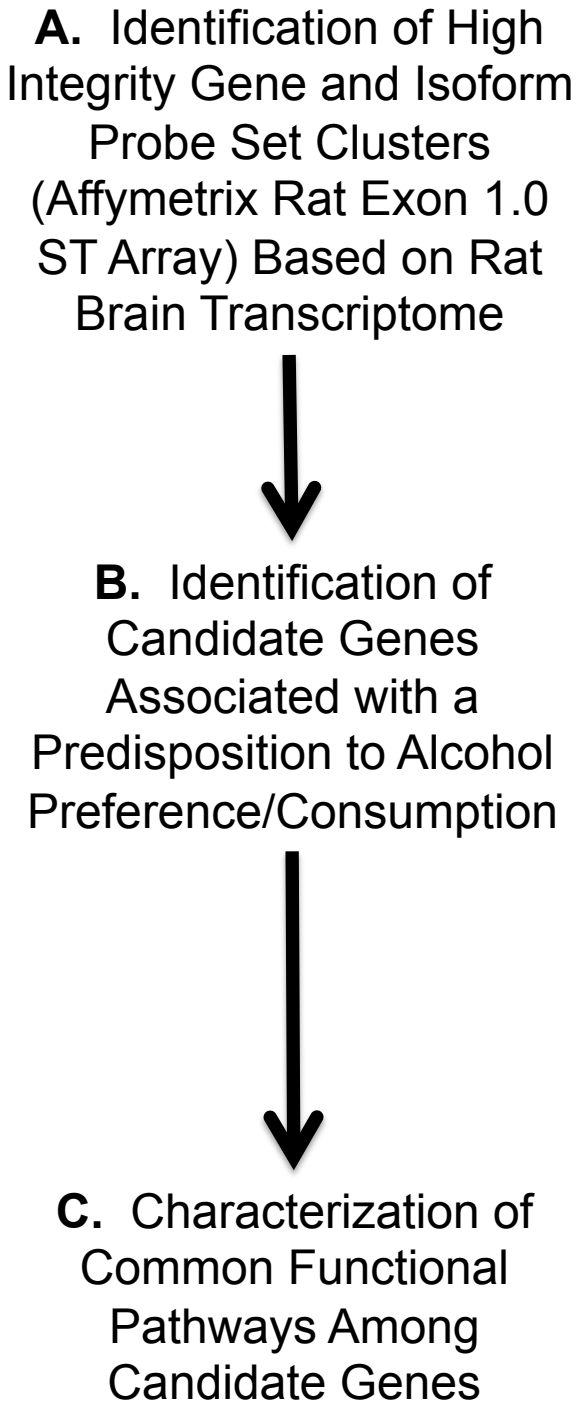
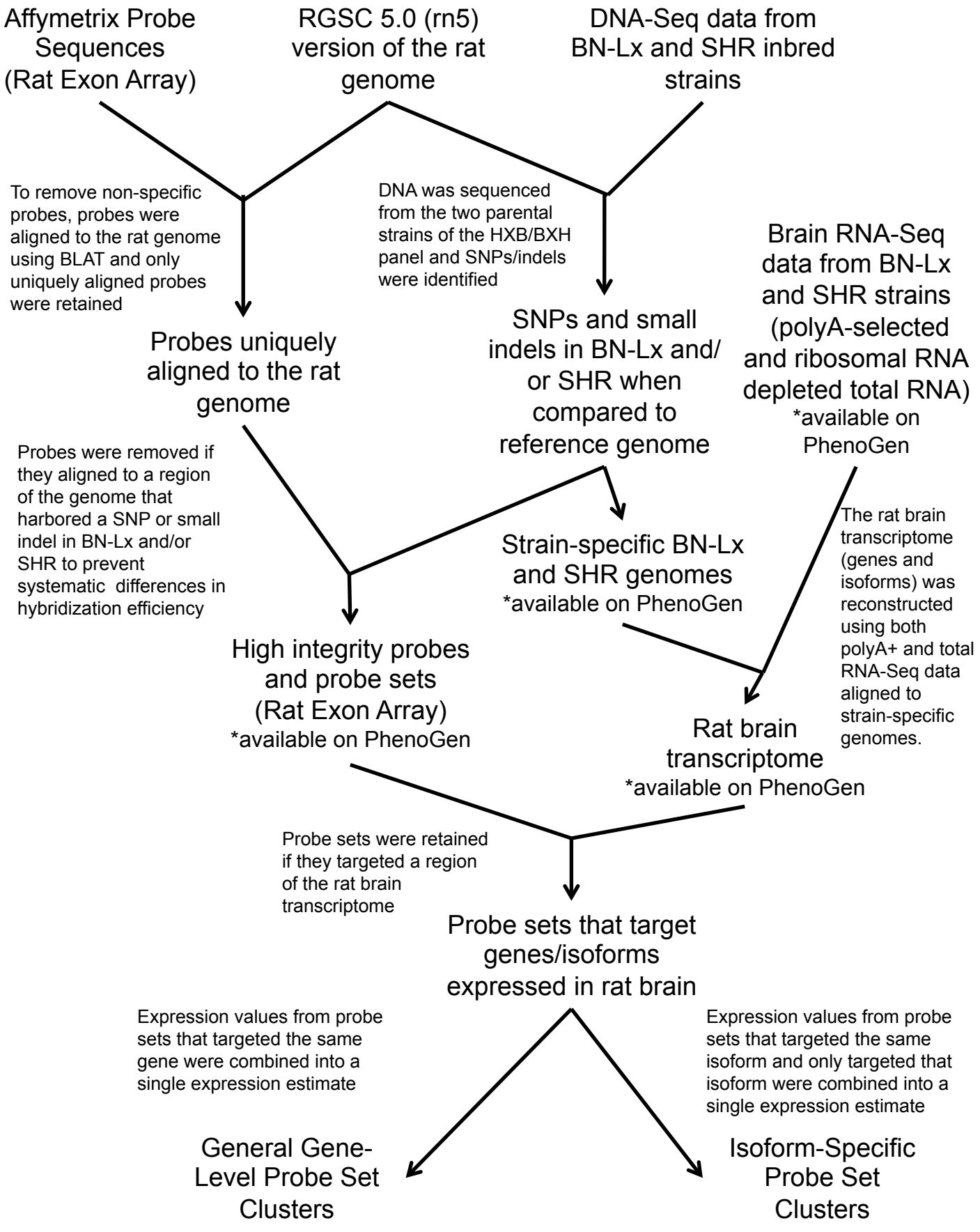
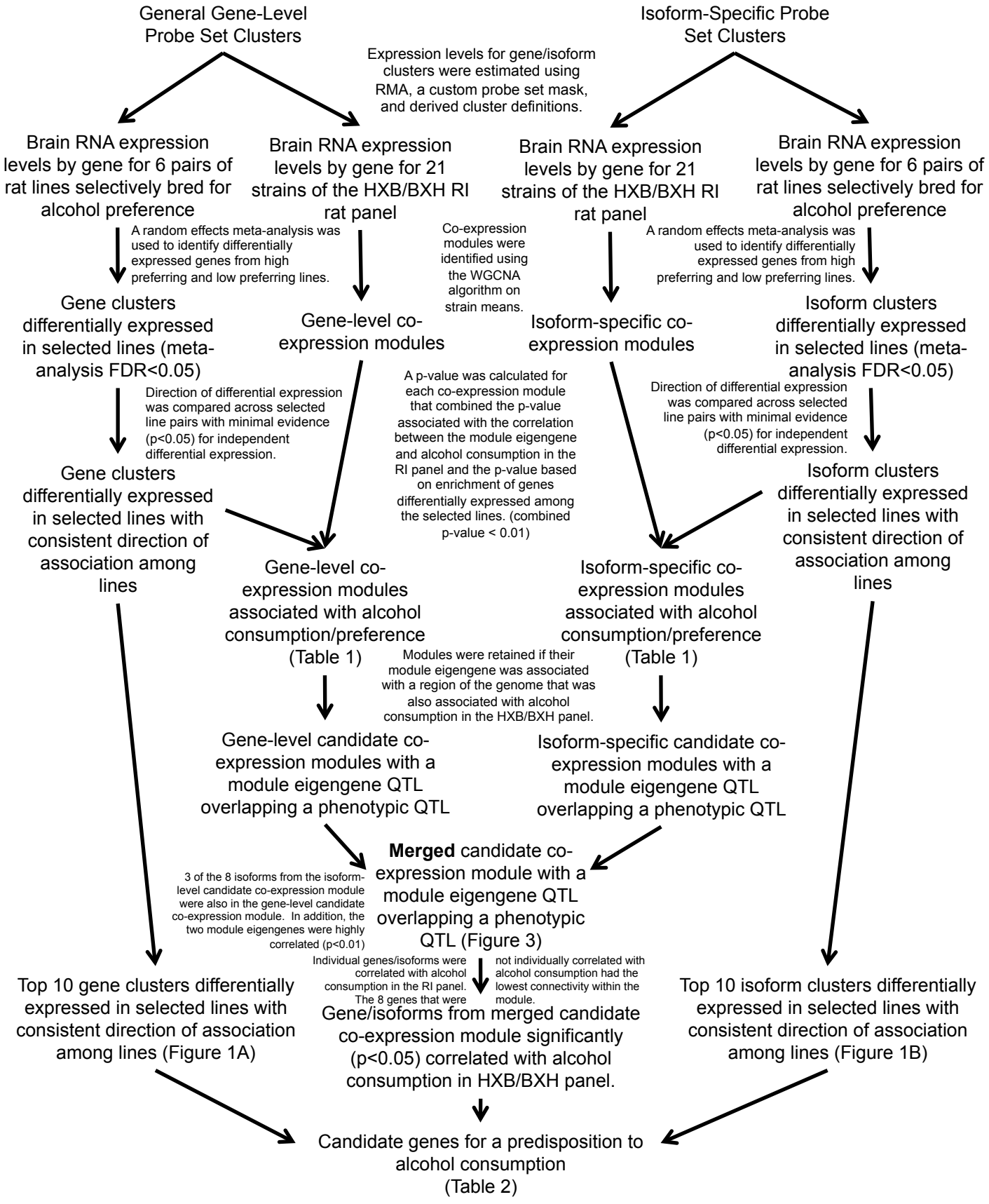
Figure S1.


A. Identification of High Integrity Gene and Isoform Probe Set Clusters (Affymetrix Rat Exon 1.0 ST Array) Based on Rat Brain Transcriptome
B. Identification of Candidate Genes Associated with a Predisposition to Alcohol Preference/Consumption
C. Characterization of Common Functional Pathways Among Candidate Genes

A. Identification of High Integrity Gene and Isoform Probe Set Clusters (Affymetrix Rat Exon 1.0 ST Array) Based on Rat Brain Transcriptome



B. Identification of Candidate Genes Associated with a Predisposition to Alcohol Preference/Consumption



C. Characterization of Common Functional Pathways Among Candidate Genes

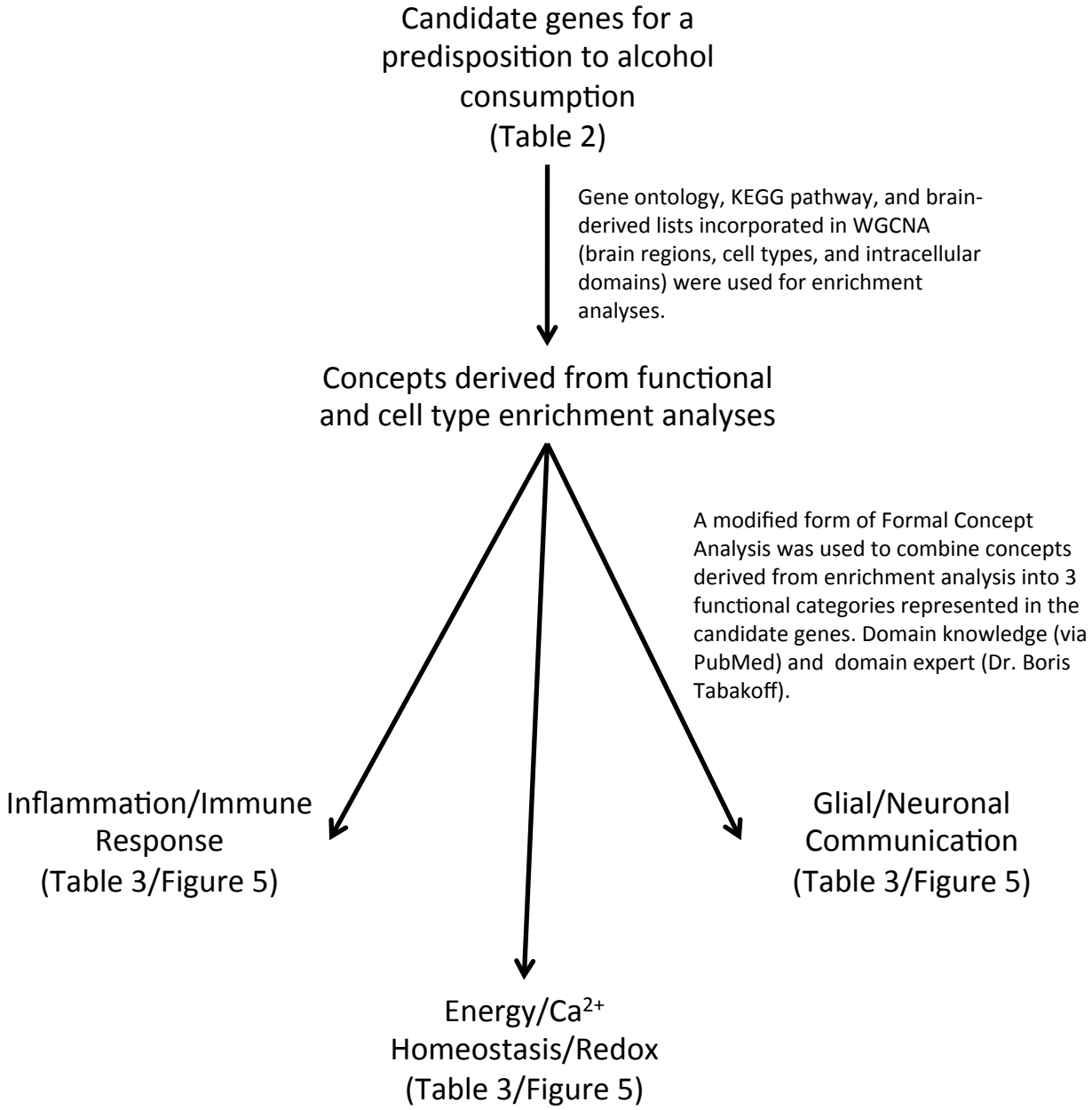


Figure S1. Complete workflow from generating probe mask and defining transcript clusters to identification of functional categories represented in candidate genes for predisposition to alcohol consumption/ preference. The work flow for the identification of biological/genetic pathways was split into 3 consecutive steps: A) identification of high integrity gene and isoform probe set clusters (Affymetrix Rat Exon 1.0 ST Array) based on rat brain transcriptome, B) identification of candidate genes associated with a predisposition to alcohol preference/consumption, and C) characterization of common functional pathways among candidate genes. Intermediate outcomes are connected by arrows and additional information about methods for obtaining these outcomes are described in smaller print next to the appropriate arrows. Intermediate outcomes that are available at http://phenogen.ucdenver.edu are labeled. Intermediate/final outcomes that are described in a figure or table within the manuscript are also labeled with the corresponding table and/or figure number in parentheses.

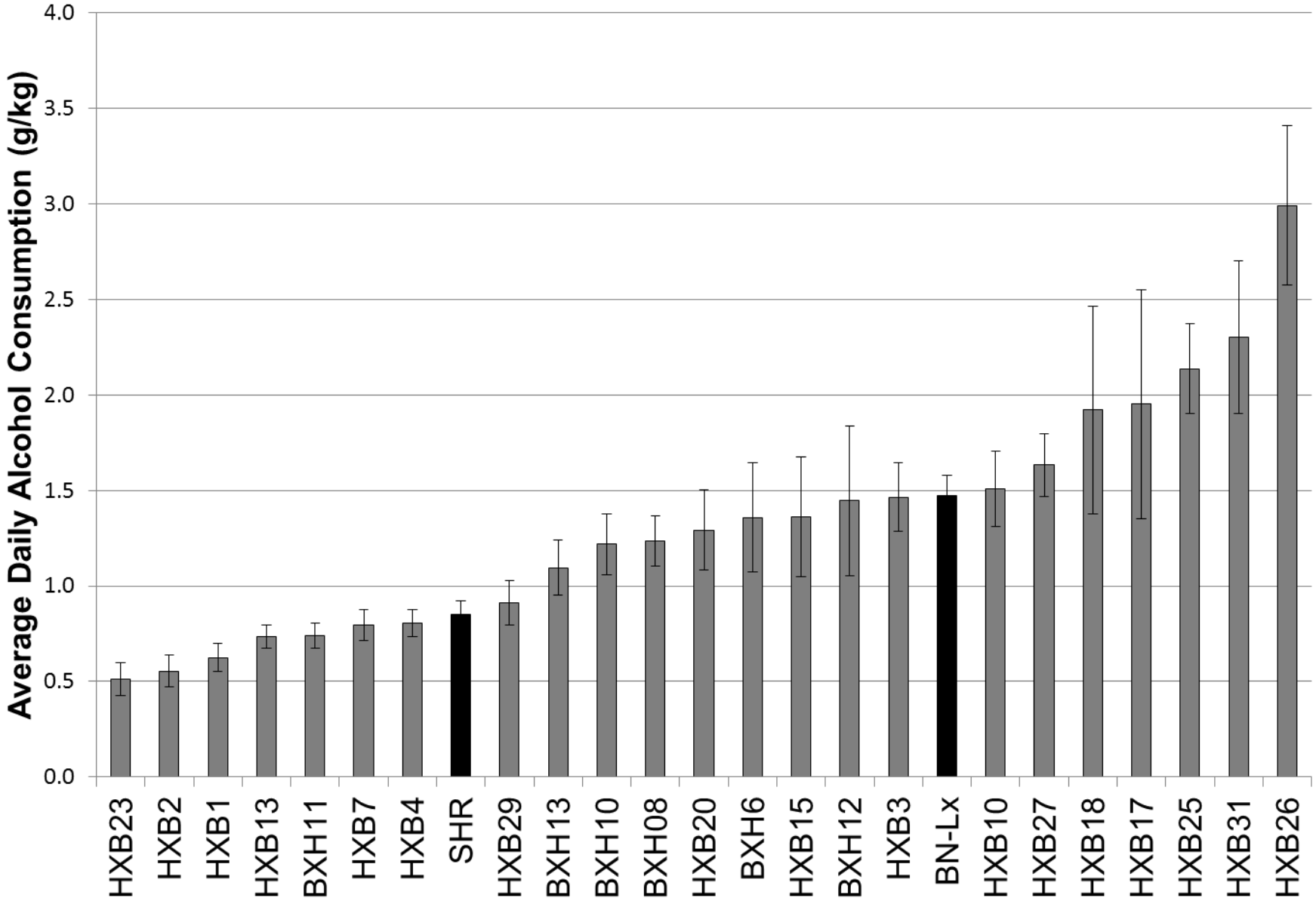


Figure S2. Distribution of average daily alcohol consumption using the 2-bottle 24-hour access paradigm in the HXB/BXH recombinant inbred rat panel. The data shown are mean ± SEM of average daily ethanol consumption (g/kg body weight) during the second week of exposure to the two-bottle choice paradigm (10% ethanol, 24-hr access). Prior exposure to the two-bottle choice paradigm, each rat was given 10% ethanol as their only choice of fluid for one week.

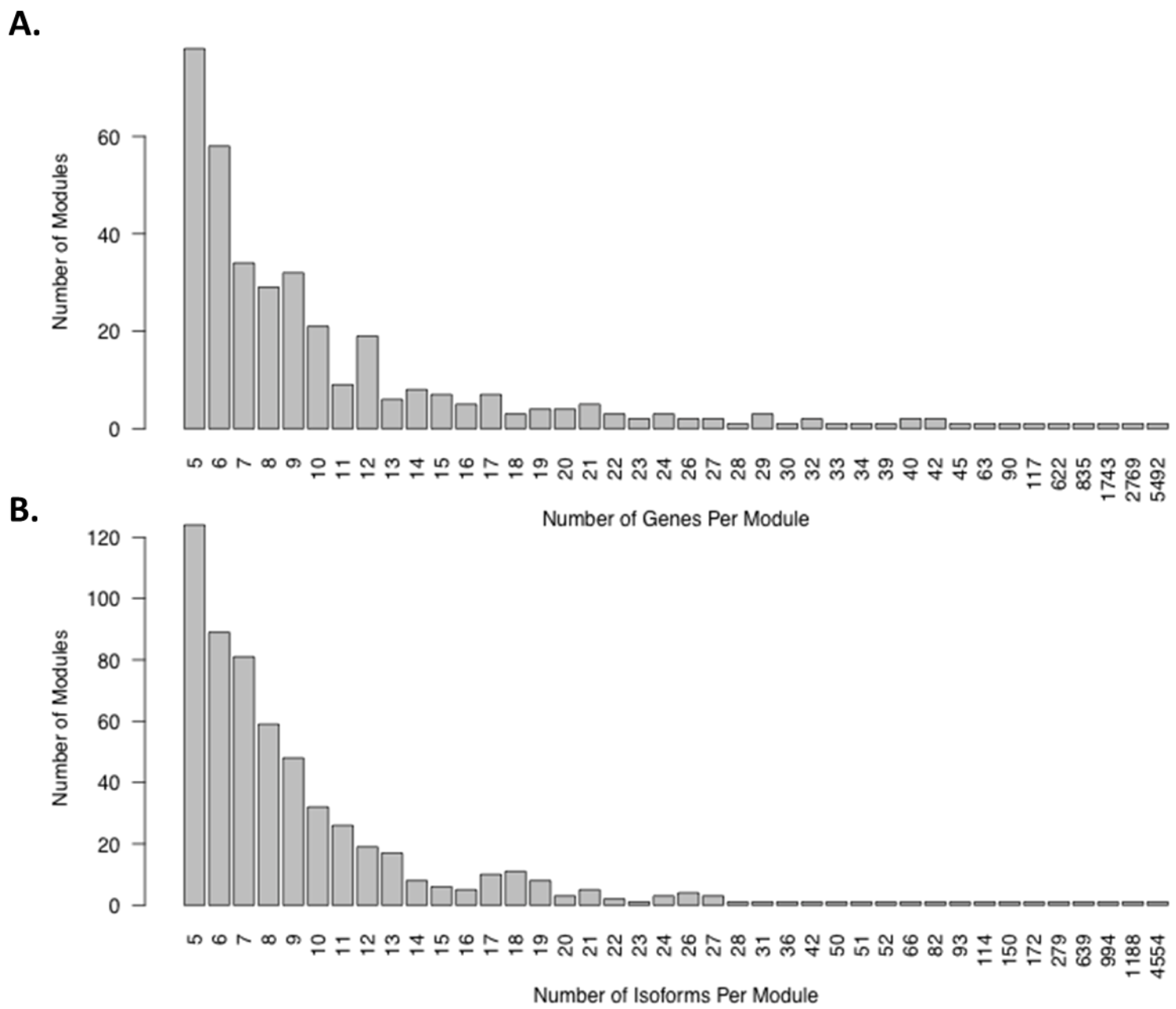


Figure S3. Distribution of module size in WGCNA. A weighted gene co-expression network analysis was executed separately for the data set that was summarized into gene clusters (A) and the data set that was summarized into isoform clusters (B). Gene/Isoform clusters with expression values above background in at least 5% of samples in the HXB/BXH RI panel were placed into modules. The default settings of the WGCNA package in R were changed to allow for the identification of smaller modules.

