Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

The basis for strain-dependent rat aldehyde dehydrogenase 1A7 (*ALDH1A7*) gene expression

Katerina Touloupi, Jenni Küblbeck, Angeliki Magklara, Ferdinand Molnár, Mika Reinisalo, Maria Konstandi, Paavo Honkakoski and Periklis Pappas

Department of Pharmacology, Faculty of Medicine, School of Health Sciences, University of Ioannina, 451 10 Ioannina, Greece (K.T., M.K., P.P.)
School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, FI-70210 Kuopio, Finland (J.K., F.M., M.R., P.H.)
Department of Clinical Chemistry, Faculty of Medicine, School of Health Sciences, University of Ioannina, 451 10 Ioannina, Greece (A.M.)
Department of Biomedical Research, Institute of Molecular Biology and Biotechnology, Foundation of Research and Technology, 451 10 Ioannina, Greece (A.M.)
Department of Biology, School of Science and Technology, Nazarbayev University, Nur-Sultan City, 010000 Kazakhstan (F.M.)
Division of Pharmacotherapy and Experimental Therapeutics, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, U.S.A. (P.H.)

This work was supported by the Greek Ministry of Education, Research and Religious Affairs - Herakleitos (K.T., A.M., M.K., P.P.) and by the Academy of Finland (J.K, F.M., M.R., P.H).

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

Running title: Strain-dependent ALDH1A7 expression in the rat

Corresponding authors:

Prof. Paavo Honkakoski

Telephone: +1 919 448 4682; email phonka@email.unc.edu

Prof. Periklis Pappas

Telephone: +30 26510 07553; email ppappas@cc.uoi.gr

Number of text pages: 26

Number of Tables: 0

Number of Figures: 6

Number of references: 38

Words in Abstract: 172

Words in Introduction: 570

Words in Discussion: 987

List of non-standard abbreviations: ALDH, aldehyde dehydrogenase; bp, base pair; CAR, constitutive androstane receptor; ChIP, chromatin immunoprecipitation; CYP, cytochrome P450; DTT, dithiothreitol; kbp, kilobase pair; NR, nuclear receptor; PB, phenobarbital; PCN, pregnenolone 16α-carbonitrile; PMSF, phenylmethyl sulfonyl fluoride; PXR, pregnane X receptor; RU486, mifepristone; RXR, retinoid X receptor; SDS, sodium dodecyl sulfate; TCPOBOP, 1,4-bis[(3,5-dichloropyridyloxy)]benzene; TF, transcription factor.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

Abstract

Aldehyde hydrogenases (ALDHs) belong to a large gene family involved in oxidation of both endogenous and exogenous compounds in mammalian tissues. Among ALDHs, the rat *ALDH1A7* gene displays a curious strain-dependence in phenobarbital (PB)-induced hepatic expression: the responsive ***RR*** strains exhibit induction of both ALDH1A7 and CYP2B mRNAs and activities while the non-responsive ***rr*** strains show induction of CYP2B only. Here, we investigated the responsiveness of *ALDH1A1*, *ALDH1A7*, *CYP2B1* and *CYP3A23* genes to prototypical CYP inducers, expression of nuclear receptors CAR and PXR, and structure of the *ALDH1A7* promoter in both rat strains. ALDH1A7 mRNA, associated protein and activity were strongly induced by PB and modestly by pregnenolone 16α-carbonitrile in the ***RR*** strain but negligibly in the ***rr*** strain while induction of ALDH1A1 and CYP mRNAs was similar between the strains. Reporter gene and chromatin immunoprecipitation assays indicated that the loss of *ALDH1A7* inducibility in the ***rr*** strain is profoundly linked with a 16-bp deletion in the proximal promoter and inability of the upstream DNA sequences to recruit CAR-retinoid X receptor heterodimers.

Significance statement

Genetic variation in rat *ALDH1A7* promoter sequences underlie the large strain-dependent differences in expression and inducibility by phenobarbital of the aldehyde dehydrogenase activity. This has implications for the design and interpretation of pharmacological and toxicological studies on the effects and disposition of aldehydes.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

Introduction

The aldehyde dehydrogenases (ALDHs) are present in all taxonomic groups and they are involved in the irreversible oxidation of endogenous and exogenous aldehydes [Sophos et al., 2001]. Their functionality extends far beyond enzymatic detoxification processes, because they participate in growth, differentiation and survival of cells and in synthesis of biomolecules such as retinoic acid, betaine and γ-aminobutyric acid. ALDHs may also serve as biomarkers for certain cancer types and Parkinson's disease [Westerlund et al., 2005; Jackson et al., 2011]. Among the 21 genes of the rat *ALDH* gene superfamily [Jackson et al., 2011], the cytosolic isoforms ALDH1A1, 2, 3 and 7 are involved in the metabolism of *e.g.* aliphatic aldehydes, retinal and cyclophosphamide [Kathmann et al., 2000; Alnouti and Klaassen, 2008]. Among these isoforms, rat ALDH1A1 and ALDH1A7 are highly homologous and display similar tissue expression patterns with high mRNA levels reported in the liver, lung and kidneys [Hsu et al., 1999; Kathmann et al., 2000; Marselos et al., 1987]. Moreover, hepatic expression of rat ALDH1A1 and 1A7 isozymes is induced by phenobarbital (PB), a prototypical inducer of several cytochrome P450 (CYP) enzymes.

In rodents, PB is known to induce many CYPs and other drug-metabolizing enzymes and transporters via the nuclear receptor (NR) constitutive androstane receptor (CAR) while many drugs induce their expression by activation of the pregnane X receptor (PXR) [Honkakoski and Negishi, 2000; Chai et al., 2013]. Others reported that selective CAR or PXR activators modestly increased (2–3.5-fold) the *Aldh1a1* and *Aldh1a7* gene expression in livers of wild-type mice but not in mice with disrupted *Car* or *Pxr* genes, while the expression of other *Aldh* family members was not appreciably affected by NR activation [Aleksunes and Klaassen, 2012]. These findings indicate that the key NRs regulating drug-metabolizing enzymes are involved in the regulation of cytosolic ALDH1A isoforms.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

The induction of cytosolic ALDH by PB has been investigated due to genetic strain difference in specific Long-Evans rat strains [Deitrich, 1971; Dunn et al., 1989; Kathmann et al., 2000]. The extent of PB-elicited induction of rat liver cytosolic ALDH activity appears to be regulated by two alleles of a single autosomal locus. The dominant allele ***R*** confers responsiveness to PB while the recessive allele ***r*** is non-responsive. Homozygous responders (***RR***) exhibit maximal ALDH induction by PB, heterozygous rats (***Rr***) display a moderate induction and homozygous non-responders (***rr***) show only a small, if any, increase of ALDH activity. Constitutive ALDH mRNA expression (initially termed as ALDH-PB, now recognized as ALDH1A7) is detectable in both strains but it is strongly inducible only in the ***RR*** rats. Interestingly, the expression of other PB-inducible enzymes such as CYP2B1 and CYP2B2 reached similar levels in these strains after exposure to PB [Dunn et al., 1989; Pappas et al., 2001]. However, the molecular basis for this ALDH-selective disparity in induction by PB remains unexplored to this day. To investigate this problem, we utilized the Wistar/Af/Han/Mol/Kuo/Io rat strain that is currently bred and maintained at University of Ioannina. This Wistar rat strain also exhibits discordant patterns of *ALDH1A* expression similar to the Long-Evans rat strains [Pappas et al., 1998; Pappas et al., 2001]. The basal ALDH1A activity is ~2-fold greater and its PB inducibility is markedly higher in the ***RR*** strain (>15-fold *versus* 2.5-fold) as compared to the ***rr*** strain. Here, we cloned and characterized *ALDH1A7* from both Wistar rat strains to identify the defect in PB induction and investigated its regulation by CAR and PXR.

Materials and methods

Chemicals. Phenobarbital (PB) and mifepristone (RU486) were from Sigma-Aldrich (St. Louis, MO). Pregnenolone 16α-carbonitrile (PCN) was bought from SantaCruz Biotechnology (Heidelberg, Germany). The synthesis of the mouse CAR agonist 1,4-bis[(3,5-

dichloropyridyloxy)]benzene (TCPOBOP) has been previously described [Honkakoski et al., 1996]. All other chemicals were of the highest analytical grade.

Experimental animal model and treatments. Wistar/Af/Han/Mol/Kuo/Io male rats (aged 3-5 months and weighing 170-230 g) were isolated and reproduced by inbreeding at the University of Ioannina, creating substrains that are responsive (***RR***) and non-responsive (***rr***) to PB [Marselos, 1976]. All animals were treated according to the European Communities Council Directive (86/609/EEC), and housed in groups of two or three in plastic cages under constant environmental conditions, on a 12-h light/dark cycle at ambient temperature and provided with chow and tap water *ad libitum*. All *in vivo* experiments were reviewed and approved by the Institutional Animal Care and Use Committee of the Medical School at the University of Ioannina (license number EL33BIO-exp01). For the multiple-dose study, three-month old male rats (four per group) were injected intraperitoneally with selective activators of CAR and PXR [Stanley et al., 2006]: three doses of PB (80 mg/kg, dissolved in saline), four doses of PCN (50 mg/kg, dissolved in olive oil) or the corresponding vehicle. Animals were sacrificed by carbon asphyxiation six hours following the PB injection. For the time-response study, three-month old male rats were divided into groups of four and treated with a single dose of PB (80 mg/kg), PCN (50 mg/kg) or vehicle. Animals were sacrificed by carbon asphyxiation following 6, 12 or 24 hours of PB or PCN injection.

Total RNA isolation, cDNA synthesis and quantitative RT-PCR. After decapitation, small liver samples were stored in RNAlater® at -80ºC until use. Total RNA was isolated with the Nucleospin® RNA II total RNA isolation kit (Macherey-Nagel, Düren, Germany). RNA samples (1 μg) served as templates for cDNA synthesis using Quanti-Tect® Reverse Transcription kit (Qiagen, Hilden, Germany). Analysis of liver mRNA levels was performed

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

with real-time RT-PCR, based on the TaqMan protocol, by use of KAPA Probe Fast qPCR Master Mix (KAPA Biosystems, Wilmington, MA) and TaqMan® Gene Expression Assay primers for rat ALDH1A1, ALDH1A7, CAR, PXR, CYP2B1, CYP3A23 and β-actin genes (Applied Biosystems, Foster City, CA). Amplification reactions (two replicate reactions per sample) were performed on the CFX96™ real-time system and the C1000™ thermal cycler (Bio-Rad Laboratories, Hercules, CA). Gene expression analysis was performed by the ΔΔCt method (Müller et al., 2002), normalized for beta-actin mRNA levels and expressed as mean ± standard deviation (N = 4) relative to the value of control ***RR*** strain sample set as 1.

Protein extraction and Western blotting. Liver samples were placed into ice-cold PBS and then homogenized either in hypotonic buffer (10 mM Hepes pH 7.9, 1.5 mM $MgCl_2$, 5 mM KCl with 0.1% NP-40) for cytosolic proteins, in high-salt extraction buffer (20 mM Hepes pH 7.9, 25% glycerol, 450 mM KCl, 1 mM EDTA) for nuclear proteins, or in RIPA buffer [50 mM Tris-HCl pH 7.5, 150 mM NaCl, 1% Triton X-100, 1% Na-deoxycholate, 0.1% sodium dodecyl sulfate (SDS)] for total proteins. Protease and phosphatase inhibitors [1 μg/ml aprotinin, 1 μg/ml leupeptin, 50 μg/ml phenylmethyl sulfonyl fluoride (PMSF), 50 μg/ml Na_3VO_4, 1 mM dithiothreitol (DTT)] were added to the above buffers just prior to use. Proteins (50–100 μg) were separated in 7.5–12% SDS-polyacrylamide gels and transferred onto a nitrocellulose membrane (Protran, Schleicher & Schuell, Dassel, Germany). Membranes were first blocked in 5% milk in Tris-buffered saline and then incubated with anti-rabbit ALDH1/2 antibody (1:500; sc-50385, Santa Cruz; likely to detect at least ALDH1A1 and 1A7 isoforms that have similar molecular masses of 54 kDa), anti-mouse CAR (1:1000; PP-N4111-00, Perseus Proteomics; Tokyo, Japan), anti-mouse PXR (1:1000; PP-H4417-00, Perseus Proteomics). Goat anti-rabbit (sc-2004; Santa Cruz) and goat anti-mouse (sc-2005; Santa Cruz) horseradish peroxidase-conjugated antibodies were used as

secondary antibodies at 1:5000 dilution. Immunodetection of bands was based on enhanced chemiluminescence kit (GE Healthcare, Little Chalfont, UK). Three independent sets of control, PB- or PCN-exposed rats were analyzed. Typical results from one set are shown in Figures 1 and 2, and results of the remaining two sets are shown in the Supplementary Figures 2 and 3.

ALDH1 activity assay. Liver samples were homogenized in ice-cold sucrose and centrifuged at 20, 000 x *g* for 30 minutes. Supernatants were used for the kinetic measurement of ALDH activity with propionaldehyde as substrate as described [Vasiliou and Marselos, 1989]. This activity cannot distinguish between the cytosolic ALDH isoforms 1A1, 2, 3 and 7. Determination of the protein concentration was carried out according to the Lowry method [Lowry, 1951]. The specific activities are expressed as mean ± standard deviation of four animals per group.

Cloning of the ALDH1A7 promoter and reporter constructs. Genomic DNA was isolated from liver tissue (25 mg) from three-month old male ***RR*** and ***rr*** rats according to the protocol of NucleoSpin® Tissue kit (Macherey-Nagel). The amplification of ~10 kbp fragments of *ALDH1A7* promoters (Rnor 5.0; Chr1: 247,897,931 - 247,887,931; reverse complement) from genomic DNA and sequencing was done by Epoch Life Science Inc. (Missouri City, TX). The ***RR*** and ***rr*** promoter sequences have been submitted to the NCBI GenBank with the accession numbers MK814117 and MK814118, respectively. The fragments were inserted into pGL3-Basic plasmid vectors (Promega, Madison, WI) at *Kpn*I and *Xho*I restriction sites. Sequential deletion fragments of *ALDH1A7* promoters were generated by amplification with Phusion high-fidelity DNA polymerase (Finnzymes, Espoo, Finland) or KAPA HiFi (Hot start) DNA polymerase (KAPA Biosystems) and forward and reverse primers (Oligomer,

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

Helsinki, Finland) shown in Supplementary Table 1. The deletion fragments were cloned into pGL3-Basic plasmid vectors and competent *E. coli* bacteria (XL-10 or DH5α) were transformed with the constructs. Colonies were screened by colony PCR using DreamTaq DNA polymerase (ThermoScientific, Vantaa, Finland). Plasmid DNAs from positive colonies were purified with QIAprep® Spin Miniprep kit (Qiagen) and verified by dideoxy sequencing (DNA Analysis Facility, Yale University, USA).

Reporter gene assays. The full-length mouse CAR and PXR constructs [Mäkinen et al., 2002] and the respective reporters (for CAR, PBREM-tk-luc; for PXR, XREM-3A4-luc) [Honkakoski et al., 2001; Goodwin et al., 1999] have been described earlier. C3A hepatoma cells were seeded on 48-well plates (~180,000 cells/cm^2). The cells were transfected with the calcium phosphate method [Küblbeck et al., 2008] with the appropriate combinations of the following three plasmids: one of the NR expression plasmids [full-length CAR, full-length PXR or the empty expression vector (100 ng/well)], one of the reporter plasmids [individual *ALDH1A7* promoter construct, negative control pGL3-Basic plasmid, positive control PBREM-tk-luc or XREM-3A4-luc reporters (450 ng/well)], *plus* the transfection control plasmid pCMVβ (900 ng/well). Untransfected cells served as blank samples for reporter enzyme assays. After transfection for four hours, the medium was removed and fresh DMEM supplemented with 5% delipidated serum (HyClone, Logan, UT) was added onto the transfected cells. The added medium also contained either the vehicle (DMSO, 0.1%) or activating NR ligands (1 μM TCPOBOP for mouse CAR; 10 μM RU486 for mouse PXR). Following the treatment with chemicals for 24 hours, the cells were lysed and assayed for luciferase and β-galactosidase activities as before [Honkakoski et al., 2001]. Blank-subtracted luciferase activities were normalized, *i.e.* divided, by blank-subtracted β-

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

galactosidase activities and expressed relative to the empty pGL3-Basic reporter set at 1. Data are mean ± standard deviation from three independent transfections.

Chromatin immunoprecipitation (ChIP) assay. ChIP assays were performed according to Magklara and Smith [2009] with some modifications. Male three-month-old ***RR*** and ***rr*** rats (two per condition) were injected with PB (80 mg/kg) or saline and killed by carbon dioxide asphyxiation three hours later. Liver samples (80-100 mg) were fixed in 1% formaldehyde solution and the cross-linking reaction was quenched by incubating the samples in 0.125 M glycine. The samples were then homogenized in sucrose buffer (0.3 M sucrose, 15 mM NaCl, 5 mM $MgCl_2$, 0.1 mM EDTA, 15 mM Tris-HCl, 0.5mM DTT, 0.1 mM PMSF) and centrifuged. The pelleted homogenates were resuspended in lysis buffer (50 mM Tris-HCl pH 8, 10 mM EDTA, 1% SDS) and incubated on ice for 10 minutes. Lysates were centrifuged and resuspended in sonication buffer (50 mM Hepes pH 7.5, 140 mM NaCl, 1 mM EDTA, 1% Triton-X, 0.1% Na-deoxycholate, 0.1% SDS). After sonication (Misonix sonicator, 80% power, 4 x 20 seconds with one-minute cooling intervals), the DNA was sheared to an average length of 300–800 bp. The supernatants were diluted in ChIP dilution buffer (1.2 mM EDTA, 167 mM NaCl, 16.7 mM Tris-HCl pH 8, 1.1% Triton X-100, 0.01% SDS) and pre-cleared with Protein A/G PLUS-Agarose beads (sc-2003; Santa Cruz) at 4 ^{o}C. Then, the beads were removed and one-tenth of the genomic DNA was retained as input control for quantitative PCR analysis. Pre-cleared chromatin samples were incubated at 4^{o}C overnight, with the addition of 5 μg polymerase II antibody (sc-899; Santa Cruz) or 5 μg of retinoid X receptor-α (RXRα) antibody (sc-553; Santa Cruz). On the following day, protein-DNA complexes were precipitated after a three-hour incubation with A/G beads. The beads were washed successively with low-salt buffer (2 mM EDTA, 20 mM Tris-HCl pH 8, 0.1% SDS, 1% Triton X-100, 150 mM NaCl), high-salt buffer (low salt buffer with 500 mM NaCl), LiCl

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

wash buffer (1 mM EDTA, 10 mM Tris-HCl pH 8, 1% Nonidet P-40, 1% Na-deoxycholate, 0.25 M LiCl) and twice with TE buffer (1 mM EDTA, 10 mM Tris-HCl pH 8). All buffers contained 1 X protease inhibitor cocktail (P-27140; Sigma). After the two elution steps, the eluates and inputs were incubated with 5 M NaCl in 65^{o}C overnight to reverse the DNA/protein cross-links. After treatment with proteinase K (two hours at 55^{o}C), DNA was purified by phenol–chloroform-isoamyl alcohol extraction and ethanol precipitation. The DNA samples were analyzed by quantitative PCR in duplicate using KAPA SYBR® FAST qPCR Master Mix(KK4601), and calculated by the $2^{\wedge-\Delta Ct}$ method after correcting for the IgG negative control. The PCR primers used were synthesized by Eurofins Genomics (Ebersberg, Germany) and listed in the Supplementary Table 2.

Statistical analysis. Differences between groups were evaluated by analysis of variance and Bonferroni's multiple comparison test. The main comparisons were between the strains and between the control and inducer groups. Probability values $p < 0.05$ were considered statistically significant, but they are shown for descriptive purposes, not for hypothesis testing.

Results

Differential expression of ALDH1A7 in RR and rr rat strains. Hepatic expression of *ALDH1A1* and *ALDH1A7* genes was studied following multiple doses of PB and PCN. Figure 1A shows that constitutive expression of rat ALDH1A7 mRNA was higher in the ***RR*** while still detectable in the ***rr*** rat strain. ALDH1A7 mRNA levels were induced over 100-fold by PB and almost 20-fold by PCN only in the responsive ***RR*** strain, while no equivalent induction was detected in the ***rr*** strain. On the other hand, constitutive rat ALDH1A1 mRNA expression was only slightly higher in the ***rr*** strain while the induction of ALDH1A1 mRNA

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

was comparable, reaching 4- to 10-fold by PB and 4- to 6-fold by PCN in both strains (Figure 1B). Western blotting of cytosolic and total protein extracts (Figure 1C and Supplementary Figure 2) indicated low expression of ALDH1 proteins in control animals while high expression was seen in PB-treated and moderate induction in PCN-treated ***RR*** rats only. It should be noted that the antibody used in Western blotting was the only one commercially available and it does not distinguish between various ALDH1A isoforms. The ALDH1 (propionaldehyde/NAD^+) activity (Figure 1D), catalyzed by all ALDH1A isozymes [Lindahl, 1992; Alnouti and Klaassen, 2008], displayed similar basal levels between the strains. It was elevated 3.5- and 4.5-fold by PB and PCN in the responsive ***RR*** rats but only marginally in the ***rr*** strain. These findings confirmed earlier reports on PB induction of both ALDH1A1 and 1A7 isoforms, and the strain difference in ALDH1A7 inducibility by PB [Dunn et al., 1989; Pappas et al., 2001].

