DMD Fast Forward. Published on March 20, 2012 as doi:10.1124/dmd.112.045088

RNA-Seq Reveals Dynamic Changes of mRNA Abundance of Cytochrome P450s and their Alternative Transcripts during Mouse Liver Development

Lai Peng, Byunggil Yoo, Sumedha S. Gunewardena, Hong Lu, Curtis D. Klaassen, and Xiao-bo Zhong

Department of Pharmacology, Toxicology, and Therapeutics, University of Kansas Medical Center, Kansas City, Kansas (L.P., C.D.K., X.B.Z); Kansas Intellectual and Developmental Disabilities Research Center, Kansas City, Kansas (B.Y., S.S.G); Department of Pharmacology, Upstate Medical University, State University of New York, Syracuse, New York (H.L.)

Running Title: Ontogeny of CYPs in mouse liver revealed by RNA-Seq

Corresponding author: Dr. Xiao-bo Zhong, Department of Pharmacology, Toxicology, and Therapeutics, The University of Kansas Medical Center, 3901 Rainbow Boulevard, Kansas City, Kansas 66160, USA. Telephone: 913-588-0401; Fax: 913-588-7501; E-mail: xzhong@kumc.edu.

Number of text pages: 40

Number of figures: 5

Number of tables: 1

Number of references: 33

Number of words in Abstract: 259

Number of words in Introduction: 661

Number of words in Discussion: 1,265

ABBREVIATIONS:

CYP, cytochrome P450; FDR-BH, Benjamini-Hochberg-adjusted False Discovery Rate; FPKM, fragments per kilobase of exon per million reads mapped; RNA-Seq, RNA sequencing;

PCR, polymerase chain reaction; 3'RACE, rapid amplification of cDNA 3' end; 3'-UTR, 3'-untranslated region.

DMD Fast Forward. Published on March 20, 2012 as DOI: 10.1124/dmd.112.045088
This article has not been copyedited and formatted. The final version may differ from this version.

Abstract

Cytochrome P450s are a superfamily of enzymes, which have critical functions in liver to catalyze the biotransformation of numerous drugs. However, the functions of most CYPs are not mature at birth, which can markedly affect the metabolism of drugs in newborns. Therefore, characterization of the developmental profiles and regulatory mechanisms of CYP expression is needed for more rational drug therapy of pediatric patients. An animal model is indispensable for studying the mechanisms of postnatal development of the CYPs. Hence we used RNA-Seq to provide a "true-quantification" of mRNA expression of all CYPs in mouse liver during development. Liver samples of male C57BL/6 mice at 12 different ages from prenatal to adulthood were used. Total mRNAs of the 103 mouse CYPs displayed two rapid increasing stages after birth, reflecting critical functional transition of liver during development. Four ontogenic expression patterns were identified among the 71 significantly expressed CYPs, which categorized genes into neonatal-, adolescent-, adolescent/adult-, and adult-enriched groups. The 10 most highly expressed subfamilies of mouse CYPs in livers of adult mice were CYP2E, 2C, 2D, 3A, 4A, 2F, 2A, 1A, 4F, and 2B, which showed diverse expression profiles during development. The expression patterns of multiple members within a CYP subfamily were often classified to different groups. RNA-Seq also enabled the quantification of known transcript variants of CYP2C44, 2C50, 2D22, 3A25, and 26B1, and to identify novel transcripts for

CYP2B10, 2D26, and 3A13. In conclusion, this study reveals the mRNA abundance of all the CYPs in mouse liver during development, and provides a foundation for mechanistic studies in the future.

Introduction

Cytochrome P450s (CYPs) are a superfamily of enzymes, which catalyze the oxidation of organic substances, including the biotransformation of numerous endobiotics (e.g. steroids, fatty acids, and eicosanoids) as well as the detoxification or bioactivation of a variety of xenobiotics (e.g. drugs, chemical carcinogens, and environmental contaminants) (Nebert and Gonzalez, 1987; Danielson, 2002). CYPs are the major enzymes involved in the metabolism and bioactivation of drugs, accounting for about 75% of drug biotransformation (Guengerich, 2008). Liver expresses the largest number of individual CYP enzymes (Hrycay and Bandiera, 2009). However, most CYPs in the liver are lowly expressed at birth. The expression of CYPs changes during liver development, which has been categorized into several different developmental patterns, and considerable interindividual variability occurs in the immediate postnatal period (Hines and McCarver, 2002; Hines, 2007). Low CYP expression in liver during postnatal development is thought to be responsible for the substantial pharmacokinetic differences between newborns and adults, and thus contributes to differences in therapeutic efficacy and adverse drug reactions in pediatric patients (Blake et al., 2005; Hines, 2008). One example is that the low CYP3A4 in neonatal livers results in a low capacity to oxidize cisapride, which can result in QT prolongation in pediatric patients (Pearce et al., 2001; Treluyer et al., 2001). An in-depth understanding of the

regulation of the ontogeny of human CYPs is needed for safer and more effective drug therapy for pediatric patients.

The paucity of suitable tissue samples and limitations due to ethical and technical issues have made it difficult to study the mechanisms controlling the ontogenic expression of CYPs in human liver (Rowell and Zlotkin, 1997). Animal models would be advantageous in overcoming these difficulties and minimizing the influence of genetic variations and the environment. In recent years, the mouse and rat have surpassed many other laboratory animals as the experimental models of choice for the study of physiology, metabolism, and disease (Muruganandan and Sinal, 2008; Hrycay and Bandiera, 2009). Advantages of these models include rapid growth, easy maintenance, and the development of genetic manipulation techniques for mechanistic studies with gain-of-function and loss-of-function strategies. Several researchers have examined the ontogenic gene expression profiles of a few CYPs in mouse or rat liver (Choudhary et al., 2004; Alcorn et al., 2007; Cherala et al., 2007; Hart et al., 2009; Li et al., 2009a). Developmental expression patterns of some CYPs in mice and rats are similar to that in humans.

The previous studies quantified CYP gene expression at the mRNA level by either microarrays or multiplex suspension arrays (Hart et al., 2009; Li et al., 2009a), which only provides relative quantification of a given CYP. These technologies detect mRNA levels by

probe hybridization and fluorescence signal intensity, which cannot compare expression levels among various CYPs, because different probes may have different hybridization efficiency. With the development of next-generation sequencing technology, RNA-sequencing (RNA-Seq) can define a transcriptome with low levels of background noise, no upper limit for quantification, and a high degree of reproducibility for both technical and biological replicates (Mortazavi et al., 2008; Nagalakshmi et al., 2008). More importantly, compared to microarrays, RNA-Seq quantifies the true abundance of mRNA molecules in biological samples, and enables the comparison of expression levels of all genes (Malone and Oliver, 2011). Furthermore, RNA-Seq has the power to quantify expression levels of alternative transcripts of the same gene and to identify novel transcripts efficiently (Pan et al., 2008; Wang et al., 2008; Malone and Oliver, 2011). In the current study, RNA-Seq was used to systematically quantify the abundance of 103 mouse CYP mRNAs in liver during development to define the ontogenic profiles of these CYP mRNAs, with an additional focus on the identification and quantification of alternative transcripts. Therefore, the purpose of this study was to generate comprehensive information on the ontogeny of the mRNAs of the CYP family of enzymes in the livers of mice, which will provide a foundation for determining the regulatory mechanisms controlling the various transcription patterns of CYPs during liver development.

DMD Fast Forward. Published on March 20, 2012 as DOI: 10.1124/dmd.112.045088
This article has not been copyedited and formatted. The final version may differ from this version.

Materials and Methods

Animals. Eight-week-old C57BL/6 breeding pairs of mice were purchased from The Jackson Laboratory (Bar Harbor, ME). Mice were housed according to the American Animal Association Laboratory Animal Care guidelines and were bred under standard conditions in the Laboratory Animal Resources Facility at the University of Kansas Medical Center. The use of these mice was approved by the Institute of Laboratory Animal Resources. Breeding pairs were established at 4:00 p.m. and separated at 8:00 a.m. the following day. The body weights of the females were recorded each day to determine pregnancy status. Livers (n = 3) from male offspring were collected at the following 12 ages: day -2 (gestational day 17), day 0 (right after birth and before the start of suckling), day 1 (exactly 24 hours after birth), and day 3, 5, 10, 15, 20, 30, 45, and 60 (collected around 9:00 a.m.), which represents periods of prenatal (day -2), neonatal (day 0 to 10), juvenile (day 15 to 30), and young adult (day 45 and 60). Livers were immediately frozen in liquid nitrogen after removal and stored at -80°C.

Total RNA Extraction. Total RNA was isolated by using RNAzol B reagent (Tel-Test Inc., Friendswood, TX) according to the manufacturer's protocol. Whole livers without gallbladders at various ages were collected for RNA isolation. RNA concentrations were quantified using a NanoDrop Spectrophotometer (NanoDrop Technologies, Wilmington, DE) at a wavelength of 260 nm. The integrity of the total RNA was evaluated on an Agilent 2100 Bioanalyzer (Agilent

Technologies Inc., Santa Clara, CA) and the samples with RNA integrity values larger than 7.0 were accepted for sequencing library construction.

cDNA Library Construction. The cDNA libraries from all samples were prepared by using an Illumina TruSeq RNA sample prep kit (Illumina, San Diego, CA). Three micrograms of total RNA were used as the RNA input based on the manufacturer's recommendation. mRNAs were isolated from the total RNAs by poly-A selection using poly-T primers. RNA fragmentation, first and second strand cDNA syntheses, end repair, adaptor ligation, and PCR amplification were performed according to the manufacturer's protocol. Fragments of the cDNA library ranged from 220 to 500 bps with an average size at 280 bp (including 120 bp adapter sequences). Quality of the cDNA libraries was evaluated using an Agilent 2100 Bioanalyzer (Agilent Technologies Inc., Santa Clara, CA) before sequencing.