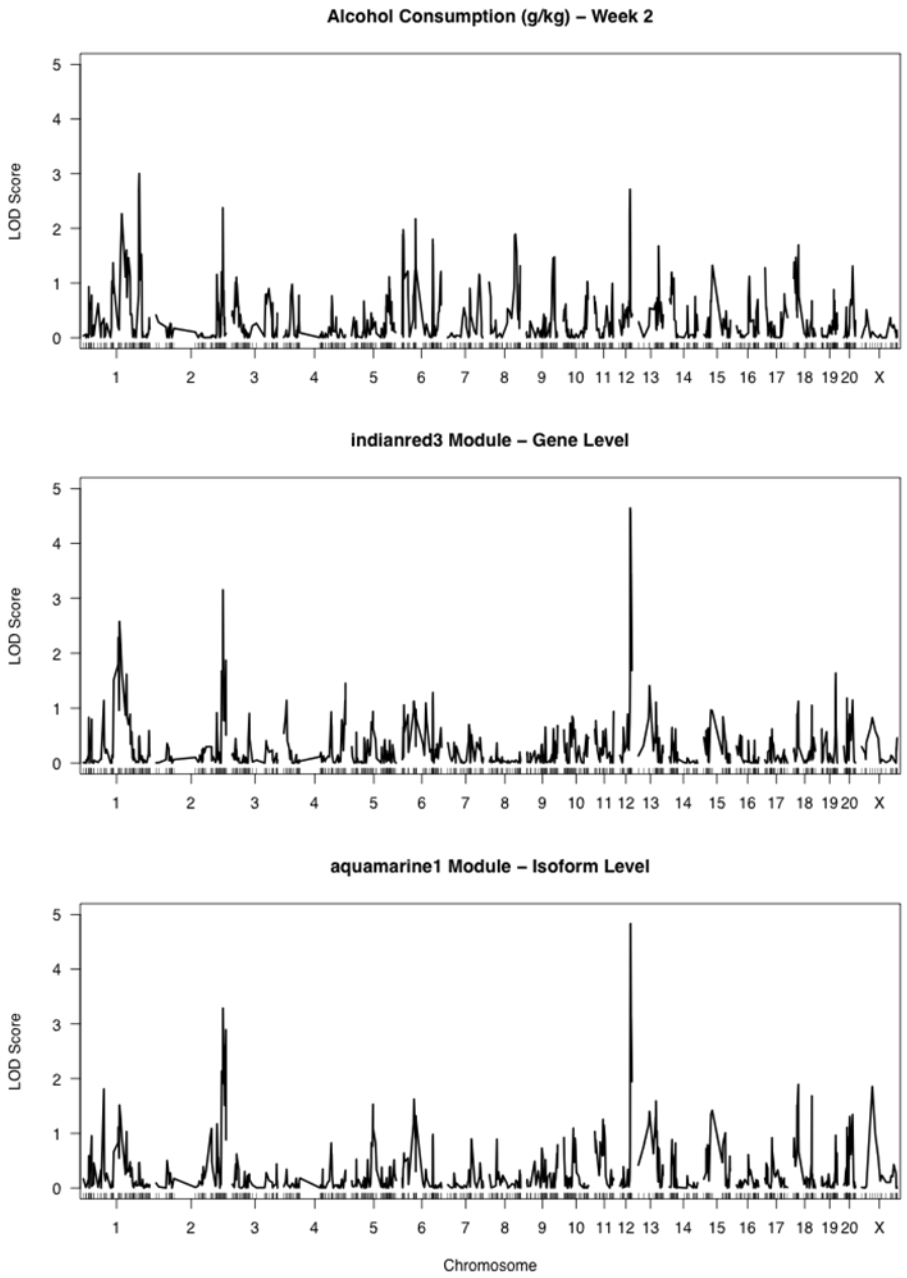


Figure S4. Comparison of LOD profiles between alcohol consumption and module eigengenes. To identify co-expression modules that were associated with a predisposition for alcohol consumption in the HXB/BXH RI panel, the module eigengene had to be correlated with alcohol consumption and the maximum QTL for the module eigengene had to overlap a bQTL for alcohol consumption. The maximum eigengene QTL for both indianred3 and aquamarine1 overlapped the bQTL on chromosome 12 for alcohol consumption. However, the LOD profiles were similar between the 3 quantitative traits at several other positions in the genome, e.g., chromosomes 1 and 2.

Doc. S1. Candidate Module Hub Gene

GENE_07346 was the hub gene (most connected gene) within the co-expression module associated with voluntary alcohol consumption. This gene is unannotated in the rn5 version of the rat genome. We examined this gene and the adjacent/overlapping gene GENE_07345, which was the second most highly connected gene in the co-expression module and was highly correlated with GENE_07346, using both RT-PCR and a variety of bioinformatics tools. Based on information gathered from these additional studies, we propose that GENE_07346 and GENE_07345 be combined into one transcript that we have labeled 'Locus R.Alc.Pref' through the remainder of this material.

qRT-PCR of Locus R.Alc.Pref

METHODS

Total RNA from brains of three SHR/Ola and three BN-Lx/Cub rats was isolated as described for RNA-Seq. Design for primers that assay the various exons of GENE_07346 were limited to specific regions of sequence unique to each proposed exon sequence. Primers were selected using Primer3 (http://frodo.wi.mit.edu/) and were evaluated for specificity and degeneracy by searching BLASTN and known SNPs (http://blast.ncbi.nlm.nih.gov/). Primers were designed to generate three PCR products that represent each exon-exon junction and both exon-exon junctions simultaneously (Table 1). Reference gene primers were as described in Langnaese et al. [1]. The cDNA first strand transcription was performed in duplicate using 1.5 ug of total RNA with the iScript RT mix for qRT-PCR following the manufacturer's protocol (BioRad, Hercules CA, USA). The synthesized cDNA was stored at -20°C until further use. Quantitative PCR was performed on a Roche LightCycler 480II Real Time PCR instrument, using Roche SYBR Green I Master (Roche, Mannheim Germany). PCR was carried out in a 15 ul volume and a final concentration of 1X reaction buffer, 500nM forward and reverse primers and 0.5 ul cDNA per reaction. All standard curve and validation reactions were performed in triplicate with no template control (NTC) reactions for all primer sets. PCR cycling parameters were as follows: hot-start at 95ºC for 5 min, 45 cycles of 95ºC for 10 sec, 60ºC for 20 sec, 72ºC for 45 sec followed by a dissociation curve measurement from 65ºC to 97ºC. Melt curve analysis for all assays was used to verify single product amplification and absence of primer dimers. NTC reactions for all primer sets were >5Cq for control and unknown samples. Advanced Relative Quantification analysis with efficiency correction (standard curves) was performed using the LC480II data collection software (release 1.5.0.39 SP4). Five log10 dilutions of plasmid template for all PCR products and plasmid template reference gene amplicon sequences were prepared and used for primer validation and standard curve reference. PCR efficiency for all target gene primer sets ranged from 90-91% with all r-squared values >0.99. Relative quantification estimates (target to reference ratios) for each sample and PCR product were normalized using three reference genes (Actb, CypA and Pgk1). Mean target to reference ratios and their associated errors for each sample and PCR product were exported to SAS statistical software for further statistical analysis. A weighted repeated

measures linear model was used to identify quantitative differences between strains and between PCR products. The inverse of the target to reference error was used to weight observations in the regression analysis.

Primer Label	Primer Sequence
Exon1.Ex1F1	CCTTCAACCACCTTTCCTCATTACC
Exon1.Ex1F2	CCTTTCGCTCATTCATTCGGAC
Exon1.Ex1R1	GTCCGAATGAATGAGCGAAAGG
Exon2.Ex2F1	TTTGCAGTCCTTGGTGCCTTG
Exon2.Ex2R1	TCCTTTCCAGCCCTTACAC
Exon3.Ex3F1	TCAGCATTTGCCTTCCAGTCTC
Exon3.Ex3R1	GAGACTGGAAGGCAAATGCTG
Exon3.Ex3R2	AGGGAAACTGAGGAAAGGGCAG

Table 1. Primer sequences for qRT-PCR of Locus R.Alc.Pref.

RESULTS

As stated in the manuscript, every PCR product was over-expressed in the SHR/Ola strain as compared to the BN-Lx/Cub strain, confirming the differences detected on the microarray (Figure 1 and Table 2). There were also significant differences in expression between the PCR products within each strain, indicating that other isoforms may be present in the samples (Table 2).

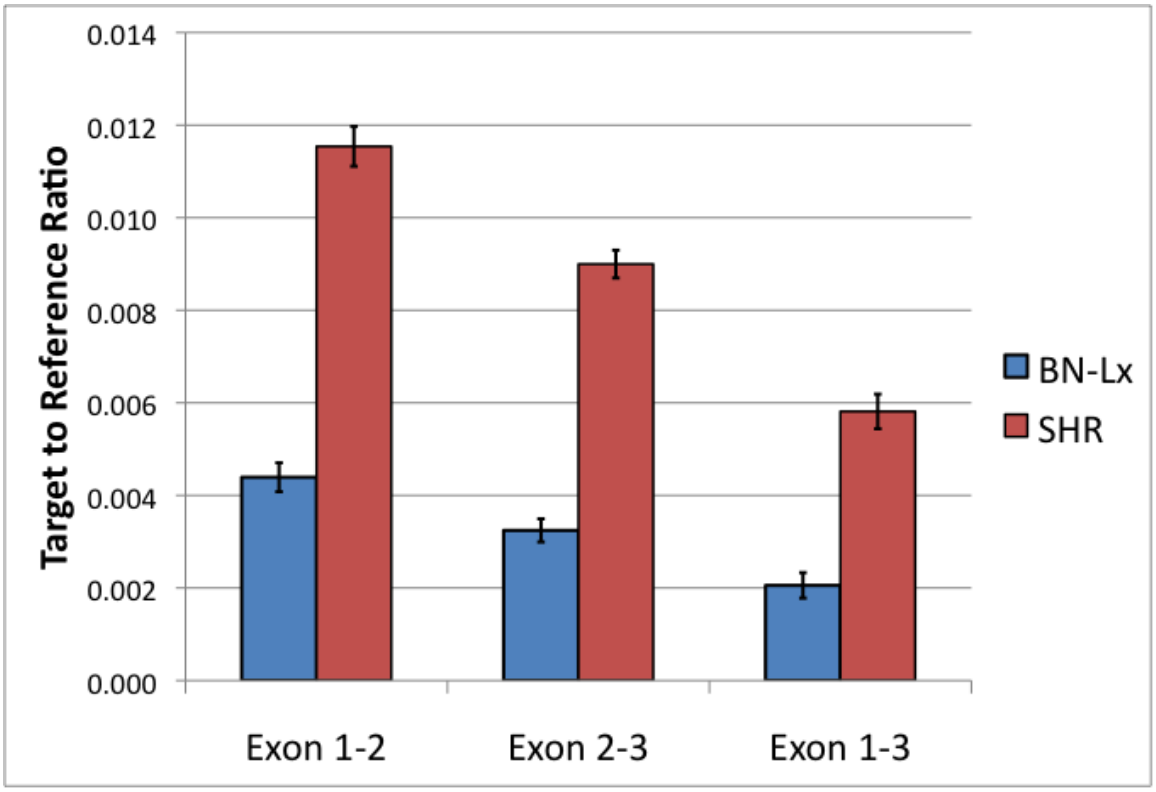


Figure 1. Relative quantification of PCR products from the Locus R.Alc.Pref expressed in rat brain. Relative transcript expression levels were normalized using three reference genes (Actb, CypA and Pgk1) and expressed as target-to-reference ratios using the LC480 Advanced Relative Quantification Software. Three biological replicates per strain were included in strain/PCR product means. Bars

represent the standard error estimates from a repeated measures linear model of target-to-reference ratios using the inverse of the target-to-reference ratio error for individual sample/PCR product estimates as weights.

	Target to Reference Ratios					
Comparison	Difference	Standard Error of Difference	Ratio	DF	t-statistic	p-value
Differences Between Strains						
Exon 1 to 2 in SHR vs. Exon 1 to 2 in BN-Lx	0.0072	0.00053	2.63	8	-13.47	<.0001
Exon 1 to 3 in SHR vs. Exon 1 to 3 in BN-Lx	0.0038	0.00046	2.83	8	-8.10	<.0001
Exon 2 to 3 in SHR vs. Exon 2 to 3 in BN-Lx	0.0058	0.00039	2.77	8	-14.73	<.0001
Differences Between PCR Products						
Exon 2 to 3 in BN-Lx vs. Exon 1 to 3 in BN-Lx	0.0012	0.00014	1.58	8	-8.38	<.0001
Exon 1 to 2 in BN-Lx vs. Exon 1 to 3 in BN-Lx	0.0023	0.00023	2.14	8	-9.97	<.0001
Exon 1 to 2 in BN-Lx vs. Exon 2 to 3 in BN-Lx	0.0012	0.00020	1.35	8	-5.62	0.0005
Exon 2 to 3 in SHR vs. Exon 1 to 3 in SHR	0.0032	0.00033	1.55	8	-9.54	<.0001
Exon 1 to 2 in SHR vs. Exon 1 to 3 in SHR	0.0057	0.00046	1.98	8	-12.60	<.0001
Exon 1 to 2 in SHR vs. Exon 2 to 3 in SHR	0.0026	0.00040	1.28	8	-6.42	0.0002

Table 2. Comparison of RNA expression estimates between strains and PCR products of Locus R.Alc.Pref expressed in rat brain. A repeated measures linear model of target-to-reference ratios was used to estimate differences in brain RNA expression levels between BN-Lx/Cub and SHR/Ola strains. Individual estimates were weighted in the model using the inverse of the target-to-reference ratio error for individual sample/PCR product estimates. Three biological replicates per strain were included in the model and three PCR products were evaluated for each biological sample.

In addition to evaluating the quantitative RNA expression levels, the PCR product that was obtained from the Exon1.Ex1F2/Exon3.Ex3R1 primer pair was sequenced.

```
>CloneSequence
CCTTCAACCACCTTTCCTCATTACCTAGGCTGGCGACCTAAATTCAAAGTATTTTGCATGTATG
ATTTCTTTTTCTGTTTTGCTTTCCTTACTGCATGTCTAGGGGGGAGGGGCCAGCCTGTCCAGAC
CCCTATCCATTTGCGAACCTGCCATATTTCCATATGGCATCCGAGGGGCTCAGAAAGAAATTAG
CGGAGTTCTGTATCTGTACAGACTGCGTAACTTTTAAACTCCCGGGCTGAAATAATATGGAGAC
ACAGGCACAGGTACAGGAAATACAACCAGTCCACAGCCTTCTCAGACTGCAACTGACCGGAAGA
AAGCCTTTCGCTCATTCATTCGGACAGCCTCGGTAGTACAGCCCTGAAACAGAAGTTGGGCTTT
TAAAAAGAAAGAAAGAAAGAAAAGTCTCTACACTCGCTGAGAAGAGATAGTCGATACACAGGAG
GGACTAGAGAATCGGATTCACCAGGCAGGCAAAAGTCCTTGGTGCCTTGTGGATTGAAGTCAAA
AGAGAGAAAACAAAACTGTACAAAGGCTGTCAGGCCTGACCCTTGGGGAAAGGCTTAAAGTTCT
ACTTAGCGTTCTGATTTTGCTTTTCAATTGTGTCTCTTCTTGTGTAAGGGCTGGAAAGGACCCT
GGAGACCACAGAACACAGACCCATGGTGGAGACATTTCCATTAATCTTGTAAAGTGATCAGTCA
CTGGGAGGATTGAATTGCATTTCCTGCCCAAAGATGTGGAGACAGATGACCTTTGCAGCAGCCC
TCAGACCCCCGGTTCACAGCTGGAGAGAGGATGTGTCAGCATTTGCCTTCCAGTCTC
```

This sequence was similar to what was expected based on the RNA-Seq data and confirmed that the three-exon transcript is expressed in rat brain.

Bioinformatic Analysis of Locus R.Alc.Pref

The proposed Locus R.Alc.Pref in rat, located on chr12:40,905,859-40,917,690 on the antisense strand in the rn5 rat genome assembly, was derived by combining information from the cloned PCR product from the qRT-PCR experiment, information on

RNA-Seq read depth patterns in the region, and information from the bioinformatic analysis of DNA/RNA sequence in the region. Locus R.Alc.Pref is homologous in some regions to transcript A930024E05 Riken in mouse.

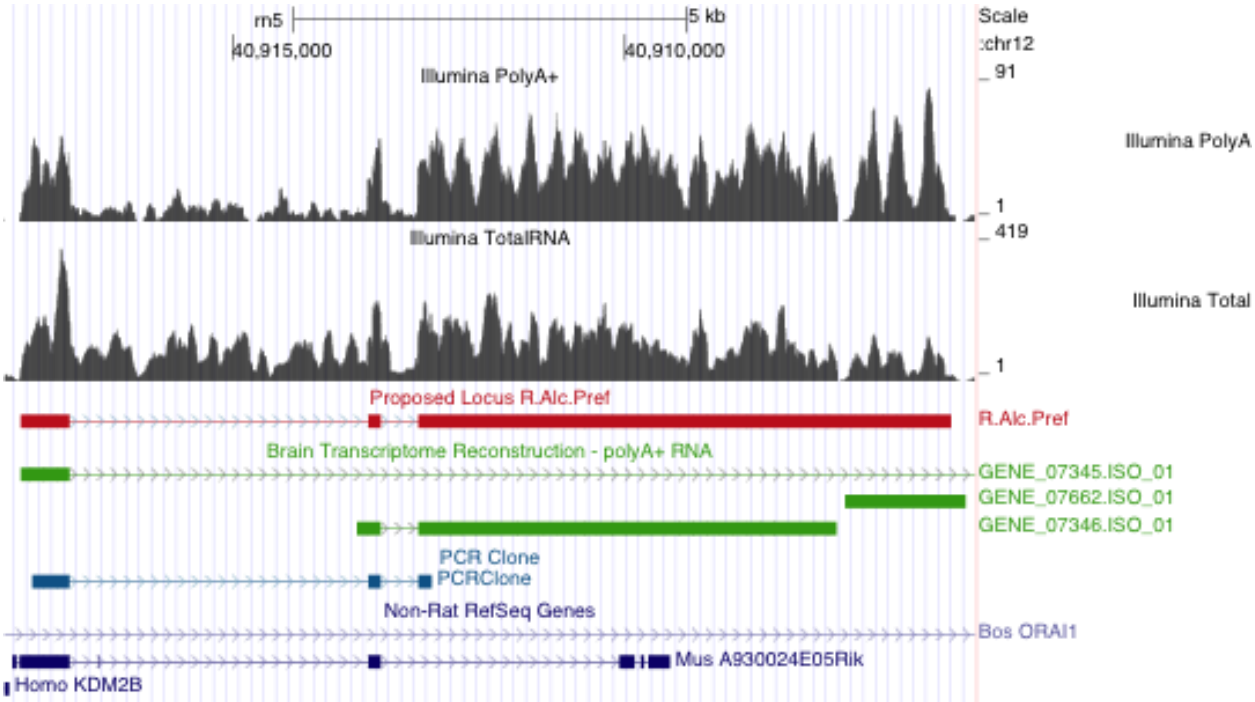


Figure 2: Proposed Locus R.Alc.Pref. The Illumina PolyA+ track displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Proposed Locus R.Alc.Pref (red) track is the proposed transcript sequence structure derived from RNA-Seq data, qRT-PCR information, and bioinformatics analyses. The Brain Transcriptome Reconstruction – polyA+ RNA (green) is the original transcript sequence structure that was derived from a transcriptome reconstruction using polyA+-selected RNA-Seq data. The PCR Clone (blue) track has the alignment of the clone sequenced from the PCR product selected using the primers in Table 1. The Non-Rat RefSeq Genes track was taken directly from the UCSC Genome Browser (http://genome.ucsc.edu). Because the transcript is on the antisense strand the graphic has been reversed so that 5' to 3' of the transcript is left to right and the base pair position along the chromosome decreases as you move from left to right.

We used bioinformatics tools to examine the sequence of the Locus R.Alc.Pref with respect to: 1) potential transcription start sites, 2) exon/intron splice site motifs, 3) start/stop codons and open reading frames, 4) SNPs and indels within transcribed regions, and 5) potential polyadenylation sites.