Expression of inducible CYPs and NRs in RR and rr rats. Similar analyses revealed that hepatic transcripts of the main rat PB- and PCN-inducible *CYP2B1* and *CYP3A23* genes [Honkakoski and Negishi, 2000; Stanley et al., 2006] are detected in the untreated ***RR*** and ***rr*** animals at similar levels and are similarly induced by PB and PCN (Figure 2A, 2B). PB, an activator of rodent CAR, preferentially induced CYP2B1 over CYP3A23 mRNA. PCN, a ligand of rodent PXR, induced CYP3A23 over CYP2B1 mRNA expression, as expected. Furthermore, there was maximally only about a two-fold change in the Ct values of β-actin mRNA between the control and induced groups (Supplementary Table 5), indicating that the observed strain differences and induction responses are not due to changes in the expression of the control gene used in normalization.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

Furthermore, the expression of rat CAR and PXR mRNA (Figure 2C) and protein levels in total liver extracts (Figure 2D and Supplementary Figure 3) were also comparable. In time-course experiments, induction of ALDH1A1, ALDH1A7 (***RR*** strain only), CYP2B1 and CYP3A23 mRNAs was clearly detectable at 6 hours following PB or PCN injection, indicating a rapid transcriptional response (Supplementary Figures 4 and 5). These results show that CAR and PXR receptors are present in both rat strains and function appropriately as regulators of *ALDH1A, CYP2B* and *CYP3A* genes.

Comparison of ALDH1A7 promoter sequences between RR and rr rats. The remarkable absence of ALDH1A7 induction by PB but normal expression and functionality of NRs in ***rr*** rats suggested that the genetic difference in inducibility between the ***RR*** and ***rr*** strains may be related to changes in DNA sequences of their respective *ALDH1A7* genes. To study this hypothesis, genomic DNA from both strains was isolated and ~10 kbp of *ALDH1A7* promoter fragments were amplified and sequenced. Alignment of the two promoter sequences showed that they are 97.6% identical (Supplementary Figure 1), with the *ALDH1A7* sequence from the ***rr*** strain matching completely the available rat *ALDH1A7* sequence for *Rattus norvegicus* (ALDH1A7; Chr1:247,847,960-247,887,931 in the Rnor 5.0 assembly). Detailed comparison of the two *ALDH1A7* promoter sequences indicated that there were two larger deletions in the ***rr*** strain compared to the ***RR*** strain: one was a 165-bp region at ~2,200 bp upstream of the transcription start site (TSS), missing in the ***rr*** strain. Another deletion (16-bp) occurred just upstream of the TATA box at -50/-34bp. In addition, several small deletions and point mutations were present (Supplementary Figure 1). With regard to transcription factor (TF) sites that could influence PB or PCN inducibility and/or hepatic expression of ALDH1A7 mRNA, we identified binding sites for CAR/RXRα sites, CCAAT box and DR3, DR4 and ER6 elements, which are shown in Supplementary Tables 3 and 4, respectively. However,

none of them showed a major difference between the ***RR*** and ***rr*** strains, and none were present in the above two deletion regions.

Activity of the full-length ALDH1A7 promoters from RR and rr rats. The 10-kbp *ALDH1A7* gene promoters were inserted into the pGL3-Basic reporter plasmid, co-transfected with mouse NR expression vectors into C3A hepatoma cells, and cells were treated with selective NR-activating ligands (Figure 3). CAR and PXR activated the PBREM- and XREM-driven positive control reporters as expected [Mäkinen et al., 2002]. The constitutively active CAR enhanced the activity of *ALDH1A7* promoter from the ***RR*** strain by ~55-fold and its ligand TCPOBOP further elevated the reporter activity over 80-fold compared to the pGL3-Basic control reporter, while the response to PCN-activated PXR was more modest, ~16-fold increase. However, the *ALDH1A7* promoter from the ***rr*** strain was not appreciably affected by CAR or PXR. These findings correlate well with induction of ALDH1A7 mRNA (Figure 1A), and indicate that the 10-kbp *ALDH1A7* ***RR*** promoter sequence harbors CAR- and PXR-responsive elements, while the ***rr*** promoter cannot be activated by these NRs.

Activity of the proximal ALDH1A7 promoters from RR and rr rats. Because of the low ALDH1A7 expression in the ***rr*** strain, we turned our attention next to the proximal *ALDH1A7* promoter that contained several TF binding sites, likely important for the constitutive activity, and the 16-bp deletion in the ***rr*** promoter. We measured the reporter activity and responsiveness to NRs of two proximal promoter fragments from both strains. Figure 4A shows that the longer (-1571/+5bp)) ***RR*** fragment conferred a reporter activity of 12.5-fold over the promoterless control construct. Deletion to shorter ***RR*** fragment (-461/+5 bp) yielded an activity of about 30% of the longer fragment. The corresponding *ALDH1A7* fragments from the **rr** strain had much weaker activities that tended to be on par or

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

marginally above the pGL3-Basic construct. This indicates that the proximal *ALDH1A7* promoter (-444/+5bp) in the ***rr*** strain contains defects that reduced the promoter activity and presumably decreased the expression of *ALDH1A7* gene.

Figure 4B shows results from the co-transfection assays with NR expression vectors and the proximal *ALDH1A7* fragments. In contrast to the 10-kbp ***RR*** promoter (Figure 3), the more proximal fragments (***RR*** -1571/+5bp; ***rr*** -1556/+5 bp, or shorter) from either rat strain were not activated by ligand-activated CAR or PXR. This indicates that the CAR/PXR -responsive element(s) lie in the far-upstream region of *ALDH1A7*.

Binding of transcriptional regulators to the ALDH1A7 promoter sequences. To shed more light on differential expression of ALDH1A7 between ***RR*** and ***rr*** rats, we studied the binding of RNA polymerase II, an essential factor for RNA transcription, to the proximal *ALDH1A7* promoter, and binding of NRs to the upstream regions. To this end, we utilized the ChIP assay with chromatin samples extracted from both strains exposed to vehicle or PB.

We detected significant RNA polymerase II binding (>6-fold over background) to the proximal promoter region (-209/-12bp) in samples from control ***RR*** rats. This binding was clearly elevated by PB (Figure 5). In contrast, RNA polymerase II showed weak or no binding to the corresponding fragment (-194/-12bp) from ***rr*** rats, regardless of treatment. This provides strong support to our hypothesis that *ALDH1A7* gene has an active and PB-responsive promoter in ***RR*** rats, while its activity is low and non-inducible in the ***rr*** strain.

Next, we sought to identify the locations of the CAR binding sites in the *ALDH1A7* far-upstream region. Because the available CAR antibody was not suitable for ChIP assays, we

used instead an antibody to RXRα, the obligate heterodimer partner of CAR and PXR for DNA binding [Honkakoski and Negishi, 2000]. A similar approach of using RXRα as a surrogate for CAR binding was taken in previous studies [Saito et al., 2013; Ohno et al., 2014]. Inspection of NR binding sites with direct or everted repeats of AGGTCA-like motifs with the NUBIScan algorithm [Podvinec et al., 2002] revealed three putative CAR/RXRα binding sites at -3075 (named CAR1), at -3331 (named CAR2) and -5202 (named CAR3) for the ***RR*** strain (the respective locations are -2907, -3167 and -5046 for the ***rr*** strain). Sequence comparison of these sites between the ***RR*** and ***rr*** strains showed no major differences, apart from a single point mutation in the CAR-RXRα binding CAR1 site (a DR4 element). There were also three other putative CAR/RXRα binding sites (upstream from -8300 in ***RR***). However, they have shown no significant ChIP binding (*data not shown*).

In Figure 6A, binding of RXRα was detectable to all three CAR/RXRα binding sites in chromatin samples isolated from control ***RR*** rats. Exposure to PB substantially enhanced RXRα binding to CAR1 and CAR2 sites (5- and 10-fold, respectively), while the increase was lower at the CAR3 site (2-fold). Experiments performed with chromatin from the ***rr*** rats (Figure 6B) showed low binding both in control and PB-treated samples.

Discussion

The present study provides evidence for the molecular basis underlying the differences in PB inducibility of ALDH activity observed in ***RR*** and ***rr*** rat strains. We showed that the low ALDH activity in the non-responsive ***rr*** strain is associated with the lack of ALDH protein and mRNA expression. This difference cannot be attributed to defects in the PB induction process itself, because expression of the CAR regulator and inducibility of the ALDH1A1, CYP2B1 and CYP3A23 mRNAs were comparable between the two strains. We also showed,

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

for the first time, that the rat *ALDH1A7* gene is regulated by PXR and its ligand, albeit to a lesser extent than by CAR.

Reporter gene assays using *ALDH1A7* promoter fragments indicated that while the largest, 10-kbp fragment from the ***RR*** rats was strongly activated by CAR and PXR, the corresponding ***rr*** fragment was practically inactive. Other experiments indicated that the proximal ***rr*** promoter is clearly weak and cannot recruit RNA polymerase as efficiently as the ***RR*** promoter. Apart from a few dispersed point mutations and deletions at positions that did not seem to affect any predicted TF binding site, there is a 16-bp ***RR*** sequence at -50/-34 bp, very close to the predicted TATA box, that is lacking in ***rr*** rats. This 16-bp sequence is quite similar to a recognition element of the general transcription factor TFIIB [Littlefield et al., 1999] and this site is essential for recruitment of the RNA polymerase II. Accordingly, RNA polymerase II failed to bind to the fragment -194/-12bp in chromatin samples from the ***rr*** rats. Based on the present findings, we suggest that the 16-bp deletion is the ultimate reason for the low ALDH1A7 activity and inducibility by PB in the non-responsive ***rr*** rats.

Upstream regions at -3050 and -3300 bp harbored sites capable of NR binding as indicated by efficient recruitment of RXRα, which is essential for the CAR-dependent transcription. While robust RXRα binding to the ***RR*** gene took place especially after PB administration, binding to the ***rr*** gene was detected only at low levels and not affected by PB. This is quite surprising given the fact that the NR binding sites in the ***rr*** gene were intact. This finding suggests that CAR binding to these distal sites may require priming, *i.e*. appropriate recruitment of factors binding to the proximal *ALDH1A7* promoter before CAR binding can occur. This implies that the inability of the proximal *ALDH1A7* ***rr*** promoter to recruit general TFs may also result in reversion of distal CAR binding sites into a transcriptionally inactive state in the ***rr*** rats.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

There is similar evidence that mutations in a single regulatory sequence can affect the accessibility to TFs of the entire gene locus [Kumasaka et al., 2016].

The genome-wide chromatin accessibility and inducibility by TCPOBOP has recently been investigated in mice. The *Cyp2b10* and *Aldh1a7* genes belong to a group termed 'Pattern 1' where TCPOBOP produces stronger induction than PCN [Cui and Klaassen, 2016]. In the *Cyp2b10* gene upstream sequence, there is low but detectable CAR/RXR binding in control animals which is enhanced by 6- to 8-fold after TCPOBOP exposure [Tian et al., 2018]. These data correlate well with our findings on the ***RR*** *ALDH1A7* gene. The murine *Aldh1a7* locus contains an active promoter, as evidenced by activating histone H3K4me3 and H3K27Ac marks [Sugathan and Waxman, 2013] and a TCPOBOP-responsive DNase I hypersensitive site [Lodato et al., 2018] overlapping the TSS, a finding that matches well with the active rat *ALDH1A7* proximal promoter in ***RR*** rats. The far-upstream DNase I hypersensitive sites are located at ~17 and ~29 kbp upstream of the *Aldh1a7* TSS but still within the same topologically associated domain [Lodato et al., 2018].

In contrast, information on genome chromatin accessibility or epigenetics in the rat liver is quite sparse. The available data [https://www.ebi.ac.uk/arrayexpress: accession number E-MTAB-2663] shows that the rat *ALDH1A1* gene possesses strong activating histone marks (H3K4me3 and H3K27Ac) that overlap the *ALDH1A1* TSS and the first exon [Villar et al., 2015]. The co-occurrence of these two histone modifications is indicative of an active promoter, in line with a robust ALDH1A1 mRNA expression in both ***rr*** and ***RR*** strains. However, in the same dataset, coinciding H3K4me3 and H3K27Ac marks closest to the *ALDH1A7* gene are more than 200 kbp away from its TSS, which correlates with the low

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

expression of ALDH1A7 mRNA. It should be noted again that the rat reference genome (Rnor 5.0) used in these studies contains the ***rr*** strain *ALDH1A7* sequence.

Collectively, our results show that PB induces ALDH1A7 expression at activity, protein and mRNA level only in the responsive ***RR*** rats via binding of CAR to upstream elements in *ALDH1A7*, which is inherently associated with enhanced recruitment of RNA polymerase II to the proximal promoter. Due to sequence variation between these rat strains, the expression and accumulation of ALDH1A7 mRNA is disrupted in the non-responsive ***rr*** rats, which likely affects also other ALDH1A7-inducing signals than PB that act via transcriptional activation.

The human relevance of our findings are still unclear. *ALDH1A7* is not present in the human genome as it is thought to be a rodent-specific duplication of the ancestral *ALDH1A1* gene. There are no direct publications on PB inducibility of human *ALDH1A1* while rodent *ALDH1A1* and *1A7* genes are both activated. We found an RNA sequencing dataset of human HepaRG cells exposed to human CAR- and PXR-activating CITCO and PB (Li et al., 2015; GeoExpress GSE71446) that showed a substantial 3.5- to 16-fold increase in CYP2B6 and CYP3A4 mRNAs but only a slight increase (< 40%) in ALDH1A1 mRNA expression.

In conclusion, we consider that the present data have shed light on the 40-year conundrum of strain differences in ALDH inducibility by PB. They also suggest that metabolism of endo- and exogenous aldehydes may be attenuated in rat strains harboring this genetic defect, and this could have consequences for the design and interpretation of biological, pharmacological and toxicological studies on functions and disposition of aldehydes.

Acknowledgments

We wish to thank Ms. Lea Pirskanen for her expert help in molecular biology assays.

Authorship contributions

Participated in research design: Touloupi, Küblbeck, Magklara, Molnár, Reinisalo, Honkakoski, Pappas.

Conducted experiments: Touloupi, Küblbeck, Molnár, Reinisalo, Konstandi, Pappas.

Performed data analysis: Touloupi, Küblbeck, Magklara, Reinisalo, Honkakoski, Pappas.

Wrote or contributed to the writing of the manuscript: Touloupi, Magklara, Honkakoski, Pappas.

References

Aleksunes DM and Klaassen CD (2012) Coordinated regulation of hepatic phase I and II drug-metabolizing genes and transporters using AhR-, CAR-, PXR-, PPARα-, and Nrf2-null mice. *Drug Metab Dispos* **40:**1366-1379.

Alnouti Y and Klaassen CD (2008) Tissue distribution, ontogeny, and regulation of aldehyde dehydrogenase (Aldh) enzymes mRNA by prototypical microsomal enzyme inducers in mice. *Toxicol Sci* **101:**51-64.

Chai X, Zeng S, and Xie W (2013) Nuclear receptors PXR and CAR: implications for drug metabolism regulation, pharmacogenomics and beyond. *Expert Opin Drug Metab Toxicol* **9:**253-266.

Cui JY and Klaassen CD (2016) RNA-Seq reveals common and unique PXR- and CAR-target gene signatures in the mouse liver transcriptome. *Biochim Biophys Acta* **1859:**1198-1217.

Deitrich RA (1971) Genetic aspects of increase in rat liver aldehyde dehydrogenase induced by phenobarbital. *Science* **173:**334-336.

Dunn TJ, Koleske AJ, Lindahl R, and Pitot HC (1989) Phenobarbital-inducible aldehyde dehydrogenase in the rat. cDNA sequence and regulation of the mRNA by phenobarbital in responsive rats. *J Biol Chem* **264:**13057-13065.

Goodwin B, Hodgson E, and Liddle C (1999) The orphan human pregnane X receptor mediates the transcriptional activation of CYP3A4 by rifampicin through a distal enhancer module. *Mol Pharmacol* **56:**1329-1339.

Honkakoski P and Negishi M (2000) Regulation of cytochrome P450 (*CYP*) genes by nuclear receptors. *Biochem J* **347:**321-337.

Honkakoski P, Jääskeläinen I, Kortelahti M, and Urtti A (2001) A novel drug-regulated gene expression system based on the nuclear receptor constitutive androstane receptor (CAR). *Pharm Res***18:**146-150.

Honkakoski P, Moore R, Gynther J, and Negishi M (1996) Characterization of phenobarbital-inducible mouse *Cyp2b10* gene transcription in primary hepatocytes. *J Biol Chem* **271:**9746-9753.

Hsu LC, Chang WC, Hoffmann I, and Duester G (1999) Molecular analysis of two closely related mouse aldehyde dehydrogenase genes: identification of a role for *Aldh1*, but not *Aldh-PB*, in the biosynthesis of retinoic acid. *Biochem J* **339:**387-395.

Jackson B, Brocker C, Thompson DC, Black W, Vasiliou K, Nebert DW, and Vasiliou V (2011) Update on the aldehyde dehydrogenase gene (*ALDH*) superfamily. *Hum Genomics* 5:283-303.

Kathmann EC, Naylor S, and Lipsky JJ (2000) Rat liver constitutive and phenobarbital-inducible cytosolic aldehyde dehydrogenases are highly homologous proteins that function as distinct isozymes. *Biochemistry* **39:**11170-11176.

Kumasaka N, Knights AJ, and Gaffney DJ (2016) Fine-mapping cellular QTLs with RASQUAL and ATAC-seq. *Nat Genet* **48:** 06-213.

Küblbeck J, Jyrkkärinne J, Poso A, Turpeinen M, Sippl W, Honkakoski P, and Windshügel B (2008) Discovery of substituted sulfonamides and thiazolidin-4-one derivatives as agonists of human constitutive androstane receptor. *Biochem Pharmacol* **76:**1288-1297.

Li D, Mackowiak B, Brayman TG, Mitchell M, Zhang L, Huang SM, and Wang H (2015) Genome-wide analysis of human constitutive androstane receptor (CAR) transcriptome in wild-type and CAR-knockout HepaRG cells. *Biochem Pharmacol* **98:**190-202.

Lindahl R (1992) Aldehyde dehydrogenases and their role in carcinogenesis. *Crit Rev Biochem Mol Biol* **27:**283-335.

Littlefield O, Korkhin Y, and Sigler PB (1999) The structural basis for the oriented assembly of a TBP/TFB/promoter complex. *Proc Natl Acad Sci USA* **96:**13668-13673.

Lodato NJ, Rampersaud A, and Waxman DJ (2018) Impact of CAR agonist ligand TCPOBOP on mouse liver chromatin accessibility. *Toxicol Sci* **164:**115-128.

Lowry OH, Rosebrough NJ, Farr AL, and Randall RJ (1951) Protein measurement with the Folin phenol reagent. *J Biol Chem* **193:**265-275.

Magklara A and Smith CL (2009) A composite intronic element directs dynamic binding of the progesterone receptor and GATA-2. *Mol Endocrinol* **23:**61-73.

Mäkinen J, Frank C, Jyrkkärinne J, Gynther J, Carlberg C, and Honkakoski P (2002) Modulation of mouse and human phenobarbital-responsive enhancer module by nuclear receptors. *Mol Pharmacol* **62:** 366-378.

Marselos M (1976) Genetic variation of drug-metabolizing enzymes in the Wistar rat. *Acta Pharmacol Toxicol* **39**:186-197.

Marselos M, Strom SC, and Michalopoulos G (1987) Effect of phenobarbital and 3-methylcholanthrene on aldehyde dehydrogenase activity in cultures of HepG2 cells and normal human hepatocytes. *Chem Biol Interact* **62:**75-88.

Müller PY, Janovjak H, Miserez AR, and Dobbie Z (2002) Processing of gene expression data generated by quantitative real-time RT-PCR. *Biotechniques* **32:**1372-1374.

Ohno M, Kanayama T, Moore R, Ray M, and Negishi M (2014) The roles of co-chaperone CCRP/DNAJC7 in *Cyp2b10* gene activation and steatosis development in mouse livers. *PLoS One* **9:**e115663.

Pappas P, Stephanou P, Karamanakos P, Vasiliou V, and Marselos M (2001) Phenobarbital inducibility and differences in protein expression of an animal model. *Chem Biol Interact* **130-132:**275-283.

Pappas P, Stephanou P, Vasiliou V, and Marselos M (1998) Zoxazolamine-induced paralysis in two rat substrains: differences in hepatic drug metabolism. *Eur J Drug Metab Pharmacokinet* **23:**461-467.

Podvinec M, Kaufmann MR, Handschin C, and Meyer UA (2002) NUBIScan, an *in silico* approach for prediction of nuclear receptor response elements. *Mol Endocrinol* **16:**1269-1279.

Pustylnyak V, Yarushkin A, Kachaylo E, Slynko N,Lyakhovich V, and Gulyaeva L (2011) Effect of several analogs of 2,4,6-triphenyldioxane-1,3 on constitutive androstane receptor activation. *Chem Biol Interact* **192:**177-183.

Saito K, Moore R, and Negishi M (2013) Nuclear receptor CAR specifically activates the two-pore K^+ channel KCNK1 gene in male mouse livers, which attenuates phenobarbital-induced hepatic hyperplasia. *Toxicol Sci* **132:**151-161.

Sophos NA, Pappa A, Ziegler TL, and Vasiliou V (2001) Aldehyde dehydrogenase gene superfamily: the 2000 update. *Chem Biol Interact* **130-132:**323-337.

Stanley LA, Horsburgh BC, Ross J, Scheer N, and Wolf CR (2006) PXR and CAR: nuclear receptors which play a pivotal role in drug disposition and chemical toxicity. *Drug Metab Rev* **38:**515-597.

Sugathan A and Waxman DJ (2013) Genome-wide analysis of chromatin states reveals distinct mechanisms of sex-dependent gene regulation in male and female mouse liver. *Mol Cell Biol* **33:** 3594-3610.

Tian J, Marino R, Johnson C, and Locker J (2018) Binding of drug-activated CAR/*Nr1i3* alters metabolic regulation in the liver. *iScience* **9:**209-228.

Vasiliou V and Marselos M (1989) Tissue distribution of inducible aldehyde dehydrogenase activity in the rat after treatment with phenobarbital or methylcholanthrene. *Pharmacol Toxicol* **64:**39-42.

Villar D, Berthelot C, Aldridge S, Rayner TF, Lukk M, Pignatelli M, Park TJ, Deaville R, Erichsen JT, Jasinska AJ, Turner JM, Bertelsen MF, Murchison EP, Flicek P, Odom DT (2015) Enhancer evolution across 20 mammalian species. *Cell* **160:** 554-566.

