID, Boutros PC, Okey AB (2011) Role of the AHR and its structure in TCDD toxicity. In: Pohjanvirta R (ed) The AH receptor in biology and toxicology. Wiley, Hoboken, pp 181–196
Pohjanvirta R, Miettinen H, Sankari S, Hegde N, Linden J (2012) Unexpected gender difference in sensitivity to the acute toxicity of dioxin in mice. Toxicol Appl Pharmacol 262:167–176
Prokopec SD, Watson JD, Lee J, Pohjanvirta R, Boutros PC (2015) Sex-related differences in murine hepatic transcriptional and proteomic responses to TCDD. Toxicol Appl Pharmacol 284:188–196
Prokopec SD, Houlahan KE, Sun RX, Watson JD, Yao CQ, Lee J, P'ng C, Pang R, Wu AH, Chong LC et al (2017) Compendium of TCDD-mediated transcriptomic response datasets in mammalian model systems. BMC Genom 18:78
Rowlands JC, Gustafsson JA (1997) Aryl hydrocarbon receptor-mediated signal transduction. Crit Rev Toxicol 27:109–134
Rushmore TH, Morton MR, Pickett CB (1991) The antioxidant responsive element. Activation by oxidative stress and identification of the DNA consensus sequence required for functional activity. J Biol Chem 266:11632–11639
Schmidt JV, Su GH, Reddy JK, Simon MC, Bradfield CA (1996) Characterization of a murine Ahr null allele: involvement of the Ah receptor in hepatic growth and development. Proc Natl Acad Sci USA 93:6731–6736
Scifo E, Szwajda A, Soliymani R, Pezzini F, Bianchi M, Dapkunas A, Debski J, Uusi-Rauva K, Dadlez M, Gingras AC et al (2015) Proteomic analysis of the palmitoyl protein thioesterase 1 interactome in SH-SY5Y human neuroblastoma cells. J Proteom 123:42–53
Shen ES, Whitlock JP Jr (1992) Protein-DNA interactions at a dioxin-responsive enhancer. Mutational analysis of the DNA-binding site for the liganded Ah receptor. J Biol Chem 267:6815–6819
Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol. https://doi.org/10.2202/1544-6115.1027
Sogawa K, Numayama-Tsuruta K, Takahashi T, Matsushita N, Miura C, Nikawa J, Gotoh O, Kikuchi Y, Fujii-Kuriyama Y (2004) A novel induction mechanism of the rat CYP1A2 gene mediated by Ah receptor-Arnt heterodimer. Biochem Biophys Res Commun 318:746–755

Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. Proc Natl Acad Sci USA 100:9440–9445

Sun RX, Chong LC, Simmons TT, Houlahan KE, Prokopec SD, Watson JD, Moffat ID, Lensu S, Linden J, P'ng C et al (2014) Cross-species transcriptomic analysis elucidates constitutive aryl hydrocarbon receptor activity. BMC Genom 15:1053

Swanson HI, Chan WK, Bradfield CA (1995) DNA binding specificities and pairing rules of the Ah receptor, ARNT, and SIM proteins. J Biol Chem 270:26292–26302

Tijet N, Boutros PC, Moffat ID, Okey AB, Tuomisto J, Pohjanvirta R (2006) Aryl hydrocarbon receptor regulates distinct dioxin-dependent and dioxin-independent gene batteries. Mol Pharmacol 69:140–153

Tuomisto JT, Viluksela M, Pohjanvirta R, Tuomisto J (1999) The AH receptor and a novel gene determine acute toxic responses to TCDD: segregation of the resistant alleles to different rat lines. Toxicol Appl Pharmacol 155:71–81

Von Burg R (1988) Tcdd. J Appl Toxicol 8:145–148

Watson JD, Prokopec SD, Smith AB, Okey AB, Pohjanvirta R, Boutros PC (2014) TCDD dysregulation of 13 AHR-target genes in rat liver. Toxicol Appl Pharmacol 274:445–454

Wisniewski JR, Zougman A, Nagaraj N, Mann M (2009) Universal sample preparation method for proteome analysis. Nat Methods 6:359–362

Yao CQ, Prokopec SD, Watson JD, Pang R, P'ng C, Chong LC, Harding NJ, Pohjanvirta R, Okey AB, Boutros PC (2012) Inter-strain heterogeneity in rat hepatic transcriptomic responses to 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD). Toxicol Appl Pharmacol 260:135–145

Yeager RL, Reisman SA, Aleksunes LM, Klaassen CD (2009) Introducing the "TCDD-inducible AhR-Nrf2 gene battery". Toxicol Sci 111:238–246

Zeeberg BR, Qin H, Narasimhan S, Sunshine M, Cao H, Kane DW, Reimers M, Stephens RM, Bryant D, Burt SK et al (2005) High-Throughput GoMiner, an 'industrial-strength' integrative gene ontology tool for interpretation of multiple-microarray experiments, with application to studies of Common Variable Immune Deficiency (CVID). BMC Bioinform 6:168

Zhang B, Duan S, Shi J, Jiang S, Feng F, Shi B, Jia Z (2018) Family-based study of association between MAFB gene polymorphisms and NSCL/P among Western Han Chinese population. Adv Clin Exp Med 27:1109–1116

Affiliations

Stephenie D. Prokopec[1] · Aileen Lu[1,2] · Sandy Che-Eun S. Lee[1] · Cindy Q. Yao[1] · Ren X. Sun[1] · John D. Watson[1] · Rabah Soliymani[3] · Richard de Borja[1] · Ada Wong[4] · Michelle Sam[4] · Philip Zuzarte[4] · John D. McPherson[4] · Allan B. Okey[2] · Raimo Pohjanvirta[5,6] · Paul C. Boutros[1,2,7,8,9,10,11]

1 Computational Biology, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada

2 Department of Pharmacology and Toxicology, University of Toronto, Toronto, ON M5S 1A8, Canada

3 Department of Biochemistry and Developmental Biology, Meilahti Clinical and Basic Proteomics Core Facility, University of Helsinki, FI-00790 Helsinki, Finland

4 Genome Technologies Program, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada

5 Laboratory of Toxicology, National Institute for Health and Welfare, FI-70701 Kuopio, Finland

6 Department of Food Hygiene and Environmental Health, University of Helsinki, FI-00790 Helsinki, Finland

7 Department of Medical Biophysics, University of Toronto, Toronto, ON M5G 1L7, Canada

8 Department of Human Genetics, University of California, Los Angeles, CA 90095, USA

9 Department of Urology, University of California, Los Angeles, CA 90095, USA

10 Institute for Precision Health, University of California, Los Angeles, CA 90095, USA

11 Jonsson Comprehensive Cancer Center, University of California, Los Angeles, CA 90024, USA