POTENTIAL TRANSCRIPTION START SITES

We examined the region around the 5' end of GENE_07345 to identify any putative transcription start sites. This region does not contain a TATA box or CCAAT box motif, but it does contain several GC-box motifs [2]. It is also near a CpG island (Figure 3) and CpG islands have been associated with gene promoters [3]. These

bioinformatic traits were used along with the RNA-Seq data to propose a transcription start site for Locus R.Alc.Pref.

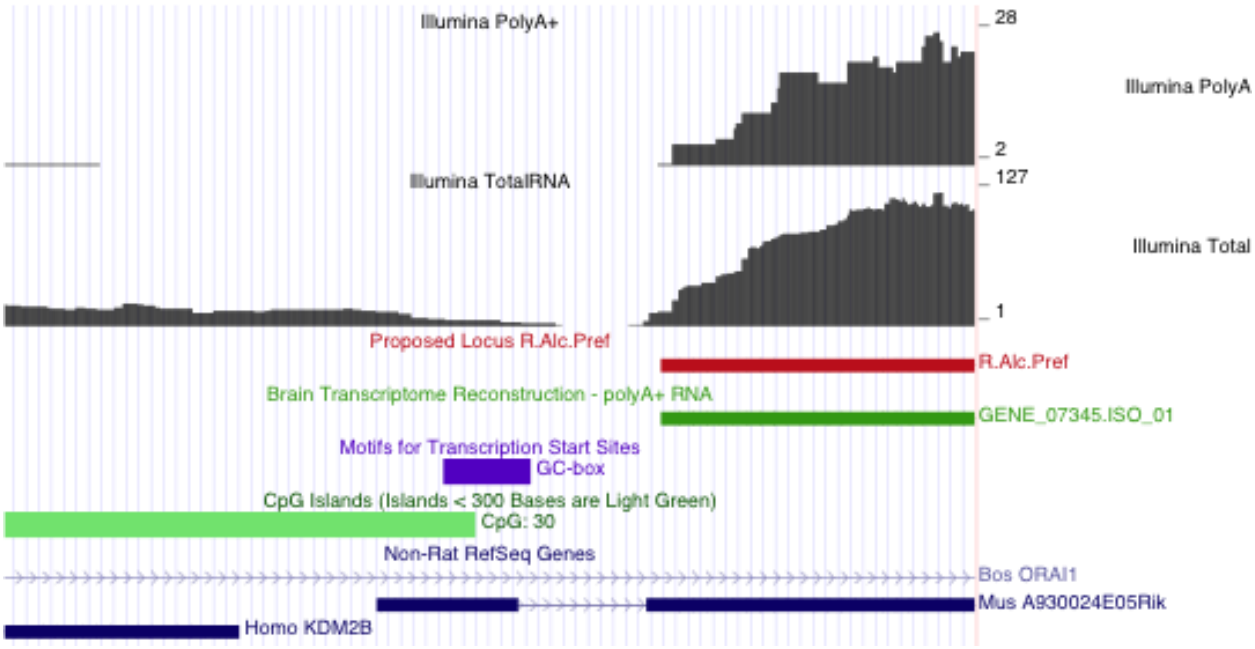


Figure 3: Transcription Start Site For Proposed Locus R.Alc.Pref. The Illumina PolyA+ track displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Proposed Locus R.Alc.Pref (red) track is the proposed transcript sequence structure derived from RNA-Seq data, qRT-PCR information, and bioinformatics analyses. The Brain Transcriptome Reconstruction – polyA+ RNA (green) is the original transcript sequence structure that was derived from a transcriptome reconstruction using polyA+-selected RNA-Seq data. The Non-Rat RefSeq Genes track was taken directly from the UCSC Genome Browser (http://genome.ucsc.edu). Because the transcript is on the antisense strand the graphic has been reversed so that 5' to 3' of the transcript is left to right and the base pair position decreases as you move from left to right.

Below is the genomic sequence near the proposed transcription start site. The bases in red correspond to the two GC-box motifs detected.

```
>chr12:40917580-40917829 (reverse
complement)AGTGCGCCTCTGCACGCAGGCGCCGCCGCCGCCCGGGAGCTGACTTTTGAAGG
GCTGAGTTGCAGGCGGCGGGCGCATCCCAGAGATGGCGGGGCAGGGAGGGAGGGAAAGAGAGGA
GGGGGTGGGGAGGGAGGTGCAGATTTGCACGGATGGCCTCAGCGGGAGGCGCGGAAATGCTAGC
CGAGCCTGCTGGCGTGAAAGCCCCCGGGGCCTGGCACCCCCAAAACCGACCCCCCCACTTCTGC
CCCTG
```

EXON/INTRON SPLICE SITE MOTIFS

There was evidence of exon/intron splice site motifs for both exon-exon junctions in Locus R.Alc.Pref that were originally identified in the transcriptome reconstruction and/or from the cloned PCR product.

Junction	Donor	Acceptor
Exon 1 to Exon 2	aagGTAAGC	TTGCAGtcc
Exon 2 to Exon 3	cagGTGAGT	TGTCAGaac

Table 3. Exon/intron splice site motifs for Locus R.Alc.Pref. Bases contained within an exon are lower case and bases contained in an intron are upper case.

START/STOP CODONS AND OPEN READING FRAMES

In the first exon, there several stop codons in every possible reading frame. Even though there are also start codons in every possible reading frame, the number of codons between start and stop codons was never more than 47. The second exon contained stop codons in every reading frame and no start codons.

In the third exon, there is an open reading frame that is 558 bases long that is located at chr12:40,911,997-40,912,548 (Figure 4).

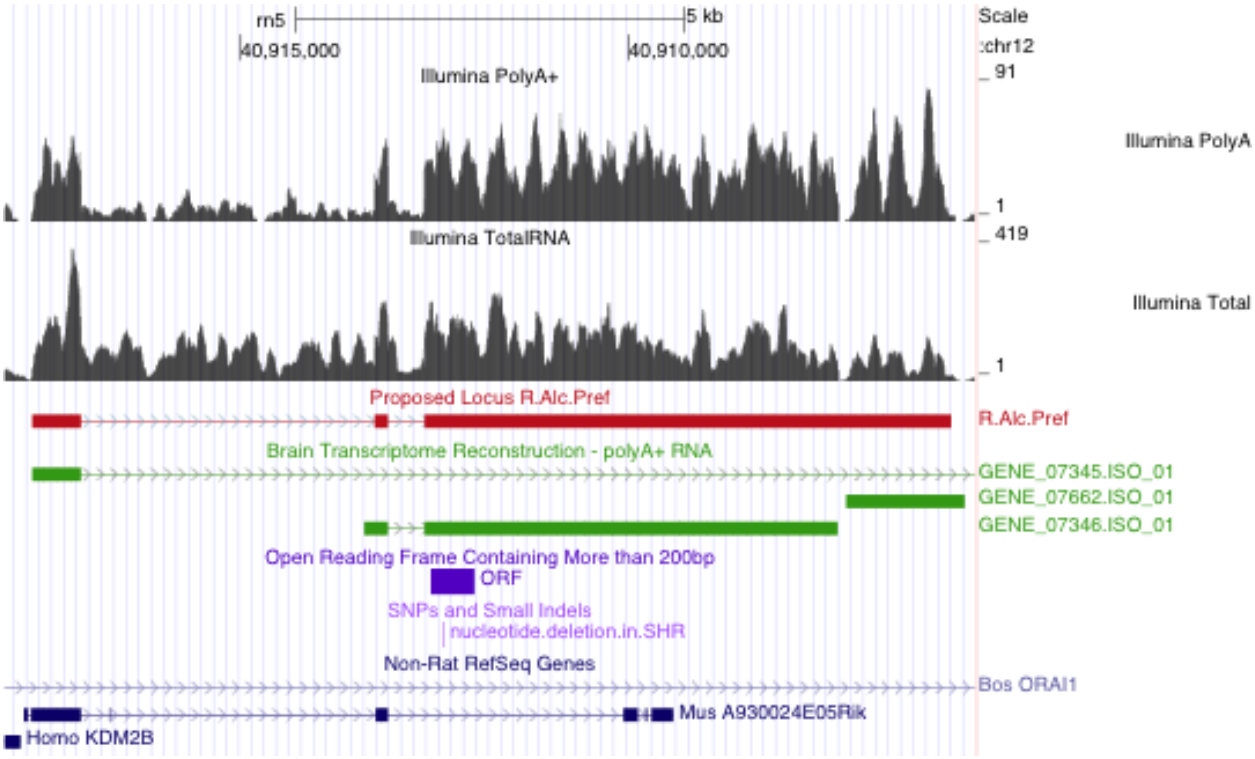


Figure 4: Open Reading Frame in Locus R.Alc.Pref. The Illumina PolyA+ track displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Proposed Locus R.Alc.Pref (red) track is the proposed transcript sequence structure derived from RNA-Seq data, qRT-PCR information, and bioinformatics analyses. The Brain Transcriptome Reconstruction – polyA+ RNA (green) is the original transcript sequence structure that was derived from a transcriptome reconstruction using polyA+-selected RNA-Seq data. The Open Reading Frame Containing More than 200bp (dark purple) is the only open reading frame in the proposed transcript that is longer than 200 bp without an intervening stop codon. The SNPs and Small Indels (light purple) track indicates a single nucleotide deletion in the SHR strain (compared to either the BN reference genome or the BN-Lx genome). The Non-Rat RefSeq Genes track was taken directly from the UCSC Genome Browser (http://genome.ucsc.edu). Because the transcript is on the antisense strand the graphic has been

reversed so that 5' to 3' of the transcript is left to right and the base pair position decreases as you move from left to right.

SNPS AND INDELS WITHIN TRANSCRIBED REGIONS

Within the ORF identified in Exon 3 (Figure 4), there is an obvious drop in read depth at a single nucleotide level. This drop corresponds to a single nucleotide deletion in the SHR/Ola that was identified in the DNA-Seq data. This deletion would have a significant impact on the protein product of the ORF and the resulting protein would be substantially shorter in the SHR strain than in the BN-Lx strain or the BN reference.

Below is the DNA sequence for the potential ORF with the start codon highlighted in green, the stop codon highlighted in red, and the single nucleotide deletion highlighted in yellow.

```
>rn5_dna range=chr12:40911892-40912554 strand=-
CCAAAGATGTGGAGACAGATGACCTTTGCAGCAGCCCTCAGACCCCCGGT
TCACAGCTGGAGAGAGGATGTGTCAGCATTTGCCTTCCAGTCTCACCCAG
CCCTTTGTTCAGAGCTCGAACCCTTTGTGCTCAGGCTGCAGCAGGGGGAT
TTTGTAATTGTTGATGCAGATTTTTTTTTTCCCTAACCTCCCACTTCTT
AACAGCTTCTGAGAAACAAACTCTTATTTGGCAGAAGATGGTATCTAAGA
GCTGCCCCAGAGGGAGAACCAGAATAGATGTGGGGAATTTTGAGAGGCCG
CTTGGCCTTTTGGCTATAACTACTGCCCTTTCCTCAGTTTCCCTGTATAC
AGTGGGCCAAACTGGACGGAGTCAGAAGTCCAACAAAGTCATAATACCTC
ATGTTAAGTGGCTCTATTTAGAAGAAGAAGAAGAAAACAAGGCTTTTACC
TATACGATCTCAAGTCAAAGGCACTTAGAACTAAAGCATTTATGTAAAAG
GACACGAAAACCTAGACAGTCCCTCAAGGTGCCGGAGAGTCTGTTTTCTC
ACCTGTGAAATGGAGGTTTTAAGGCAAAACCCACTGCAGTCTCACTATGT
GTCCTATATGCTCACAAAAGGGCTCACCCAGTAATGGTATTCAAAGGAAG
CTGATGCTCTGGG
```

The protein product of the above sequence in the BN-Lx strain and the BN reference is below.

```
>R.Alc.Pref_putative_protein_in BN_reference
MWRQMTFAAALRPPVHSWREDVSAFAFQSHPALCSELEPFVLRLQQGDFVIVDADFFFSLTSHF
LTASEKQTLIWQKMVSKSCPRGRTRIDVGNFERPLGLLAITTALSSVSLYTVGQTGRSQKSNKV
IIPHVKWLYLEEEEENKAFTYTISSQRHLELKHLCKRTRKPRQSLKVPESLFSHL
```

The protein product of the above sequence in the SHR strain with the highlighted thymine removed is below.

```
>R.Alc.Pref_putative_protein_SHR_variant
MWRQMTFAAALRPPVHSWREDVSAFAFQSHPALCSELEPFVLRLQQGDFVIVDADFFFP
```

POTENTIAL POLYADENYLATION SITES

Although the original transcript reconstruction indicated that the transcript ends at 40,907,309 bp, the read depth drops to zero at this point and then immediately increases to the same depth as the transcript. There is a strong candidate for a transcript cleavage and polyadenylation site farther downstream. This site was detected by PolyA-Seq [4] and was found in both rat brain and rat testis. The site, located within chr12:40,905,859-40,905,948, contains a canonical polyadenylation motif ATTAAA and TG- and T-rich downstream regions. Therefore, this site was used as the end of the proposed Locus R.Alc.Pref transcript.

Summary

Locus R.Alc.Pref shows transcription in both polyA+ RNA-Seq data and the total RNA-Seq data. Transcription in this region follows the antisense strand and begins near the 5' end of a sense-strand gene, Kdm2b. We could locate no canonical transcription start site motifs (TATA box or CAAT box sequences) near the start of transcription, but sequences associated with transcription start sites, namely a CpG island and GC box motifs, [2,3] are located at this point.

RT-PCR for the region gives evidence for the existence of three exons. The donor and receptor exon splice site pairs both contain canonical splice site motifs, supporting the PCR result. The first two exons contain stop codons in all reading frames, while the third exon contains an open reading frame of 558 bases. There is at least one reasonable mRNA cleavage/polyadenylation site, located approximately 6,500 bp downstream from the 3' end of the open reading frame.

References

1. Langnaese K, John R, Schweizer H, Ebmeyer U, Keilhoff G. Selection of reference genes for quantitative real-time PCR in a rat asphyxial cardiac arrest model. BMC Mol Bio 2008, 9:53.
2. Narang V, Sung WK, Mittal A. Computational modeling of oligonucleotide positional densities for human promoter prediction. Artif Intell Med. 2005 Sep-Oct;35(1-2):107-19.
3. Deaton, AM and Bird, A. CpG islands and the regulation of transcription. Genes & Dev. 2011. 25:1010-1022
4. Derti A, Garrett-Engele P, Macisaac KD, Stevens RC, Sriram S, Chen R, Rohl CA, Johnson JM, Babak T. A quantitative atlas of polyadenylation in five mammals. Genome Res. 2012 Jun;22(6):1173-83.

Doc. S2. Individual Gene Reports

For the individual gene reports, data for each candidate gene in Table 2 are outlined. Included in each report is information to evaluate concordance with regard to the gene's transcript sequence structure, detailed information on the gene's association with alcohol consumption, and detailed information on the expression levels of the gene and its associated isoform in the RNA-Seq data.

1. **Concordance of transcript structure** - the transcript structure derived from the transcriptome reconstruction was compared to the observed distribution of reads across the corresponding genomic area, including the number of reads that cover exon-exon junctions in the transcript, and to the correlation structure among the probe sets on the exon array that correspond to the transcript, using data from the HXB/BXH recombinant inbred panel.

2. **Association with voluntary alcohol consumption** - the evidence for positive or negative association between expression of the transcript and a predisposition to drinking was measured across the six pairs of selected lines and the HXB/BXH recombinant inbred panel for each gene individually. Meta-analysis results were derived as described in the text.

3. **Quantification of RNA-Seq expression** - expression estimates (FPKM) from the ribosomal RNA-depleted total RNA were calculated for each splice variant of the gene in the two progenitor strains of the RI panel.

1 5830418K08Rik (RIKEN cDNA 5830418K08 gene)

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis; isoform-level selected lines meta-analysis

Gene ID: GENE_31062

1.1 Concordance of Transcript Structure

Figure 1.1: Reconstructed Transcripts of 5830418K08Rik. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

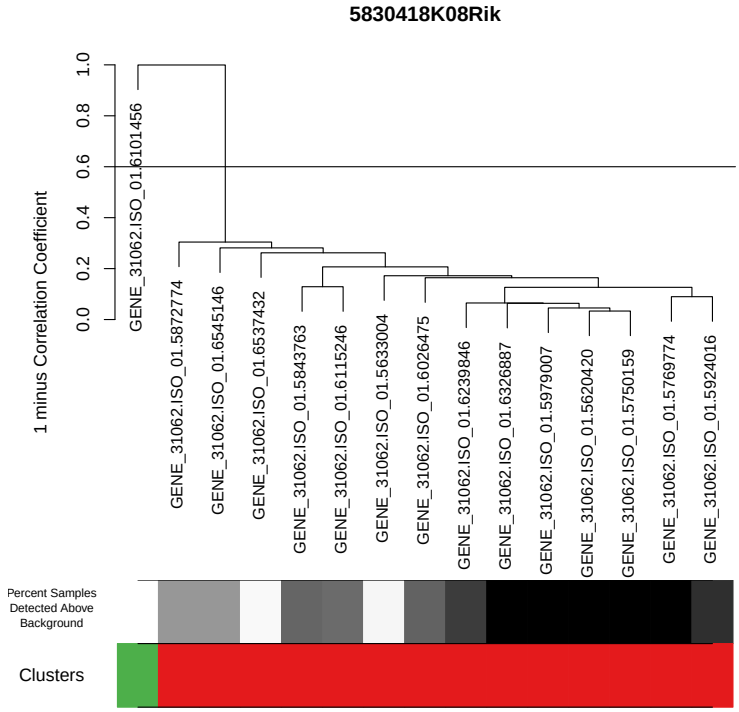


Figure 1.2: Correlation Structure - All Probe Sets Related To 5830418K08Rik. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

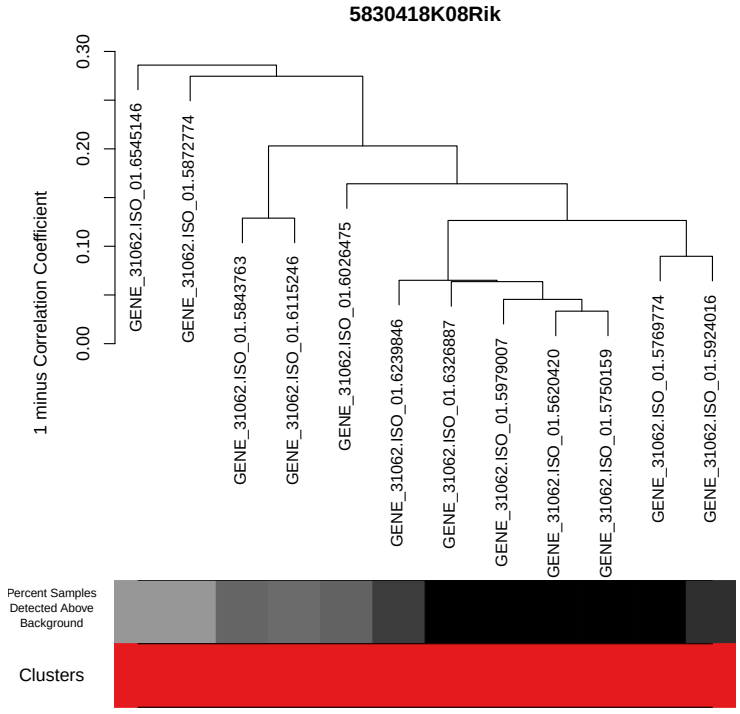


Figure 1.3: Correlation Structure - Probe Sets for 5830418K08Rik Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

The correlation structure among the probe sets for this gene matched the structure identified in the transcriptome reconstruction. The only probe set (green cluster) that did not correlate with the other probe sets was not detected above background on the Affymetrix Exon Array in more than 5% of RI samples.