Westerlund M, Galter D, Carmine A, and Olson L (2005) Tissue- and species-specific expression patterns of class I, III and IV Adh and Aldh1 mRNAs in rodent embryos. *Cell Tissue Res* **322:**227-236.

Footnotes

This study is dedicated to Prof. Marios Marselos on his discovery of the strain difference and subsequent work on ALDH induction in the Ioannina rat strains. This work was supported by the Greek Ministry of Education, Research and Religious Affairs - Herakleitos (K.T., A.M., M.K., P.P.) and by the Academy of Finland (J.K, F.M., M.R., P.H). The ALDH1A7 promoter sequences have been submitted to the Genbank with accession numbers MK814117 (***RR***) and MK814118 (***rr***).

Conflicts of interest

All authors declare no conflicts of interest associated with this publication.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

Figure legends

Figure 1. Strain-dependent expression and inducibility of ALDH1A isoforms. Normalized expression of ALDH1A7 (A) and ALDH1A1 (B) mRNA in control, PB- or PCN-treated ***RR*** and ***rr*** rat livers. Levels are expressed as scatter plots and means ± SD from four animals, each sample with two technical replicates. Cytosolic and total protein levels of ALDH1 protein (C) and activity (D) in control, PB- or PCN-treated ***RR*** and ***rr*** rat livers. Statistically significant differences ($p < 0.05$) as compared to the respective control group or between the strains are marked by * and # signs, respectively. The data in panels A, B and D are also shown relative to the control RR values, set at 1.0, and expressed as means ± SD of -fold change below the X axis.

Figure 2. Similar expression of CYP2Bs and NRs in both rat substrains. Normalized expression of CYP2B1 (A) and CYP3A23 (B) mRNA in control, PB- or PCN-treated ***RR*** and ***rr*** rat livers. (C) Comparison of hepatic CAR and PXR mRNA expression between control rat strains. The ratio of target gene to β-actin mRNA expression (Y axis) are shown as scatter plots and means ± SD from four animals, each sample with two technical replicates. These normalized levels are also depicted as relative to the mean ± SD of the control ***RR*** strain, set at 1.0, and indicated below the X axis. Statistically significant differences ($p < 0.05$) as compared to the respective control group or between the strains are marked by * and # signs, respectively. (D) Expression of CAR and PXR proteins in control, PB- or PCN-treated ***RR*** and ***rr*** rat livers.

Figure 3. Strain-dependent activation of *ALDH1A7* promoter constructs by NRs. Luciferase activity of the full-length **RR**-*ALDH1A7* (*grey columns*) and **rr**-*ALDH1A7* (*black columns*) constructs co-transfected with empty (-) or indicated CAR or PXR expression

vectors (+), and treated with DMSO (ligand -) or selective CAR or PXR activators (ligand +). Positive control plasmids were PBREM-tk-luc for CAR and XREM-3A4-luc for PXR (*white columns*). Data are means ± SD from three independent transfections, and expressed relative to the empty vector (set at 1.0). (*) denotes a statistically significant difference from the empty vector at $p < 0.05$.

Figure 4. Basal and NR-elicited activation of proximal *ALDH1A7* promoter constructs. (A) Basal luciferase reporter activity from **RR**-*ALDH1A7* and **rr**-*ALDH1A7* longer (-1571/+5 bp) and shorter (-461/+5 bp) promoter fragments. (#) denotes a statistically significant different from the pGL3-Basic construct at $p < 0.05$; (B) NR-dependent luciferase activity from **RR**-*ALDH1A7* and **rr**-*ALDH1A7* longer and shorter promoter fragments in co-transfection assays with empty or indicated NR expression vectors. (ns) denotes no statistically significant differences between constructs with empty or NR expression vectors.

Figure 5. Polymerase II binding to RR-*ALDH1A7* and rr-*ALDH1A7* proximal promoters. Both strains were dosed with saline (C) or PB for three hours, liver nuclear fractions were isolated and cross-linked, and RNA polymerase-associated DNA was analysed by ChIP assay as detailed in Materials and Methods. The amount of recovered DNA is shown relative to the negative control (nc) region (set at 1.0). Data are scatter plots from two independent experiments, each sample with two technical replicates.

Figure 6. RXRα binding to putative CAR binding sites in the distal regions of RR-*ALDH1A7* (A) and rr-*ALDH1A7* (B) promoters. Both strains were dosed with saline (C) or PB for three hours, liver nuclear fractions were isolated and cross-linked, and RXRα-associated DNA was analysed by ChIP assay as detailed in Materials and Methods. The

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

amount of recovered DNA is shown relative to the negative control region (set at 1.0). Data are scatter plots from two independent experiments, each sample with two technical replicates.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

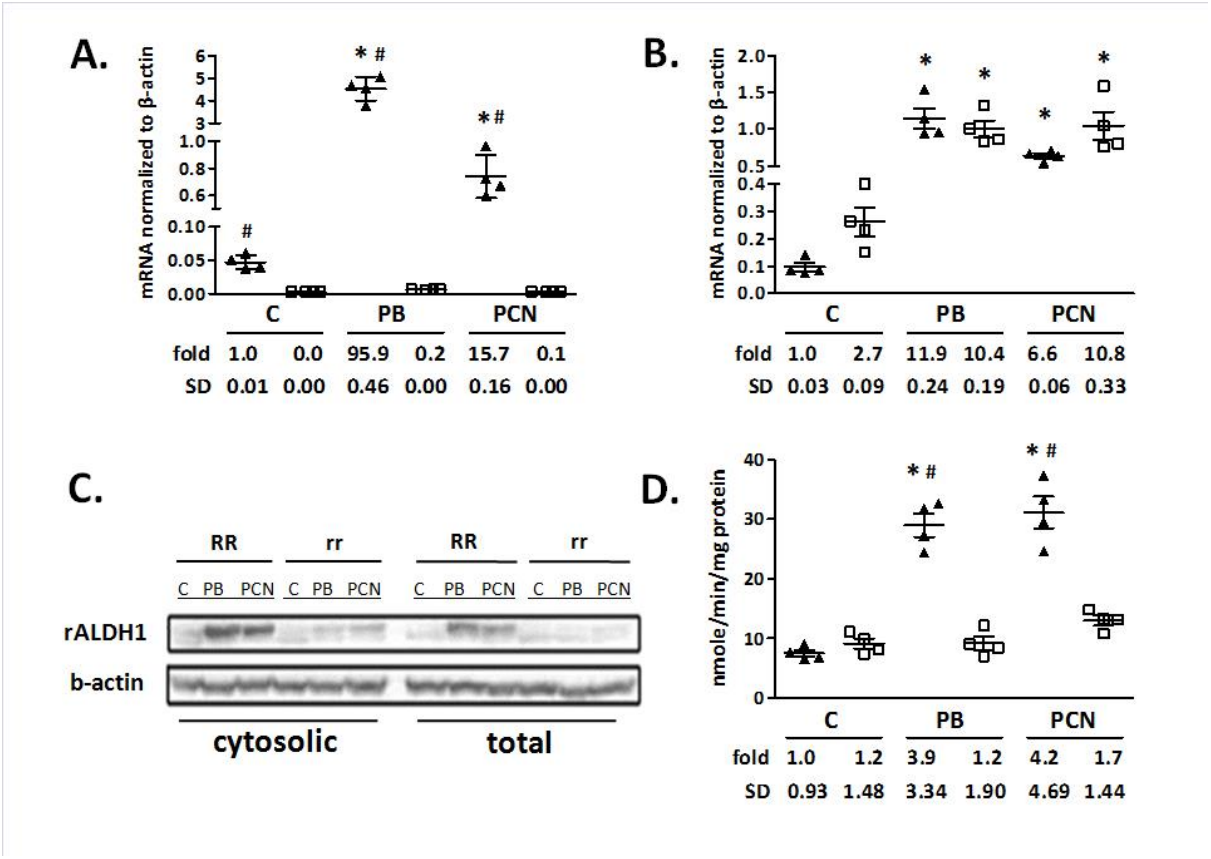


Figure 1.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

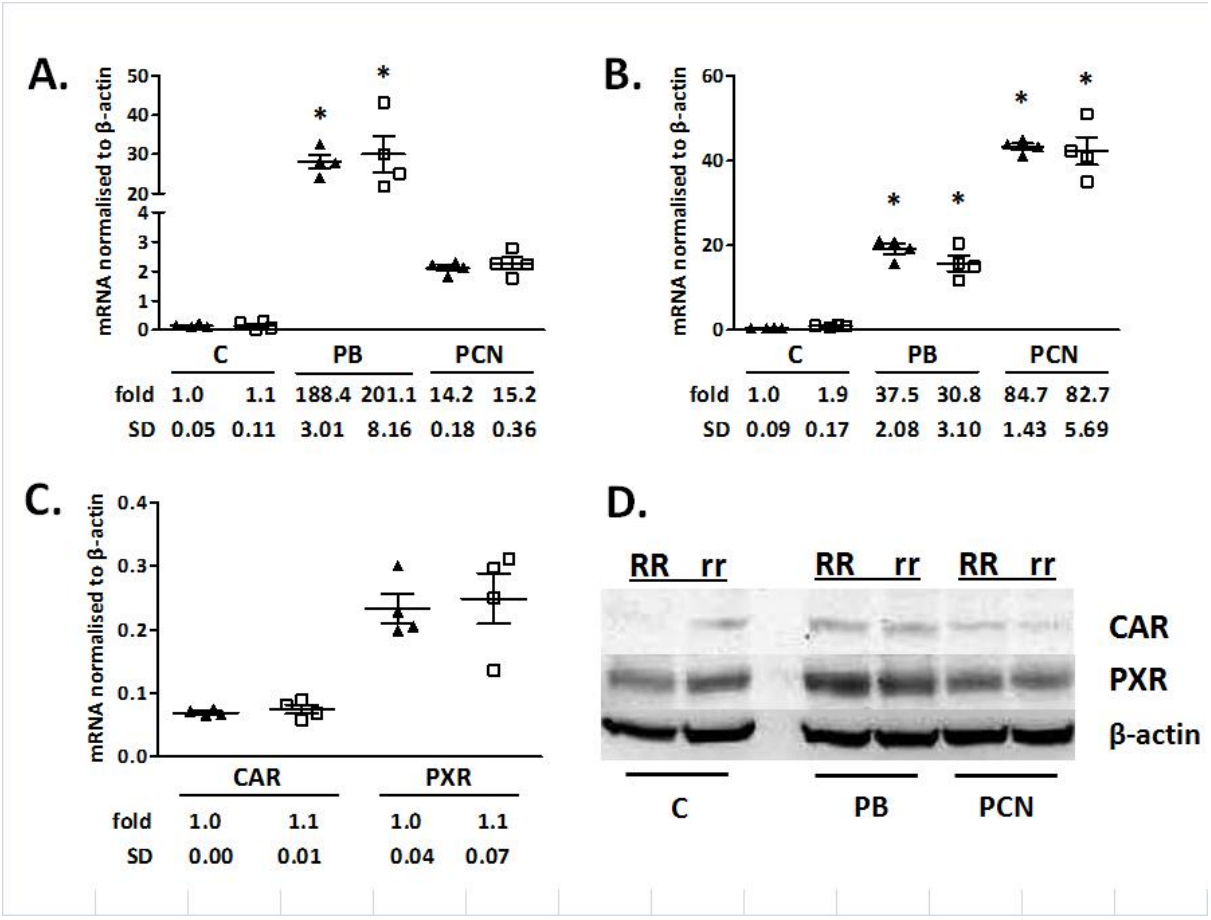


Figure 2.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

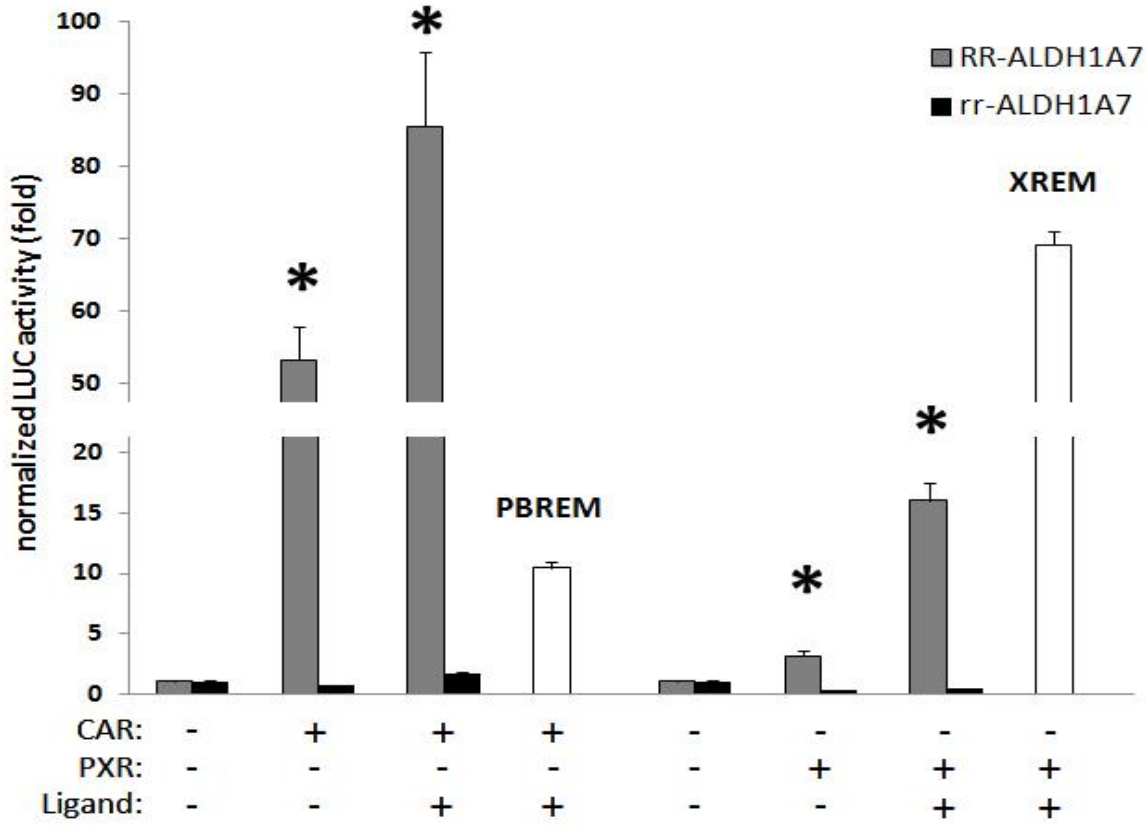


Figure 3.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

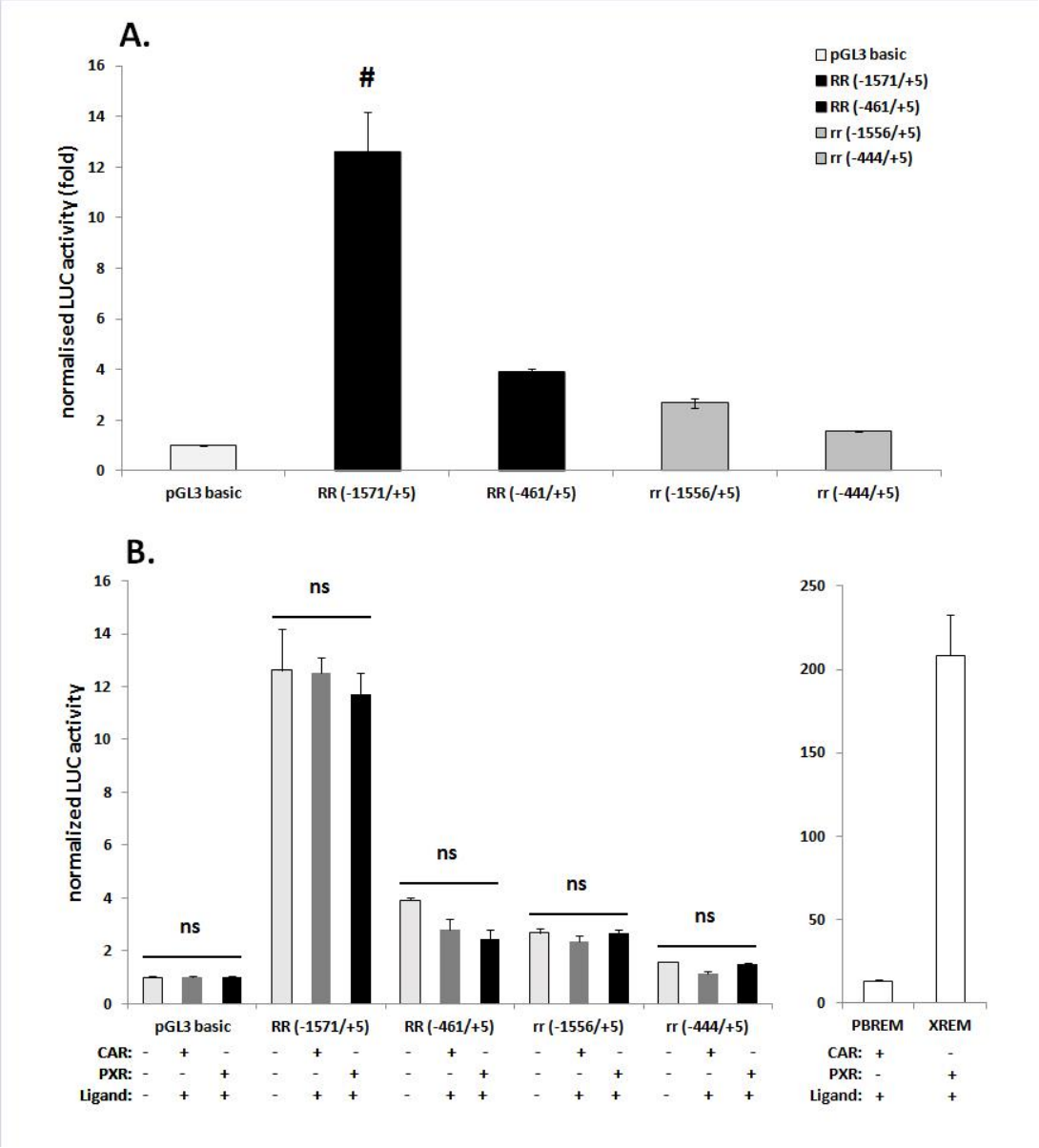


Figure 4.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

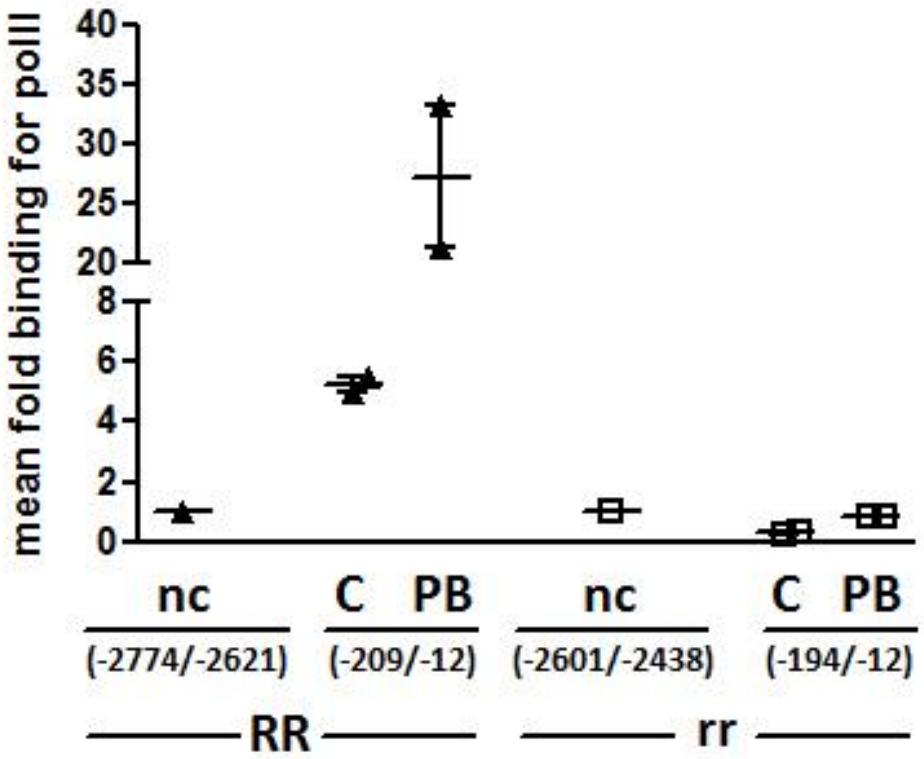


Figure 5.

Molecular Pharmacology Fast Forward. Published on October 1, 2019 as DOI: 10.1124/mol.119.117424
This article has not been copyedited and formatted. The final version may differ from this version.

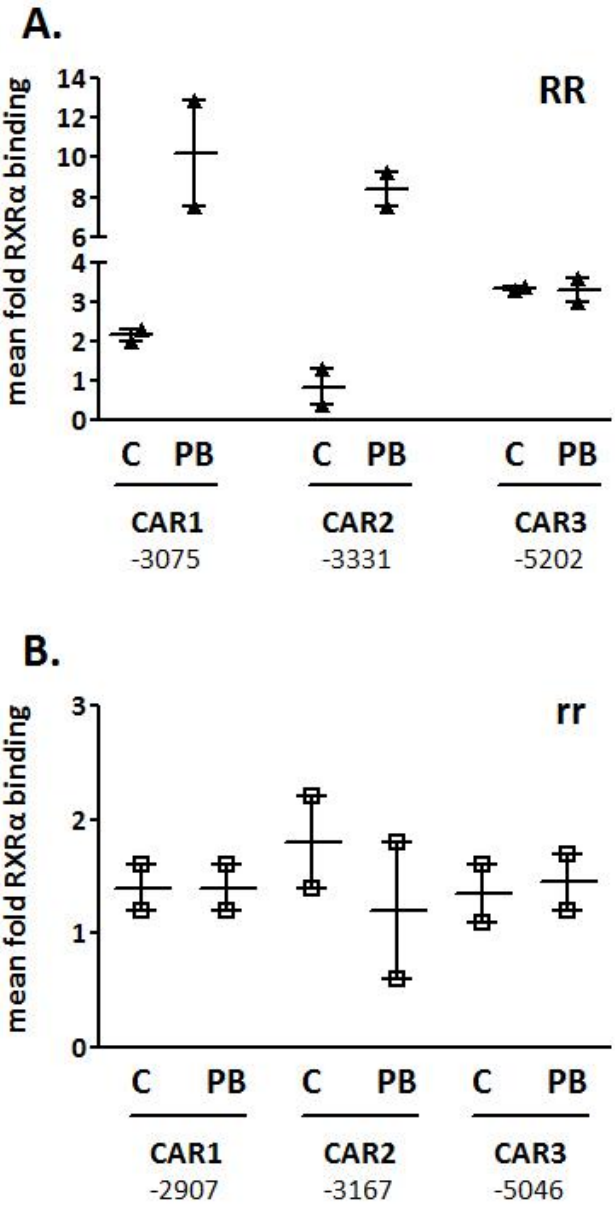


Figure 6.