RNA-Seq. RNA-Seq was performed on an Illumina HiSeq 2000 instrument by the Genome Sequencing Facility at KUMC. Clusters of the cDNA libraries were generated on a TruSeq flow cell and sequenced for 100 bp paired end reads (2 × 100) with a TruSeq 200 cycle SBS kit (Illumina, San Diego, CA). A PhiX bacteriophage genome and a universal human reference RNA were used as controls and sequenced in parallel with other samples to ensure that the data generated for each run were accurately calibrated during the image analysis and data analysis. In

addition, the PhiX was spiked into each cDNA sample at approximately 1% as an internal quality control.

RNA-Seq Data Analysis. After the sequencing images were generated by the sequencing platform, the pixel-level raw data collection, image analysis, and base calling were performed by Illumina's Real Time Analysis software. The output bcl files were converted to qseq files by Illumina's BCL Converter 1.7 software and subsequently converted to FASTQ files for downstream analysis. The RNA-Seq reads from the FASTQ files were mapped to the mouse reference genome (NCBI37/mm9) and the splice junctions were identified by TopHat 1.2. The output files in BAM (binary sequence alignment) format were analyzed by Cufflinks 1.0.3 to estimate the transcript abundance (Trapnell et al., 2010). The mRNA abundance was expressed as the number of fragments per kilobase of exon per million reads mapped (FPKM).

Data Visualization and Statistics. The *Cyp* genes (103 in total) were retrieved from Cufflinks output for further analysis. Significant gene expression was determined by the drop-in-deviance F test of the fitted FPKM data to a generalized linear model with a poisson link function, a statistic designed to measure the significance of a gene's measured FPKM value relative to a zero FPKM value. The p-values were adjusted for extra Poisson variation and corrected for false discovery by the Benjamini-Hochberg method (FDR-BH) with a threshold set at 0.05 for significant expression (Benjamini and Hochberg, 1995). A two-way hierarchical

clustering dendrogram was generated by JMP v. 9.0 (SAS, Cary, NC) to determine the expression patterns of the CYP mRNAs during liver development.

Validation of alternative mRNA transcripts. Alternative splicing events in the CYP2B10 and CYP2D26 were validated by reverse transcription and end-point PCR. Total RNA (5 μg) was reverse-transcribed with M-MLV reverse transcriptase (Invitrogen, Carlsbad, CA) and random primers in a final volume of 50 μl. End-point PCRs with reaction volumes of 20 μl were performed. Primer sequences are listed in Table 1. The PCR products were separated on a 2% agarose gel. Bands of the expected sizes were excised and purified with a Qiagen QIAquick gel extraction kit (Valencia, CA). Subsequently, DNA was sequenced with sequencing primers (Table 1) by ACGT Inc. (Wheeling, IL). The alternative polyadenylation of CYP3A13 mRNA was validated by Rapid Amplification of cDNA 3'-End (3'RACE) using the FirstChoice RNA ligase mediated rapid amplification of cDNA ends (RLM-RACE, Ambion Inc., Austin, TX) as described before (Li et al., 2012). Primer sequences are shown in Table 1.

DMD Fast Forward. Published on March 20, 2012 as DOI: 10.1124/dmd.112.045088
This article has not been copyedited and formatted. The final version may differ from this version.

Results

Total Expression and Proportions of Individual CYPs during Liver Development.

RNA-Seq generated an average of 175 million 100 bp paired-end reads per sample for the 36 samples from 12 different ages (n=3). More than 80% of the reads from each sample were mapped to the mouse reference genome (NCBI37/mm9) by TopHat 1.2 (data not shown). Transcript abundances of the 103 mouse *Cyp* genes (including 9 pseudogenes) were calculated by Cufflinks and presented as FPKM values (Supplemental Table 1). The FPKM values were highly reproducible in the triplicate samples at each age as indicated by the small standard errors. If the Benjamini-Hochberg adjusted drop-in-deviance F-test p-value (FDR-BH) in at least one of the 12 ages were less than 0.05 for a *Cyp* gene, then that gene was considered to be expressed in liver during development. Of the 103 mouse *Cyps*, 71 genes were expressed during liver development.

Total FPKM values of all CYP mRNAs increased about 25-fold during postnatal liver development from day -2 (479) to day 60 (12,345), with two surges followed by a plateau stage after each surge (Fig. 1A). Fig. 1B shows the composition of the CYPs, represented as percentage of the total CYP mRNA FPKM, at the beginning and end ages of these surge and plateau stages, i.e. day -2, 1, 10, 20, and 60. The first surge occurred from day -2 (FPKM, 479) to day 1 (5,926) with more than a 10-fold increase of total CYP mRNA FPKM values. During

the first surge, the top 10 highly expressed CYPs changed dramatically, with only CYP2D22 and CYP2D10 in the top 10 CYPs at both day -2 and day 1. During the first plateau stage, although the total CYP mRNA FPKM values were relatively constant between day 1 and 10, the individual CYP mRNAs were altered, with CYP2E1 the most abundant at day 10. The second surge occurred between day 10 and 20, with nearly a 2-fold increase of total CYP FPKM values from 6,139 to 11,087. The individual CYP mRNAs, which increased during the second surge from day 10 to 20, were similar, as 6 CYPs appeared at both ages in the top 10 highly expressed CYPs, and the top 3 are the same in an order of CYP2E1, 3A11, followed by 2D26. During the second plateau stage of day 20 to 60, the individual CYP mRNAs continued to change, and became more diverse. At day 60, the top 10 highly expressed CYPs consisted of only 74% of the total CYP mRNAs, and many additional CYPs were expressed.

Expression Patterns of CYPs during Liver Development. *Cyp* genes with a maximum over minimum expression ratio less than two among the twelve studied ages were to be excluded for this party of the study. The data indicate that all 71 significantly expressed *Cyp* genes had larger than two-fold differences in expression during development. Two-way hierarchical clustering analysis of the 71 *Cyp* genes revealed 4 distinct patterns of gene expression with respect to age, which are defined as neonatal-enriched (Group 1), adolescent-enriched (Group 2), adolescent/adult-enriched (Group 3), and adult-enriched (Group 4) (Fig. 2A). The sum of the

FPKM values in each group was plotted against age in Fig. 2B. Group 1 (neonatal-enriched) contained CYP2D26, 3A16, 3A41a/b, 3A44, 4A10, 4A14, 4A31, 4A32, 4F16, 4F18, 20A1, and 39A1. The CYPs in this group increased markedly after birth and reached a peak around day 1 of age, and then decreased and remained relatively low after day 30. Group 2 CYPs (adolescent-enriched) consisted of CYP2A22, 2B9, 2B13, 2B23, 2C37, 2C40, 2C68, 2C69, 2D37-ps, 2E1, 2G1, 4F15, 7A1, and 26B1. The CYPs in this group increased rapidly after birth, reached maximal expression around day 20, and then decreased to lower levels in adult mice. Group 3 CYPs (adolescent/adult-enriched) included CYP1A2, 2A4, 2A5, 2A12, 2B10, 2C29, 2C39, 2C50, 2C54, 2C55, 2C70, 2D22, 2J6, 3A13, 3A25, 3A59, 4B1, 4F14, and 17A1. The mRNA of the CYPs in this group were expressed at very low levels at birth, but increased substantially between day 10 and 25, and reached a plateau in adult mice. Group 4 CYPs (adult-enriched) consisted of CYP2C38, 2C44; 2C67, 2D9, 2D10, 2D11, 2D12, 2D13, 2D34, 2D40, 2D44, 2F2, 2J5, 2R1, 2U1, 3A11, 4A12A/B, 4F13, 4V3, 7B1, 8B1, 26A1, 27A1, and 51. The CYPs in this group were expressed lowly 2-days before birth, gradually increased after birth with a rapid increase around 25-days of age, and reached highest expression in adults (45 and 60 days of ages). The largest correlation distance with respect to neighboring ages for the four groups (as shown by the clade separation on the dendrogram scale) was observed between 20- and 25-days of age.

Comparison of Transcript Abundance of Some CYP Subfamilies during Liver Development. The 103 mouse *Cyp* genes belong to 16 families and 39 subfamilies. Ontogenic patterns of the major CYP subfamilies in liver are summarized in Fig. 3.

CYP1A subfamily. The CYP1A subfamily has two members, CYP1A1 and 2. CYP1A1 mRNA was expressed at very low levels in liver at all ages. CYP1A2 had a developmental pattern belong to group 3 (adolescent/adult-enriched) (Fig. 3A).

CYP2A subfamily. The CYP2A subfamily consists of five members, CYP2A4, 5, 12, 22, and 2AB1. Except for CYP2AB1, the other four genes in liver were expressed with the adolescent or adolescent/adult-enriched patterns (Fig. 3B).

CYP2B subfamily. The CYP2B subfamily contains five members, CYP2B9, 10, 13, 19, and 23. CYP2B9 and 13 contributed the most to the expressed levels of this subfamily with an adolescent-enriched pattern (Fig. 3C).

CYP2C subfamily. The CYP2C subfamily includes 16 members, CYP2C29, 37, 38, 39, 40, 44, 50, 53-ps, 54, 55, 65, 66, 67, 68, 69, and 70. CYP2C29, 37, 50, 69, and 70 were the most highly expressed members in this CYP subfamily in livers of mice. CYP2C69 was neonatal-enriched, CYP2C37 was adolescent-enriched, and CYP2C29, 50, and 70 were adolescent/adult-enriched (Fig. 3D).

DMD Fast Forward. Published on March 20, 2012 as DOI: 10.1124/dmd.112.045088
This article has not been copyedited and formatted. The final version may differ from this version.