1.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.14	0.5115
HAD2 vs. LAD2	0.16	0.5082
P vs. NP	-0.23	0.1150
AA vs. ANA	-0.37	0.0860
sP vs. sNP	-0.57	<0.0001
UChB vs. UChA	-0.63	<0.0001
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	-0.24	0.2950

Table 1.1: Association of Expression Levels with Alcohol Consumption - 5830418K08Rik (GENE_31062). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

1.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_31062	4.17	2.19
GENE_31062.ISO_01	4.17	2.19

Table 1.2: RNA-Seq - 5830418K08Rik. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

2 Cd74 (Cd74 molecule, major histocompatibility complex, class II invariant chain)

Association With Alcohol Consumption in:
isoform-level selected lines meta-analysis

Associated Isoform ID: GENE_13690.ISO_02
Gene ID: GENE_13690

2.1 Concordance of Transcript Structure

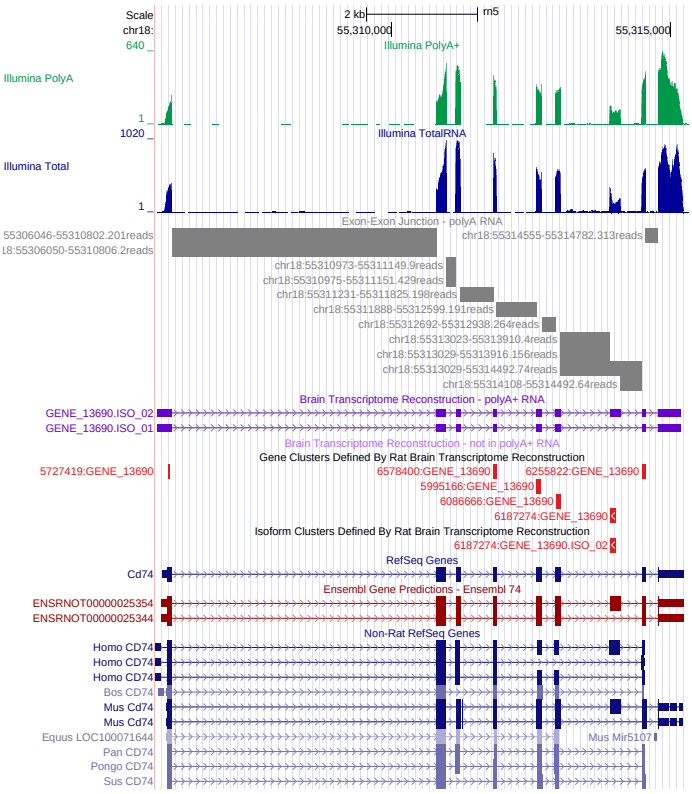


Figure 2.1: Reconstructed Transcripts of Cd74. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

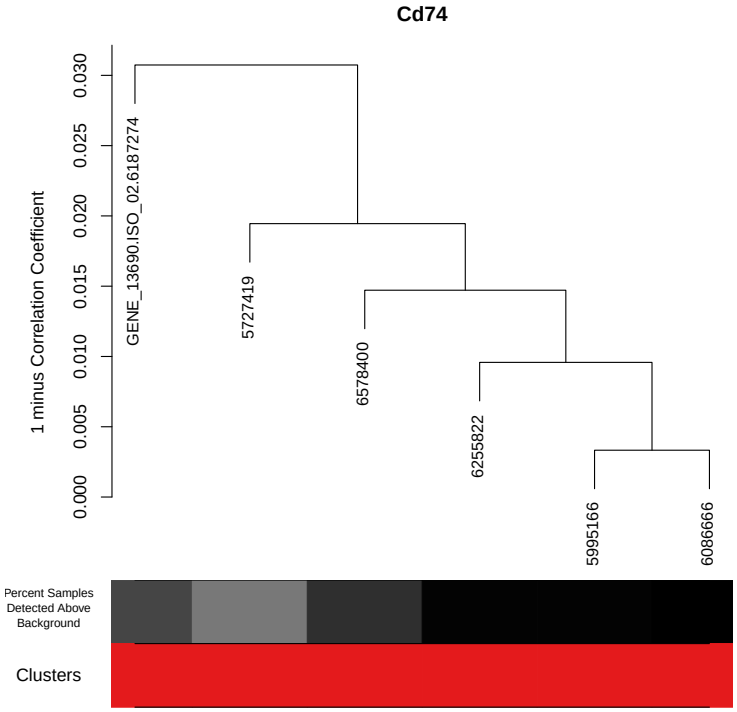


Figure 2.2: Correlation Structure - All Probe Sets Related To Cd74. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

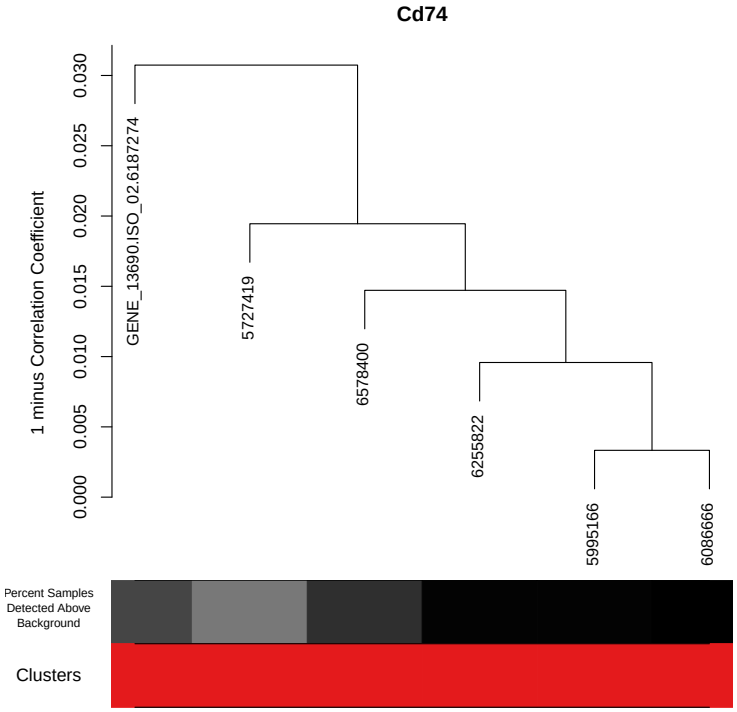


Figure 2.3: Correlation Structure - Probe Sets for Cd74 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

All the probe sets for this gene were highly correlated with each other (correlation coefficient>0.96). The second isoform was identified as associated with alcohol consumption. It was represented by only one probe set that aligns to a single exon. This probe set was the least correlated probe set in the dendrogram.

2.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	-0.30	0.1095
HAD2 vs. LAD2	-0.19	0.4216
P vs. NP	-0.05	0.7353
AA vs. ANA	-0.37	<0.0001
sP vs. sNP	0.24	0.4969
UChB vs. UChA	0.09	0.6992
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	-0.08	0.7188

Table 2.1: Association of Expression Levels with Alcohol Consumption - Cd74 (GENE_13690.ISO_02). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

2.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_13690	3.95	16.55
GENE_13690.ISO_01	2.33	8.88
GENE_13690.ISO_02	1.62	7.67

Table 2.2: RNA-Seq - Cd74. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

3 Coq5 (coenzyme Q5 homolog, methyltransferase (S. cerevisiae))

Association With Alcohol Consumption in:
gene-level WGCNA

Gene ID: GENE_07386

3.1 Concordance of Transcript Structure

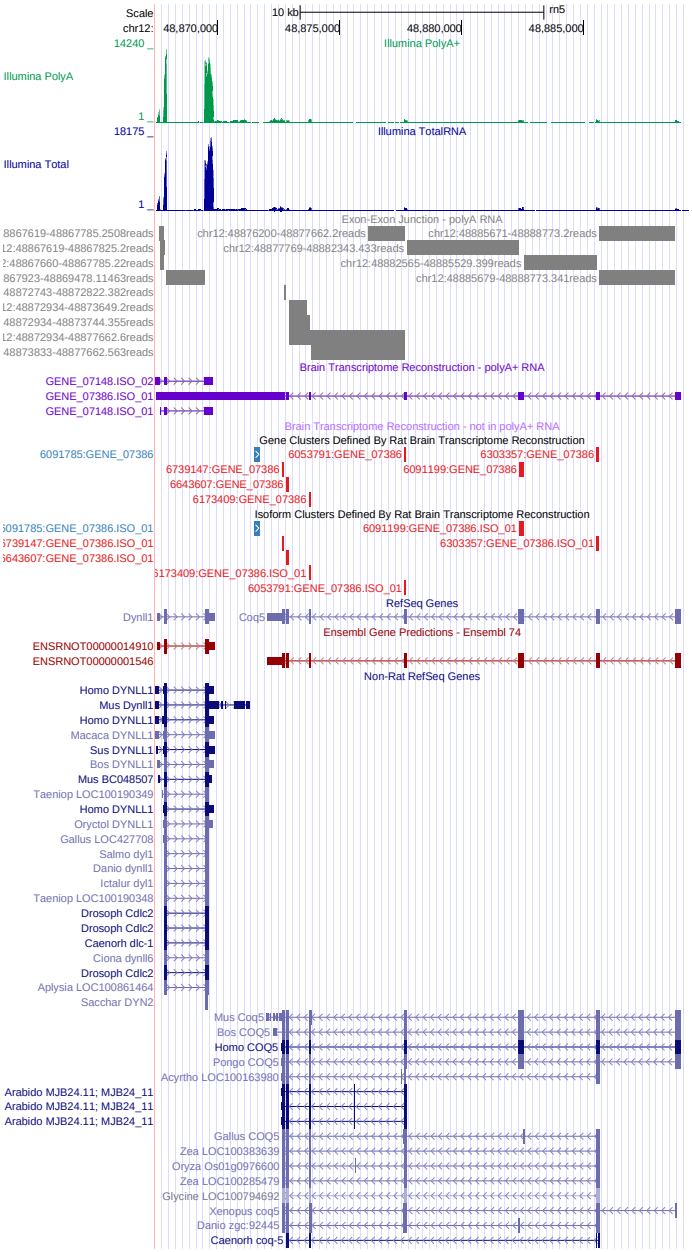


Figure 3.1: Reconstructed Transcripts of Coq5. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

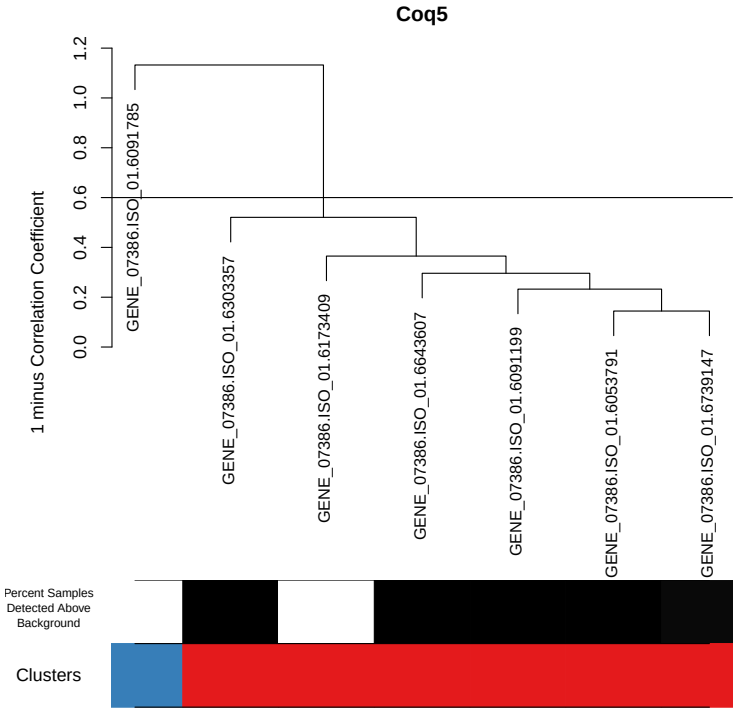


Figure 3.2: Correlation Structure - All Probe Sets Related To Coq5. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

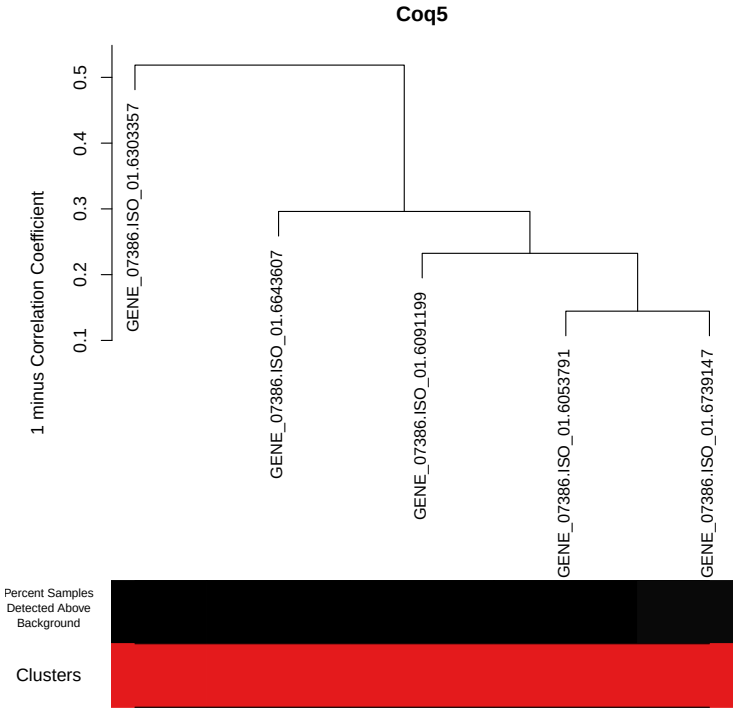


Figure 3.3: Correlation Structure - Probe Sets for Coq5 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

The correlation structure among the probe sets for this gene closely matched the transcriptome reconstruction with the exception of one probe set. This probe set aligned to the 3' region of the gene that was present in the reconstruction but not in the current annotation for the gene. The probe set was not detected above background on the Affymetrix Exon Array in the RI panel.

3.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	-0.08	0.3861
HAD2 vs. LAD2	-0.12	0.1910
P vs. NP	0.21	0.0240
AA vs. ANA	0.03	0.7311
sP vs. sNP	-0.09	0.3242
UChB vs. UChA	-0.18	0.0416
meta-analysis	-	0.3328
HXB/BXH Panel		
Correlation*	-0.50	0.0214

Table 3.1: Association of Expression Levels with Alcohol Consumption - Coq5 (GENE_07386). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

3.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_07386	7.73	8.72
GENE_07386.ISO_01	7.73	8.72

Table 3.2: RNA-Seq - Coq5. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

4 Ctss (cathepsin S)

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis

Gene ID: GENE_15787

4.1 Concordance of Transcript Structure

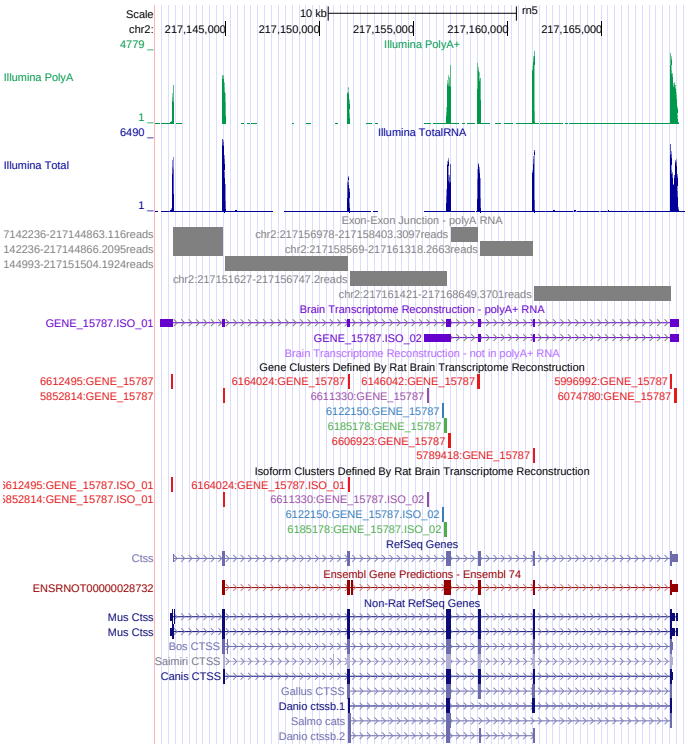


Figure 4.1: Reconstructed Transcripts of Ctss. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

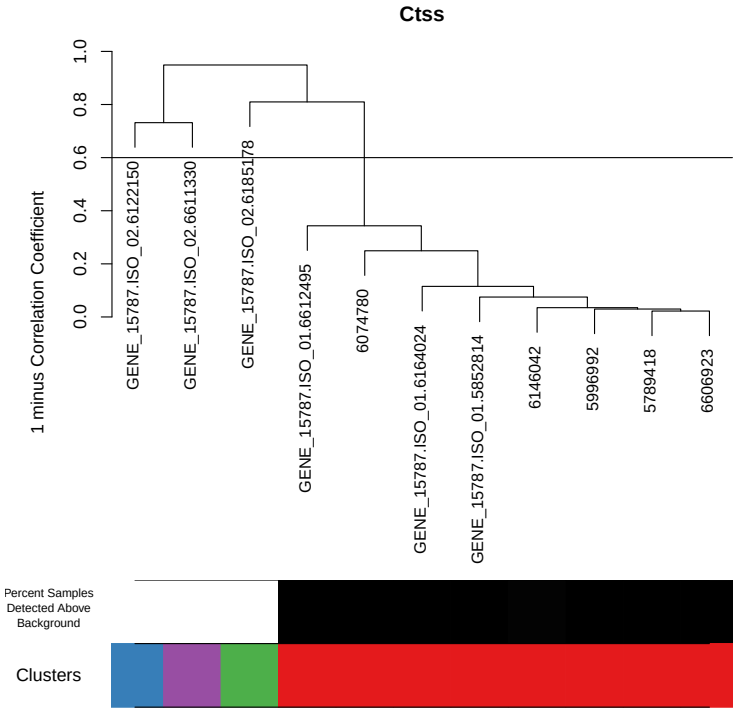


Figure 4.2: Correlation Structure - All Probe Sets Related To Ctss. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

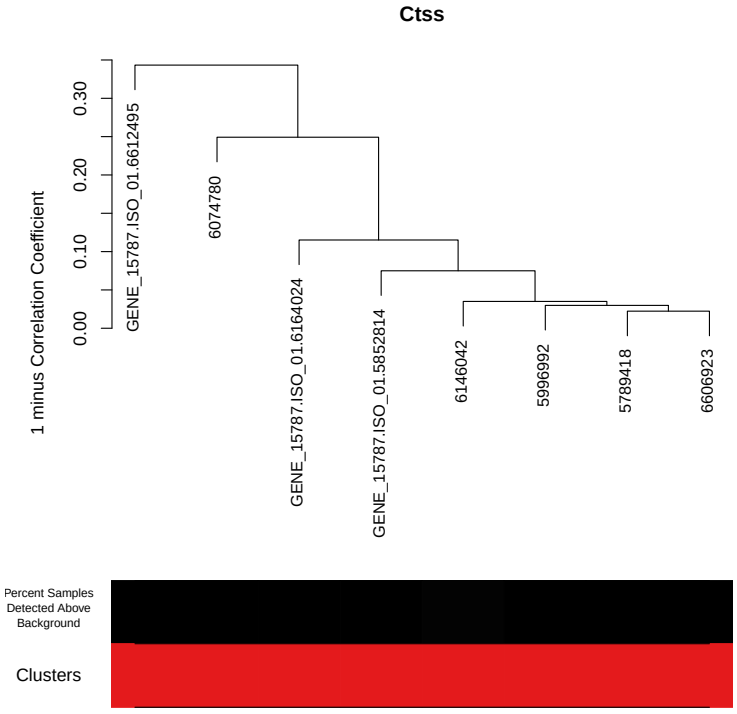


Figure 4.3: Correlation Structure - Probe Sets for Ctss Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

The correlation structure among the probe sets that were detected above background in the RI panel matched the structure of the longer isoform of Ctss with higher expression (GENE_15787.ISO_1). The probe sets that aligned to the alternative 5' region for isoform 2 were not detected above background on the Affymetrix Exon Array and were not correlated with the other probe sets.

4.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.11	0.3058
HAD2 vs. LAD2	0.16	0.4486
P vs. NP	-0.17	0.0347
AA vs. ANA	-0.23	0.0339
sP vs. sNP	-0.27	<0.0001
UChB vs. UChA	-0.08	0.3401
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	-0.05	0.8196

Table 4.1: Association of Expression Levels with Alcohol Consumption - Ctss (GENE_15787). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

4.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_15787	43.73	50.52
GENE_15787.ISO_01	43.30	50.02
GENE_15787.ISO_02	0.44	0.50

Table 4.2: RNA-Seq - Ctss. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

5 Fbln1 (fibulin 1)

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis

Gene ID: GENE_28831

5.1 Concordance of Transcript Structure

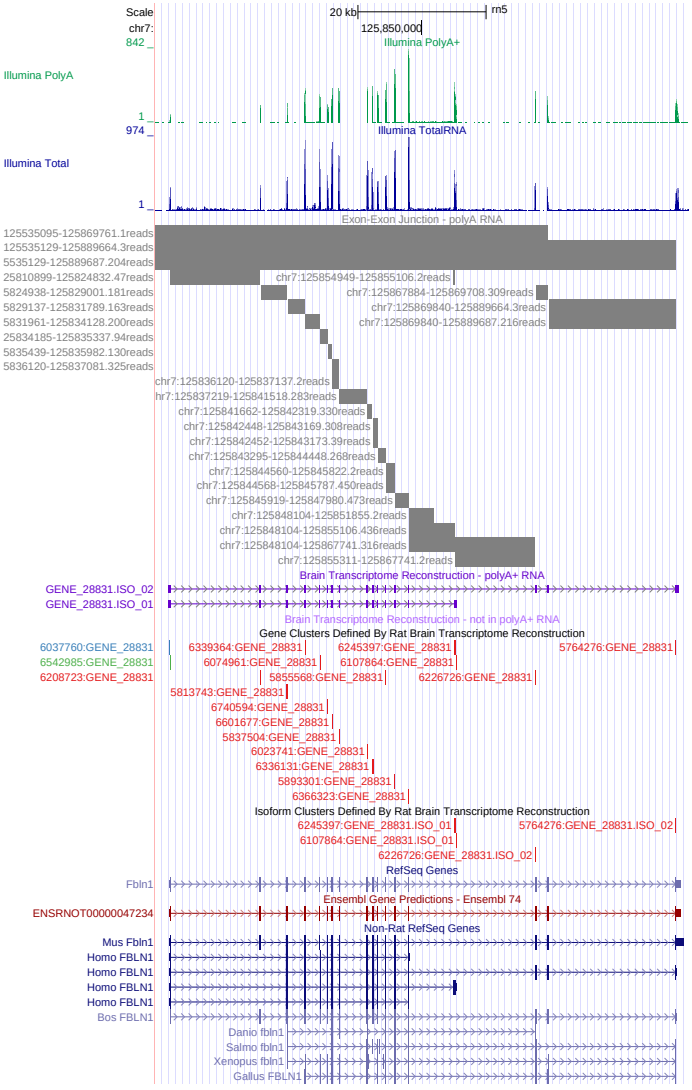


Figure 5.1: Reconstructed Transcripts of Fbln1. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

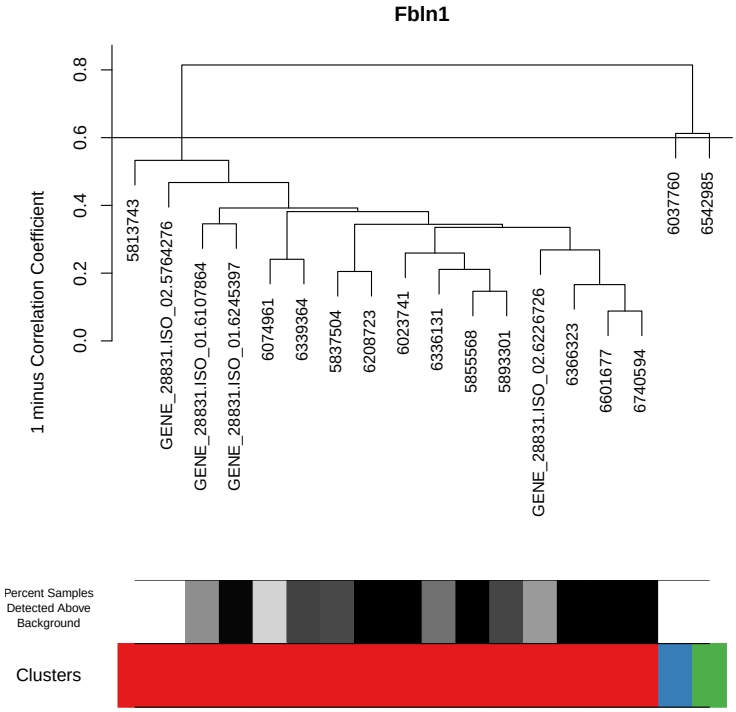


Figure 5.2: Correlation Structure - All Probe Sets Related To Fbln1. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

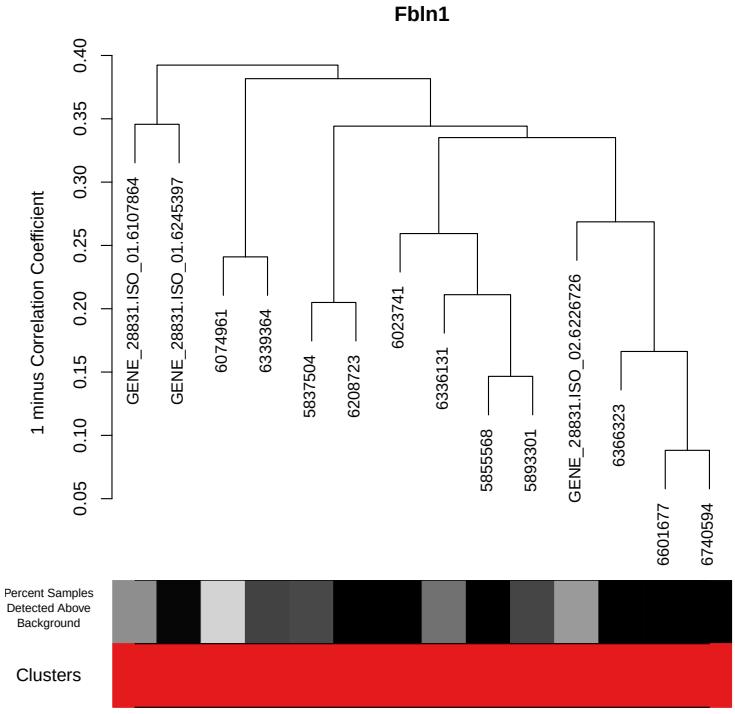


Figure 5.3: Correlation Structure - Probe Sets for Fbln1 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

The majority of probe sets for this gene were highly correlated and were detected above background on the Affymetrix Exon Array in the RI panel. The two probe sets that were not correlated were not detected above background and aligned to the 5' area of the transcript.