Supplementary Tables and Figures

MOL # 117424

The basis for strain-dependent rat aldehyde dehydrogenase 1A7 (*ALDH1A7*) gene expression

Katerina Touloupi, Jenni Küblbeck, Angeliki Magklara, Ferdinand Molnár, Mika Reinisalo, Maria Konstandi, Paavo Honkakoski and Periklis Pappas

Supplementary Table 1. Primers for ALDH1A7 cloning.

Supplementary Table 2. Primers for ChIP assay.

Supplementary Table 3. Predicted transcription factor sites in the ***RR*** and ***rr*** *ALDH1A7* upstream sequences.

Supplementary Table 4. Predicted nuclear receptor binding elements in the ***RR*** and ***rr*** *ALDH1A7* upstream sequences.

Supplementary Table 5. Group mean Ct values of β-actin, ALDH and CYP mRNAs in the multiple dose study.

Supplementary Figure 1. Alignment of the *ALDH1A7* promoters from the ***RR*** and ***rr*** strains.

Supplementary Figure 2. Cytosolic and total ALDH1 protein levels in control, PB- and PCN-treated ***RR*** and ***rr*** rat livers.

Supplementary Figure 3. Total CAR and PXR protein levels in control, PB- and PCN-treated ***RR*** and ***rr*** rat livers.

Supplementary Figure 4. Time course of PB induction in ***RR*** and ***rr*** rats.

Supplementary Figure 5. Time course of PCN induction in ***RR*** and ***rr*** rats.

Supplementary Table 1. Primers used for cloning of the ALDH1A7 promoter from *RR* and *rr* strains into plasmid vector pGL3-Basic.

Deletion Fragment (location)	**Forward Primer** **(5'-**flanking nucleotides-*RE site*- **ALDH1A7 sequence -3')**
Longer promoter fragment (-1571 bp for ***RR*** and -1556 bp for ***rr***)	5'-CTCTC*ACGCGT***CACTTATAA**-3'
Shorter promoter fragment (-461 bp for ***RR*** and -444 bp for ***rr***)	5'-AAA*ACGCGT***ACTCTGATCTG**-3'
+5 bp for both ***RR*** and ***rr***	**Reverse primer** 5'-TGTGT*CTCGAG***TTCTGAAAG**-3'

Serial Cloner® and OLIGO-6 software were used in order to select preferential REs and design the desired primers, respectively. Subcloning was accomplished by digestion of the amplicons with restriction enzymes *Mlu*I (Thermo Scientific, ER0561) and *Xho*I (Thermo Scientific, ER0691) producing 5' and 3' overhangs. The pGL3-Basic vector were also digested with the above enzymes and ligation reactions were carried out by use of T4 DNA ligase (Thermo Scientific, EL0014).

Supplementary Table 2. Forward and reverse primers used for amplification of immunoprecipitated DNA of ALDH1A7 promoter from *RR* and *rr* strains.

Name of promoter site	*Primer sequence and location*
ALDH1A7 proximal promoter	Forward primer (-143 for ***RR***, -128 for ***rr***): 5'-CATTTAAAGGCAAAGGCTCCC-3' Reverse primer (-12 for ***RR***, -12 for ***rr***): 5'-GCACTTGCTCCTTTTTATCTGC -3'
CAR1 site	Forward primer (***RR***, -3158): 5'-TGTGTTGACCTTCATAAAAAGTTCT-3' Forward primer (***rr***, -2968): 5'-TCTGTGATTTACCCATGACAATAAT-3' Reverse primer (-3038 for ***RR***, -2871 for ***rr***): 5'-TTCAAGACCTACCCTTACAGCC-3'
CAR2 site	Forward primer (***RR***, - 3352): 5'-AAAGAGTTCACACACACACAAATG-3' Reverse primer (***RR***, -3194): 5'-CCTACCTCTTGTGAACATTTTTTT-3' Forward primer (***rr***, -3188): 5'-ACTTCACACACACACACACAAATG-3' Reverse primer (***rr***, -3026): 5'-CCTACCTCTTGTTGAACATTTCTTT-3'
CAR3 site	Forward primer (-5260 for ***RR***, -5104 for ***rr***): 5'-TTCTCACCTCATTCTCTACAAGT-3' Reverse primer (-5173 for ***RR***, -5017 for ***rr***): 5'-GCAGGCATTGTGGTCAGTGA-3'

Negative control	Forward primer (***RR***, -2774): 5'-GAGAAAATGACTGAAAGAAGCTGAT-3' Forward primer (***rr***, -2601): 5'-GAAAAGGACTGAAAGAAGCTGATAA -3' Reverse primer (-2621 for ***RR***, -2438 for ***rr***): 5'-TTCAGGAATATGATGTCTTTTGGTC-3'
Actin **(proximal promoter)**	Forward primer: 5'-GGTGACCCCCAGAATACAGG-3' Reverse primer: 5'-GAAGAGTTTGGCGATGGGTG -3'
CYP2B2 PBREM **(positive control)**	Forward primer (-2320): 5'-CGTGGACACAACCTTCAAG-3' Reverse primer (-2197): 5'-GAGCAAGGTCCTGGTGTC-3' [Pustylnyak et al., 2011]

Supplementary Table 3. Location of binding sites for CAR/RXRα heterodimers, HNF4, and CAAT boxes on ALDH1A7-*RR* and ALDH1A7-*rr* gene promoters.

CAR/RXRα	HNF4	CAAT box
-2566/-2542 (*rr*)	-572/-548 (*RR*) -555/-531(*rr*)	-72/-66 (*RR*) -55/-50 (*rr*)
-3075/-3051 (*RR*) -2907/-2883 (*rr*)	-1596/-1572 (*rr*)	-197/-192 (*RR*) -177/-171 (*rr*)
-3331/-3307 (*RR*) -3167/-3143 (*rr*)	-1628/-1604 (*RR*) -1613/-1589 (*rr*)	
-5202/-5178 (*RR*) -5046/-5022 (*rr*)	-1649/-1625 (*RR*) -1634/-1610 (*rr*)	
-8539/-8515 (*RR*) -8387/-8363 (*rr*)	-2076/-2052 (*RR*) -2061/-2037 (*rr*)	
-9646/-9622 (*RR*) -9494/-9470 (*rr*)	-2592/-2568 (*RR*) -2409/-2385 (*rr*)	
-9815/-9791 (*RR*) -9663/-9639 (*rr*)	-3067/-3043 (*RR*) -2899/-2875 (*rr*)	

There were additionally over 40 predicted HNF1 sites in both ***RR*** and ***rr*** upstream promoters, that were mostly clustered at the approximate positions of -9800, -8250, -7200, -6400, -5100, -2900, -2100 and -1800 bp in the ***RR*** *ALDH1A7* sequence. Due to sequence variation between the strains, two HNF1 sites were present in the ***RR*** strain only while four were unique to the ***rr*** strain. None of the predicted HNF1 sites were within the CAR1, CAR2 or CAR3 sites used for ChIP analysis.

Supplementary Table 4. Location of DR3-, DR4- and ER6-type nuclear receptor binding elements on the ALDH1A7-*RR* and ALDH1A7-*rr* gene promoters.

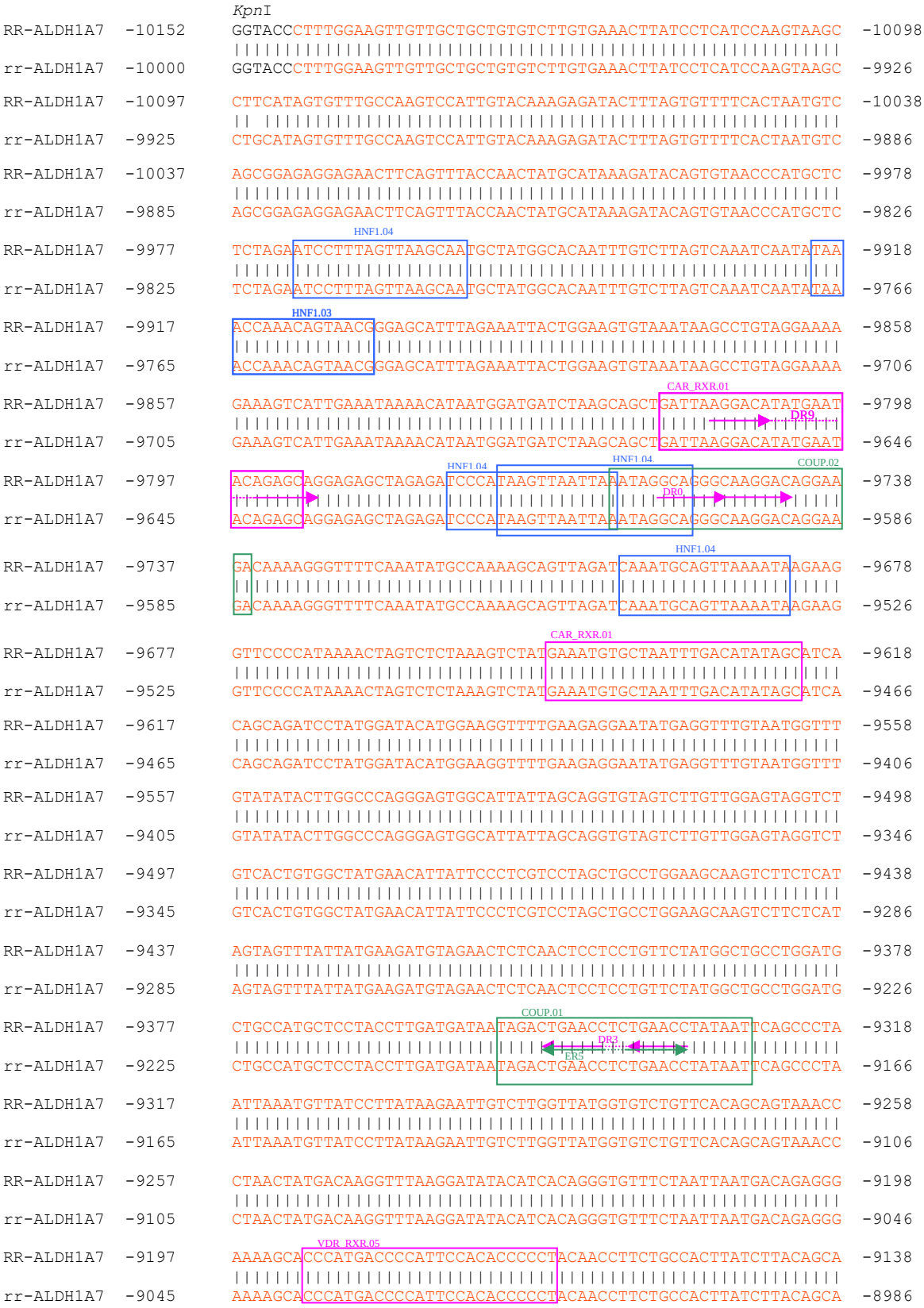
DR3	DR4	ER6
-87/-73 (*RR*) -71/-57 (*rr*)	-2596/-2581 (*RR*) -2413/-2398 (*rr*)	-1586/-1569 (*RR*) -1571/-1554 (*rr*)
-96/-82 (*RR*) -80/-66 (*rr*)	-2561/-2545 (*rr*)	
-2842/-2828 (*RR*)	-2852/-2837(*RR*)	
-7208/-7194 (*RR*) -7052/-7038 (*rr*)	-3071/-3056 (*RR*) -2903/-2888 (*rr*)	
-8263/-8249 (*rr*)		
-9347/-9333 (*RR*) -9195/-9181 (*rr*)		

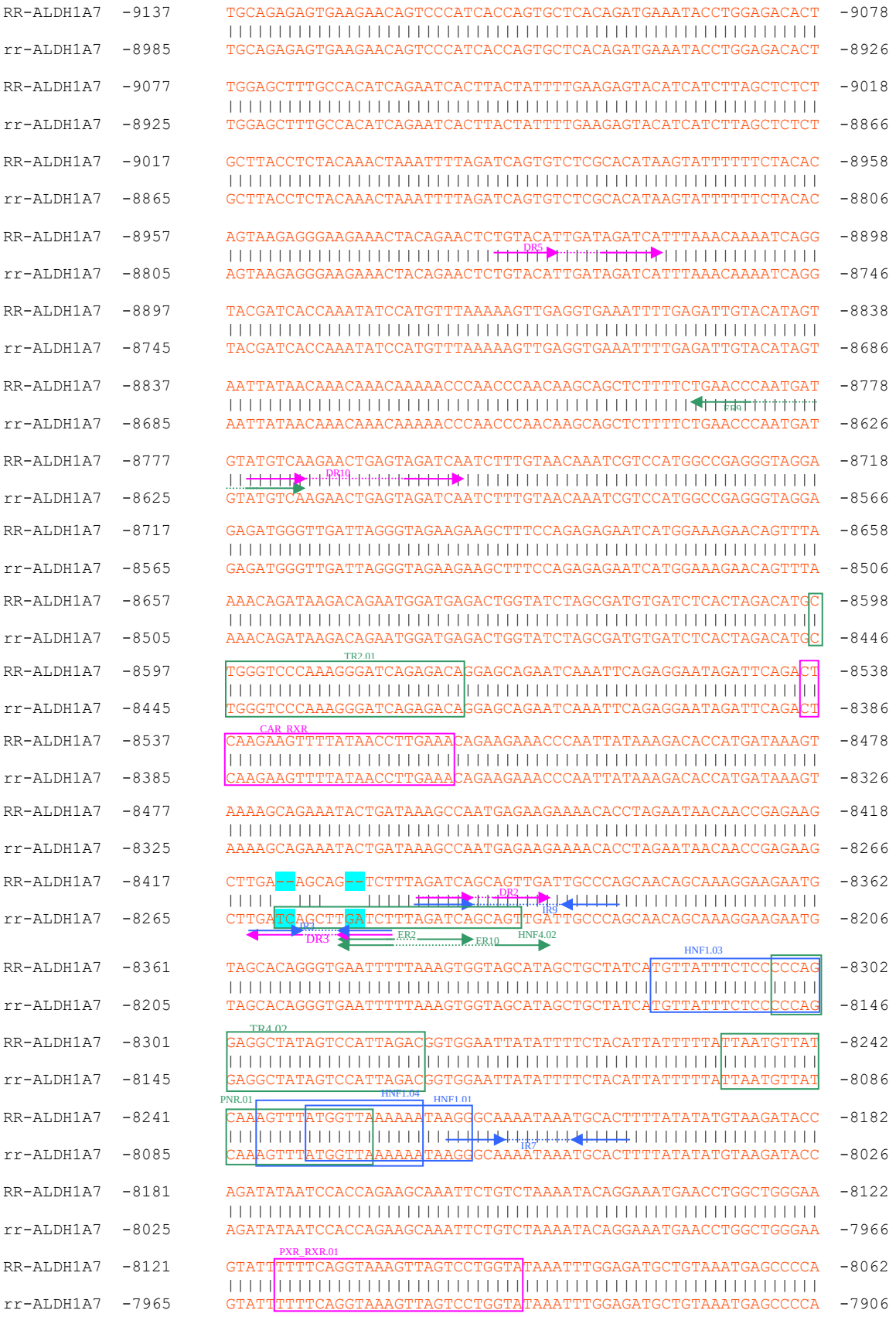
Supplementary Table 5. Group mean Ct values of β-actin, ALDH and CYP mRNAs in the multiple dose study.

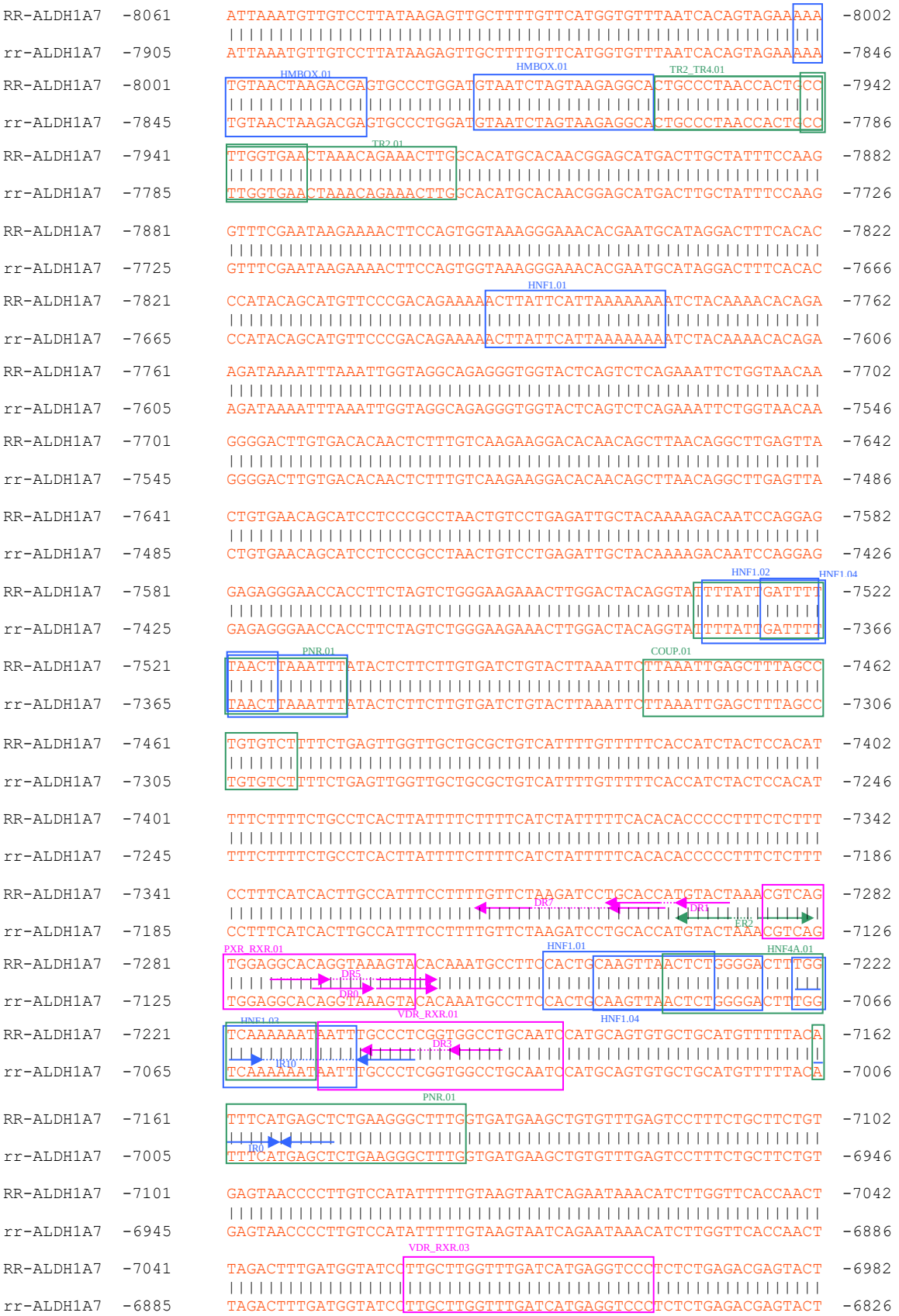
	Mean Ct values per group				
	ALDH1A7	ALDH1A1	CYP2B1	CYP3A23	β-actin
RR strain					
control saline	24,48	23,58	23,57	21,23	20,30
PB	18,29	20,30	15,66	16,21	20,46
control oil	24,12	22,99	21,83	20,49	19,44
PCN	20,12	20,32	18,59	14,23	19,66
rr strain					
control saline	32,09	22,87	24,64	20,79	19,96
PB	27,31	20,22	15,35	16,25	20,19
control oil	32,64	21,68	22,13	19,92	20,05
PCN	28,05	19,79	18,62	14,39	19,77

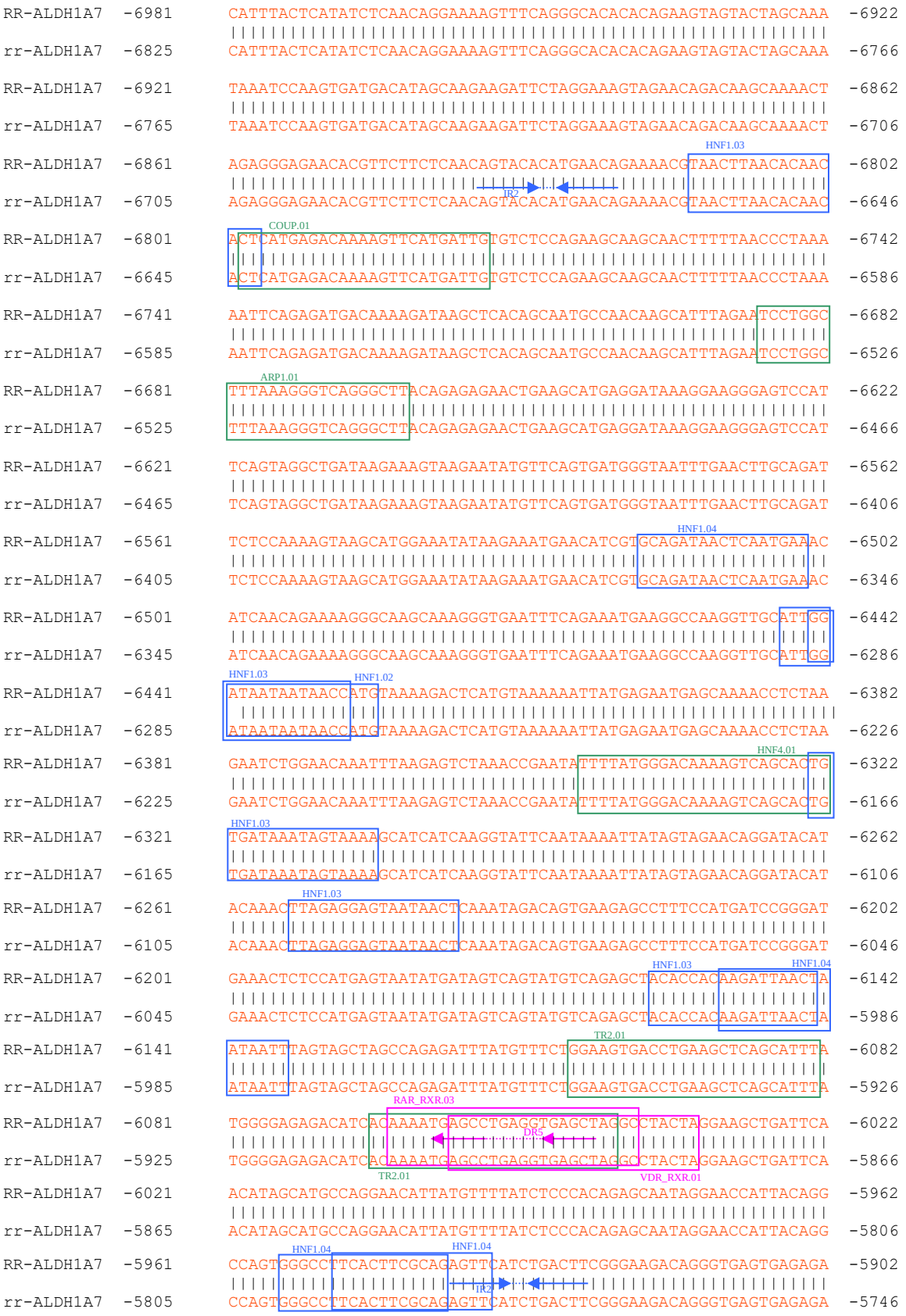
Indicated mRNAs were analyzed as described in Materials and Methods using Taqman chemistry and observed mean Ct values for each group are listed.