CYP2D subfamily. The CYP2D subfamily consists of 10 members, CYP2D9, 10, 11, 12, 13, 22, 26, 34, 37-ps, and 40. Three genes, CYP2D9, 10, and 26, were expressed to a higher extend than the other CYP2D genes. CYP2D26 increased remarkably after birth and remained at a consistent level until day 20, and decreased after 30-days of age (neonatal-enriched). CYP2D9 was expressed lowly until day 20 and increased markedly to reach a plateau at 30-days of age (adolescent/adult-enriched). CYP2D10 increased gradually throughout development (Fig. 3E).

CYP2E subfamily. The CYP2E subfamily has only one member, CYP2E1, which was one of the most abundant CYPs expressed in liver. CYP2E1 increased sharply from day -2 to day 5, and then continually increased to a peak at day 20, followed by a decrease and then increase to a consistently high level thereafter (Fig. 3F).

CYP2F subfamily. The CYP2F subfamily also contains only one member, CYP2F2, which was expressed lowly until 15-days of age, and thereafter gradually increased (Fig. 3G).

CYP3A subfamily. The CYP3A subfamily includes 9 members, CYP3A11, 13, 16, 25, 41A, 41B, 44, 57, and 59. CYP3A11 was the most abundantly expressed gene in this subfamily, which gradually increased throughout postnatal development (Fig. 3H). Several CYP3A members, including CYP3A16, 41A, 41B, and 44, were neonatal-enriched with the highest expressed levels around 5-days of age (Fig. 3H).

DMD Fast Forward. Published on March 20, 2012 as DOI: 10.1124/dmd.112.045088
This article has not been copyedited and formatted. The final version may differ from this version.

CYP4A subfamily. The CYP4A subfamily has 8 members, CYP4A10, 12A, 12B, 14, 29-ps, 30B-ps, 31, and 32. Four members in this subfamily, CYP4A10, 14, 31, and 32, had a neonatal-enriched expression pattern with the highest expression at 1-day of age. The other two members, CYP4A12A and 12B, were not expressed until day 30, and then increased rapidly in adult mice (Fig. 3I).

CYP4F subfamily. The CYP4F subfamily contains 9 genes, CYP4F13, 14, 15, 16, 17, 18, 39, 40, and 41-ps. CYP4F13, 14, and 15 were expressed in liver and they appeared to have an adolescent or adolescent/adult enriched pattern (Fig. 3J). CYP4F16, 18, and 39 were also expressed, but at very low levels.

Ontogenic patterns of the other CYP subfamilies are also presented (Supplemental Table 1).

Ontogeny of Known Transcript Variants of CYPs during Liver Development. Among the 103 mouse *Cyp* genes, six have known alternative transcripts annotated in the NCBI database, including *Cyp1a1* (two isoforms), *Cyp2c44* (two isoforms), *Cyp2c50* (three isoforms), *Cyp3a25* (two isoforms), and *Cyp26b1* (two isoforms). CYP1A1 was expressed very lowly in liver. RNA-Seq was able to provide the ontogenic patterns of the annotated alternative transcripts of the five *Cyp* genes that were significantly expressed in liver (Fig. 4A-E, Supplemental Table 2).

For the two CYP2C44 transcripts, NM_001167905 lacks the last exon of NM_001001446 and has an alternative polyadenylation site. NM_001001446 was the major isoform in mouse

liver throughout development. Its mRNA was high at birth and reached a peak around 5- and 10-days of age, decreased between 10- and 20-days of age, and increased again and reached a high expression plateau after 45-days of age (Fig. 4A). NM_001167905 was not expressed throughout liver development.

Cyp2c50 has three transcripts. NM_001167875 lacks the fifth exon of NM_134144. NM_001167877 is a truncated isoform with the coding region containing only the first 4 exons of NM_134144. NM_134144 was the major isoform in mouse liver throughout development. Its expression profile was a typical Group 3 pattern (adolescent/adult-enriched) (Fig. 4B). NM_001167875 and NM_001167877 were very lowly expressed throughout liver development.

The two CYP2D22 transcripts encode the same protein, but NM_019823 has 6 less nucleotides at the end of the first exon than NM_001163472, which results in a shorter 5'-UTR. NM_019823 was the major transcript of CYP2D22 expressed throughout liver development. mRNA of NM_019823 increased rapidly from 2-days before birth to 3-days after birth, and then remained relatively consistent with a small peak at 20-days of age. mRNA of NM_001163472 remained consistently low at all ages (Fig. 4C).

Cyp3a25 has two transcripts. NR_030782 is a non-coding RNA with a skipped sixth exon of NM_019792. The expression profiles of these two transcripts were similar during development, but NM_019792 had higher expression levels at all postnatal ages. Both transcripts had a "day 1

surge", which decreased markedly by 5-days of age and then rapidly increased between 10- and 20-days of age, and reached a plateau at 20-days of age (Fig. 4D).

The two CYP26B1 transcripts encode the same protein, but have different transcription start sites. NM_001177713 is shorter than NM_175475 in the 5'-UTR. Both transcripts were expressed very lowly during mouse liver development (Fig. 4E).

Identification of Novel Transcripts of CYPs in Liver by RNA-Seq. Potential novel transcripts of CYPs, which are not annotated in the NCBI database, were identified by examining sequencing results in the UCSC genome browser (http://genome.ucsc.edu/cgi-bin/hgGateway). Alignment results from TopHat were converted to BigWig files and uploaded to the browser to display the distribution of reads in the genome. Junction files were also used to determine alternative splicing events. In the browser, each junction consists of two connected blocks, where each block is as long as the maximal overhang of any read spanning the junction. The score is the number of alignments spanning the junction.

Cyp2b10 RefSeq gene in assembly of NCBI37/mm9 was located on chromosome 7 between 26,682,683 and 26,711,643 with 9 exons. A tiny peak of read alignment between the peaks of exon 8 and 9 was observed around chromosome position 26,710,000 in the Bigwig view (Fig. 5A). Two junctions with a score of 3 indicated alternative splicing events that could generate a cassette exon in this region. The cassette exon was validated by endpoint PCR using cDNAs

from two liver samples of 60-days old mice. The forward PCR primer mapped to exon 6 and the reverse PCR primer mapped across the junction between the cassette exon and exon 9 (Table 1), which produced an amplicon containing the cassette exon with a size of 411 bp in all samples examined (Fig. 5A). Sequencing of the PCR product revealed the presence of a 31 bp in the cassette exon, which could alter the protein translation reading frame when retained in the transcript.

A similar cassette exon peak and alternative splicing junctions were also found in the browser view of the *Cyp2d26* gene, which is located on the minus strand of chromosome 15 between 82,620,537 and 82,624,675 (Fig. 5B). The cassette exon was located between exon 7 and 8, and validated by endpoint PCR with two sets of primers. Primer set 1 bound to exon 5 and the cassette exon, which amplified a fragment with an expected size of 520 bp. Primer set 2 bound to the cassette exon and exon 9, which amplified a fragment with an expected size of 327 bp. The cassette exon sequence contained 42 nucleotides, which would change the protein product if retained in the transcript.

Thirteen peaks of sequencing reads were observed along the RefSeq *Cyp3a13* gene. The first twelve exons were covered by a single peak. However, in exon 13, a region with high-level reads followed by a region with low-level reads was found. A sharp decrease in the sequencing reads occurred around chromosome position 138,335,200 in the 3'-UTR of *Cyp3a13* (Fig. 5C).

DMD Fast Forward. Published on March 20, 2012 as DOI: 10.1124/dmd.112.045088
This article has not been copyedited and formatted. The final version may differ from this version.

The RNA-Seq reads ended at coordinate 138,334,161 of chromosome 5, which matched with the termination of this gene in the reference sequence. This suggested an alternative polyadenylated mRNA transcript in the 3'-UTR region of CYP3A13. To validate the shorter 3'-UTR transcript, we performed 3'RACE-PCR with two RNA samples from livers of 60-day old mice. The forward PCR primers, shown in Fig. 9, together with reverse primers provided by the RLM-RACE kit, generated a band that was sequenced using the *Cyp3a13* specific sequencing primer (Table 1). Sequencing results revealed a poly(A)-containing sequence that uniquely mapped to the 3'-UTR of CYP3A13. Termination of the shorter 3'-UTR transcript occurred at chromosome location 138, 336,180, which was more than 2kb upstream of the end of the reference gene sequence. In addition, an AATAAA sequence was noted at 23 bases upstream of the termination site, which might serve as a recognition signal for the poly(A) polymerase complex.

Discussion

In the current study, RNA-Seq was used to truly quantify mRNA abundance of the 103 *Cyp* genes in mouse liver during postnatal development. Compared to other commonly used methods for RNA quantification, such as microarray, branched DNA, or real-time PCR, RNA-Seq directly counts sequence reads of the nucleotide molecules in biological samples, which is the only method for true quantification of mRNA expression. Expression of a transcript is represented by FPKM (fragments per kilobase of exon per million reads mapped), which normalizes sequencing depths between different samples and sizes between various genes, allowing direct comparison of mRNAs among various transcripts in a genome-wide scale. As shown in Fig. 1, the quantitative abundance of mRNAs encoded by the 103 *Cyp* genes was determined in mice at 12 ages, which has not been reported previously.

From perinatal through neonatal to adult life, the total FPKM values of all CYP transcripts increased about 25 fold with two rapidly increasing stages (Fig. 1A). The first surge occurred from 2-days before birth to 1-day after birth. Not only the total abundance, but also the composition of expressed CYPs was markedly altered during this period, when the pup was born and lost the direct physiological connection with the mother. Multiple hormones were known to influence the expression of drug metabolizing enzymes. During pregnancy, maternal hormones cross the placenta into the baby. After birth, the loss of maternal hormones may be a contributing

factor to the changes in the expression of some CYPs. The new born mice also began to be exposed to the environment, which contains all kinds of xenobiotics. A critical function of CYPs is xenobiotic metabolism, thus the first surge of CYP expression may provide protections to the newborns against environmental toxic compounds. The second surge occurred from day 10 to day 20, with a rapid increase in abundance of the CYP mRNAs. But the expression of each CYP was similar from 10- to 20-days of age, as the expression of most CYPs increased simultaneously. This period is also the most rapidly growing stage of postnatal liver development, as studies have shown that cell proliferation in mouse liver is high until 20-days of age, and then the hepatic architecture begins to resemble the adult liver (Apte et al., 2007).