5.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.09	0.4086
HAD2 vs. LAD2	-0.03	0.7583
P vs. NP	0.22	0.0007
AA vs. ANA	-0.19	0.0892
sP vs. sNP	0.05	0.2744
UChB vs. UChA	0.22	<0.0001
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	-0.28	0.2231

Table 5.1: Association of Expression Levels with Alcohol Consumption - Fbln1 (GENE_28831). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

5.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_28831	8.75	7.37
GENE_28831.ISO_01	4.91	4.66
GENE_28831.ISO_02	3.84	2.71

Table 5.2: RNA-Seq - Fbln1. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

6 Fbxo45 (F-box protein 45)

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis; isoform-level selected lines meta-analysis

Gene ID: GENE_05781

6.1 Concordance of Transcript Structure

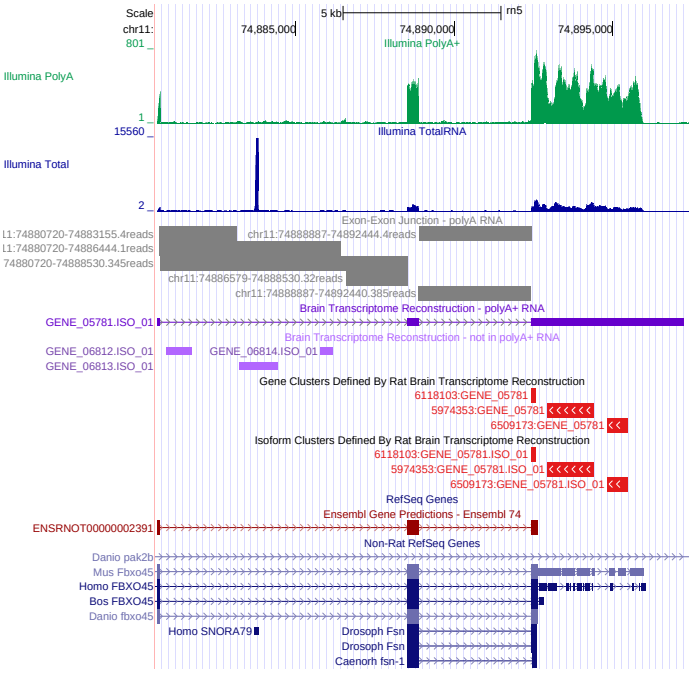


Figure 6.1: Reconstructed Transcripts of Fbxo45. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

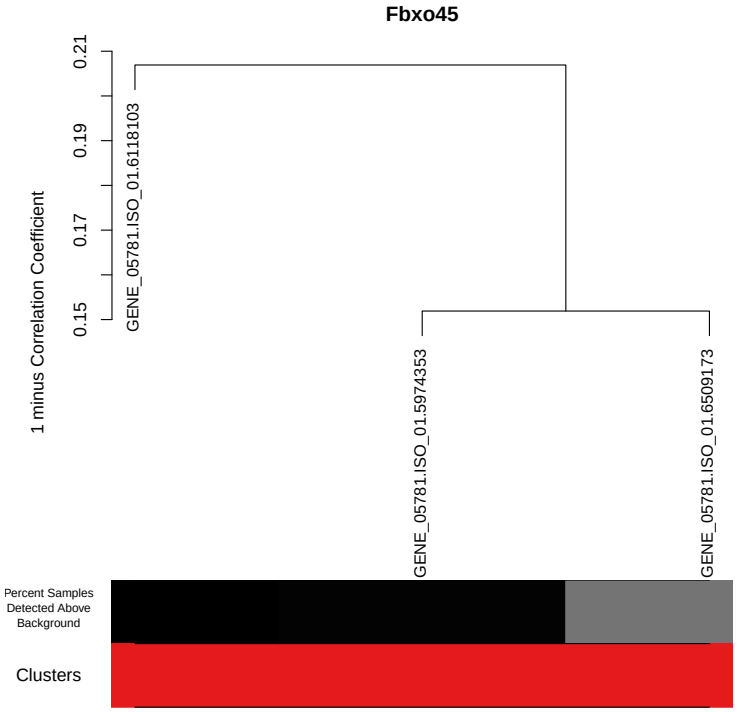


Figure 6.2: Correlation Structure - All Probe Sets Related To Fbxo45. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

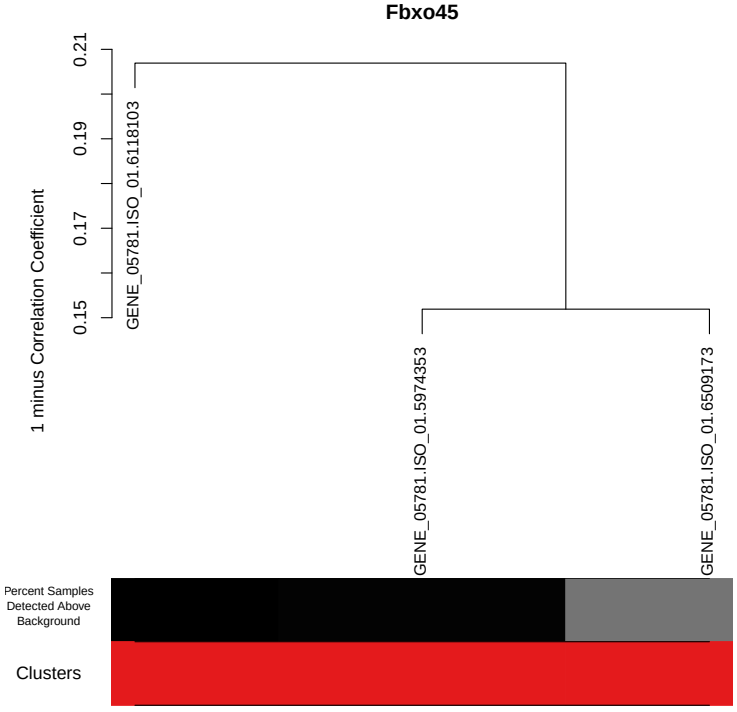


Figure 6.3: Correlation Structure - Probe Sets for Fbxo45 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

The three probe sets that aligned to this gene were highly correlated and were detected above background on the exon array in the RI panel. The three probe sets all aligned to the 3' area of the gene that was extended in the transcriptome reconstruction beyond the transcript end denoted in the Ensembl database. Neither of the other two exons of the gene were probed by the exon array.

6.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.13	0.5490
HAD2 vs. LAD2	0.18	0.2621
P vs. NP	-0.15	0.2382
AA vs. ANA	-0.04	0.6306
sP vs. sNP	-0.22	<0.0001
UChB vs. UChA	-0.17	0.2274
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	-0.12	0.6025

Table 6.1: Association of Expression Levels with Alcohol Consumption - Fbxo45 (GENE_05781). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

6.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_05781	11.67	16.15
GENE_05781.ISO_01	11.67	16.15

Table 6.2: RNA-Seq - Fbxo45. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

7 GENE_07345 (partial overlap with Orai1 and mouse A930024E05Rik)

Association With Alcohol Consumption in:
gene-level WGCNA

Gene ID: GENE_07345

7.1 Concordance of Transcript Structure

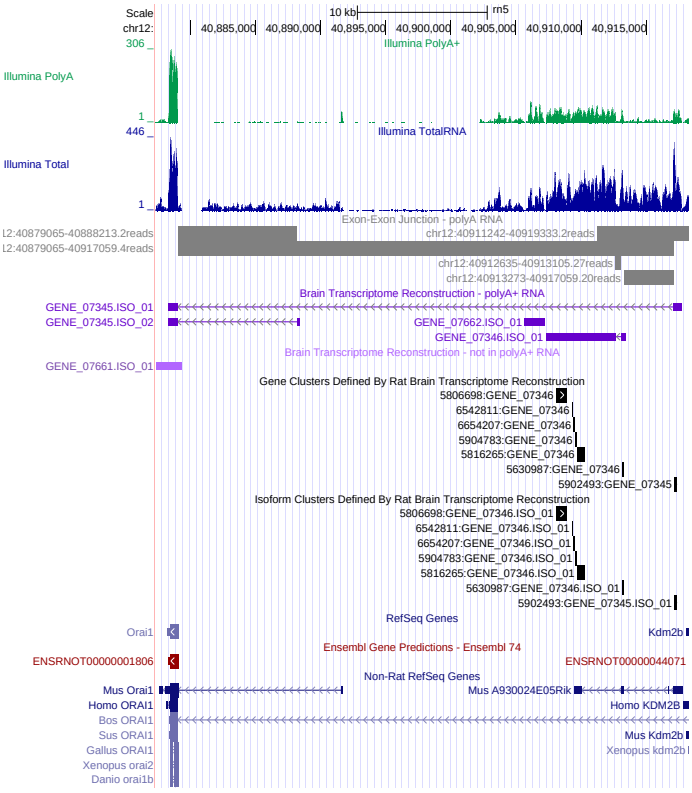


Figure 7.1: Reconstructed Transcripts of GENE_07345. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

Comments

This gene cluster consisted of only one probe set. That probe set, GENE_07345.ISO_01.5902493, was above detection limits in 44% percent of the samples.

Because this gene cluster only contained a single probe set, the correlation structure was not evaluated. See additional supplemental information for full interrogation of this gene cluster and the adjacent gene cluster, GENE_07346, for their relationship with each other and with A930024E05Rik.

7.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.10	0.5964
HAD2 vs. LAD2	0.01	0.9406
P vs. NP	-0.32	0.0488
AA vs. ANA	0.98	<0.0001
sP vs. sNP	-0.19	0.5966
UChB vs. UChA	-1.12	<0.0001
meta-analysis	-	0.1829
HXB/BXH Panel		
Correlation*	-0.63	0.0023

Table 7.1: Association of Expression Levels with Alcohol Consumption - GENE_07345 (GENE_07345). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

7.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_07345	4.23	4.12
GENE_07345.ISO_01	3.80	2.70
GENE_07345.ISO_02	0.43	1.41

Table 7.2: RNA-Seq - GENE_07345. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

8 GENE_07346 (homologous with mouse A930024E05Rik)

Association With Alcohol Consumption in:
gene-level WGCNA

Gene ID: GENE_07346

8.1 Concordance of Transcript Structure

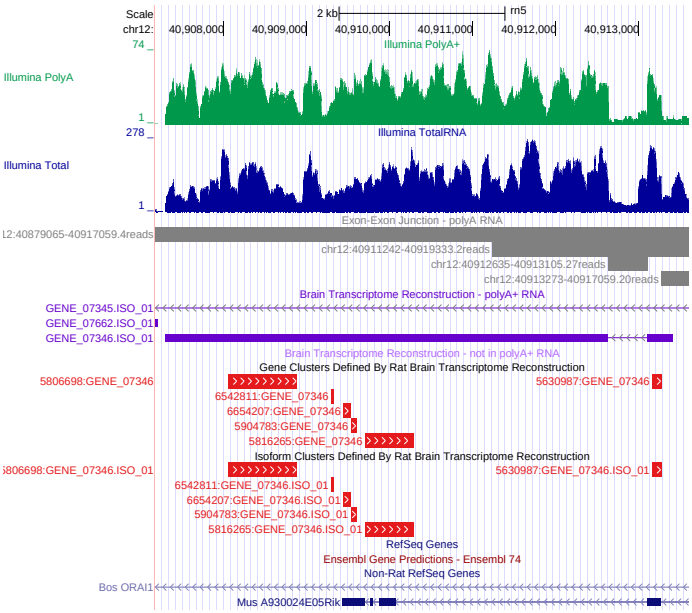


Figure 8.1: Reconstructed Transcripts of GENE_07346. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

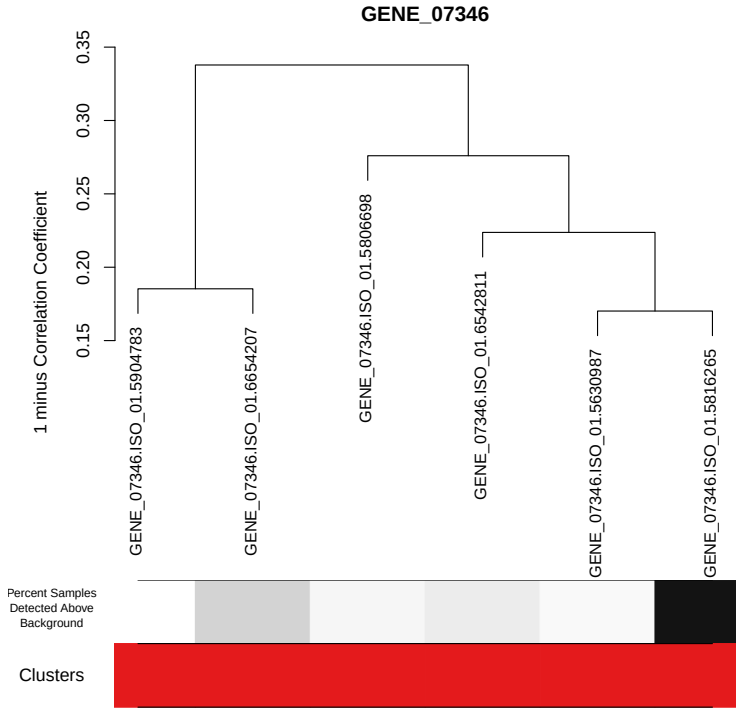


Figure 8.2: Correlation Structure - All Probe Sets Related To GENE_07346. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

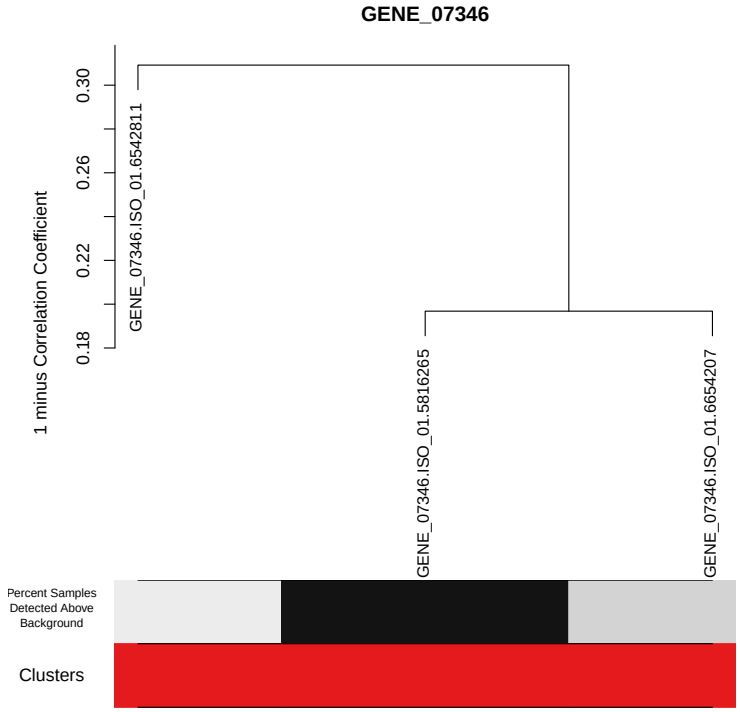


Figure 8.3: Correlation Structure - Probe Sets for GENE_07346 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

All probe sets from this gene cluster were highly correlated, but only 3 of the 6 were detected above background in the RI panel. See additional supplemental information for full interrogation of this gene cluster and the adjacent gene cluster, GENE_07345, for their relationship with each other and with A930024E05Rik.

8.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.08	0.3042
HAD2 vs. LAD2	0.38	0.0838
P vs. NP	-0.15	0.1838
AA vs. ANA	0.43	0.0298
sP vs. sNP	-0.05	0.7139
UChB vs. UChA	-0.42	<0.0001
meta-analysis	-	0.6752
HXB/BXH Panel		
Correlation*	-0.55	0.0106

Table 8.1: Association of Expression Levels with Alcohol Consumption - GENE_07346 (GENE_07346). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

8.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_07346	2.66	1.83
GENE_07346.ISO_01	2.66	1.83

Table 8.2: RNA-Seq - GENE_07346. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

9 GENE_09839 and GENE_09839.ISO_01 (no annotation)

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis; isoform-level selected lines meta-analysis

Gene ID: GENE_09839

9.1 Concordance of Transcript Structure

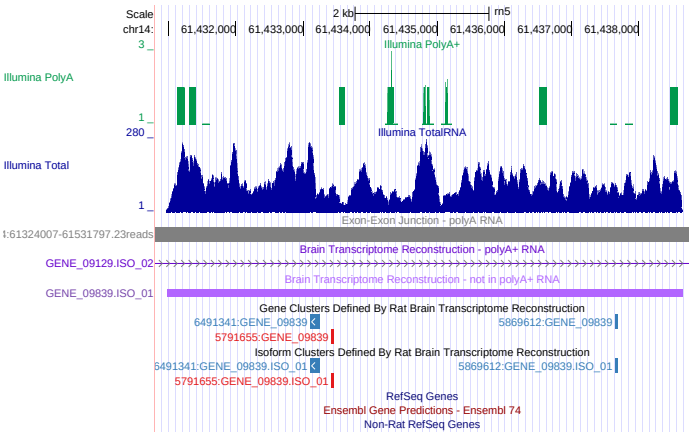


Figure 9.1: Reconstructed Transcripts of GENE_09839 and GENE_09839.ISO_01. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

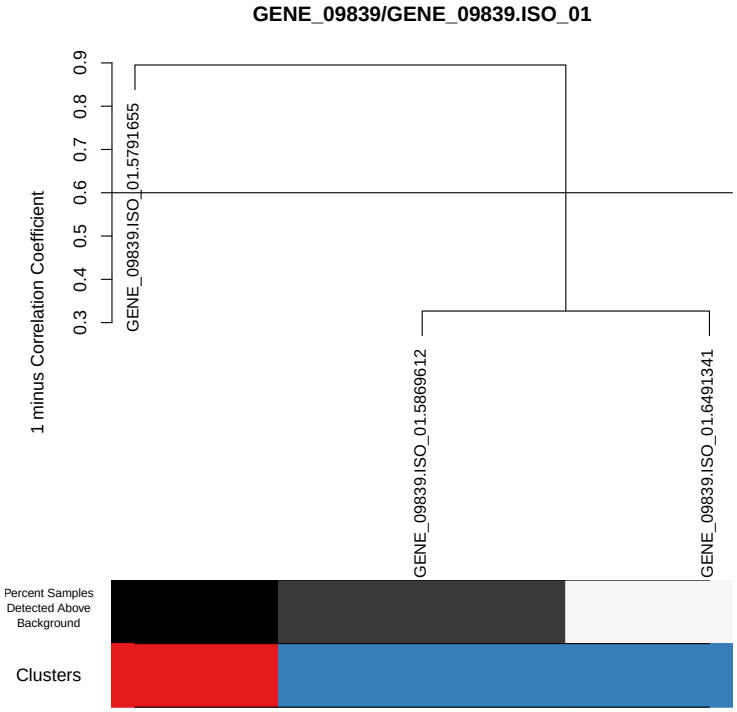


Figure 9.2: Correlation Structure - All Probe Sets Related To GENE_09839 and GENE_09839.ISO_01. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

Only 2 of the 3 probe sets for this gene cluster were expressed above background on the Affymetrix Exon Array in more than 5% of RI samples, GENE_09839.ISO_01.5791655 and GENE_09839.ISO_01.5869612. The probe sets detected above background were not correlated (correlation coefficient=0.17).