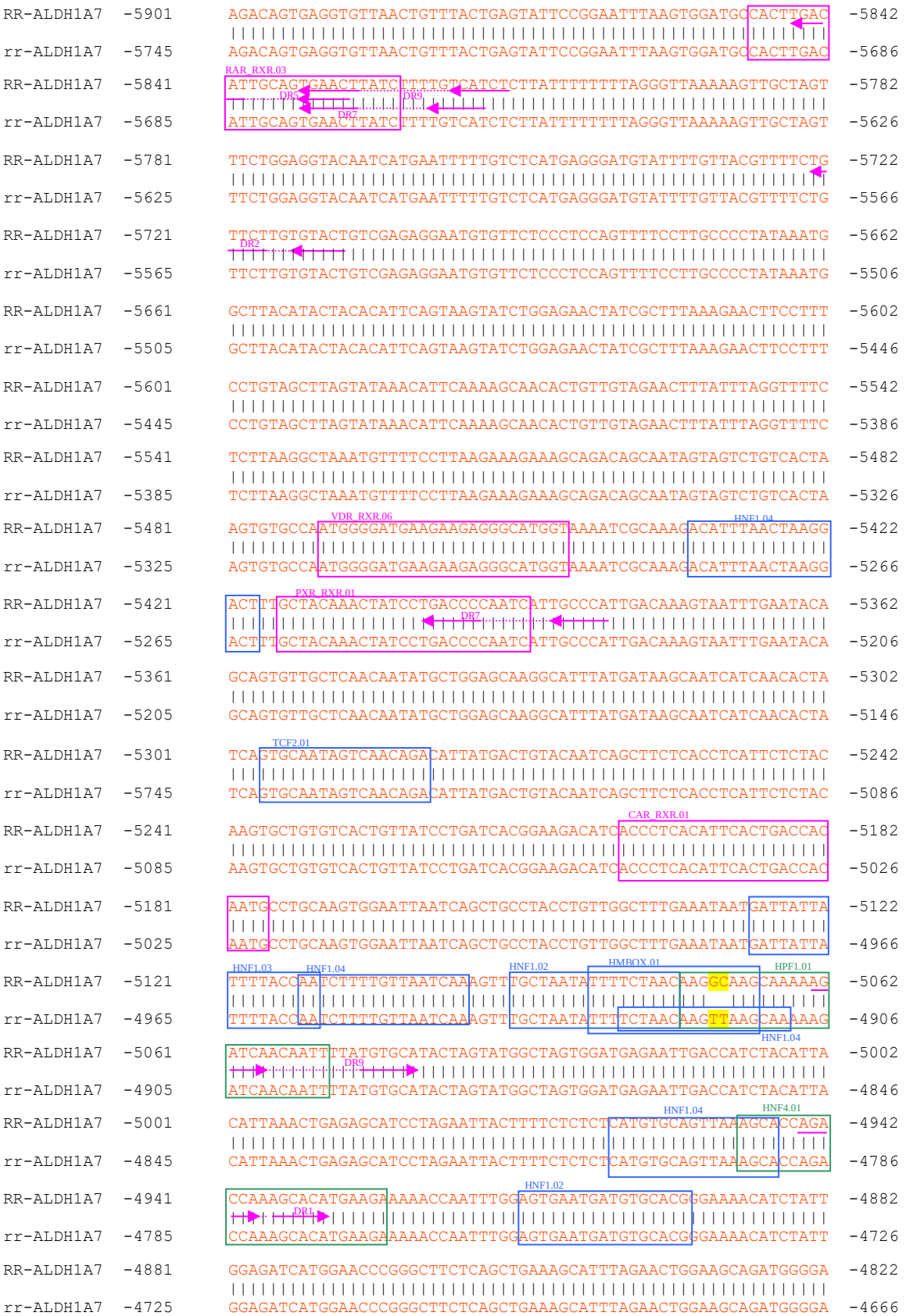
Supplementary Figure 1. Alignment of ~10 kbp *ALDH1A7* promoter sequences.

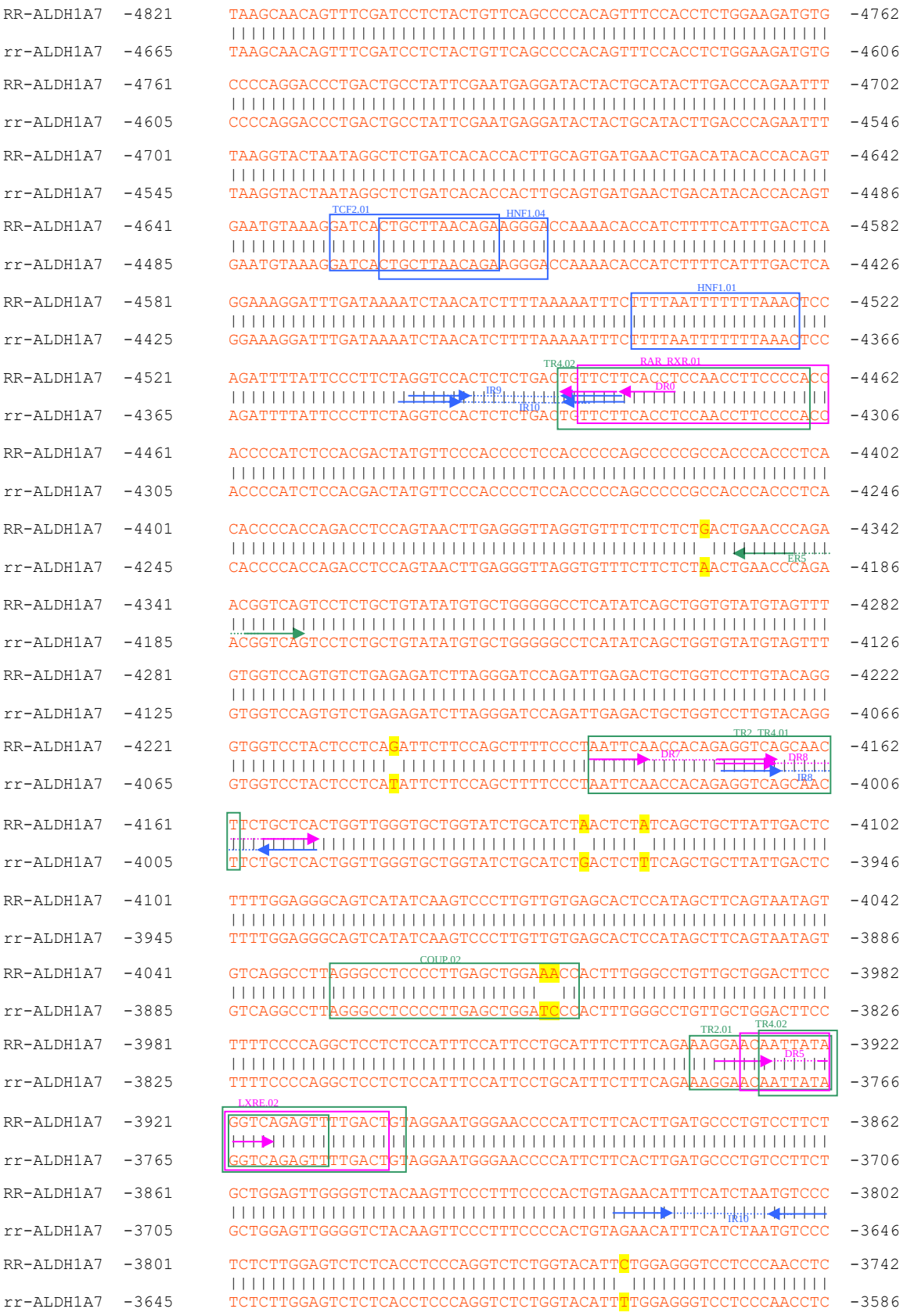


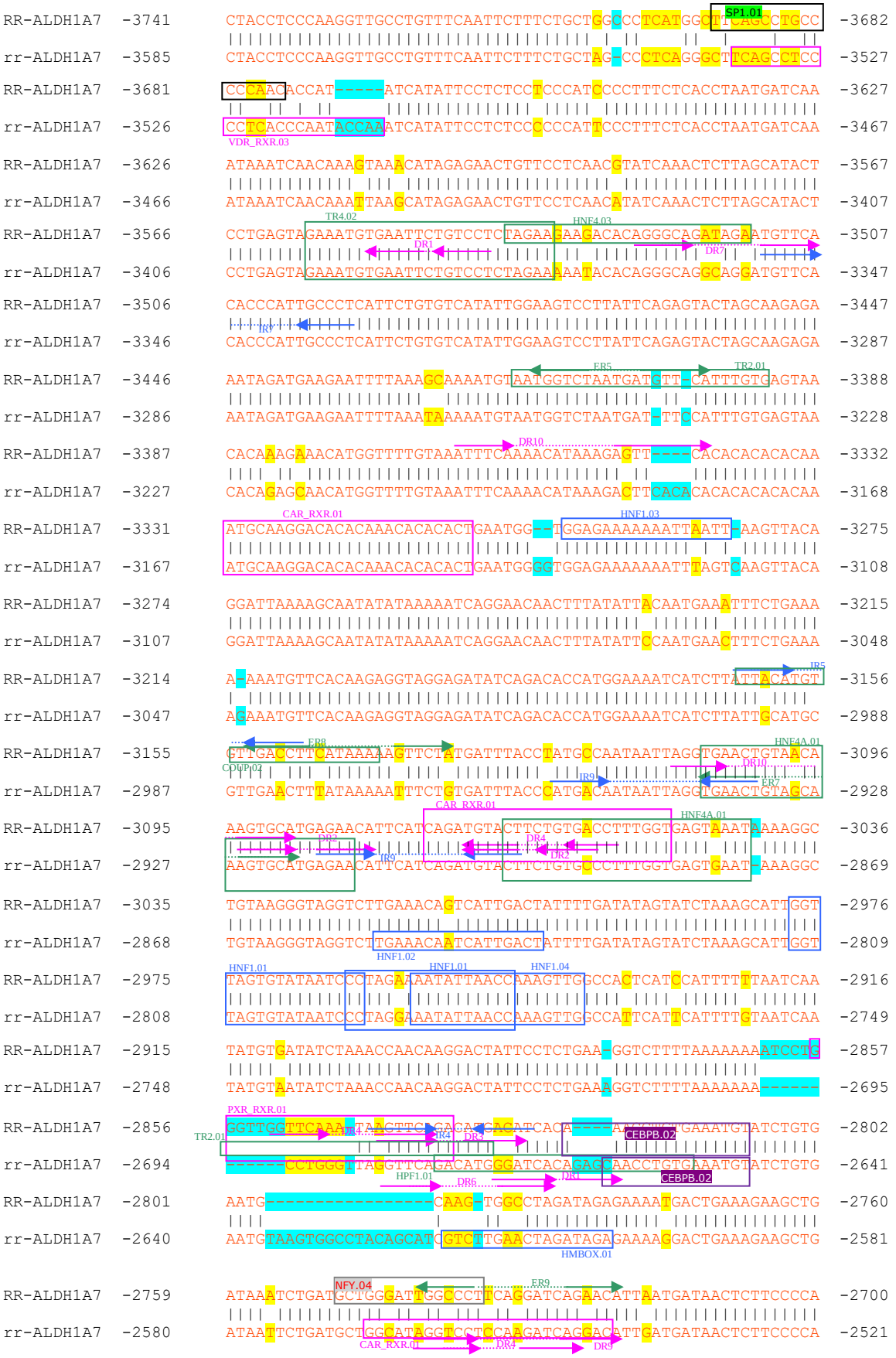

RR-ALDH1A7 -9137 TGCAGAGAGTGAAGAACAGTCCCATCACCAGTGCTCACAGATGAAATACCTGGAGACACT -9078
rr-ALDH1A7 -8985 TGCAGAGAGTGAAGAACAGTCCCATCACCAGTGCTCACAGATGAAATACCTGGAGACACT -8926
RR-ALDH1A7 -9077 TGGAGCTTTGCCACATCAGAATCACTTACTATTTTGAAGAGTACATCATCTTAGCTCTCT -9018
rr-ALDH1A7 -8925 TGGAGCTTTGCCACATCAGAATCACTTACTATTTTGAAGAGTACATCATCTTAGCTCTCT -8866
RR-ALDH1A7 -9017 GCTTACCTCTACAAACTAAATTTTAGATCAGTGTCTCGCACATAAGTATTTTTTCTACAC -8958
rr-ALDH1A7 -8865 GCTTACCTCTACAAACTAAATTTTAGATCAGTGTCTCGCACATAAGTATTTTTTCTACAC -8806
RR-ALDH1A7 -8957 AGTAAGAGGGAAGAAACTACAGAACTCTGTACATTGATAGATCATTTAAACAAAATCAGG -8898
DR5
rr-ALDH1A7 -8805 AGTAAGAGGGAAGAAACTACAGAACTCTGTACATTGATAGATCATTTAAACAAAATCAGG -8746
RR-ALDH1A7 -8897 TACGATCACCAAATATCCATGTTTAAAAAGTTGAGGTGAAATTTTGAGATTGTACATAGT -8838
rr-ALDH1A7 -8745 TACGATCACCAAATATCCATGTTTAAAAAGTTGAGGTGAAATTTTGAGATTGTACATAGT -8686
RR-ALDH1A7 -8837 AATTATAACAAACAAACAAAAACCCAACCCAACAAGCAGCTCTTTTCTGAACCCAATGAT -8778
ER1
rr-ALDH1A7 -8685 AATTATAACAAACAAACAAAAACCCAACCCAACAAGCAGCTCTTTTCTGAACCCAATGAT -8626
RR-ALDH1A7 -8777 GTATGTCAAGAACTGAGTAGATCAATCTTTGTAACAAATCGTCCATGGCCGAGGGTAGGA -8718
DR10
rr-ALDH1A7 -8625 GTATGTCAAGAACTGAGTAGATCAATCTTTGTAACAAATCGTCCATGGCCGAGGGTAGGA -8566
RR-ALDH1A7 -8717 GAGATGGGTTGATTAGGGTAGAAGAAGCTTTCCAGAGAGAATCATGGAAAGAACAGTTTA -8658
rr-ALDH1A7 -8565 GAGATGGGTTGATTAGGGTAGAAGAAGCTTTCCAGAGAGAATCATGGAAAGAACAGTTTA -8506
RR-ALDH1A7 -8657 AAACAGATAAGACAGAATGGATGAGACTGGTATCTAGCGATGTGATCTCACTAGACATGC -8598
rr-ALDH1A7 -8505 AAACAGATAAGACAGAATGGATGAGACTGGTATCTAGCGATGTGATCTCACTAGACATGC -8446
TR2.01
RR-ALDH1A7 -8597 TGGGTCCCAAAGGGATCAGAGACAGGAGCAGAATCAAATTCAGAGGAATAGATTCAGACT -8538
rr-ALDH1A7 -8445 TGGGTCCCAAAGGGATCAGAGACAGGAGCAGAATCAAATTCAGAGGAATAGATTCAGACT -8386
CAR_RXR
RR-ALDH1A7 -8537 CAAGAAGTTTTATAACCTTGAAACAGAAGAAACCCAATTATAAAGACACCATGATAAAGT -8478
rr-ALDH1A7 -8385 CAAGAAGTTTTATAACCTTGAAACAGAAGAAACCCAATTATAAAGACACCATGATAAAGT -8326
RR-ALDH1A7 -8477 AAAAGCAGAAATACTGATAAAGCCAATGAGAAGAAAACACCTAGAATAACAACCGAGAAG -8418
rr-ALDH1A7 -8325 AAAAGCAGAAATACTGATAAAGCCAATGAGAAGAAAACACCTAGAATAACAACCGAGAAG -8266
RR-ALDH1A7 -8417 CTTGA--AGCAG--TCTTTAGATCAGCAGTTGATTGCCCAGCAACAGCAAAGGAAGAATG -8362
DR2
IR9
rr-ALDH1A7 -8265 CTTGATCAGCTTGATCTTTAGATCAGCAGTTGATTGCCCAGCAACAGCAAAGGAAGAATG -8206
DR1
DR3
ER2
ER10
HNF4.02
HNF1.03
RR-ALDH1A7 -8361 TAGCACAGGGTGAATTTTTAAAGTGGTAGCATAGCTGCTATCATGTTATTTCTCCCCCAG -8302
rr-ALDH1A7 -8205 TAGCACAGGGTGAATTTTTAAAGTGGTAGCATAGCTGCTATCATGTTATTTCTCCCCCAG -8146
TR4.02
RR-ALDH1A7 -8301 GAGGCTATAGTCCATTAGACGGTGGAATTATATTTTCTACATTATTTTTATTAATGTTAT -8242
rr-ALDH1A7 -8145 GAGGCTATAGTCCATTAGACGGTGGAATTATATTTTCTACATTATTTTTATTAATGTTAT -8086
PNR.01
HNF1.04
HNF1.01
RR-ALDH1A7 -8241 CAAAGTTTATGGTTAAAAAATAAGGGCAAAATAAATGCACTTTTATATATGTAAGATACC -8182
IR7
rr-ALDH1A7 -8085 CAAAGTTTATGGTTAAAAAATAAGGGCAAAATAAATGCACTTTTATATATGTAAGATACC -8026
RR-ALDH1A7 -8181 AGATATAATCCACCAGAAGCAAATTCTGTCTAAAATACAGGAAATGAACCTGGCTGGGAA -8122
rr-ALDH1A7 -8025 AGATATAATCCACCAGAAGCAAATTCTGTCTAAAATACAGGAAATGAACCTGGCTGGGAA -7966
PXR_RXR.01
RR-ALDH1A7 -8121 GTATTTTTTCAGGTAAAGTTAGTCCTGGTATAAATTTGGAGATGCTGTAAATGAGCCCCA -8062
rr-ALDH1A7 -7965 GTATTTTTTCAGGTAAAGTTAGTCCTGGTATAAATTTGGAGATGCTGTAAATGAGCCCCA -7906