Hierarchical clustering of significantly expressed CYPs revealed four hepatic ontogenic expression patterns (Fig. 2). The expression profiles of these patterns were similar to a previous study of CYP ontogeny in mice (Hart et al., 2009), with some genes enriched at younger ages and some enriched at older ages. However, the previous study examined only a few CYPs, whereas the present study included all the expressed CYPs in the liver of mice, and therefore generated a more comprehensive summary of ontogenic patterns. Whereas, the earlier study concluded there were three patterns of CYP ontogeny in livers of mice (Hart et al., 2009), the present study concludes there are four patterns. Clustering of the CYPs from 12 developmental ages indicates that the largest correlation distance with respect to neighboring ages is between

20- and 25-days of age, which spanned the period when the diet changes from milk to chow. Because CYPs have important functions in metabolizing chemicals absorbed from the gastrointestinal tract, it is reasonable to anticipate this change in CYP expression with a change in food intake.

The most highly expressed CYP mRNAs in liver of mice were from the CYP1-4 families, which belong to the xenobiotic-metabolizing CYPs. They displayed changes in expression during development. Some CYP subfamilies were primarily expressed during the neonatal stage, whereas others were mainly expressed during the adolescent or adult stage of development. Even within each subfamily, individual members were expressed with various developmental patterns. In the CYP2D subfamily, CYP2D26 mRNA reached its maximum expression during the neonatal period, whereas CYP2D9 reached its maximum expression during the adolescent/adult period (Fig. 3E). Similarly in the CYP3A subfamily, CYP3A16, 41A, 41B, and 44 were maximally expressed during the neonatal period, whereas CYP3A11 was maximally expressed in the adult mice (Fig. 3H). A similar developmental switch also exists in human CYP3A enzymes (Schuetz et al., 1994; Stevens et al., 2003; Leeder et al., 2005). Thus mice might serve as a good model to study the mechanisms controlling the developmental switches of CYP expression in liver. Notably, it is difficult to directly compare our data in mice with other previous published data in human. Our study determined levels of mRNA transcripts of CYPs,

whereas most other published studies determined protein concentrations and enzyme activities of human CYPs.

As RNA-Seq has the ability to examine the transcriptome at the resolution of a single base (Wang et al., 2009), it can quantify the mRNAs of genes as well as their individual transcript variants (Fig. 4). The known transcript variants of the CYPs might be translated into different proteins, as in CYP2C44 and CYP2C50, or those that differ only in the UTR regions of the transcripts, as in CYP2D22 and CYP26B1. Because transcript variants can alter protein functions, mRNA stability, subcellular localization, and translation efficiency of the transcripts (Pesole et al., 2000), identification of the predominant transcripts of CYPs and their developmental expression patterns might help to predict variations in protein activity.

A research has shown that 92-94% of human genes undergo alternative splicing, promoter usage, and polyadenylation (Wang et al., 2008). Only six out of the 103 mouse *Cyp* genes in the NCBI database have annotated transcript variants, so alternative splicing CYP transcripts in mice is probably an underexplored area. With the advantage of RNA-Seq, the present study identified several novel transcript variants in CYP2B10 and 2D26, even though they are expressed at very low levels (very small peaks in BigWig view of Fig. 5A and 5B). When the cassette exons are translated, the resultant CYP enzymes will have different protein sequences, which might lead to a difference in enzyme activities. The CYP3A13 transcripts, with the shorter 3'-UTR, appeared

to be the predominant transcript of this gene. A similar shorter 3'-UTR transcript has been reported for the human *CYP3A4* gene (Li et al., 2012). Although the CYP3A4 transcript with shorter 3'-UTR doesn't change amino acid sequences, it is more stable and can translate to protein more efficiently than the CYP3A4 transcript with the longer 3'-UTR (Li et al., 2012). The mouse *Cyp3a13* gene might serve as a model to study how alternative polyadenylation events are regulated. Further studies are needed to investigate the functional significance of these novel transcripts.

CYP proteins and enzyme activities were not determined in this study, due to technique limitations. Specific antibodies against most of the mouse CYP proteins, specific substrates, and inhibitors of many individual mouse CYP enzymes are not available. Technological breakthroughs in proteomics and metabolomics are essential to study the ontogeny of CYP proteins.

In summary, the present study describes the ontogeny of the mRNAs of all CYPs in the mouse liver during development. These data provide the fundamental knowledge for studying mechanisms of regulation of the transcription of *Cyp* genes in liver during development. A previous study from our laboratories indicates that the developmental regulation of *Cyp3a* gene expression in mouse liver may be due to dynamic changes of histone modifications during postnatal liver maturation (Li et al., 2009b). No doubt, epigenetics and various hepatic

transcription factors play an important role in regulating ontogenic expression of CYP enzymes (Kyrmizi et al., 2006). Studies on the networking of transcription factors and epigenetic signatures in liver may shed light on the regulatory mechanisms of the observed CYP developmental patterns. These mechanisms may be of considerable importance in understanding the kinetics of xenobiotic metabolism during the neonatal period.

Acknowledgement:

We thank Drs. Julia Yue Cui, Helen Renaud, and Dan Li for help on sample collection and data analysis. We would like to thank Mr. Clark Bloomer from the KUMC Sequencing Core Facilities for his technical assistance on RNA-Seq.

Authorship Contribution

Participated in research design: Zhong, Klaassen, Lu, Gunewardena, Peng, Yoo.

Conducted experiments: Peng.

Contributed new reagents or analytic tools: n/a.

Performed data analysis: Peng, Gunewardena, Yoo, Zhong.

Wrote or contributed to the writing of the manuscript: Zhong, Klaassen, Lu, Gunewardena, Peng, Yoo.

References

Alcorn J, Elbarbry FA, Allouh MZ and McNamara PJ (2007) Evaluation of the assumptions of an ontogeny model of rat hepatic cytochrome P450 activity. *Drug Metab Dispos* **35:**2225-2231.

Apte U, Zeng G, Thompson MD, Muller P, Micsenyi A, Cieply B, Kaestner KH and Monga SP (2007) beta-Catenin is critical for early postnatal liver growth. *Am J Physiol Gastrointest Liver Physiol* **292:**G1578-1585.

Benjamini Y and Hochberg Y (1995) Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B-Methodological* **57:**289-300.

Blake MJ, Castro L, Leeder JS and Kearns GL (2005) Ontogeny of drug metabolizing enzymes in the neonate. *Semin Fetal Neonatal Med* **10:**123-138.

Cherala G, Shapiro BH and D'Mello A P (2007) Effect of perinatal low protein diets on the ontogeny of select hepatic cytochrome p450 enzymes and cytochrome p450 reductase in the rat. *Drug Metab Dispos* **35:**1057-1063.

Choudhary D, Jansson I, Sarfarazi M and Schenkman JB (2004) Xenobiotic-metabolizing cytochromes P450 in ontogeny: evolving perspective. *Drug Metabolism Reviews* **36:**549-568.

Danielson PB (2002) The cytochrome P450 superfamily: biochemistry, evolution and drug metabolism in humans. *Curr Drug Metab* **3:**561-597.

Guengerich FP (2008) Cytochrome p450 and chemical toxicology. *Chem Res Toxicol* **21:**70-83.

Hart SN, Cui Y, Klaassen CD and Zhong XB (2009) Three patterns of cytochrome P450 gene expression during liver maturation in mice. *Drug Metab Dispos* **37:**116-121.

Hines RN (2007) Ontogeny of human hepatic cytochromes P450. *J Biochem Mol Toxicol* **21:**169-175.

Hines RN (2008) The ontogeny of drug metabolism enzymes and implications for adverse drug events. *Pharmacol Ther* **118:**250-267.

Hines RN and McCarver DG (2002) The ontogeny of human drug-metabolizing enzymes: phase I oxidative enzymes. *J Pharmacol Exp Ther* **300:**355-360.

Hrycay EG and Bandiera SM (2009) Expression, function and regulation of mouse cytochrome P450 enzymes: comparison with human P450 enzymes. *Curr Drug Metab* **10:**1151-1183.

Kyrmizi I, Hatzis P, Katrakili N, Tronche F, Gonzalez FJ and Talianidis I (2006) Plasticity and expanding complexity of the hepatic transcription factor network during liver development. *Genes Dev* **20:**2293-2305.

Leeder JS, Gaedigk R, Marcucci KA, Gaedigk A, Vyhlidal CA, Schindel BP and Pearce RE (2005) Variability of CYP3A7 expression in human fetal liver. *J Pharmacol Exp Ther* **314:**626-635.

Li D, Gaedigk R, Hart SN, Leeder JS and Zhong XB (2012) The Role of CYP3A4 mRNA Transcript with Shortened 3'-Untranslated Region in Hepatocyte Differentiation, Liver Development, and Response to Drug Induction. *Mol Pharmacol* **81:**86-96.

Li T, Huang J, Jiang Y, Zeng Y, He F, Zhang MQ, Han Z and Zhang X (2009a) Multi-stage analysis of gene expression and transcription regulation in C57/B6 mouse liver development. *Genomics* **93:**235-242.

Li Y, Cui Y, Hart SN, Klaassen CD and Zhong XB (2009b) Dynamic patterns of histone methylation are associated with ontogenic expression of the Cyp3a genes during mouse liver maturation. *Mol Pharmacol* **75:**1171-1179.

Malone JH and Oliver B (2011) Microarrays, deep sequencing and the true measure of the transcriptome. *BMC Biol* **9:**34.

Mortazavi A, Williams BA, McCue K, Schaeffer L and Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* **5:**621-628.