For this one-exon transcript detected only in the total RNA data, the correlation structure among the probe sets of the cluster did not match the transcriptome reconstruction. Two of the probe sets were correlated, but one of the pair is not detected above background. The probe set that did not correlate with the other two probe sets was located in a 'valley' as compared to the two correlated probe sets when examining the profile of RNA-Seq reads from the total RNA samples.

9.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	-0.21	0.0830
HAD2 vs. LAD2	-0.00	0.9975
P vs. NP	-0.46	0.0110
AA vs. ANA	0.20	0.3837
sP vs. sNP	-0.29	<0.0001
UChB vs. UChA	-0.06	0.4837
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	-0.12	0.6175

Table 9.1: Association of Expression Levels with Alcohol Consumption - GENE_09839 and GENE_09839.ISO_01 (GENE_09839). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

9.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_09839	2.18	1.98
GENE_09839.ISO_01	2.18	1.98

Table 9.2: RNA-Seq - GENE_09839 and GENE_09839.ISO_01. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

10 GENE_18351.ISO_01 (no annotation)

Association With Alcohol Consumption in:
isoform-level selected lines meta-analysis

Associated Isoform ID: GENE_18351.ISO_01
Gene ID: GENE_18351

10.1 Concordance of Transcript Structure

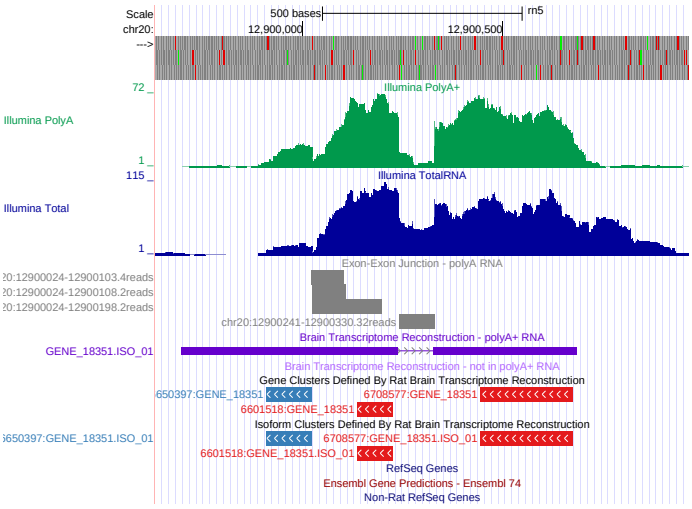


Figure 10.1: Reconstructed Transcripts of GENE_18351.ISO_01. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

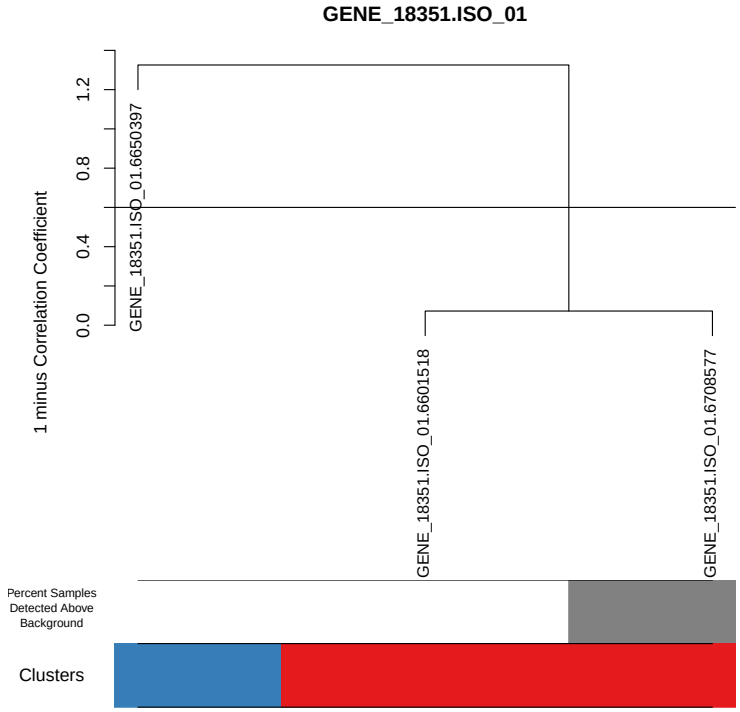


Figure 10.2: Correlation Structure - All Probe Sets Related To GENE_18351.ISO_01. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

Only one of the three probesets for this isoform cluster was detected above background on the exon array in more than 5% of the RI samples, GENE_18351.ISO_01.6708577.

In this isoform cluster, 2 of the 3 probe sets are correlated although one of those was not detected above background in the RI panel, however this isoform cluster was associated in the selected lines not the RI panel. The third probe set that does not correlate with the other two is from the 5' region of the gene. This region has a much lower read count according to the read count profiles on the browser image than the regions interrogated by the other two probe sets.

10.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.51	<0.0001
HAD2 vs. LAD2	0.24	0.1347
P vs. NP	0.12	0.0847
AA vs. ANA	0.42	0.0017
sP vs. sNP	-0.29	0.0599
UChB vs. UChA	0.32	0.0005
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	-0.18	0.4221

Table 10.1: Association of Expression Levels with Alcohol Consumption - GENE_18351.ISO_01 (GENE_18351.ISO_01). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

10.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_18351	0.31	1.53
GENE_18351.ISO_01	0.31	1.53

Table 10.2: RNA-Seq - GENE_18351.ISO_01. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

11 GENE_27603 (no annotation)

Association With Alcohol Consumption in:
gene-level WGCNA

Gene ID: GENE_27603

11.1 Concordance of Transcript Structure

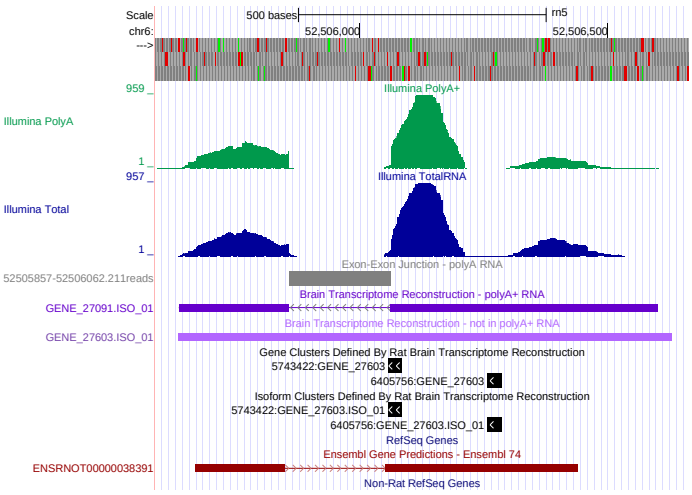


Figure 11.1: Reconstructed Transcripts of GENE_27603. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

Comments

For this gene, only 2 probe sets were included in the gene cluster. GENE_27603.ISO_01.5743422 and GENE_27603.ISO_01.6405756 had a relatively small negative correlation (correlation coefficient=-0.31). GENE_27603.ISO_01.5743422 was expressed above detection limits on the exon array in 100% of the RI

samples and GENE_27603.ISO_01.6405756 was expressed above detection limits on the exon array in 0% of the RI samples.

The probe set that was not expressed above background is located in a 'valley' in the RNA-Seq read count profiles.

11.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	-0.02	0.9000
HAD2 vs. LAD2	-0.24	0.1029
P vs. NP	0.05	0.7181
AA vs. ANA	0.12	0.4185
sP vs. sNP	0.11	0.4288
UChB vs. UChA	-0.38	0.0082
meta-analysis	-	0.3247
HXB/BXH Panel		
Correlation*	-0.50	0.0208

Table 11.1: Association of Expression Levels with Alcohol Consumption - GENE_27603 (GENE_27603). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

11.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_27603	0.27	0.27
GENE_27603.ISO_01	0.27	0.27

Table 11.2: RNA-Seq - GENE_27603. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

12 Ift81 (intraflagellar transport 81 homolog)

Association With Alcohol Consumption in:
gene-level WGCNA

Gene ID: GENE_07113

12.1 Concordance of Transcript Structure

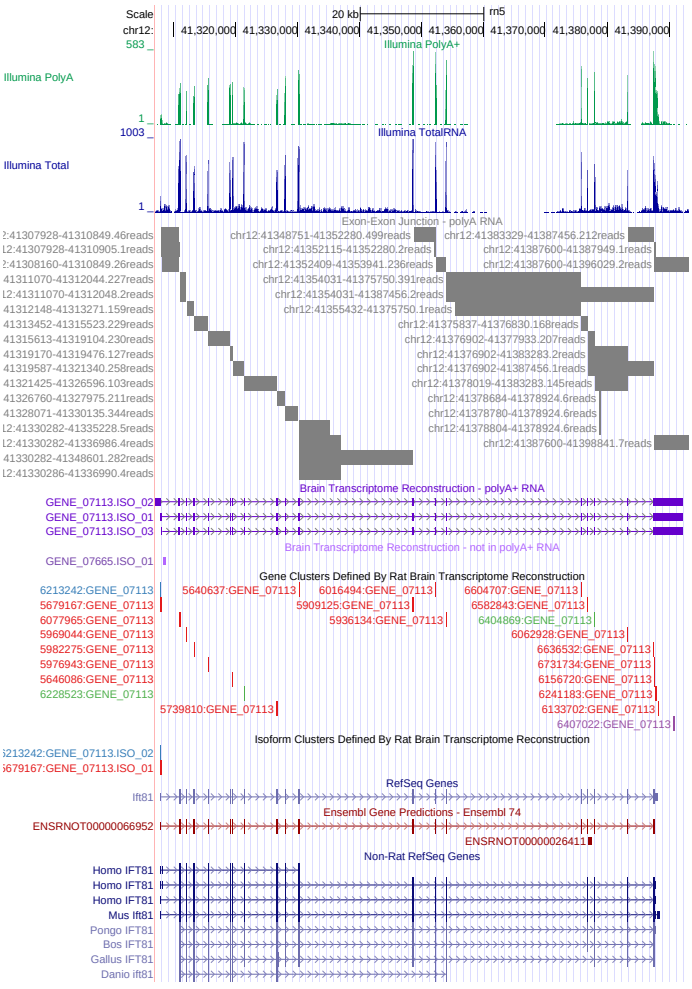


Figure 12.1: Reconstructed Transcripts of Ift81. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

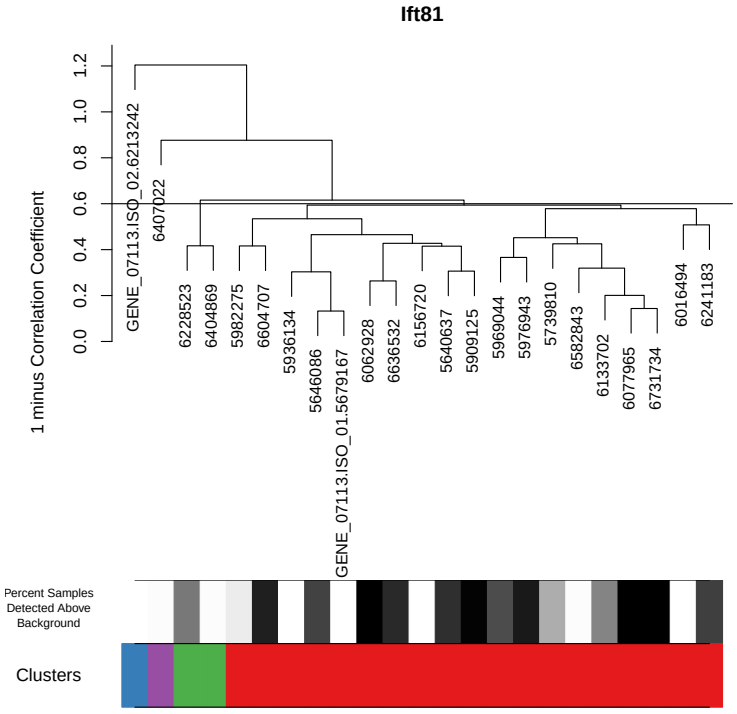


Figure 12.2: Correlation Structure - All Probe Sets Related To Ift81. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

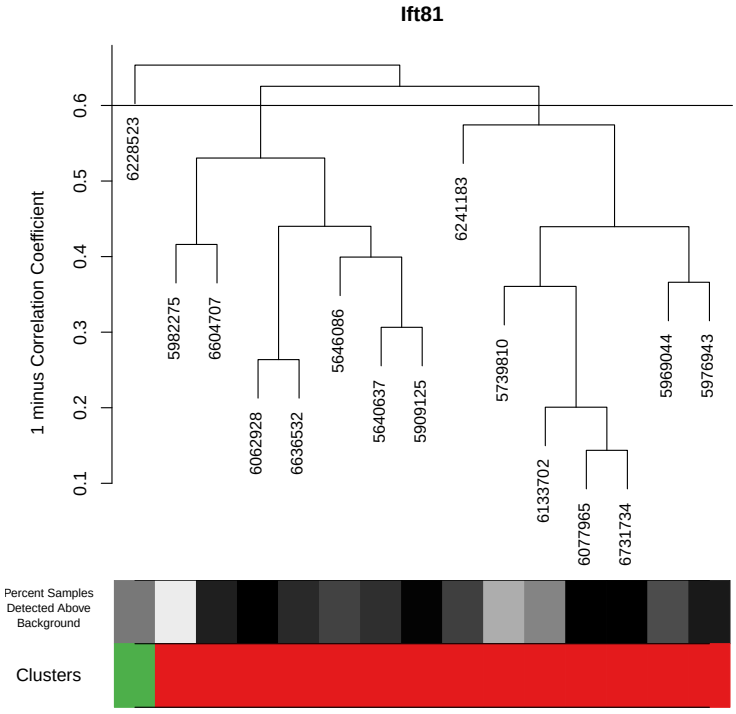


Figure 12.3: Correlation Structure - Probe Sets for Ift81 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

Among the probe sets for this gene cluster that were expressed above background, only one was not correlated at the chosen threshold of 0.6 with the other probe sets. However, it barely missed this threshold.

12.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.17	0.0677
HAD2 vs. LAD2	0.10	0.3298
P vs. NP	-0.18	0.0542
AA vs. ANA	0.10	0.2958
sP vs. sNP	0.03	0.7884
UChB vs. UChA	-0.33	0.0010
meta-analysis	-	0.6024
HXB/BXH Panel		
Correlation*	-0.43	0.0507

Table 12.1: Association of Expression Levels with Alcohol Consumption - Ift81 (GENE_07113). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

12.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_07113	4.22	3.67
GENE_07113.ISO_01	3.25	2.42
GENE_07113.ISO_02	0.84	1.14
GENE_07113.ISO_03	0.13	0.12

Table 12.2: RNA-Seq - Ift81. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

13 Maats1 (MYCBP-associated, testis expressed 1)

Association With Alcohol Consumption in:
gene-level WGCNA

Gene ID: GENE_05917

13.1 Concordance of Transcript Structure

Figure 13.1: Reconstructed Transcripts of Maats1. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

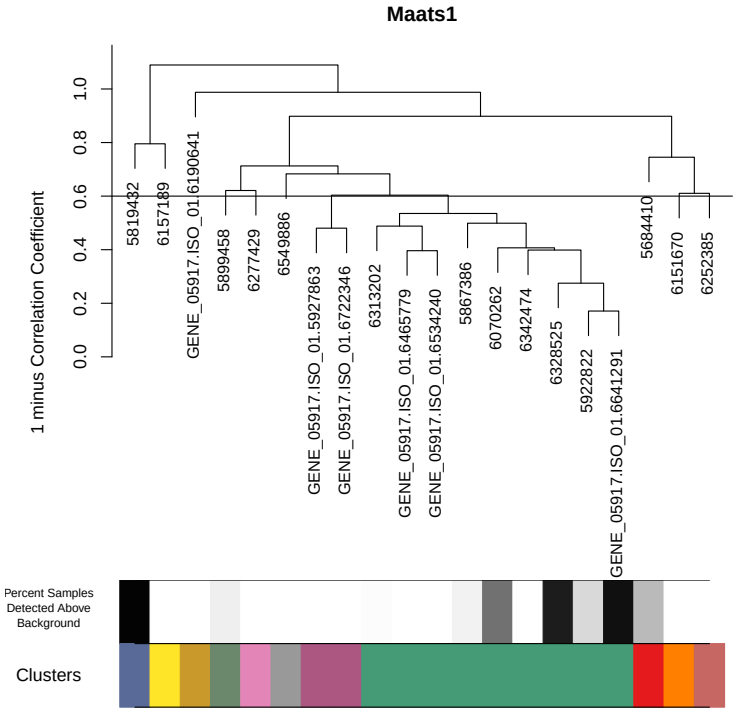


Figure 13.2: Correlation Structure - All Probe Sets Related To Maats1. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

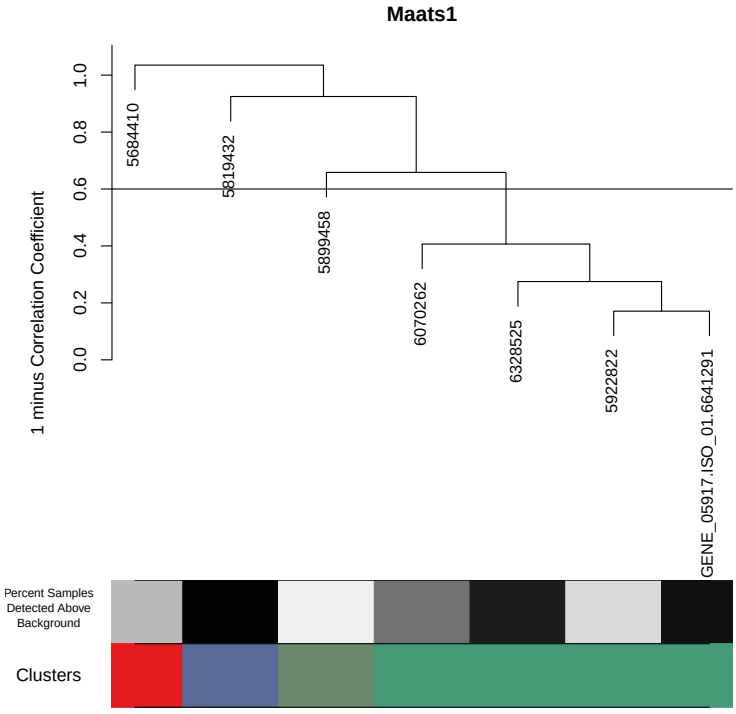


Figure 13.3: Correlation Structure - Probe Sets for Maats1 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

This gene had low expression levels in both the RNA-Seq data and in the exon array data. Only 7 of the 20 probe sets assigned to this cluster had expression levels above background. Four of the seven were correlated. Two of the uncorrelated probe sets were from the 3' area of the gene that differed from the annotated version of the gene.