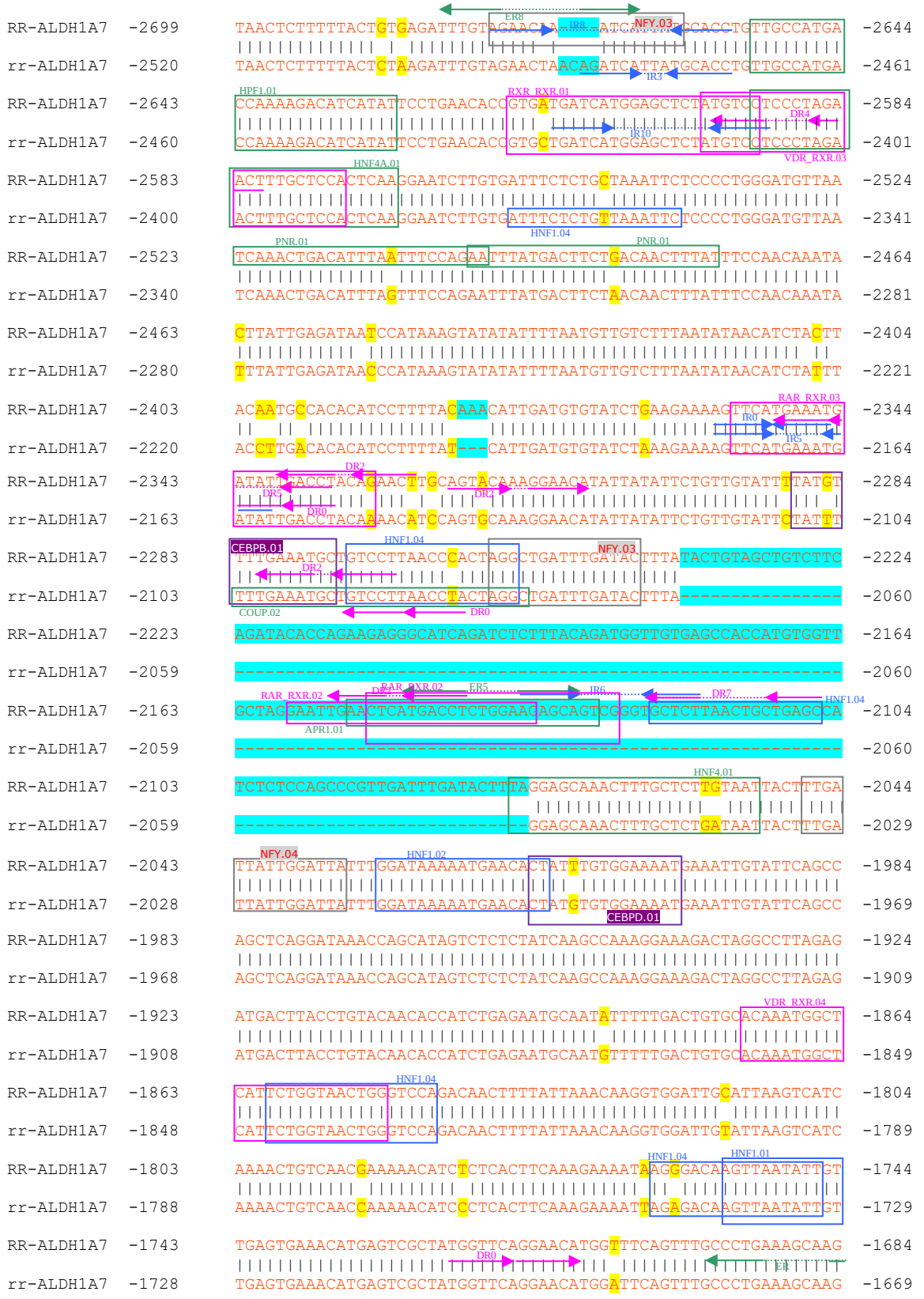

RR-ALDH1A7 -8061 ATTAAATGTTGTCCTTATAAGAGTTGCTTTTGTTCATGGTGTTTAATCACAGTAGAAAAA -8002
rr-ALDH1A7 -7905 ATTAAATGTTGTCCTTATAAGAGTTGCTTTTGTTCATGGTGTTTAATCACAGTAGAAAAA -7846
HMBOX.01
HMBOX.01
TR2_TR4.01
RR-ALDH1A7 -8001 TGTAACTAAGACGAGTGCCCTGGATGTAATCTAGTAAGAGGCACTGCCCTAACCACTGCC -7942
rr-ALDH1A7 -7845 TGTAACTAAGACGAGTGCCCTGGATGTAATCTAGTAAGAGGCACTGCCCTAACCACTGCC -7786
TR2.01
RR-ALDH1A7 -7941 TTGGTGAACTAAACAGAAACTTGGCACATGCACAACGGAGCATGACTTGCTATTTCCAAG -7882
rr-ALDH1A7 -7785 TTGGTGAACTAAACAGAAACTTGGCACATGCACAACGGAGCATGACTTGCTATTTCCAAG -7726
RR-ALDH1A7 -7881 GTTTCGAATAAGAAAACTTCCAGTGGTAAAGGGAAACACGAATGCATAGGACTTTCACAC -7822
rr-ALDH1A7 -7725 GTTTCGAATAAGAAAACTTCCAGTGGTAAAGGGAAACACGAATGCATAGGACTTTCACAC -7666
HNF1.01
RR-ALDH1A7 -7821 CCATACAGCATGTTCCCGACAGAAAAACTTATTCATTAAAAAAAATCTACAAAACACAGA -7762
rr-ALDH1A7 -7665 CCATACAGCATGTTCCCGACAGAAAAACTTATTCATTAAAAAAAATCTACAAAACACAGA -7606
RR-ALDH1A7 -7761 AGATAAAATTTAAATTGGTAGGCAGAGGGTGGTACTCAGTCTCAGAAATTCTGGTAACAA -7702
rr-ALDH1A7 -7605 AGATAAAATTTAAATTGGTAGGCAGAGGGTGGTACTCAGTCTCAGAAATTCTGGTAACAA -7546
RR-ALDH1A7 -7701 GGGGACTTGTGACACAACTCTTTGTCAAGAAGGACACAACAGCTTAACAGGCTTGAGTTA -7642
rr-ALDH1A7 -7545 GGGGACTTGTGACACAACTCTTTGTCAAGAAGGACACAACAGCTTAACAGGCTTGAGTTA -7486
RR-ALDH1A7 -7641 CTGTGAACAGCATCCTCCCGCCTAACTGTCCTGAGATTGCTACAAAAGACAATCCAGGAG -7582
rr-ALDH1A7 -7485 CTGTGAACAGCATCCTCCCGCCTAACTGTCCTGAGATTGCTACAAAAGACAATCCAGGAG -7426
HNF1.02
HNF1.04
RR-ALDH1A7 -7581 GAGAGGGAACCACCTTCTAGTCTGGGAAGAAACTTGGACTACAGGTATTTTATTGATTTT -7522
rr-ALDH1A7 -7425 GAGAGGGAACCACCTTCTAGTCTGGGAAGAAACTTGGACTACAGGTATTTTATTGATTTT -7366
PNR.01
COUP.01
RR-ALDH1A7 -7521 TAACTTAAATTTATACTCTTCTTGTGATCTGTACTTAAATTCTTAAATTGAGCTTTAGCC -7462
rr-ALDH1A7 -7365 TAACTTAAATTTATACTCTTCTTGTGATCTGTACTTAAATTCTTAAATTGAGCTTTAGCC -7306
RR-ALDH1A7 -7461 TGTGTCTTTTCTGAGTTGGTTGCTGCGCTGTCATTTTGTTTTTCACCATCTACTCCACAT -7402
rr-ALDH1A7 -7305 TGTGTCTTTTCTGAGTTGGTTGCTGCGCTGTCATTTTGTTTTTCACCATCTACTCCACAT -7246
RR-ALDH1A7 -7401 TTTCTTTTCTGCCTCACTTATTTTCTTTTCATCTATTTTTCACACACCCCCTTTCTCTTT -7342
rr-ALDH1A7 -7245 TTTCTTTTCTGCCTCACTTATTTTCTTTTCATCTATTTTTCACACACCCCCTTTCTCTTT -7186
DR7
DR1
ER2
RR-ALDH1A7 -7341 CCTTTCATCACTTGCCATTTCCTTTTGTTCTAAGATCCTGCACCATGTACTAAACGTCAG -7282
rr-ALDH1A7 -7185 CCTTTCATCACTTGCCATTTCCTTTTGTTCTAAGATCCTGCACCATGTACTAAACGTCAG -7126
PXR_RXR.01
HNF1.01
HNF4A.01
DR5
DR0
RR-ALDH1A7 -7281 TGGAGGCACAGGTAAAGTACACAAATGCCTTCCACTGCAAGTTAACTCTGGGGACTTTGG -7222
rr-ALDH1A7 -7125 TGGAGGCACAGGTAAAGTACACAAATGCCTTCCACTGCAAGTTAACTCTGGGGACTTTGG -7066
HNF1.03
VDR_RXR.01
HNF1.04
DR3
DR10
RR-ALDH1A7 -7221 TCAAAAAATAATTTGCCCTCGGTGGCCTGCAATCCATGCAGTGTGCTGCATGTTTTTACA -7162
rr-ALDH1A7 -7065 TCAAAAAATAATTTGCCCTCGGTGGCCTGCAATCCATGCAGTGTGCTGCATGTTTTTACA -7006
PNR.01
IR0
RR-ALDH1A7 -7161 TTTCATGAGCTCTGAAGGGCTTTGGTGATGAAGCTGTGTTTGAGTCCTTTCTGCTTCTGT -7102
rr-ALDH1A7 -7005 TTTCATGAGCTCTGAAGGGCTTTGGTGATGAAGCTGTGTTTGAGTCCTTTCTGCTTCTGT -6946
RR-ALDH1A7 -7101 GAGTAACCCCTTGTCCATATTTTTGTAAGTAATCAGAATAAACATCTTGGTTCACCAACT -7042
rr-ALDH1A7 -6945 GAGTAACCCCTTGTCCATATTTTTGTAAGTAATCAGAATAAACATCTTGGTTCACCAACT -6886
VDR_RXR.03
RR-ALDH1A7 -7041 TAGACTTTGATGGTATCCTTGCTTGGTTTGATCATGAGGTCCCTCTCTGAGACGAGTACT -6982
rr-ALDH1A7 -6885 TAGACTTTGATGGTATCCTTGCTTGGTTTGATCATGAGGTCCCTCTCTGAGACGAGTACT -6826

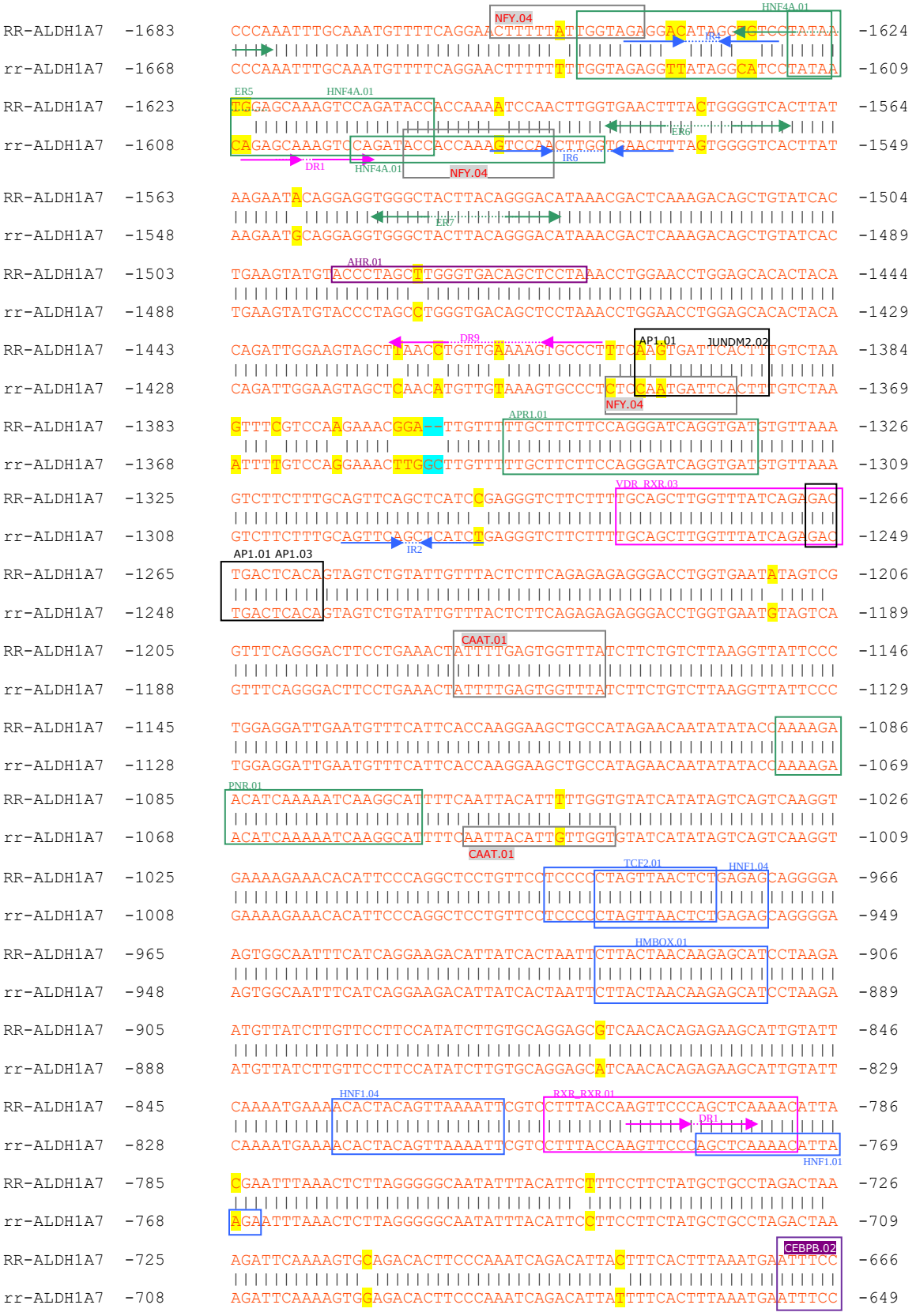

RR-ALDH1A7 -6981 CATTTACTCATATCTCAACAGGAAAAGTTTCAGGGCACACACAGAAGTAGTACTAGCAAA -6922
rr-ALDH1A7 -6825 CATTTACTCATATCTCAACAGGAAAAGTTTCAGGGCACACACAGAAGTAGTACTAGCAAA -6766
RR-ALDH1A7 -6921 TAAATCCAAGTGATGACATAGCAAGAAGATTCTAGGAAAGTAGAACAGACAAGCAAAACT -6862
rr-ALDH1A7 -6765 TAAATCCAAGTGATGACATAGCAAGAAGATTCTAGGAAAGTAGAACAGACAAGCAAAACT -6706
HNF1.03
RR-ALDH1A7 -6861 AGAGGGAGAACACGTTCTTCTCAACAGTACACATGAACAGAAAACGTAACTTAACACAAC -6802
IR2
rr-ALDH1A7 -6705 AGAGGGAGAACACGTTCTTCTCAACAGTACACATGAACAGAAAACGTAACTTAACACAAC -6646
COUP.01
RR-ALDH1A7 -6801 ACTCATGAGACAAAAGTTCATGATTGTGTCTCCAGAAGCAAGCAACTTTTTAACCCTAAA -6742
rr-ALDH1A7 -6645 ACTCATGAGACAAAAGTTCATGATTGTGTCTCCAGAAGCAAGCAACTTTTTAACCCTAAA -6586
RR-ALDH1A7 -6741 AATTCAGAGATGACAAAAGATAAGCTCACAGCAATGCCAACAAGCATTTAGAATCCTGGC -6682
rr-ALDH1A7 -6585 AATTCAGAGATGACAAAAGATAAGCTCACAGCAATGCCAACAAGCATTTAGAATCCTGGC -6526
ARP1.01
RR-ALDH1A7 -6681 TTTAAAGGGTCAGGGCTTACAGAGAGAACTGAAGCATGAGGATAAAGGAAGGGAGTCCAT -6622
rr-ALDH1A7 -6525 TTTAAAGGGTCAGGGCTTACAGAGAGAACTGAAGCATGAGGATAAAGGAAGGGAGTCCAT -6466
RR-ALDH1A7 -6621 TCAGTAGGCTGATAAGAAAGTAAGAATATGTTCAGTGATGGGTAATTTGAACTTGCAGAT -6562
rr-ALDH1A7 -6465 TCAGTAGGCTGATAAGAAAGTAAGAATATGTTCAGTGATGGGTAATTTGAACTTGCAGAT -6406
HNF1.04
RR-ALDH1A7 -6561 TCTCCAAAAGTAAGCATGGAAATATAAGAAATGAACATCGTGCAGATAACTCAATGAAAC -6502
rr-ALDH1A7 -6405 TCTCCAAAAGTAAGCATGGAAATATAAGAAATGAACATCGTGCAGATAACTCAATGAAAC -6346
RR-ALDH1A7 -6501 ATCAACAGAAAAGGGCAAGCAAAGGGTGAATTTCAGAAATGAAGGCCAAGGTTGCATTGG -6442
rr-ALDH1A7 -6345 ATCAACAGAAAAGGGCAAGCAAAGGGTGAATTTCAGAAATGAAGGCCAAGGTTGCATTGG -6286
HNF1.03
HNF1.02
RR-ALDH1A7 -6441 ATAATAATAACCATGTAAAAGACTCATGTAAAAAATTATGAGAATGAGCAAAACCTCTAA -6382
rr-ALDH1A7 -6285 ATAATAATAACCATGTAAAAGACTCATGTAAAAAATTATGAGAATGAGCAAAACCTCTAA -6226
HNF4.01
RR-ALDH1A7 -6381 GAATCTGGAACAAATTTAAGAGTCTAAACCGAATATTTTATGGGACAAAAGTCAGCACTG -6322
rr-ALDH1A7 -6225 GAATCTGGAACAAATTTAAGAGTCTAAACCGAATATTTTATGGGACAAAAGTCAGCACTG -6166
HNF1.03
RR-ALDH1A7 -6321 TGATAAATAGTAAAAGCATCATCAAGGTATTCAATAAAATTATAGTAGAACAGGATACAT -6262
rr-ALDH1A7 -6165 TGATAAATAGTAAAAGCATCATCAAGGTATTCAATAAAATTATAGTAGAACAGGATACAT -6106
HNF1.03
RR-ALDH1A7 -6261 ACAAACTTAGAGGAGTAATAACTCAAATAGACAGTGAAGAGCCTTTCCATGATCCGGGAT -6202
rr-ALDH1A7 -6105 ACAAACTTAGAGGAGTAATAACTCAAATAGACAGTGAAGAGCCTTTCCATGATCCGGGAT -6046
HNF1.03
HNF1.04
RR-ALDH1A7 -6201 GAAACTCTCCATGAGTAATATGATAGTCAGTATGTCAGAGCTACACCACAAGATTAACTA -6142
rr-ALDH1A7 -6045 GAAACTCTCCATGAGTAATATGATAGTCAGTATGTCAGAGCTACACCACAAGATTAACTA -5986
TR2.01
RR-ALDH1A7 -6141 ATAATTTAGTAGCTAGCCAGAGATTTATGTTTCTGGAAGTGACCTGAAGCTCAGCATTTA -6082
rr-ALDH1A7 -5985 ATAATTTAGTAGCTAGCCAGAGATTTATGTTTCTGGAAGTGACCTGAAGCTCAGCATTTA -5926
RAR_RXR.03
RR-ALDH1A7 -6081 TGGGGAGAGACATCACAAAATGAGCCTGAGGTGAGCTAGGCCTACTAGGAAGCTGATTCA -6022
DR5
rr-ALDH1A7 -5925 TGGGGAGAGACATCACAAAATGAGCCTGAGGTGAGCTAGGCCTACTAGGAAGCTGATTCA -5866
TR2.01
VDR_RXR.01
RR-ALDH1A7 -6021 ACATAGCATGCCAGGAACATTATGTTTTATCTCCCACAGAGCAATAGGAACCATTACAGG -5962
rr-ALDH1A7 -5865 ACATAGCATGCCAGGAACATTATGTTTTATCTCCCACAGAGCAATAGGAACCATTACAGG -5806
HNF1.04
HNF1.04
RR-ALDH1A7 -5961 CCAGTGGGCCTTCACTTCGCAGAGTTCATCTGACTTCGGGAAGACAGGGTGAGTGAGAGA -5902
IR2
rr-ALDH1A7 -5805 CCAGTGGGCCTTCACTTCGCAGAGTTCATCTGACTTCGGGAAGACAGGGTGAGTGAGAGA -5746