Muruganandan S and Sinal CJ (2008) Mice as clinically relevant models for the study of cytochrome P450-dependent metabolism. *Clin Pharmacol Ther* **83:**818-828.

Nagalakshmi U, Wang Z, Waern K, Shou C, Raha D, Gerstein M and Snyder M (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* **320:**1344-1349.

Nebert DW and Gonzalez FJ (1987) P450 genes: structure, evolution, and regulation. *Annu Rev Biochem* **56:**945-993.

Pan Q, Shai O, Lee LJ, Frey BJ and Blencowe BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **40:**1413-1415.

Pearce RE, Gotschall RR, Kearns GL and Leeder JS (2001) Cytochrome P450 Involvement in the biotransformation of cisapride and racemic norcisapride in vitro: differential activity of individual human CYP3A isoforms. *Drug Metab Dispos* **29:**1548-1554.

Pesole G, Grillo G, Larizza A and Liuni S (2000) The untranslated regions of eukaryotic mRNAs: structure, function, evolution and bioinformatic tools for their analysis. *Brief Bioinform* **1:**236-249.

Rowell M and Zlotkin S (1997) The ethical boundaries of drug research in pediatrics. *Pediatr Clin North Am* **44:**27-40.

Schuetz JD, Beach DL and Guzelian PS (1994) Selective expression of cytochrome P450 CYP3A mRNAs in embryonic and adult human liver. *Pharmacogenetics* **4:**11-20.

Stevens JC, Hines RN, Gu C, Koukouritaki SB, Manro JR, Tandler PJ and Zaya MJ (2003) Developmental expression of the major human hepatic CYP3A enzymes. *J Pharmacol Exp Ther* **307:**573-582.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ and Pachter L (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28:**511-515.

Treluyer JM, Rey E, Sonnier M, Pons G and Cresteil T (2001) Evidence of impaired cisapride metabolism in neonates. *Br J Clin Pharmacol* **52:**419-425.

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP and Burge CB (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* **456:**470-476.

Wang Z, Gerstein M and Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* **10:**57-63.

Footnotes

This study was supported by the National Institute of Health National Institute for Environmental Health Science [Grant ES-019487] (to X.B.Z, C.D.K, and H.L); the National Institute of Health National Institute of General Medical Sciences [Grant GM-087376] (to X.B.Z); the National Institute of Health National Institute for Environmental Health Science [Grant ES-009649] (to C.D.K); and the National Institute of Health National Center for Research Resources [Grant RR-021940] (to C.D.K and X.B.Z).

DMD Fast Forward. Published on March 20, 2012 as DOI: 10.1124/dmd.112.045088
This article has not been copyedited and formatted. The final version may differ from this version.

Legends for figures

FIG. 1. A. Total mRNAs of the 103 mouse *Cyp* genes in liver during development. RNA-Seq was done for liver mRNAs of male C57BL/6 mice at 12 ages from 2-days before birth to 60-days after birth. The FPKM values of all 103 mouse *Cyp* genes at each age were added and plotted to show the developmental pattern of total CYP mRNAs. Bars represented the mean ± SEM of three individual animals. **B.** Individual CYP mRNAs (shown as percentage of total CYP mRNAs) at 2-days before birth, 1-, 10-, 20-, and 60-days after birth. Each gene was presented in a unique color for all ages. Only the top 10 CYP mRNAs at each age were listed in alphabetic order, and the rest were grouped as "Others".

FIG. 2. A. Hierarchical clustering of expression profiles for 71 significantly expressed CYP mRNA in livers of male C57BL/6 mice. Significant expression was determined by the Bejamini-Hochberg-adjusted False Discovery Rate (FDR-BH) with a threshold of <0.05. The two trees described the relationship between different gene expression profiles (right tree) and various ages (bottom tree). The dendrogram scale represented the correlation distances. Average FPKM values of 3 replicates per age were given by colored squares. Red: relatively high expression; blue: relatively low expression. The dashed lines categorize the expression profiles into 4 groups. **B.** Changes of total FPKM values of the CYP mRNAs in each group

based on the hierarchical clustering in 2A through different ages during development. Bars represented the mean ± SEM of three individual samples.

FIG 3. A. Expression patterns of the selected CYP subfamilies during liver development. **A**. CYP1A. **B**. CYP2A. **C**. CYP2B. **D**. CYP2C. **E**. CYP2D. **F**. CYP2E. **G**. CYP2F. **H**. CYP3A. **I**. CYP4A. **J**. CYP4F. Data were presented as mean FPKM ± SEM of three individual samples.

FIG 4. The mRNA of known alternative transcripts of 5 CYPs that were significantly expressed during liver development. The transcripts were labeled with their NCBI transcript ID. Data were presented as mean FPKM ± SEM of three individual samples.

FIG 5. **A**. Identification of a cassette exon in CYP2B10 mRNA. **B**. Identification of a cassette exon in CYP2D26 mRNA. **C**. Identification of an alternative polyadenylation site in CYP3A13 mRNA. RNA-Seq reads of a day 60 sample were aligned to the reference genome mm9 by TopHat and viewed as distribution peaks by the UCSC genome browser. Output junction reads from TopHat were also shown. The cassette exons or alternative polyadenylation were validated by end-point PCR or 3'-RACE. The PCR products were shown by gel electrophoresis and

further sequenced by single pass DNA sequencing with a designed primer. The sequence was aligned back to the *Cyp* gene regions.

DMD Fast Forward. Published on March 20, 2012 as DOI: 10.1124/dmd.112.045088
This article has not been copyedited and formatted. The final version may differ from this version.

Table 1. Primer sequences for validation of novel transcripts identified by RNA-Seq

Gene	Method	Amplification	Sequence (5' to 3')	PCR product
Cyp2b10	End-point PCR	From exon 6 to exon 9 (including cassette exon)	F: ATGGCTTCCTGCTCATGCTCAAGT	411 bp
			R: GACAAATGCGCTTTCCCACAGACT	
	Sequencing	From exon 7 to exon 9	AGTGCCACACAGAGTGACCAAAGA	
Cyp2d26	End-point PCR	From exon 5 to cassette exon	F: ACTACACATCCCTGGTTTGCCTGA	520 bp
			R: AGCCTCTGAGCACCTTCTCTTGTA	
	Sequencing	From exon 6 to cassette exon	TGATTGACCTGTTCATGGCAGGGA	
	End-point PCR	From cassette exon to exon 9	F: TACAAGAGAAGGTGCTCAGAGGCT	327 bp
			R: TAGGGCTCTGGAGTAACTGGCATT	
	Sequencing	From exon 8 to cassette exon	CATGAAGGCCTCGTGCTTCACAAA	
Cyp3a13	3' RACE	cDNA synthesis	GCGAGCACAGAATTAATACGACTC ACTATAGGT12VN	
		Outer PCR	F: AAGTTGCTCTTGTCAGAGTCCTGC	
			R: GCGAGCACAGAATTAATACGACT	
		Inner PCR	F: CACTGTCCAGCCTTGTAAGGAAAC	
			R: CGCGGATCCGAATTAATACGACTC ACTATAGG	
		Sequencing	AAGTTGCTCTTGTCAGAGTCCTGC	

Figure 1.