13.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.06	0.3007
HAD2 vs. LAD2	0.02	0.7202
P vs. NP	0.10	0.1149
AA vs. ANA	-0.04	0.5818
sP vs. sNP	-0.12	0.0481
UChB vs. UChA	-0.03	0.6501
meta-analysis	-	0.9975
HXB/BXH Panel		
Correlation*	-0.56	0.0080

Table 13.1: Association of Expression Levels with Alcohol Consumption - Maats1 (GENE_05917). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

13.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_05917	1.14	0.98
GENE_05917.ISO_01	0.91	0.75
GENE_05917.ISO_02	0.23	0.23

Table 13.2: RNA-Seq - Maats1. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

14 Nxph1 (neurexophilin 1)

Association With Alcohol Consumption in:
isoform-level selected lines meta-analysis

Associated Isoform ID: GENE_21565.ISO_02
Gene ID: GENE_21565

14.1 Concordance of Transcript Structure

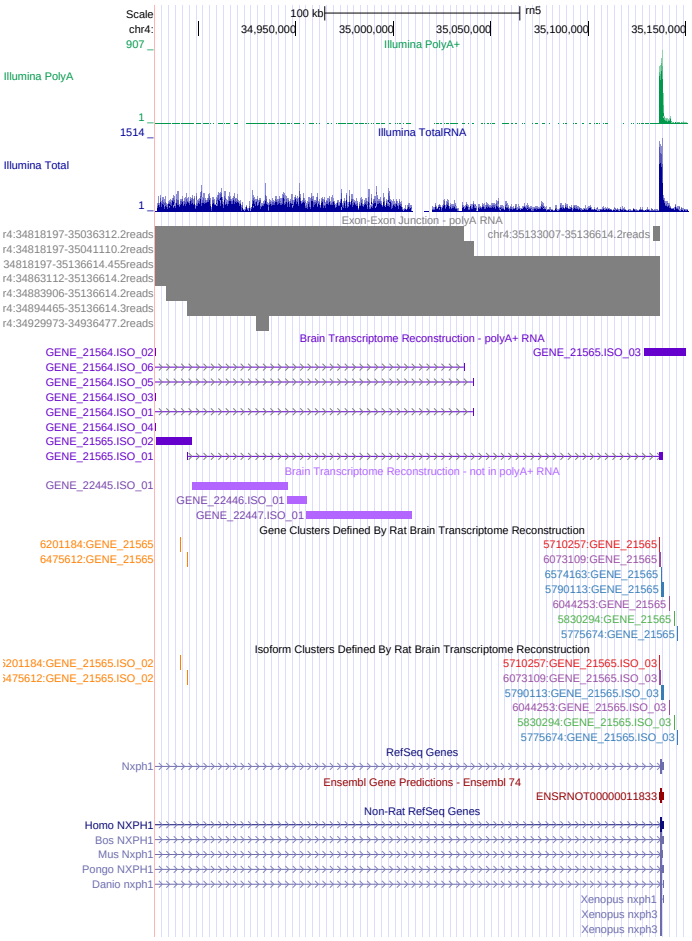


Figure 14.1: Reconstructed Transcripts of Nxph1. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

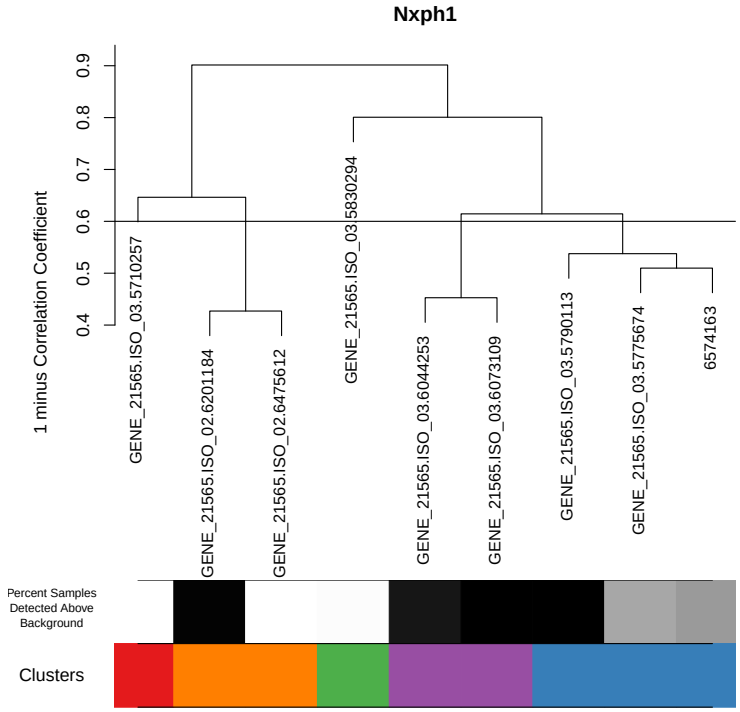


Figure 14.2: Correlation Structure - All Probe Sets Related To Nxph1. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

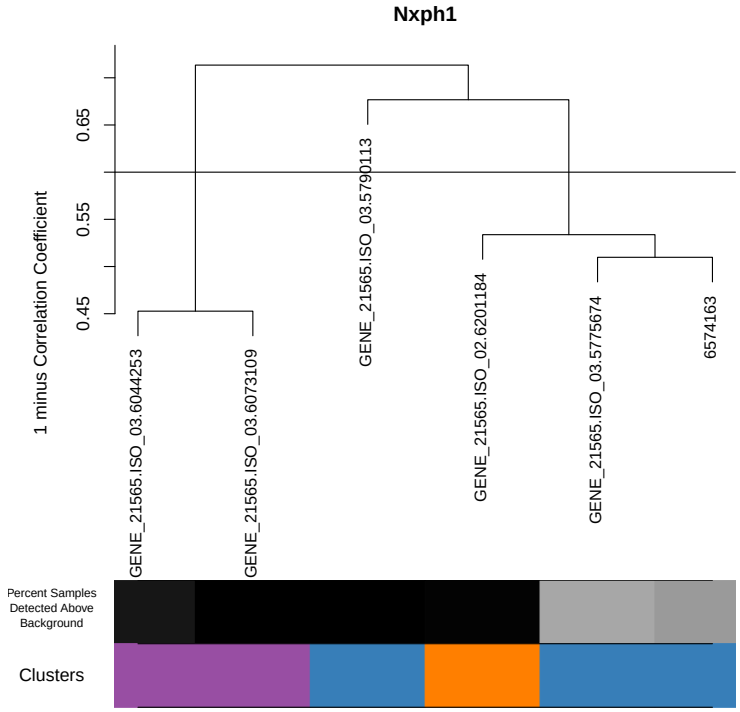


Figure 14.3: Correlation Structure - Probe Sets for Nxph1 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

Isoform 2 was the only isoform of this gene that was associated with drinking. Two probe sets were included in this isoform cluster. The two probe sets were highly correlated with each other, although one probe set was not detected above background in the RI panel (the isoform was associated in the selected lines).

14.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	-0.04	0.8808
HAD2 vs. LAD2	0.13	0.5887
P vs. NP	-0.85	0.0002
AA vs. ANA	0.26	0.5080
sP vs. sNP	-0.15	0.1901
UChB vs. UChA	-0.49	<0.0001
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	-0.09	0.6858

Table 14.1: Association of Expression Levels with Alcohol Consumption - Nxph1 (GENE_21565.ISO_02). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

14.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_21565	19.45	18.98
GENE_21565.ISO_01	15.86	15.04
GENE_21565.ISO_02	2.53	2.83
GENE_21565.ISO_03	1.06	1.11

Table 14.2: RNA-Seq - Nxph1. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

15 Oas1b (2-5 oligoadenylate synthetase 1B)

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis; gene-level WGCNA; isoform-level selected lines meta-analysis; isoform-level WGCNA

Gene ID: GENE_07363

15.1 Concordance of Transcript Structure

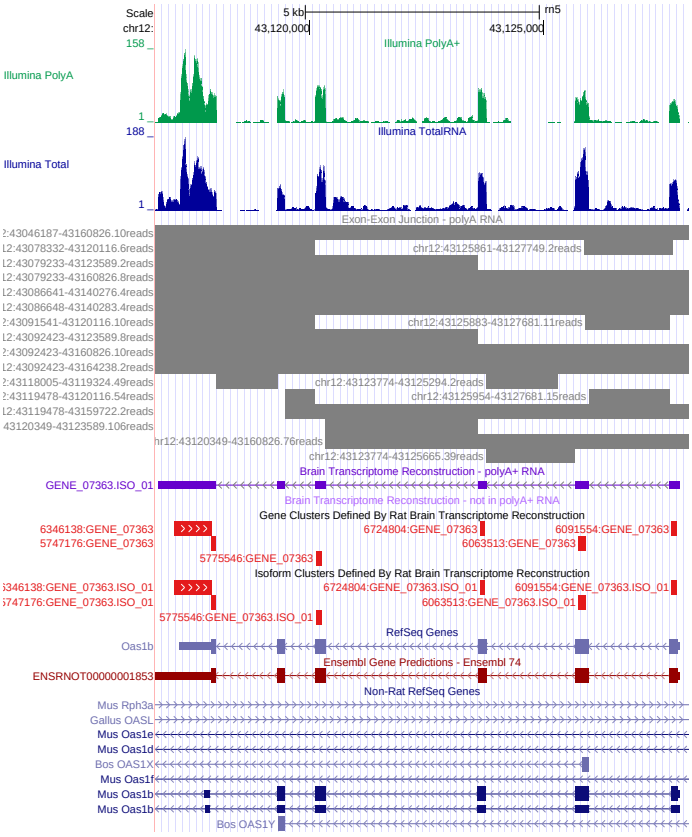


Figure 15.1: Reconstructed Transcripts of Oas1b. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

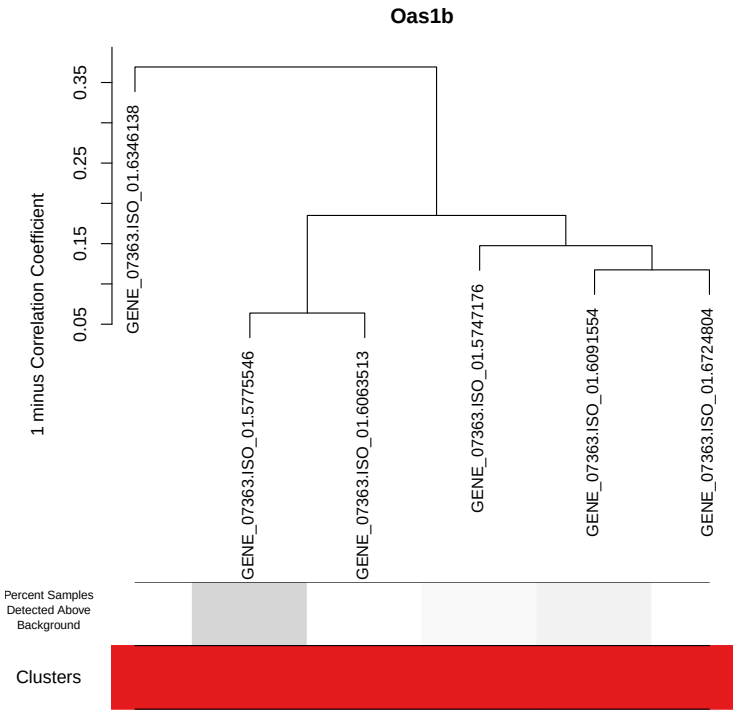


Figure 15.2: Correlation Structure - All Probe Sets Related To Oas1b. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

Only 1 of the 6 probe sets for this gene/isoform cluster was detected above background on the exon array in more than 5% of the RI samples, GENE_07363.ISO_01.5775546.

All of the probe sets for this gene cluster were highly correlated, but only one was detected above background in the RI panel.

15.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	1.16	<0.0001
HAD2 vs. LAD2	0.02	0.9411
P vs. NP	0.48	<0.0001
AA vs. ANA	-0.14	0.6345
sP vs. sNP	0.54	0.0208
UChB vs. UChA	0.26	0.0003
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	0.34	0.1365

Table 15.1: Association of Expression Levels with Alcohol Consumption - Oas1b (GENE_07363). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

15.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_07363	1.28	1.43
GENE_07363.ISO_01	1.28	1.43

Table 15.2: RNA-Seq - Oas1b. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

16 P2rx4 (purinergic receptor P2X, ligand-gated ion channel 4)

Association With Alcohol Consumption in:
gene-level WGCNA

Gene ID: GENE_07349

16.1 Concordance of Transcript Structure

Figure 16.1: Reconstructed Transcripts of P2rx4. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

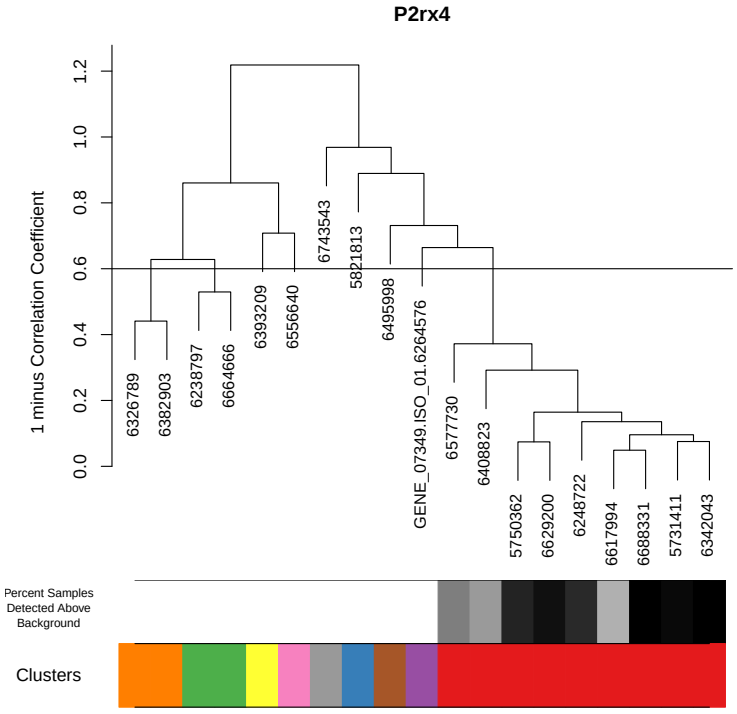


Figure 16.2: Correlation Structure - All Probe Sets Related To P2rx4. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

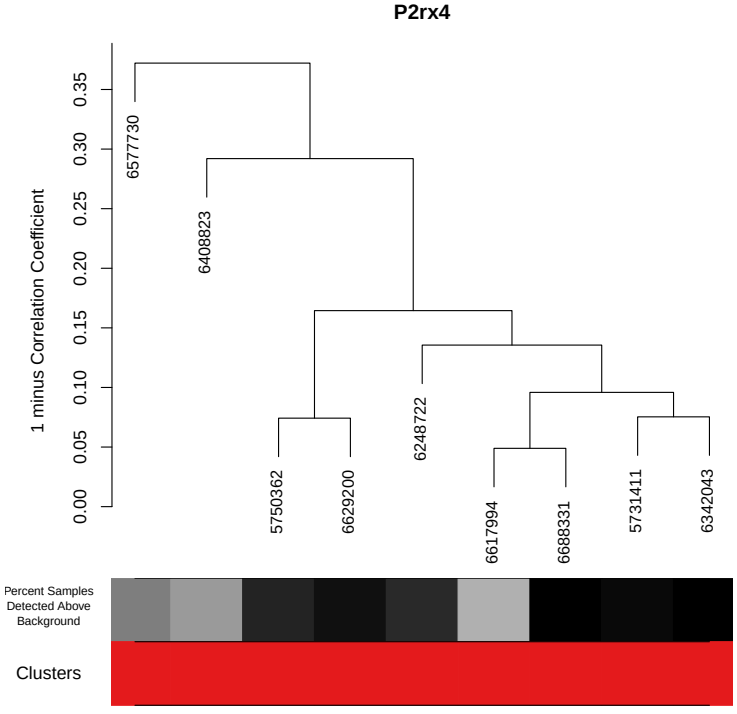


Figure 16.3: Correlation Structure - Probe Sets for P2rx4 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

All of the probe sets for this gene cluster that were expressed above background on the exon array in the RI panel were highly correlated. All but one of the probe sets that were below detection limits were from the 3' area of the reconstructed transcript that extended far beyond the end of the transcript in Ensembl and RefSeq. This extension of the 3' area in the reconstructed transcript was not supported by the read count profile. The other probe set that was not detected above background was from the 5' area of the transcript.

16.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	-0.17	0.0194
HAD2 vs. LAD2	-0.20	0.0106
P vs. NP	0.16	0.0291
AA vs. ANA	0.09	0.2503
sP vs. sNP	-0.07	0.3025
UChB vs. UChA	-0.31	<0.0001
meta-analysis	-	0.0215
HXB/BXH Panel		
Correlation*	-0.63	0.0024

Table 16.1: Association of Expression Levels with Alcohol Consumption - P2rx4 (GENE_07349). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

16.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_07349	2.16	1.94
GENE_07349.ISO_01	1.92	1.42
GENE_07349.ISO_02	0.12	0.34
GENE_07349.ISO_03	0.12	0.19

Table 16.2: RNA-Seq - P2rx4. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

17 Plcd4 (phospholipase C, delta 4)

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis

Gene ID: GENE_32863

17.1 Concordance of Transcript Structure

Figure 17.1: Reconstructed Transcripts of Plcd4. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

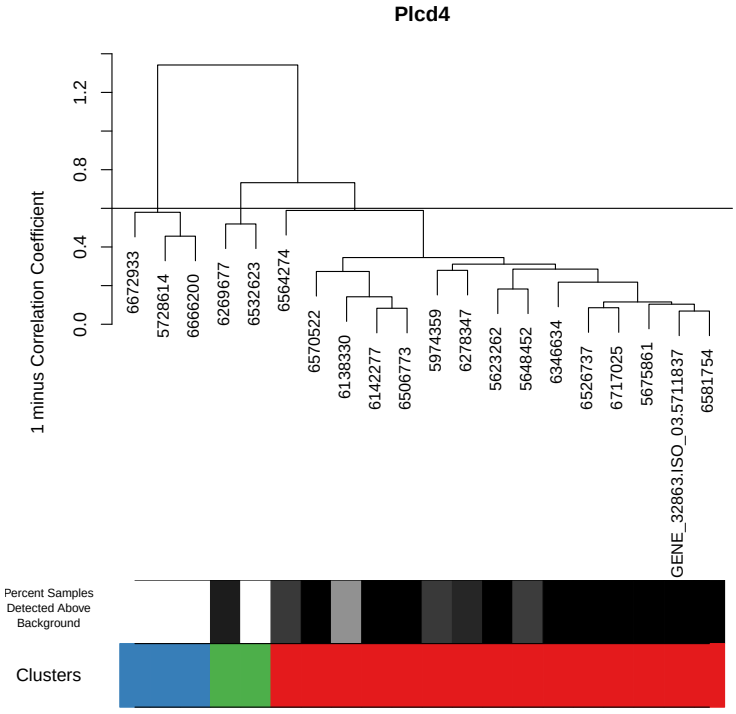


Figure 17.2: Correlation Structure - All Probe Sets Related To Plcd4. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

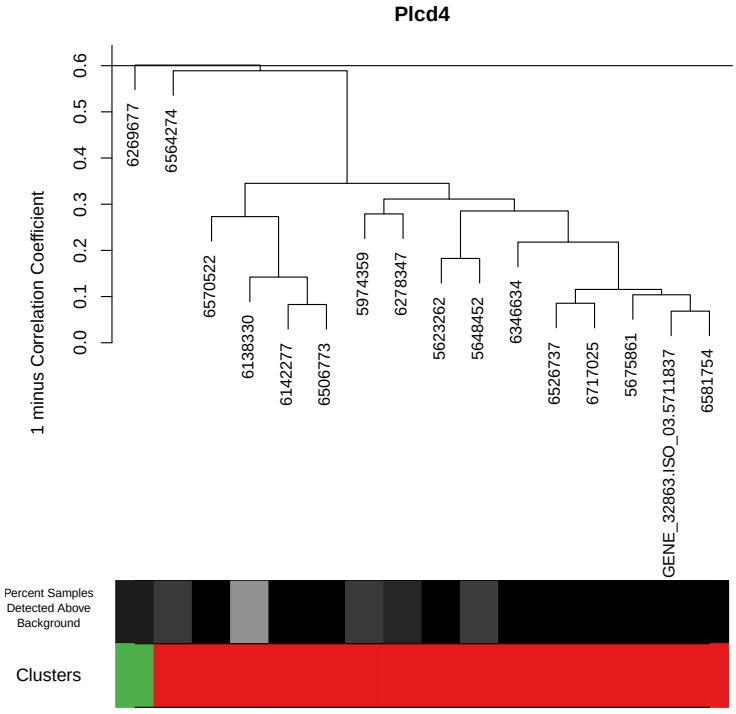


Figure 17.3: Correlation Structure - Probe Sets for Plcd4 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

All but one of the probe sets assigned to this gene cluster that were expressed above background in the RI panel reached the criterion of a distance measure (1 minus the correlation coefficient) less than 0.6. The one that did not barely missed this criterion.

17.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.07	0.1244
HAD2 vs. LAD2	0.05	0.4977
P vs. NP	0.15	0.0593
AA vs. ANA	-0.08	0.3016
sP vs. sNP	0.09	0.1830
UChB vs. UChA	0.14	<0.0001
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	0.10	0.6644

Table 17.1: Association of Expression Levels with Alcohol Consumption - Plcd4 (GENE_32863). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

17.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_32863	4.83	4.36
GENE_32863.ISO_01	3.23	2.68
GENE_32863.ISO_02	1.10	1.56
GENE_32863.ISO_03	0.48	0.08
GENE_32863.ISO_04	0.03	0.04

Table 17.2: RNA-Seq - Plcd4. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

18 RT1-CE10 (RT1-CE10 RT1 class I, locus CE10)

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis; isoform-level selected lines meta-analysis

Associated Isoform ID: GENE_18451.ISO_01
Gene ID: GENE_18451

18.1 Concordance of Transcript Structure

Figure 18.1: Reconstructed Transcripts of RT1-CE10. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

Comments

For this gene, only 2 probe sets were included in the gene cluster. GENE_18451.ISO_01.6021180 and GENE_18451.ISO_01.6392290 were not correlated (correlation coefficient=0.06). GENE_18451.ISO_01.6021180 was above detection limits on the exon array in 0% of the RI samples and GENE_18451.ISO_01.6392290 was above detection limits on the exon array in 15% of the RI samples.

Although the two probesets for this gene were not correlated in the RI panel and only one probe set was detected above background in the RI panel, this cluster was identified as associated with alcohol consumption in the selected lines and may represent a gene that is unique to that population.