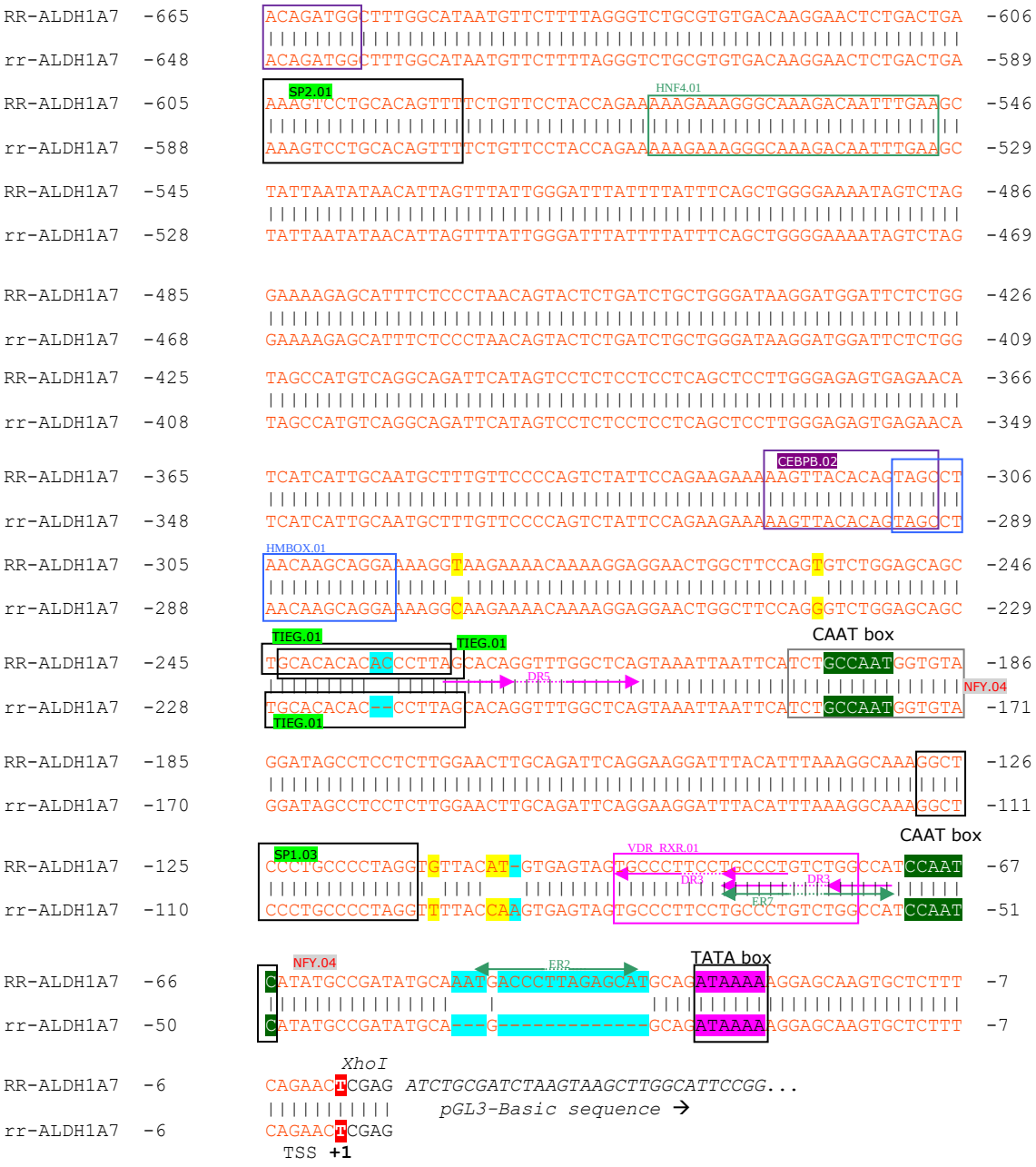

RR-ALDH1A7 -5901 AGACAGTGAGGTGTTAACTGTTTACTGAGTATTCCGGAATTTAAGTGGATGCCACTTGAC -5842
rr-ALDH1A7 -5745 AGACAGTGAGGTGTTAACTGTTTACTGAGTATTCCGGAATTTAAGTGGATGCCACTTGAC -5686
RAR_RXR.03
DR5
DR8
DR7
RR-ALDH1A7 -5841 ATTGCAGTGAACTTATCTTTTGTCATCTCTTATTTTTTTTAGGGTTAAAAAGTTGCTAGT -5782
rr-ALDH1A7 -5685 ATTGCAGTGAACTTATCTTTTGTCATCTCTTATTTTTTTTAGGGTTAAAAAGTTGCTAGT -5626
RR-ALDH1A7 -5781 TTCTGGAGGTACAATCATGAATTTTTGTCTCATGAGGGATGTATTTTGTTACGTTTTCTG -5722
rr-ALDH1A7 -5625 TTCTGGAGGTACAATCATGAATTTTTGTCTCATGAGGGATGTATTTTGTTACGTTTTCTG -5566
RR-ALDH1A7 -5721 TTCTTGTGTACTGTCGAGAGGAATGTGTTCTCCCTCCAGTTTTCCTTGCCCCTATAAATG -5662
DR2
rr-ALDH1A7 -5565 TTCTTGTGTACTGTCGAGAGGAATGTGTTCTCCCTCCAGTTTTCCTTGCCCCTATAAATG -5506
RR-ALDH1A7 -5661 GCTTACATACTACACATTCAGTAAGTATCTGGAGAACTATCGCTTTAAAGAACTTCCTTT -5602
rr-ALDH1A7 -5505 GCTTACATACTACACATTCAGTAAGTATCTGGAGAACTATCGCTTTAAAGAACTTCCTTT -5446
RR-ALDH1A7 -5601 CCTGTAGCTTAGTATAAACATTCAAAAGCAACACTGTTGTAGAACTTTATTTAGGTTTTC -5542
rr-ALDH1A7 -5445 CCTGTAGCTTAGTATAAACATTCAAAAGCAACACTGTTGTAGAACTTTATTTAGGTTTTC -5386
RR-ALDH1A7 -5541 TCTTAAGGCTAAATGTTTTCCTTAAGAAAGAAAGCAGACAGCAATAGTAGTCTGTCACTA -5482
rr-ALDH1A7 -5385 TCTTAAGGCTAAATGTTTTCCTTAAGAAAGAAAGCAGACAGCAATAGTAGTCTGTCACTA -5326
VDR_RXR.06
HNF1.04
RR-ALDH1A7 -5481 AGTGTGCCAATGGGGATGAAGAAGAGGGCATGGTAAAATCGCAAAGACATTTAACTAAGG -5422
rr-ALDH1A7 -5325 AGTGTGCCAATGGGGATGAAGAAGAGGGCATGGTAAAATCGCAAAGACATTTAACTAAGG -5266
PXR_RXR.01
DR7
RR-ALDH1A7 -5421 ACTTTGCTACAAACTATCCTGACCCCAATCATTGCCCATTGACAAAGTAATTTGAATACA -5362
rr-ALDH1A7 -5265 ACTTTGCTACAAACTATCCTGACCCCAATCATTGCCCATTGACAAAGTAATTTGAATACA -5206
RR-ALDH1A7 -5361 GCAGTGTTGCTCAACAATATGCTGGAGCAAGGCATTTATGATAAGCAATCATCAACACTA -5302
rr-ALDH1A7 -5205 GCAGTGTTGCTCAACAATATGCTGGAGCAAGGCATTTATGATAAGCAATCATCAACACTA -5146
TCF2.01
RR-ALDH1A7 -5301 TCAGTGCAATAGTCAACAGACATTATGACTGTACAATCAGCTTCTCACCTCATTCTCTAC -5242
rr-ALDH1A7 -5745 TCAGTGCAATAGTCAACAGACATTATGACTGTACAATCAGCTTCTCACCTCATTCTCTAC -5086
CAR_RXR.01
RR-ALDH1A7 -5241 AAGTGCTGTGTCACTGTTATCCTGATCACGGAAGACATCACCCTCACATTCACTGACCAC -5182
rr-ALDH1A7 -5085 AAGTGCTGTGTCACTGTTATCCTGATCACGGAAGACATCACCCTCACATTCACTGACCAC -5026
RR-ALDH1A7 -5181 AATGCCTGCAAGTGGAATTAATCAGCTGCCTACCTGTTGGCTTTGAAATAATGATTATTA -5122
rr-ALDH1A7 -5025 AATGCCTGCAAGTGGAATTAATCAGCTGCCTACCTGTTGGCTTTGAAATAATGATTATTA -4966
HNF1.03
HNF1.04
HNF1.02
HMBOX.01
HPF1.01
RR-ALDH1A7 -5121 TTTTACCAATCTTTTGTTAATCAAAGTTTGCTAATATTTTCTAACAAGGCAAGCAAAAAG -5062
rr-ALDH1A7 -4965 TTTTACCAATCTTTTGTTAATCAAAGTTTGCTAATATTTTCTAACAAGTTAAGCAAAAAG -4906
HNF1.04
RR-ALDH1A7 -5061 ATCAACAATTTTATGTGCATACTAGTATGGCTAGTGGATGAGAATTGACCATCTACATTA -5002
DR5
rr-ALDH1A7 -4905 ATCAACAATTTTATGTGCATACTAGTATGGCTAGTGGATGAGAATTGACCATCTACATTA -4846
HNF1.04
HNF4.01
RR-ALDH1A7 -5001 CATTAAACTGAGAGCATCCTAGAATTACTTTTCTCTCTCATGTGCAGTTAAAGCACCAGA -4942
rr-ALDH1A7 -4845 CATTAAACTGAGAGCATCCTAGAATTACTTTTCTCTCTCATGTGCAGTTAAAGCACCAGA -4786
HNF1.02
RR-ALDH1A7 -4941 CCAAAGCACATGAAGAAAAACCAATTTGGAGTGAATGATGTGCACGGGAAAACATCTATT -4882
DR1
rr-ALDH1A7 -4785 CCAAAGCACATGAAGAAAAACCAATTTGGAGTGAATGATGTGCACGGGAAAACATCTATT -4726
RR-ALDH1A7 -4881 GGAGATCATGGAACCCGGGCTTCTCAGCTGAAAGCATTTAGAACTGGAAGCAGATGGGGA -4822
rr-ALDH1A7 -4725 GGAGATCATGGAACCCGGGCTTCTCAGCTGAAAGCATTTAGAACTGGAAGCAGATGGGGA -4666


RR-ALDH1A7 -4821 TAAGCAACAGTTTCGATCCTCTACTGTTCAGCCCCACAGTTTCCACCTCTGGAAGATGTG -4762
rr-ALDH1A7 -4665 TAAGCAACAGTTTCGATCCTCTACTGTTCAGCCCCACAGTTTCCACCTCTGGAAGATGTG -4606
RR-ALDH1A7 -4761 CCCCAGGACCCTGACTGCCTATTCGAATGAGGATACTACTGCATACTTGACCCAGAATTT -4702
rr-ALDH1A7 -4605 CCCCAGGACCCTGACTGCCTATTCGAATGAGGATACTACTGCATACTTGACCCAGAATTT -4546
RR-ALDH1A7 -4701 TAAGGTACTAATAGGCTCTGATCACACCACTTGCAGTGATGAACTGACATACACCACAGT -4642
rr-ALDH1A7 -4545 TAAGGTACTAATAGGCTCTGATCACACCACTTGCAGTGATGAACTGACATACACCACAGT -4486
TCF2.01
HNF1.04
RR-ALDH1A7 -4641 GAATGTAAAGGATCACTGCTTAACAGAAGGGACCAAAACACCATCTTTTCATTTGACTCA -4582
rr-ALDH1A7 -4485 GAATGTAAAGGATCACTGCTTAACAGAAGGGACCAAAACACCATCTTTTCATTTGACTCA -4426
HNF1.01
RR-ALDH1A7 -4581 GGAAAGGATTTGATAAAATCTAACATCTTTTAAAAATTTCTTTTAATTTTTTTAAACTCC -4522
rr-ALDH1A7 -4425 GGAAAGGATTTGATAAAATCTAACATCTTTTAAAAATTTCTTTTAATTTTTTTAAACTCC -4366
TR4.02
RAR_RXR.01
RR-ALDH1A7 -4521 AGATTTTATTCCCTTCTAGGTCCACTCTCTGACTGTTCTTCACCTCCAACCTTCCCCACC -4462
IR9
DR0
IR10
rr-ALDH1A7 -4365 AGATTTTATTCCCTTCTAGGTCCACTCTCTGACTGTTCTTCACCTCCAACCTTCCCCACC -4306
RR-ALDH1A7 -4461 ACCCCATCTCCACGACTATGTTCCCACCCCTCCACCCCCAGCCCCCGCCACCCACCCTCA -4402
rr-ALDH1A7 -4305 ACCCCATCTCCACGACTATGTTCCCACCCCTCCACCCCCAGCCCCCGCCACCCACCCTCA -4246
RR-ALDH1A7 -4401 CACCCCACCAGACCTCCAGTAACTTGAGGGTTAGGTGTTTCTTCTCTGACTGAACCCAGA -4342
ER5
rr-ALDH1A7 -4245 CACCCCACCAGACCTCCAGTAACTTGAGGGTTAGGTGTTTCTTCTCTAACTGAACCCAGA -4186
RR-ALDH1A7 -4341 ACGGTCAGTCCTCTGCTGTATATGTGCTGGGGGCCTCATATCAGCTGGTGTATGTAGTTT -4282
rr-ALDH1A7 -4185 ACGGTCAGTCCTCTGCTGTATATGTGCTGGGGGCCTCATATCAGCTGGTGTATGTAGTTT -4126
RR-ALDH1A7 -4281 GTGGTCCAGTGTCTGAGAGATCTTAGGGATCCAGATTGAGACTGCTGGTCCTTGTACAGG -4222
rr-ALDH1A7 -4125 GTGGTCCAGTGTCTGAGAGATCTTAGGGATCCAGATTGAGACTGCTGGTCCTTGTACAGG -4066
TR2_TR4.01
RR-ALDH1A7 -4221 GTGGTCCTACTCCTCAGATTCTTCCAGCTTTTCCCTAATTCAACCACAGAGGTCAGCAAC -4162
DR7
DR8
IR8
rr-ALDH1A7 -4065 GTGGTCCTACTCCTCATATTCTTCCAGCTTTTCCCTAATTCAACCACAGAGGTCAGCAAC -4006
RR-ALDH1A7 -4161 TTCTGCTCACTGGTTGGGTGCTGGTATCTGCATCTAACTCTATCAGCTGCTTATTGACTC -4102
rr-ALDH1A7 -4005 TTCTGCTCACTGGTTGGGTGCTGGTATCTGCATCTGACTCTTTCAGCTGCTTATTGACTC -3946
RR-ALDH1A7 -4101 TTTTGGAGGGCAGTCATATCAAGTCCCTTGTTGTGAGCACTCCATAGCTTCAGTAATAGT -4042
rr-ALDH1A7 -3945 TTTTGGAGGGCAGTCATATCAAGTCCCTTGTTGTGAGCACTCCATAGCTTCAGTAATAGT -3886
COUP.02
RR-ALDH1A7 -4041 GTCAGGCCTTAGGGCCTCCCCTTGAGCTGGAAACCACTTTGGGCCTGTTGCTGGACTTCC -3982
rr-ALDH1A7 -3885 GTCAGGCCTTAGGGCCTCCCCTTGAGCTGGATCCCACTTTGGGCCTGTTGCTGGACTTCC -3826
TR2.01
TR4.02
RR-ALDH1A7 -3981 TTTTCCCCAGGCTCCTCTCCATTTCCATTCCTGCATTTCTTTCAGAAAGGAACAATTATA -3922
DR5
rr-ALDH1A7 -3825 TTTTCCCCAGGCTCCTCTCCATTTCCATTCCTGCATTTCTTTCAGAAAGGAACAATTATA -3766
LXRE.02
RR-ALDH1A7 -3921 GGTCAGAGTTTTGACTGTAGGAATGGGAACCCCATTCTTCACTTGATGCCCTGTCCTTCT -3862
rr-ALDH1A7 -3765 GGTCAGAGTTTTGACTGTAGGAATGGGAACCCCATTCTTCACTTGATGCCCTGTCCTTCT -3706
RR-ALDH1A7 -3861 GCTGGAGTTGGGGTCTACAAGTTCCCTTTCCCCACTGTAGAACATTTCATCTAATGTCCC -3802
IR10
rr-ALDH1A7 -3705 GCTGGAGTTGGGGTCTACAAGTTCCCTTTCCCCACTGTAGAACATTTCATCTAATGTCCC -3646
RR-ALDH1A7 -3801 TCTCTTGGAGTCTCTCACCTCCCAGGTCTCTGGTACATTCTGGAGGGTCCTCCCAACCTC -3742
rr-ALDH1A7 -3645 TCTCTTGGAGTCTCTCACCTCCCAGGTCTCTGGTACATTTTGGAGGGTCCTCCCAACCTC -3586

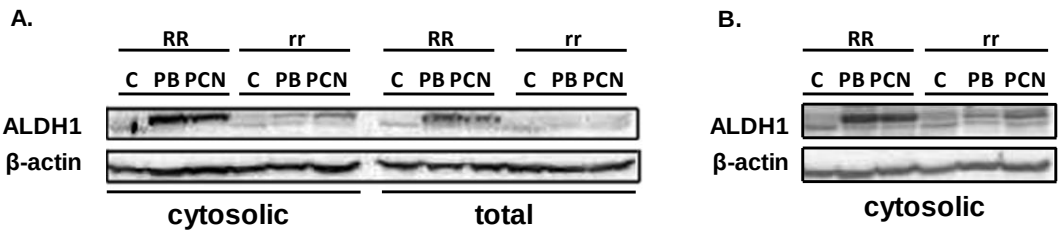

RR-ALDH1A7 -3741 CTACCTCCCAAGGTTGCCTGTTTCAATTCTTTCTGCTGGCCTCATGGCTTCAGCCTGCC -3682
SP1.01
rr-ALDH1A7 -3585 CTACCTCCCAAGGTTGCCTGTTTCAATTCTTTCTGCTAG-CCCTCAGGGCTTCAGCCTCC -3527
RR-ALDH1A7 -3681 CCCAACACCAT-----ATCATATTCCTCTCCTCCCATCCCCTTTCTCACCTAATGATCAA -3627
rr-ALDH1A7 -3526 CCTCACCCAATACCAAATCATATTCCTCTCCCCCATTCCCTTTCTCACCTAATGATCAA -3467
VDR_RXR.03
RR-ALDH1A7 -3626 ATAAATCAACAAAGTAAACATAGAGAACTGTTCCTCAACGTATCAAACTCTTAGCATACT -3567
rr-ALDH1A7 -3466 ATAAATCAACAAATTAAGCATAGAGAACTGTTCCTCAACATATCAAACTCTTAGCATACT -3407
TR4.02
HNF4.03
RR-ALDH1A7 -3566 CCTGAGTAGAAATGTGAATTCTGTCCTCTAGAAGAAGACACAGGGCAGATAGAATGTTCA -3507
DR1
DR7
rr-ALDH1A7 -3406 CCTGAGTAGAAATGTGAATTCTGTCCTCTAGAAAAATACACAGGGCAGGCAGGATGTTCA -3347
RR-ALDH1A7 -3506 CACCCATTGCCCTCATTCTGTGTCATATTGGAAGTCCTTATTCAGAGTACTAGCAAGAGA -3447
IR7
rr-ALDH1A7 -3346 CACCCATTGCCCTCATTCTGTGTCATATTGGAAGTCCTTATTCAGAGTACTAGCAAGAGA -3287
ER5
TR2.01
RR-ALDH1A7 -3446 AATAGATGAAGAATTTTAAAGCAAAATGTAATGGTCTAATGATGTT-CATTTGTCAGTAA -3388
rr-ALDH1A7 -3286 AATAGATGAAGAATTTTAAATAAAAATGTAATGGTCTAATGAT-TTCCATTTGTGAGTAA -3228
DR10
RR-ALDH1A7 -3387 CACAAAGAAACATGGTTTTGTAAATTTCAAAACATAAAGAGTT----CACACACACACAA -3332
rr-ALDH1A7 -3227 CACAGAGCAACATGGTTTTGTAAATTTCAAAACATAAAGACTTCACACACACACACAA -3168
CAR_RXR.01
HNF1.03
RR-ALDH1A7 -3331 ATGCAAGGACACACAAACACACACTGAATGG--TGGAGAAAAAAATTAATT-AAGTTACA -3275
rr-ALDH1A7 -3167 ATGCAAGGACACACAAACACACACTGAATGGGGTGGAGAAAAAAATTTAGTCAAGTTACA -3108
RR-ALDH1A7 -3274 GGATTAAAAGCAATATATAAAAATCAGGAACAACTTTATATTACAATGAAATTTCTGAAA -3215
rr-ALDH1A7 -3107 GGATTAAAAGCAATATATAAAAATCAGGAACAACTTTATATTCCAATGAACTTTCTGAAA -3048
IR5
RR-ALDH1A7 -3214 A-AAATGTTCACAAGAGGTAGGAGATATCAGACACCATGGAAAATCATCTTATTACATGT -3156
rr-ALDH1A7 -3047 AGAAATGTTCACAAGAGGTAGGAGATATCAGACACCATGGAAAATCATCTTATTGCATGC -2988
ER8
HNF4A.01
RR-ALDH1A7 -3155 GTTGACCTTCATAAAAAGTTCTATGATTTACCTATGCCAATAATTAGGTGAACTGTAACA -3096
COUP.02
DR10
IR9
ER7
rr-ALDH1A7 -2987 GTTGAACTTTATAAAAATTTCTGTGATTTACCCATGACAATAATTAGGTGAACTGTAGCA -2928
CAR_RXR.01
HNF4A.01
RR-ALDH1A7 -3095 AAGTGCATGAGAACATTCATCAGATGTACTTCTGTGACCTTTGGTGAGTAAATAAAAGGC -3036
DR2
DR4
DR2
IR9
DR4
DR2
rr-ALDH1A7 -2927 AAGTGCATGAGAACATTCATCAGATGTACTTCTGTGCCCTTTGGTGAGTGAAT-AAAGGC -2869
RR-ALDH1A7 -3035 TGTAAGGGTAGGTCTTGAAACAGTCATTGACTATTTTGATATAGTATCTAAAGCATTGGT -2976
rr-ALDH1A7 -2868 TGTAAGGGTAGGTCTTGAAACAATCATTGACTATTTTGATATAGTATCTAAAGCATTGGT -2809
HNF1.02
HNF1.01
HNF1.01
HNF1.04
RR-ALDH1A7 -2975 TAGTGTATAATCCCTAGAAAATATTAACCAAAGTTGGCCACTCATCCATTTTTTAATCAA -2916
rr-ALDH1A7 -2808 TAGTGTATAATCCCTAGGAAATATTAACCAAAGTTGGCCATTCATTCATTTTGTAATCAA -2749
RR-ALDH1A7 -2915 TATGTGATATCTAAACCAACAAGGACTATTCCTCTGAA-GGTCTTTTAAAAAAAATCCTG -2857
rr-ALDH1A7 -2748 TATGTAATATCTAAACCAACAAGGACTATTCCTCTGAAAGGTCTTTTAAAAAAA------ -2695
PXR_RXR.01
RR-ALDH1A7 -2856 GGTTGGTTCAAA-TAAGTTGGAGAAGACATCACA-----ATGTTGTGAAATGTATCTGTG -2802
TR2.01
DR4
DR4
CEBPB.02
rr-ALDH1A7 -2694 ------CCTGGGTTAGGTTCAGACATGGGATCACAGAGCAACCTGTGAAATGTATCTGTG -2641
HPF1.01
DR6
DR1
CEBPB.02
RR-ALDH1A7 -2801 AATG----------CAAG-TGGCCTAGATAGAGAAAATGACTGAAAGAAGCTG -2760
rr-ALDH1A7 -2640 AATGTAAGTGGCCTACAGCATCGTCTTGAACTAGATAGAGAAAAGGACTGAAAGAAGCTG -2581
HMBOX.01
NFY.04
ER9
RR-ALDH1A7 -2759 ATAAATCTGATGCTGGGATTGGCCCTTCAGGATCAGAACATTAATGATAACTCTTCCCCA -2700
rr-ALDH1A7 -2580 ATAATTCTGATGCTGGCATAGGTCCTCCAAGATCAGGACATTGATGATAACTCTTCCCCA -2521
CAR_RXR.01
DR4
DR9