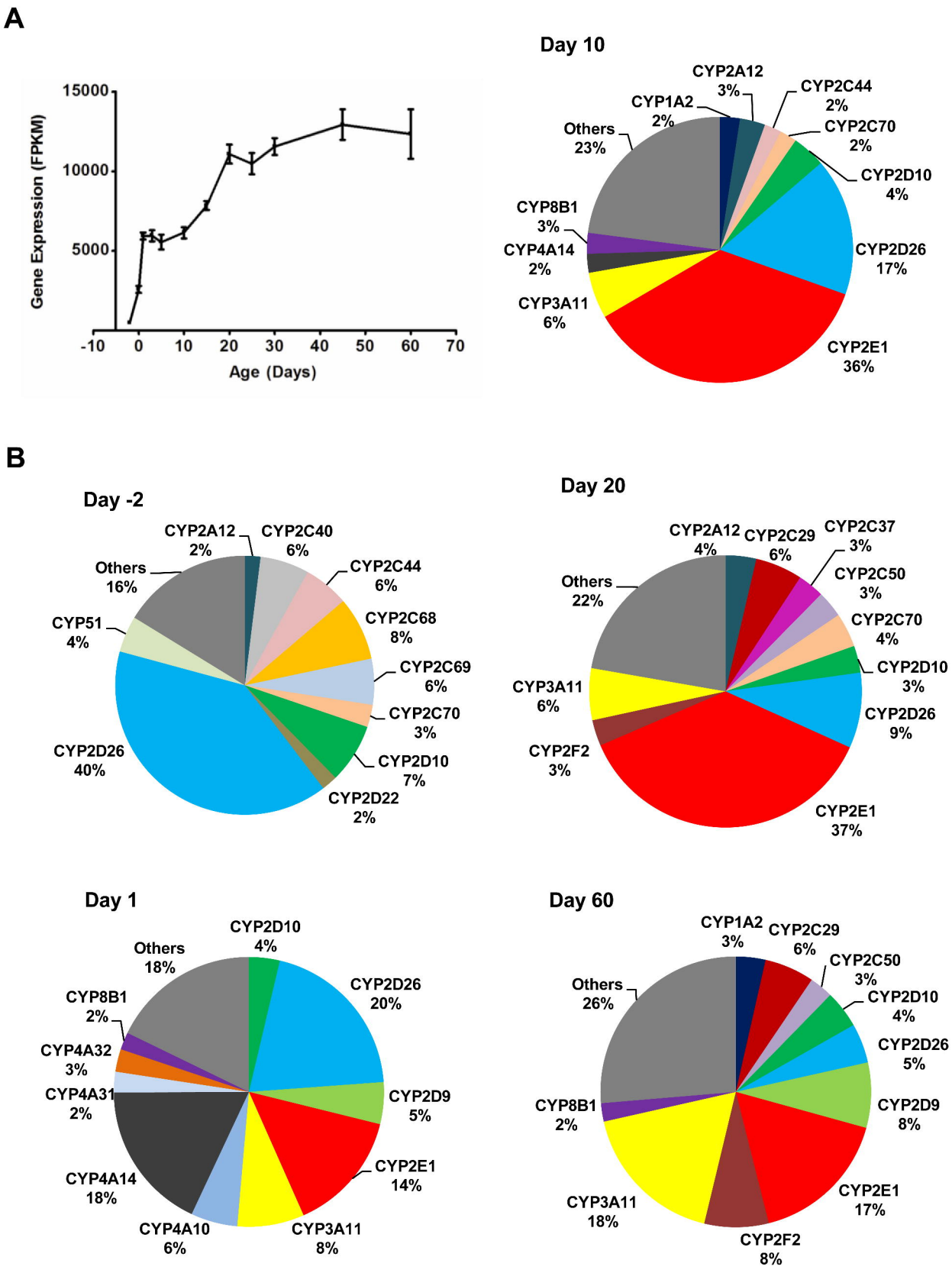

A
Gene Expression (FPKM)
15000
10000
5000
-10
0
10
20
30
40
50
60
70
Age (Days)
Day 10
CYP2A12 3%
CYP2C44 2%
CYP1A2 2%
CYP2C70 2%
Others 23%
CYP2D10 4%
CYP8B1 3%
CYP2D26 17%
CYP4A14 2%
CYP3A11 6%
CYP2E1 36%
B
Day -2
CYP2A12 2%
CYP2C40 6%
Others 16%
CYP2C44 6%
CYP51 4%
CYP2C68 8%
CYP2C69 6%
CYP2C70 3%
CYP2D26 40%
CYP2D10 7%
CYP2D22 2%
Day 20
CYP2A12 4%
CYP2C29 6%
CYP2C37 3%
Others 22%
CYP2C50 3%
CYP2C70 4%
CYP2D10 3%
CYP3A11 6%
CYP2D26 9%
CYP2F2 3%
CYP2E1 37%
Day 1
CYP2D10 4%
Others 18%
CYP2D26 20%
CYP8B1 2%
CYP4A32 3%
CYP4A31 2%
CYP2D9 5%
CYP2E1 14%
CYP4A14 18%
CYP4A10 6%
CYP3A11 8%
Day 60
CYP1A2 3%
CYP2C29 6%
CYP2C50 3%
Others 26%
CYP2D10 4%
CYP2D26 5%
CYP8B1 2%
CYP2D9 8%
CYP2E1 17%
CYP3A11 18%
CYP2F2 8%

Figure 2.

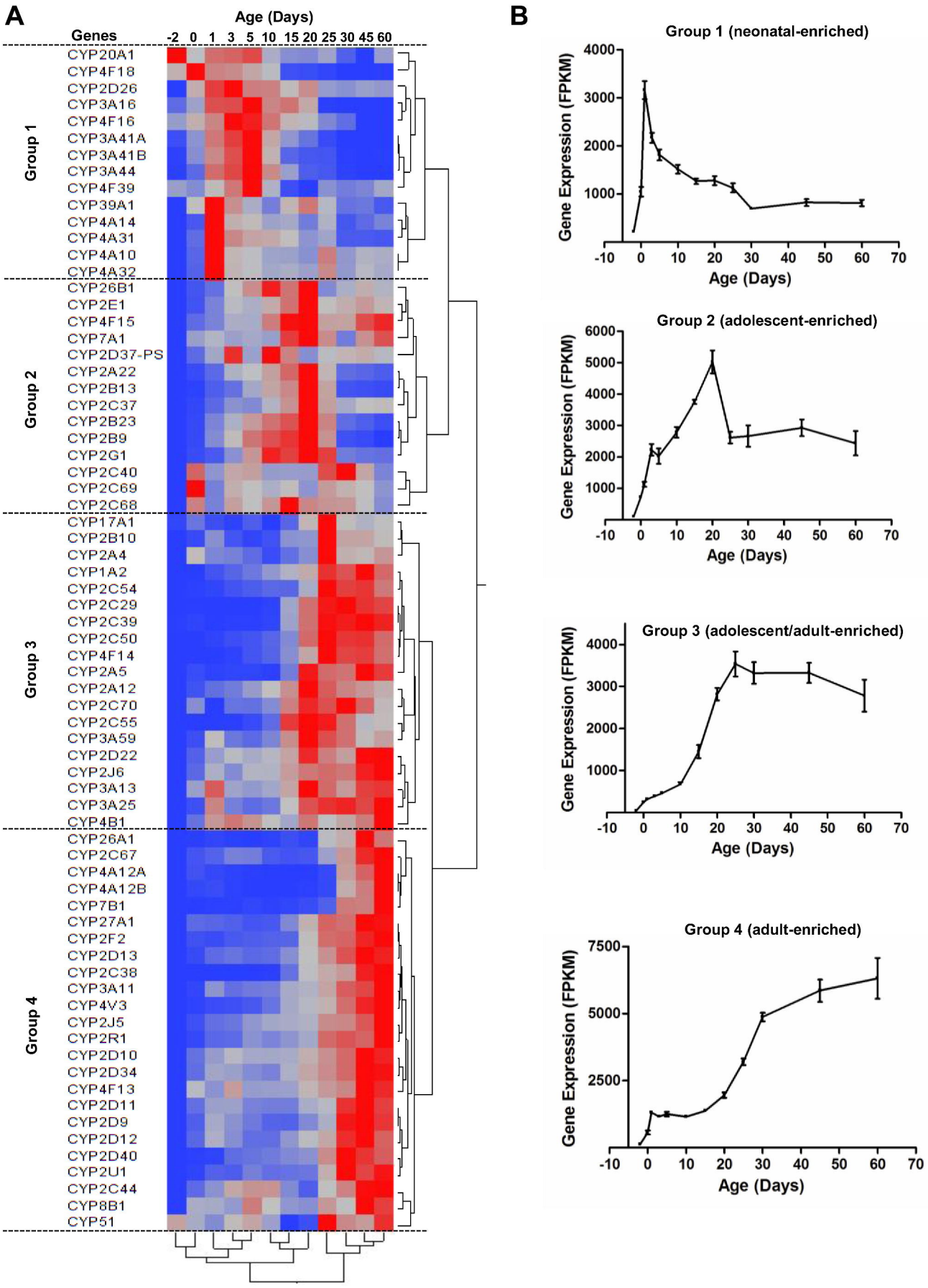

A
Genes
Age (Days)
-2 0 1 3 5 10 15 20 25 30 45 60
Group 1
CYP20A1
CYP4F18
CYP2D26
CYP3A16
CYP4F16
CYP3A41A
CYP3A41B
CYP3A44
CYP4F39
CYP39A1
CYP4A14
CYP4A31
CYP4A10
CYP4A32
Group 2
CYP26B1
CYP2E1
CYP4F15
CYP7A1
CYP2D37-PS
CYP2A22
CYP2B13
CYP2C37
CYP2B23
CYP2B9
CYP2G1
CYP2C40
CYP2C69
CYP2C68
Group 3
CYP17A1
CYP2B10
CYP2A4
CYP1A2
CYP2C54
CYP2C29
CYP2C39
CYP2C50
CYP4F14
CYP2A5
CYP2A12
CYP2C70
CYP2C55
CYP3A59
CYP2D22
CYP2J6
CYP3A13
CYP3A25
CYP4B1
Group 4
CYP26A1
CYP2C67
CYP4A12A
CYP4A12B
CYP7B1
CYP27A1
CYP2F2
CYP2D13
CYP2C38
CYP3A11
CYP4V3
CYP2J5
CYP2R1
CYP2D10
CYP2D34
CYP4F13
CYP2D11
CYP2D9
CYP2D12
CYP2D40
CYP2U1
CYP2C44
CYP8B1
CYP51
B
Group 1 (neonatal-enriched)
Gene Expression (FPKM)
Age (Days)
Group 2 (adolescent-enriched)
Gene Expression (FPKM)
Age (Days)
Group 3 (adolescent/adult-enriched)
Gene Expression (FPKM)
Age (Days)
Group 4 (adult-enriched)
Gene Expression (FPKM)
Age (Days)

Figure 3.

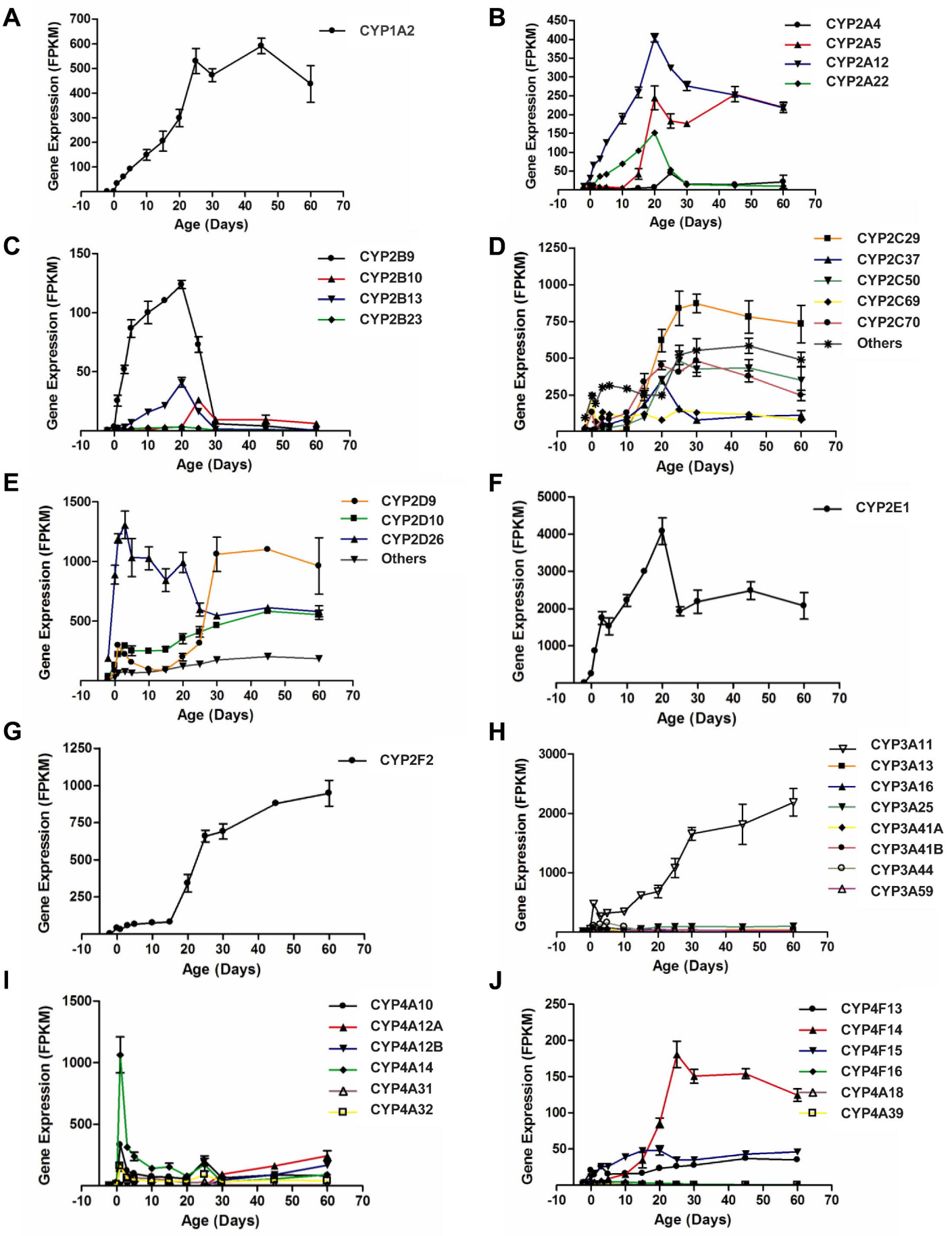

A
Gene Expression (FPKM)
Age (Days)
CYP1A2
B
CYP2A4
CYP2A5
CYP2A12
CYP2A22
C
CYP2B9
CYP2B10
CYP2B13
CYP2B23
D
CYP2C29
CYP2C37
CYP2C50
CYP2C69
CYP2C70
Others
E
CYP2D9
CYP2D10
CYP2D26
Others
F
CYP2E1
G
CYP2F2
H
CYP3A11
CYP3A13
CYP3A16
CYP3A25
CYP3A41A
CYP3A41B
CYP3A44
CYP3A59
I
CYP4A10
CYP4A12A
CYP4A12B
CYP4A14
CYP4A31
CYP4A32
J
CYP4F13
CYP4F14
CYP4F15
CYP4F16
CYP4A18
CYP4A39

Figure 4

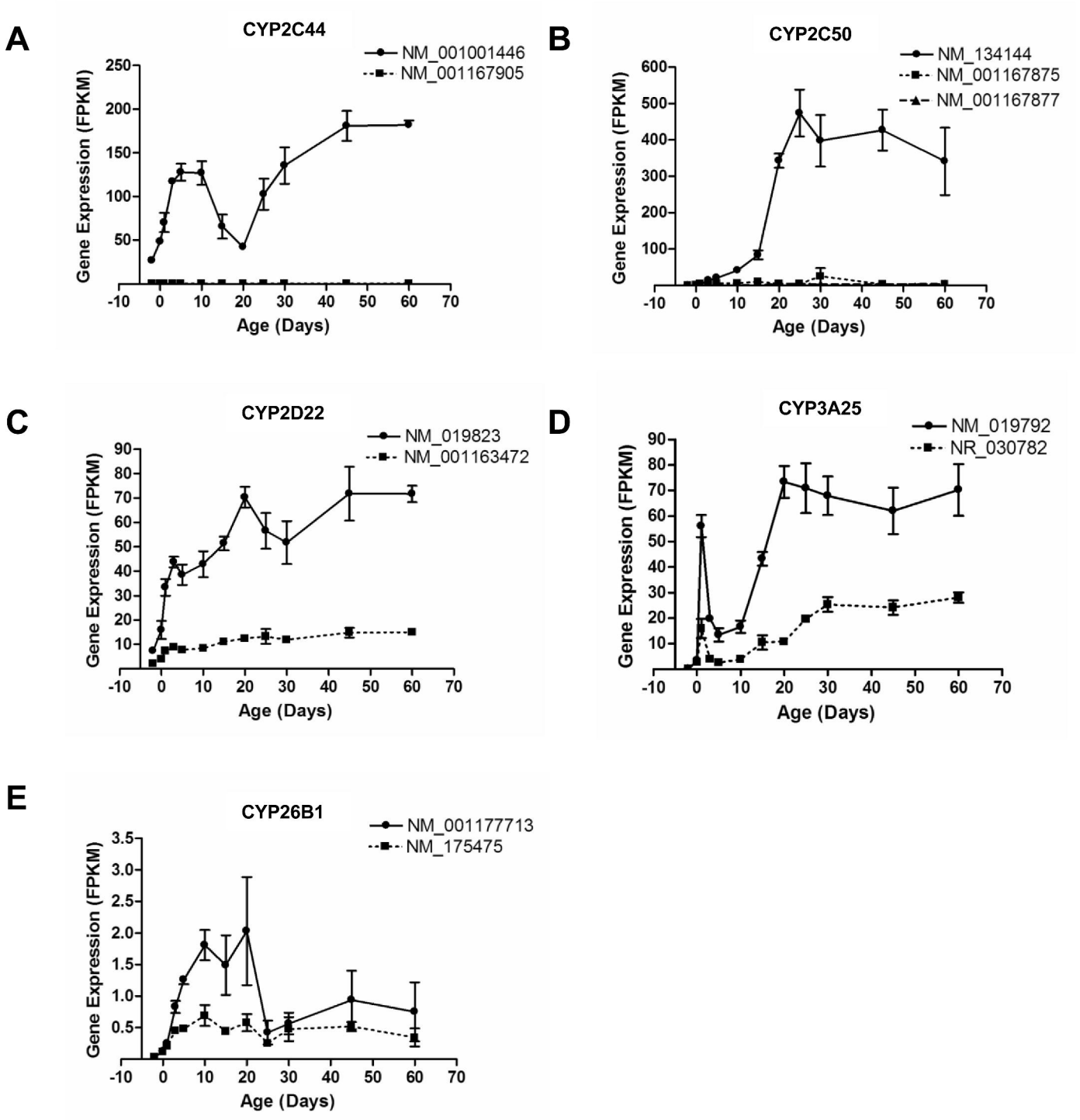
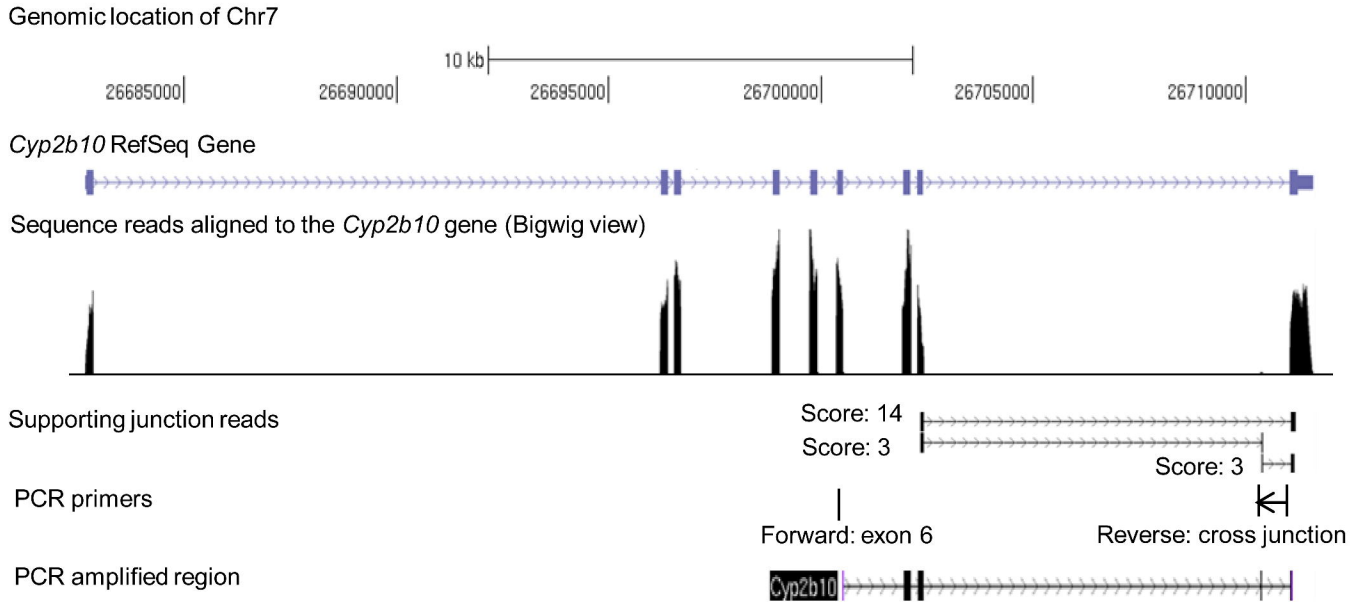
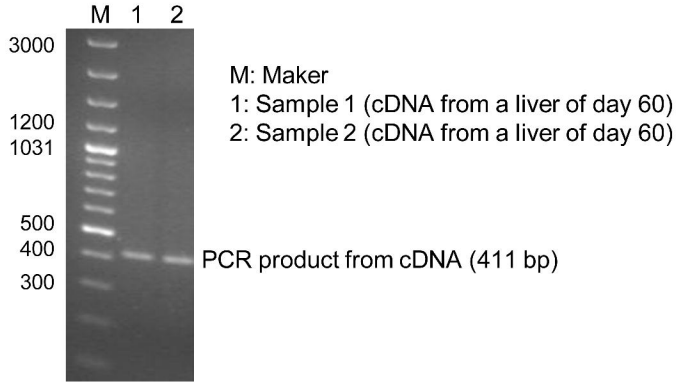