RR-ALDH1A7 -2699 TAACTCTTTTTACTGTGAGATTTGTAGAACTAACACATCATTATGCACCTGTTGCCATGA -2644
rr-ALDH1A7 -2520 TAACTCTTTTTACTCTAAGATTTGTAGAACTAACACATCATTATGCACCTGTTGCCATGA -2461
RR-ALDH1A7 -2643 CCAAAAGACATCATATTCCTGAACACCGTGATGATCATGGAGCTCTATGTCCTCCCTAGA -2584
rr-ALDH1A7 -2460 CCAAAAGACATCATATTCCTGAACACCGTGCTGATCATGGAGCTCTATGTCCTCCCTAGA -2401
RR-ALDH1A7 -2583 ACTTTGCTCCACTCAAGGAATCTTGTGATTTCTCTGCTAAATTCTCCCCTGGGATGTTAA -2524
rr-ALDH1A7 -2400 ACTTTGCTCCACTCAAGGAATCTTGTGATTTCTCTGTTAAATTCTCCCCTGGGATGTTAA -2341
RR-ALDH1A7 -2523 TCAAACTGACATTTAATTTCCAGAATTTATGACTTCTGACAACTTTATTTCCAACAAATA -2464
rr-ALDH1A7 -2340 TCAAACTGACATTTAGTTTCCAGAATTTATGACTTCTAACAACTTTATTTCCAACAAATA -2281
RR-ALDH1A7 -2463 CTTATTGAGATAATCCATAAAGTATATATTTTAATGTTGTCTTTAATATAACATCTACTT -2404
rr-ALDH1A7 -2280 TTTATTGAGATAACCCATAAAGTATATATTTTAATGTTGTCTTTAATATAACATCTATTT -2221
RR-ALDH1A7 -2403 ACAATGCCACACATCCTTTTACAAACATTGATGTGTATCTGAAGAAAAGTTCATGAAATG -2344
rr-ALDH1A7 -2220 ACCTTGACACACATCCTTTTAT---CATTGATGTGTATCTAAAGAAAAGTTCATGAAATG -2164
RR-ALDH1A7 -2343 ATATTGACCTACAGAACTTGCAGTACAAAGGAACATATTATATTCTGTTGTATTTTATGT -2284
rr-ALDH1A7 -2163 ATATTGACCTACAAAACATCCAGTGCAAAGGAACATATTATATTCTGTTGTATTCTATTT -2104
RR-ALDH1A7 -2283 TTTGAAATGCTGTCCTTAACCCACTAGGCTGATTTGATACTTTATACTGTAGCTGTCTTC -2224
rr-ALDH1A7 -2103 TTTGAAATGCTGTCCTTAACCTACTAGGCTGATTTGATACTTTA---------- -2060
RR-ALDH1A7 -2223 AGATACACCAGAAGAGGGCATCAGATCTCTTTACAGATGGTTGTGAGCCACCATGTGGTT -2164
rr-ALDH1A7 -2059 ---------- -2060
RR-ALDH1A7 -2163 GCTAGGAATTGAACTCATGACCTCTGGAAGAGCAGTCGGGTGCTCTTAACTGCTGAGCCA -2104
rr-ALDH1A7 -2059 ---------- -2060
RR-ALDH1A7 -2103 TCTCTCCAGCCCGTTGATTTGATACTTTAGGAGCAAACTTTGCTCTTGTAATTACTTTGA -2044
rr-ALDH1A7 -2059 ----------GGAGCAAACTTTGCTCTGATAATTACTTTGA -2029
RR-ALDH1A7 -2043 TTATTGGATTATTTGGATAAAAATGAACACTATTTGTGGAAAATGAAATTGTATTCAGCC -1984
rr-ALDH1A7 -2028 TTATTGGATTATTTGGATAAAAATGAACACTATGTGTGGAAAATGAAATTGTATTCAGCC -1969
RR-ALDH1A7 -1983 AGCTCAGGATAAACCAGCATAGTCTCTCTATCAAGCCAAAGGAAAGACTAGGCCTTAGAG -1924
rr-ALDH1A7 -1968 AGCTCAGGATAAACCAGCATAGTCTCTCTATCAAGCCAAAGGAAAGACTAGGCCTTAGAG -1909
RR-ALDH1A7 -1923 ATGACTTACCTGTACAACACCATCTGAGAATGCAATATTTTTGACTGTGCACAAATGGCT -1864
rr-ALDH1A7 -1908 ATGACTTACCTGTACAACACCATCTGAGAATGCAATGTTTTTGACTGTGCACAAATGGCT -1849
RR-ALDH1A7 -1863 CATTCTGGTAACTGGGTCCAGACAACTTTTATTAAACAAGGTGGATTGCATTAAGTCATC -1804
rr-ALDH1A7 -1848 CATTCTGGTAACTGGGTCCAGACAACTTTTATTAAACAAGGTGGATTGTATTAAGTCATC -1789
RR-ALDH1A7 -1803 AAAACTGTCAACGAAAAACATCTCTCACTTCAAAGAAAATAAGGGACAAGTTAATATTGT -1744
rr-ALDH1A7 -1788 AAAACTGTCAACCAAAAACATCCCTCACTTCAAAGAAAATTAGAGACAAGTTAATATTGT -1729
RR-ALDH1A7 -1743 TGAGTGAAACATGAGTCGCTATGGTTCAGGAACATGGTTTCAGTTTGCCCTGAAAGCAAG -1684
rr-ALDH1A7 -1728 TGAGTGAAACATGAGTCGCTATGGTTCAGGAACATGGATTCAGTTTGCCCTGAAAGCAAG -1669

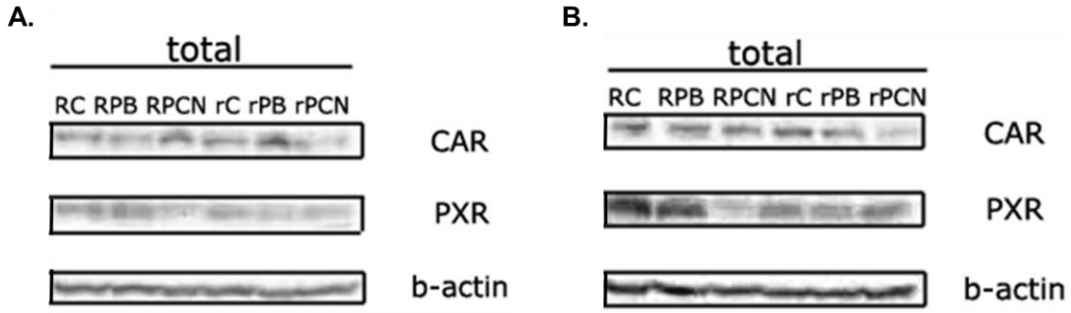

RR-ALDH1A7 -1683 CCCAAATTTGCAAATGTTTTCAGGAACTTTTTATTGGTAGAGGACATAGCAGTCCTATAA -1624
rr-ALDH1A7 -1668 CCCAAATTTGCAAATGTTTTCAGGAACTTTTTTTGGTAGAGGTTATAGGCATCCTATAA -1609
RR-ALDH1A7 -1623 TGGAGCAAAGTCCAGATACCACCAAAATCCAACTTGGTGAACTTTACTGGGGTCACTTAT -1564
rr-ALDH1A7 -1608 CAGAGCAAAGTCCAGATACCACCAAAGTCCAACTTGGTGAACTTTAGTGGGGTCACTTAT -1549
RR-ALDH1A7 -1563 AAGAATACAGGAGGTGGGCTACTTACAGGGACATAAACGACTCAAAGACAGCTGTATCAC -1504
rr-ALDH1A7 -1548 AAGAATGCAGGAGGTGGGCTACTTACAGGGACATAAACGACTCAAAGACAGCTGTATCAC -1489
RR-ALDH1A7 -1503 TGAAGTATGTACCCTAGCTTGGGTGACAGCTCCTAAACCTGGAACCTGGAGCACACTACA -1444
rr-ALDH1A7 -1488 TGAAGTATGTACCCTAGCCTGGGTGACAGCTCCTAAACCTGGAACCTGGAGCACACTACA -1429
RR-ALDH1A7 -1443 CAGATTGGAAGTAGCTTAACCTGTTGAAAAGTGCCCTTTCAAGTGATTCACTTTGTCTAA -1384
rr-ALDH1A7 -1428 CAGATTGGAAGTAGCTCAACATGTTGTAAAGTGCCCTCTCCAATGATTCACTTTGTCTAA -1369
RR-ALDH1A7 -1383 GTTTCGTCCAAGAAACGGA--TTGTTTTTGCTTCTTCCAGGGATCAGGTGATGTGTTAAA -1326
rr-ALDH1A7 -1368 ATTTTGTCCAGGAAACTTGGCTTGTTTTTGCTTCTTCCAGGGATCAGGTGATGTGTTAAA -1309
RR-ALDH1A7 -1325 GTCTTCTTTGCAGTTCAGCTCATCCGAGGGTCTTCTTTTGCAGCTTGGTTTATCAGAGAC -1266
rr-ALDH1A7 -1308 GTCTTCTTTGCAGTTCAGCTCATCTGAGGGTCTTCTTTTGCAGCTTGGTTTATCAGAGAC -1249
RR-ALDH1A7 -1265 TGACTCACAGTAGTCTGTATTGTTTACTCTTCAGAGAGAGGGACCTGGTGAATATAGTCG -1206
rr-ALDH1A7 -1248 TGACTCACAGTAGTCTGTATTGTTTACTCTTCAGAGAGAGGGACCTGGTGAATGTAGTCA -1189
RR-ALDH1A7 -1205 GTTTCAGGGACTTCCTGAAACTATTTTGAGTGGTTTATCTTCTGTCTTAAGGTTATTCCC -1146
rr-ALDH1A7 -1188 GTTTCAGGGACTTCCTGAAACTATTTTGAGTGGTTTATCTTCTGTCTTAAGGTTATTCCC -1129
RR-ALDH1A7 -1145 TGGAGGATTGAATGTTTCATTCACCAAGGAAGCTGCCATAGAACAATATATACCAAAAGA -1086
rr-ALDH1A7 -1128 TGGAGGATTGAATGTTTCATTCACCAAGGAAGCTGCCATAGAACAATATATACCAAAAGA -1069
RR-ALDH1A7 -1085 ACATCAAAAATCAAGGCATTTTCAATTACATTTTTGGTGTATCATATAGTCAGTCAAGGT -1026
rr-ALDH1A7 -1068 ACATCAAAAATCAAGGCATTTTCAATTACATTGTTGGTGTATCATATAGTCAGTCAAGGT -1009
RR-ALDH1A7 -1025 GAAAAGAAACACATTCCCAGGCTCCTGTTCCTCCCCCTAGTTAACTCTGAGAGCAGGGGA -966
rr-ALDH1A7 -1008 GAAAAGAAACACATTCCCAGGCTCCTGTTCCTCCCCCTAGTTAACTCTGAGAGCAGGGGA -949
RR-ALDH1A7 -965 AGTGGCAATTTCATCAGGAAGACATTATCACTAATTCTTACTAACAAGAGCATCCTAAGA -906
rr-ALDH1A7 -948 AGTGGCAATTTCATCAGGAAGACATTATCACTAATTCTTACTAACAAGAGCATCCTAAGA -889
RR-ALDH1A7 -905 ATGTTATCTTGTTCCTTCCATATCTTGTGCAGGAGCGTCAACACAGAGAAGCATTGTATT -846
rr-ALDH1A7 -888 ATGTTATCTTGTTCCTTCCATATCTTGTGCAGGAGCATCAACACAGAGAAGCATTGTATT -829
RR-ALDH1A7 -845 CAAAATGAAAACACTACAGTTAAAATTCGTCCTTTACCAAGTTCCCAGCTCAAAACATTA -786
rr-ALDH1A7 -828 CAAAATGAAAACACTACAGTTAAAATTCGTCCTTTACCAAGTTCCCAGCTCAAAACATTA -769
RR-ALDH1A7 -785 CGAATTTAAACTCTTAGGGGGCAATATTTACATTCTTTCCTTCTATGCTGCCTAGACTAA -726
rr-ALDH1A7 -768 AGAATTTAAACTCTTAGGGGGCAATATTTACATTCCTTCCTTCTATGCTGCCTAGACTAA -709
RR-ALDH1A7 -725 AGATTCAAAAGTGCAGACACTTCCCAAATCAGACATTACTTTCACTTTAAATGAATTTCC -666
rr-ALDH1A7 -708 AGATTCAAAAGTGGAGACACTTCCCAAATCAGACATTATTTTCACTTTAAATGAATTTCC -649
NFY.04
HNF4A.01
ER5
HNF4A.01
ER6
DR1
HNF4A.01
IR6
NFY.04
ER7
AHR.01
DR9
AP1.01
JUNDM2.02
NFY.04
APR1.01
VDR_RXR.03
AP1.01 AP1.03
IR2
CAAT.01
CAAT.01
PNR.01
TCF2.01
HNF1.04
HMBOX.01
HNF1.04
RXR_RXR.01
DR1
HNF1.01
CEBPB.02



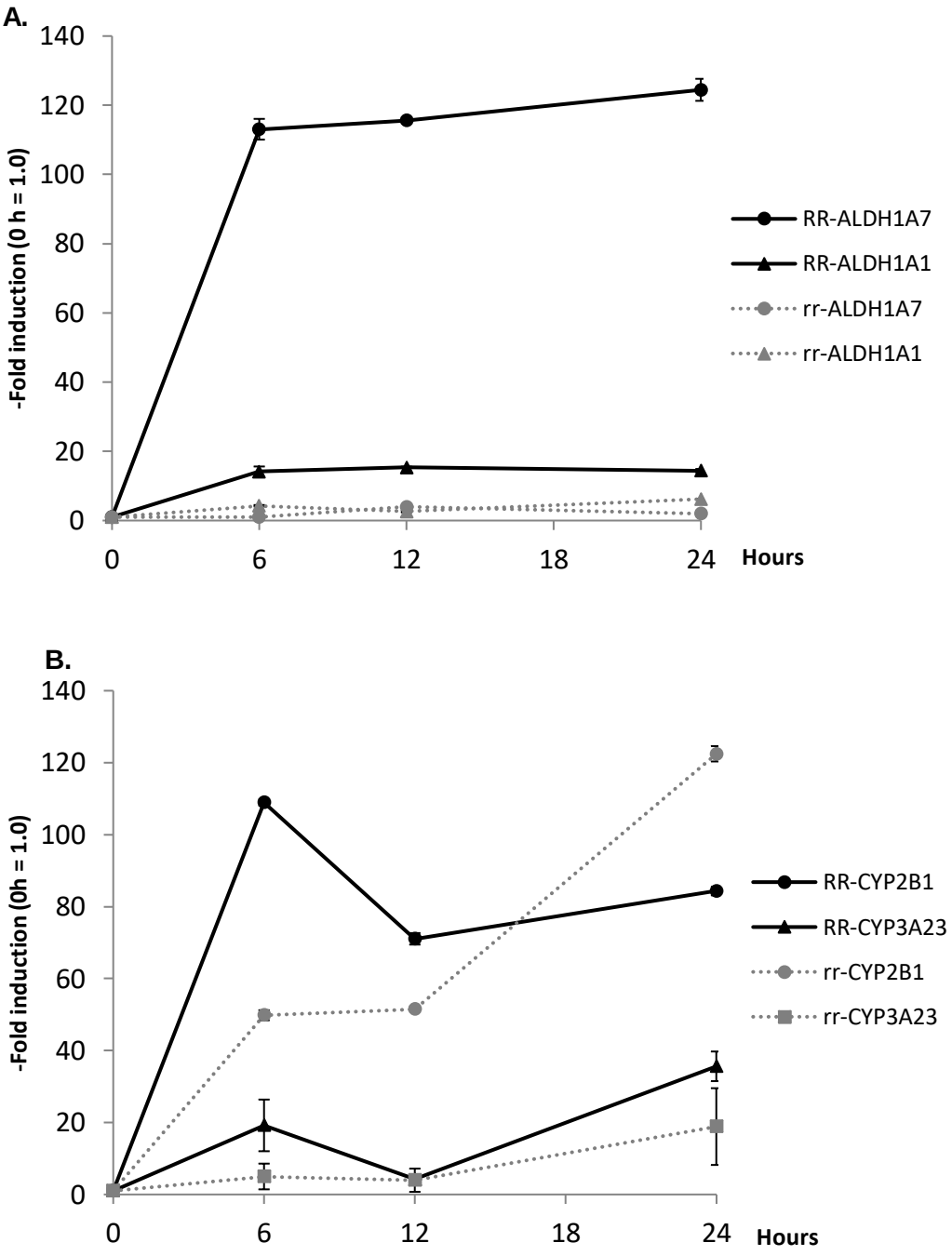
Supplementary Figure 1. Alignment of ~10 kbp *ALDH1A7* promoter sequences from RR and rr rat strains. RR and rr prefixes indicate the strain and yellow and blue highlights correspond to point mutations and deletions, respectively. The 165-bp deletion begins at position -2239 and the 16-bp deletion at position -50 of the *RR-ALDH1A7* sequence. Putative TF sites and the TATA box are indicated by the colored boxes and TF abbreviations, and direct, inverted and everted repeats by arrows. Transcription start site is shown by a red box and marked TSS +1. The 5' *Kpn*I (GGTACC) and 3' *Xho*I (CTCGAG) cloning sites and pGL3-basic sequences are indicated by black nucleotides and *italics*. The numbering relative to the TSS is based on the Rnor 5.0 genome assembly that is derived from a single rat strain (BN/SsNHsdMCW) while the current assembly Rnor 6.0 is a composite from mixed strains.



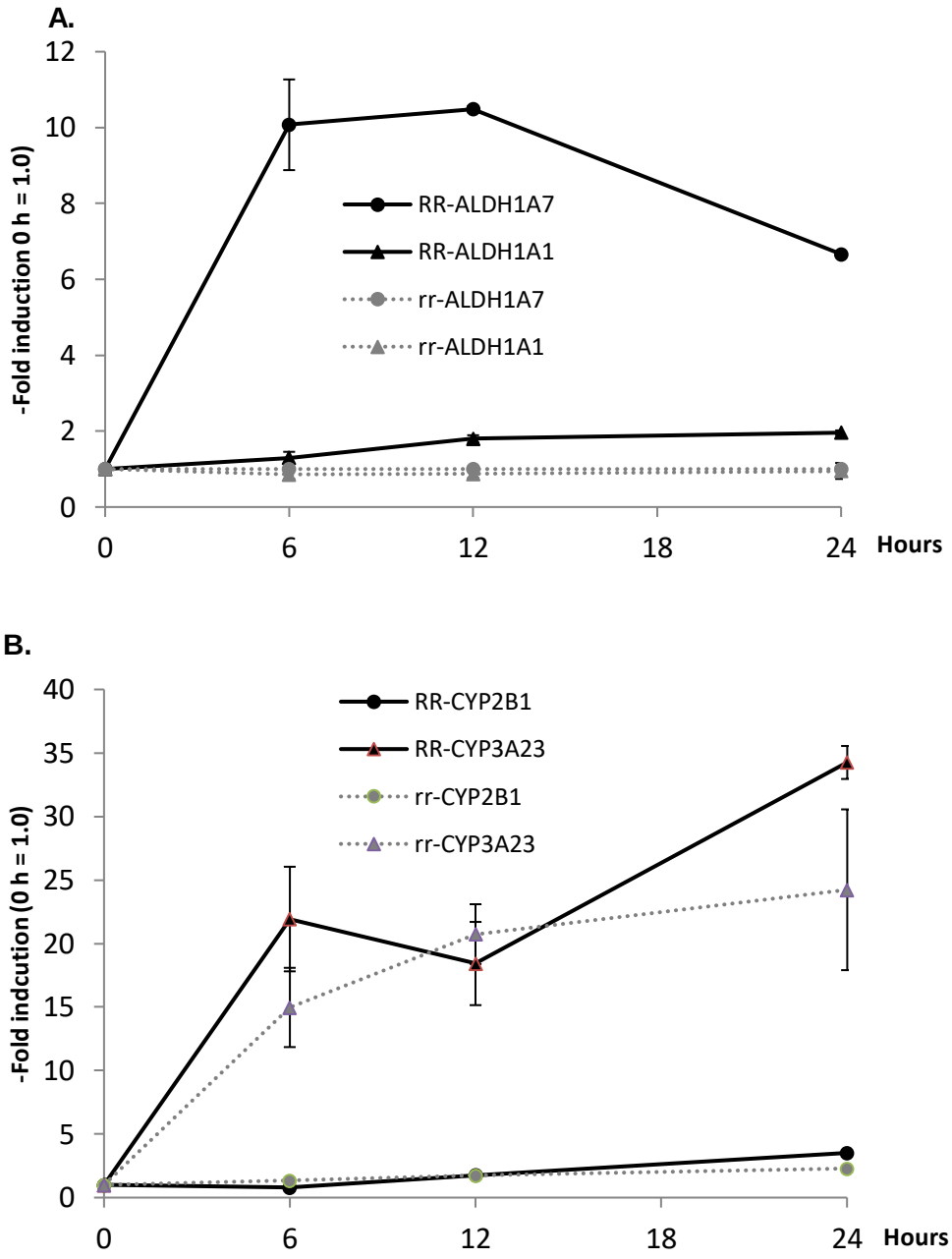
Supplementary Figure 2. Cytosolic and total ALDH1 protein levels in control, PB- and PCN-treated ***RR*** and ***rr*** rat livers. Panels A and B represent data from individual animals from independent experiments.



Supplementary Figure 3. Total CAR and PXR protein levels in control, PB- and PCN-treated ***RR*** and ***rr*** rat livers. Panels A and B represent data from individual animals from independent experiments. Labels R and r denote ***RR*** and ***rr*** strains, respectively.



Supplementary Figure 4. Time course of PB induction in ***RR*** and ***rr*** rats. Normalized ALDH1A7 and ALDH1A1 (A) and CYP2B1 and CYP3A23 (B) mRNA expression at 0, 6, 12 and 24 hours after injection with a single dose of 80/mg/kg PB. N = 4, each sample was measured in duplicates. The time point 0 hours was set at 1.00 for all genes.



Supplementary Figure 5. Time course of PCN induction in ***RR*** and ***rr*** rats. Normalized ALDH1A7 and ALDH1A1 (A) and CYP2B1 and CYP3A23 (B) mRNA expression at 0, 6, 12 and 24 hours after injection with a single dose of 50 mg/kg PCN. N = 4, each sample was measured in duplicates. The time point 0 hours was set at 1.0 for all genes.