A
CYP2C44
NM_001001446
NM_001167905
Gene Expression (FPKM)
Age (Days)
B
CYP2C50
NM_134144
NM_001167875
NM_001167877
Gene Expression (FPKM)
Age (Days)
C
CYP2D22
NM_019823
NM_001163472
Gene Expression (FPKM)
Age (Days)
D
CYP3A25
NM_019792
NR_030782
Gene Expression (FPKM)
Age (Days)
E
CYP26B1
NM_001177713
NM_175475
Gene Expression (FPKM)
Age (Days)

Figure 5

A





Sequence and position of the cassette exon:

TTTTCTGCCCTTCTCAACAGACAGGACAAATTTTTGATCAAAAGTCTGTGGGAAAGCGCATTTGTCTTGGT

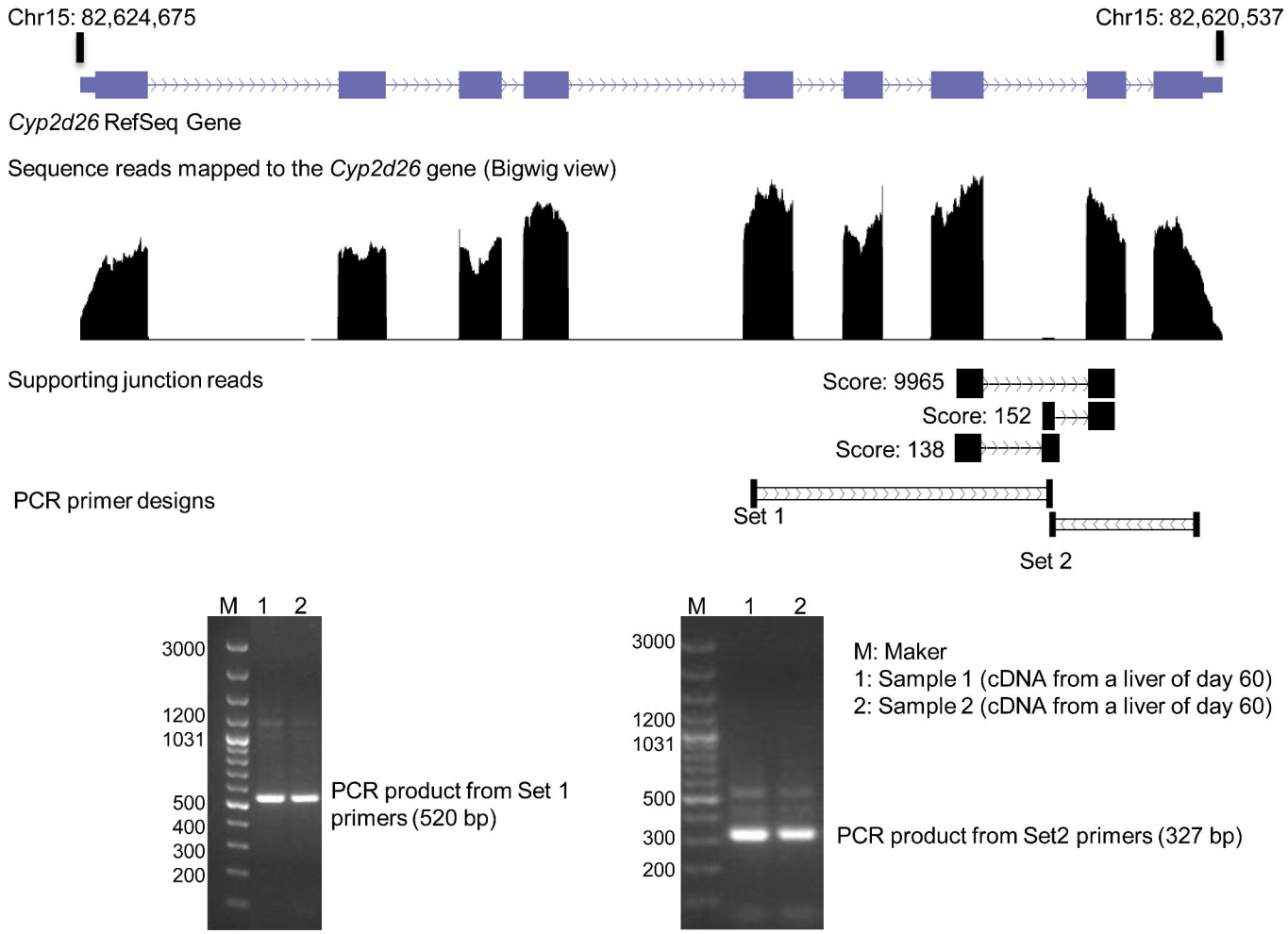
end of exon 8 | cassette exon | start of exon 9

Chr7: 26,710,388

Chr7: 26,710,418

Figure 5

B



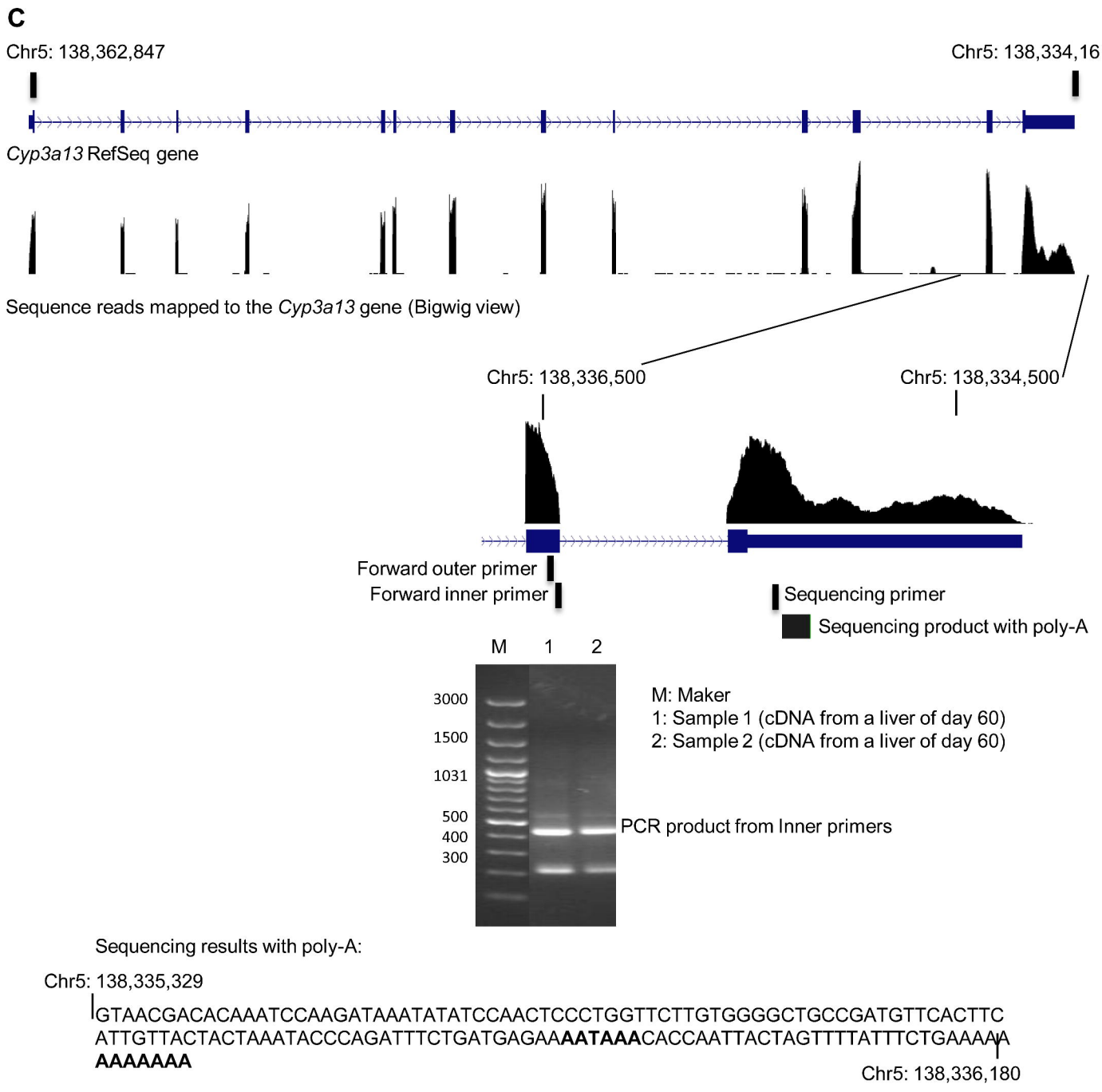
Sequence and position of the cassette exon:

AAGACTTCTTCATCCCCAAGGGCCTTGGATGGTACAAGAGAAGGTGCTCAGAGGCTCCACGGGGGACGACACTCATTCCC

end of exon 7 | cassette exon | start of exon 8

Chr15: 82,621,185 Chr15: 82,621,144

Figure 5



Sequencing results with poly-A:

Chr5: 138,335,329

GTAACGACACAAATCCAAGATAAATATATCCAACTCCCTGGTTCTTGTGGGGCTGCCGATGTTCACTTC
ATTGTTACTACTAAATACCCAGATTTCTGATGAGAA**AATAAA**CACCAATTACTAGTTTTATTTCTGAAAAA
AAAAAAA

Chr5: 138,336,180