18.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.64	<0.0001
HAD2 vs. LAD2	0.45	<0.0001
P vs. NP	0.34	<0.0001
AA vs. ANA	0.07	0.6011
sP vs. sNP	0.19	0.2908
UChB vs. UChA	-0.10	0.5869
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	-0.24	0.2850

Table 18.1: Association of Expression Levels with Alcohol Consumption - RT1-CE10 (GENE_18451.ISO_01). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

18.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_18451	0.47	0.87
GENE_18451.ISO_01	0.47	0.87

Table 18.2: RNA-Seq - RT1-CE10. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

19 Tgm2 (transglutaminase 2, C polypeptide)

Association With Alcohol Consumption in:
isoform-level selected lines meta-analysis

Associated Isoform ID: GENE_19407.ISO_01
Gene ID: GENE_19407

19.1 Concordance of Transcript Structure

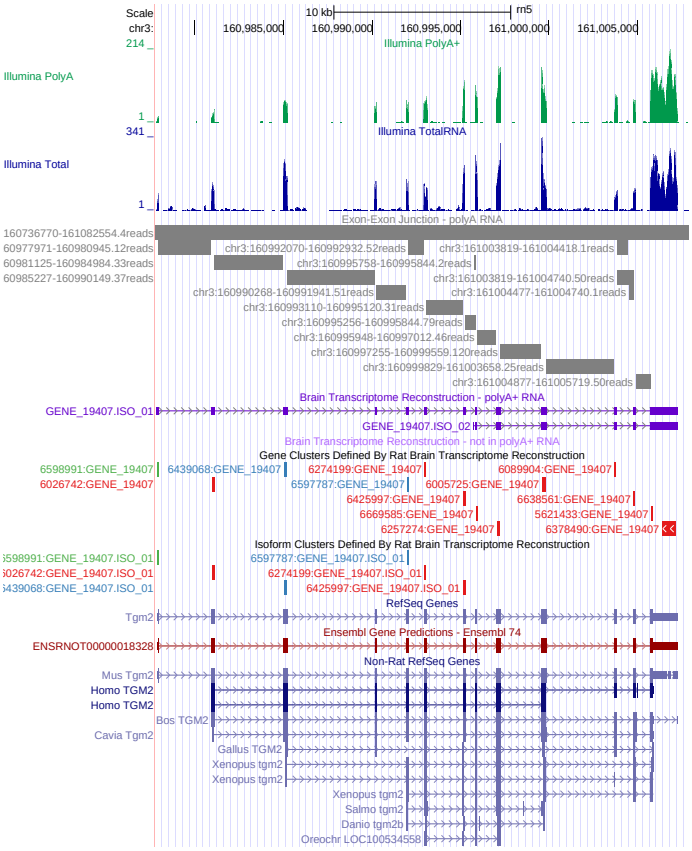


Figure 19.1: Reconstructed Transcripts of Tgm2. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

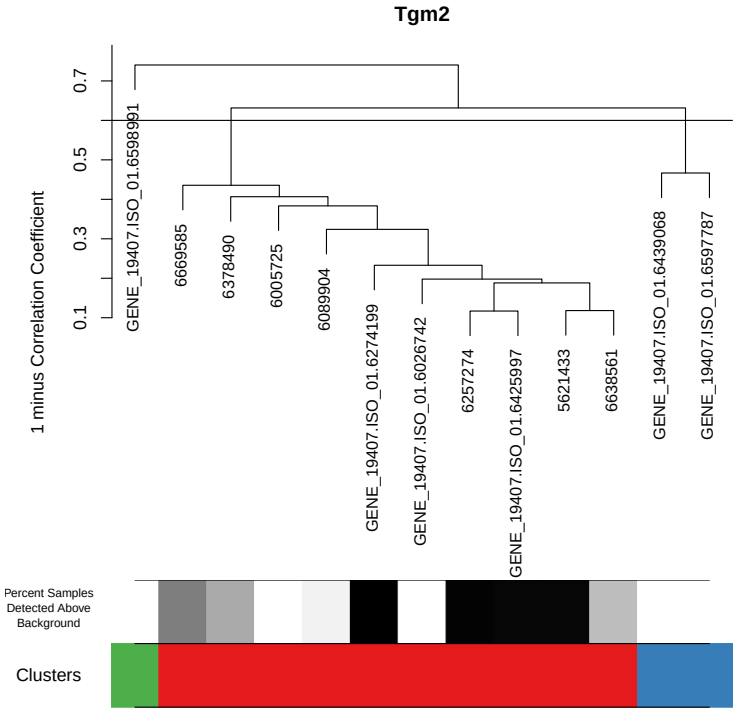


Figure 19.2: Correlation Structure - All Probe Sets Related To Tgm2. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

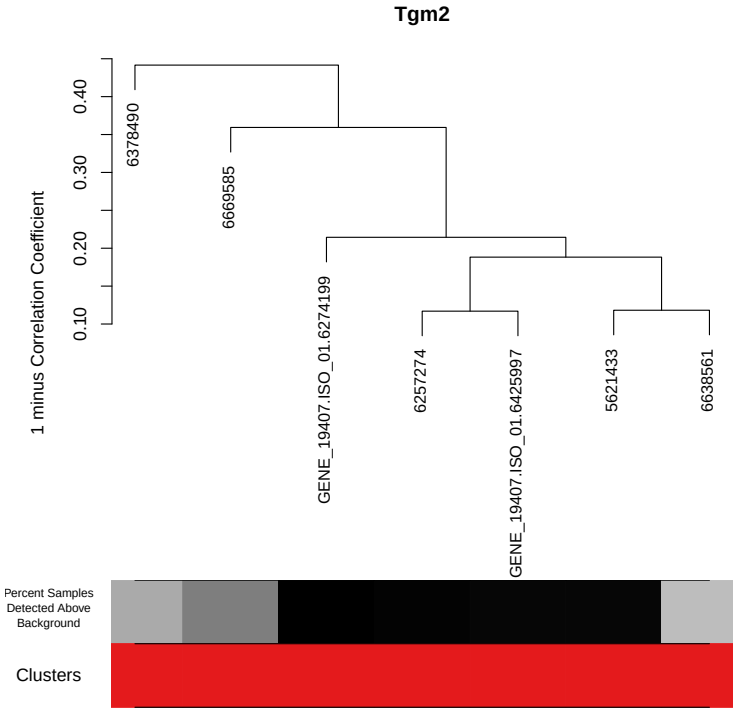


Figure 19.3: Correlation Structure - Probe Sets for Tgm2 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

Of this six probe sets assigned to this isoform cluster (only Isoform 1 was associated with drinking in the selected lines), only two were expressed above background in the RI panel. All the probe sets for the entire gene that were expressed above background were highly correlated in the RI panel.

19.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	-0.10	0.2477
HAD2 vs. LAD2	-0.17	0.1729
P vs. NP	0.36	0.0002
AA vs. ANA	-0.06	0.7861
sP vs. sNP	-0.00	0.9583
UChB vs. UChA	0.21	<0.0001
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	0.06	0.7907

Table 19.1: Association of Expression Levels with Alcohol Consumption - Tgm2 (GENE_19407.ISO_01). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

19.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_19407	3.17	2.07
GENE_19407.ISO_01	2.50	1.75
GENE_19407.ISO_02	0.68	0.32

Table 19.2: RNA-Seq - Tgm2. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

20 Tmem116 (transmembrane protein 116)

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis; gene-level WGCNA; isoform-level selected lines meta-analysis; isoform-level WGCNA

Gene ID: GENE_07359

20.1 Concordance of Transcript Structure

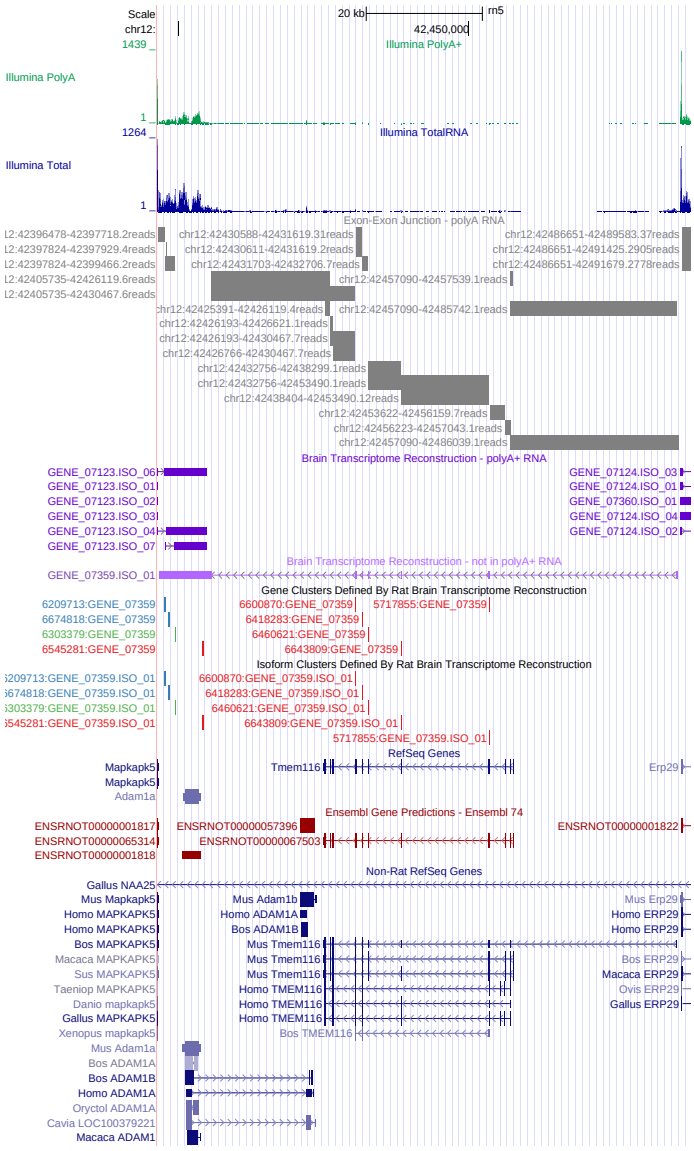


Figure 20.1: Reconstructed Transcripts of Tmem116. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

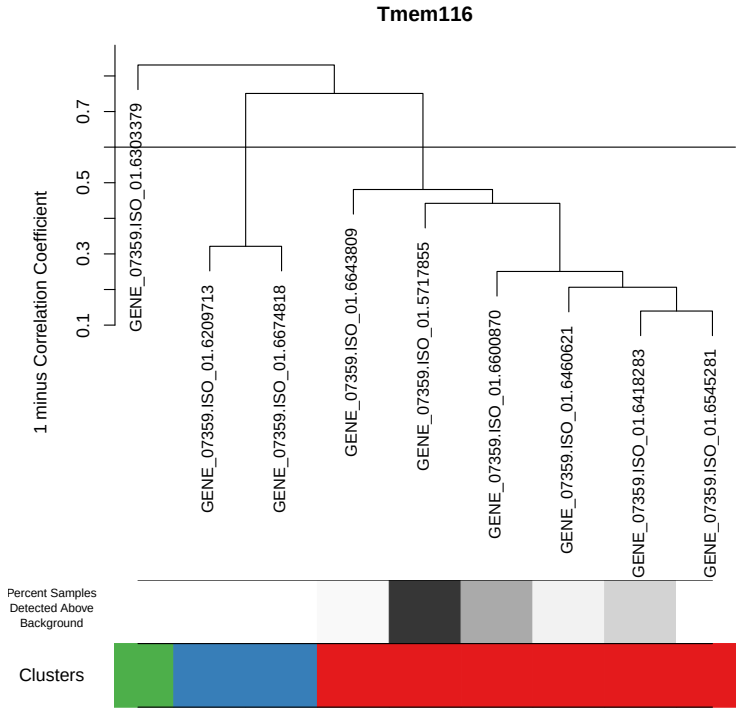


Figure 20.2: Correlation Structure - All Probe Sets Related To Tmem116. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

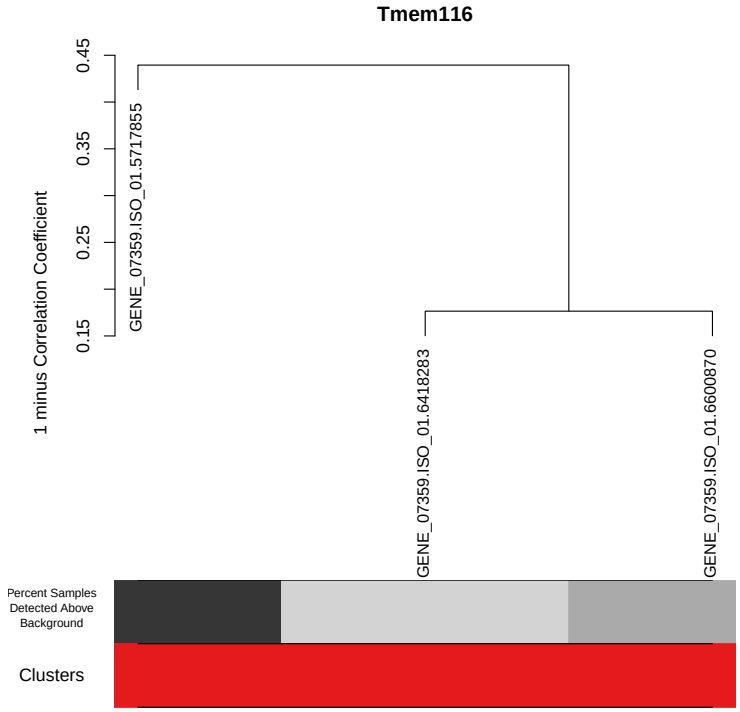


Figure 20.3: Correlation Structure - Probe Sets for Tmem116 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

Six of the nine probe sets assigned to this gene/isoform cluster were highly correlated, although only 3 of those were expressed above background in the RI panel. The 3 probe sets that were not correlated were from the 3' end of the gene that was different from the annotated 3' end of Tmem116.

20.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.30	<0.0001
HAD2 vs. LAD2	0.01	0.8400
P vs. NP	0.07	0.3637
AA vs. ANA	0.09	0.5234
sP vs. sNP	0.16	0.0052
UChB vs. UChA	0.08	0.0923
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	0.34	0.1326

Table 20.1: Association of Expression Levels with Alcohol Consumption - Tmem116 (GENE_07359). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

20.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_07359	1.56	1.22
GENE_07359.ISO_01	1.56	1.22

Table 20.2: RNA-Seq - Tmem116. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

21 Tmem14a (transmembrane protein 14A)

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis; isoform-level selected lines meta-analysis

Associated Isoform ID: GENE_32764.ISO_01
Gene ID: GENE_32764

21.1 Concordance of Transcript Structure

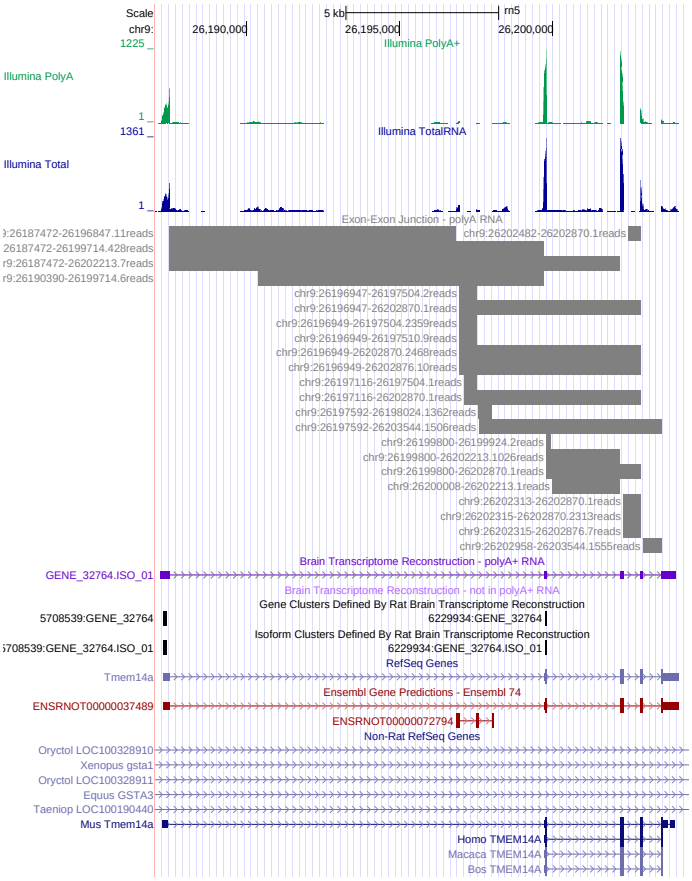


Figure 21.1: Reconstructed Transcripts of Tmem14a. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

Comments

For this gene, only 2 probe sets were included in the gene cluster. GENE_32764.ISO_01.5708539 and GENE_32764.ISO_01.6229934 were highly correlated (correlation coefficient=0.97). GENE_32764.ISO_01.5708539

was above detection limits on the exon array in 21% of the RI samples and GENE_32764.ISO_01.6229934 was above detection limits on the exon array in 100% of the RI samples.

Both probe sets for this gene were highly correlated and expressed above background on the exon array. However, they represented only 2 of the 5 exons for this gene. The other 3 exons were not probed by the array.

21.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	-1.15	<0.0001
HAD2 vs. LAD2	-1.21	<0.0001
P vs. NP	0.12	0.3048
AA vs. ANA	-0.17	0.1640
sP vs. sNP	0.33	0.2719
UChB vs. UChA	-1.09	<0.0001
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	0.03	0.8870

Table 21.1: Association of Expression Levels with Alcohol Consumption - Tmem14a (GENE_32764.ISO_01). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

21.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_32764	3.64	7.73
GENE_32764.ISO_01	3.64	7.73

Table 21.2: RNA-Seq - Tmem14a. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

22 Txnip (thioredoxin interacting protein)

Association With Alcohol Consumption in:
gene-level WGCNA

Gene ID: GENE_15796

22.1 Concordance of Transcript Structure

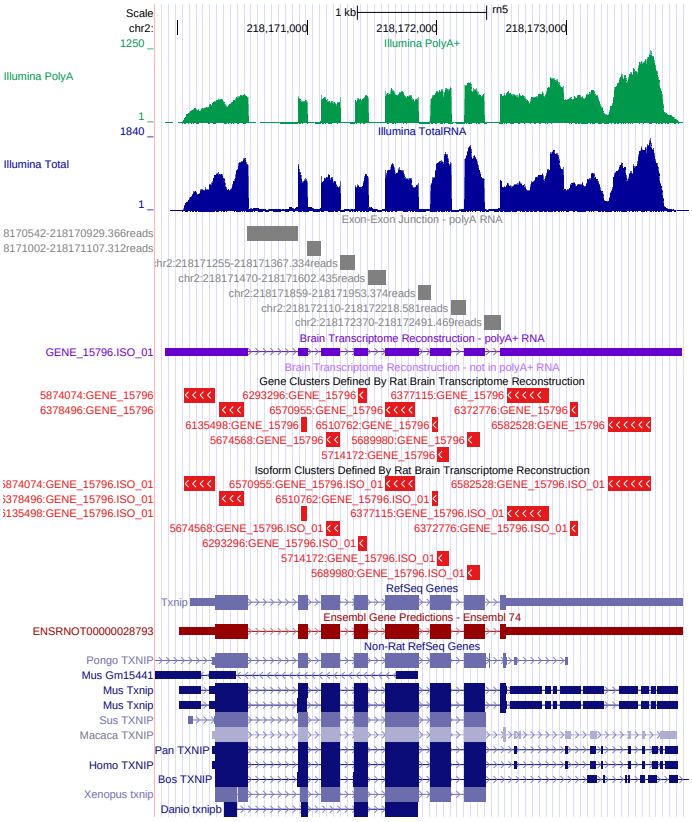


Figure 22.1: Reconstructed Transcripts of Txnip. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

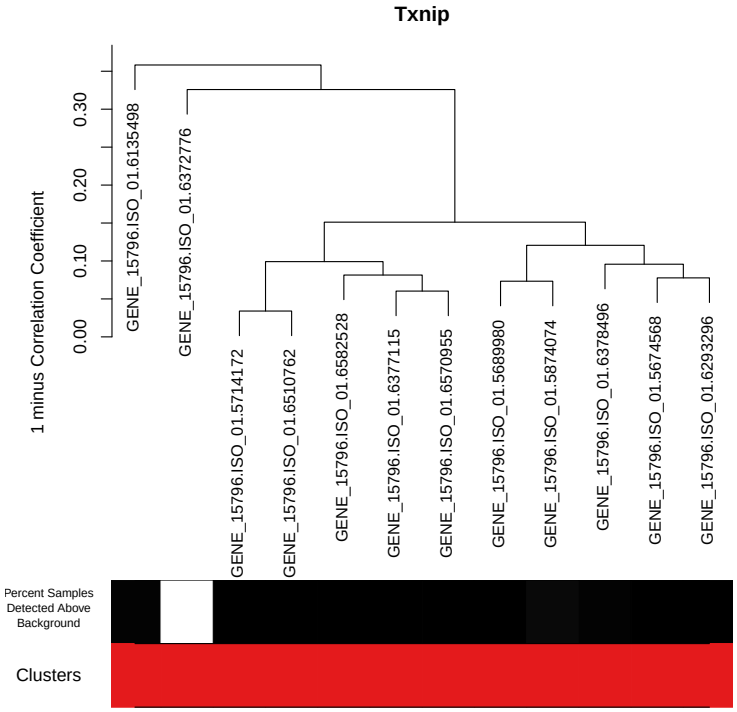


Figure 22.2: Correlation Structure - All Probe Sets Related To Txnip. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

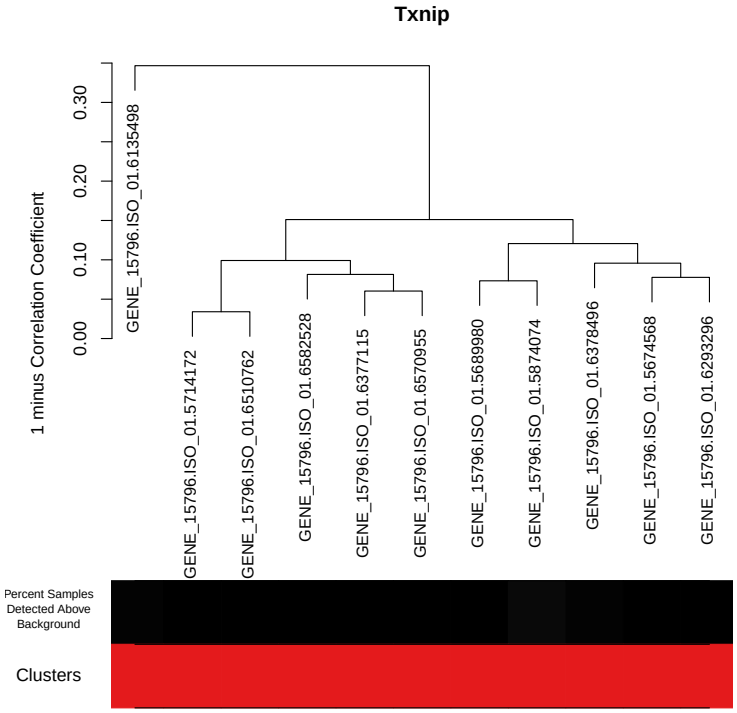


Figure 22.3: Correlation Structure - Probe Sets for Txnip Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

All of the probe sets for this gene cluster were highly correlated.

22.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	0.15	0.0281
HAD2 vs. LAD2	-0.18	0.0149
P vs. NP	0.19	0.0086
AA vs. ANA	-0.05	0.5277
sP vs. sNP	-0.01	0.8689
UChB vs. UChA	0.32	<0.0001
meta-analysis	-	0.0359
HXB/BXH Panel		
Correlation*	0.61	0.0033

Table 22.1: Association of Expression Levels with Alcohol Consumption - Txnip (GENE_15796). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

22.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_15796	12.16	15.98
GENE_15796.ISO_01	12.16	15.98

Table 22.2: RNA-Seq - Txnip. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

23 Vps52 (vacuolar protein sorting 52 homolog (S. cerevisiae))

Association With Alcohol Consumption in:
gene-level selected lines meta-analysis

Gene ID: GENE_18492

23.1 Concordance of Transcript Structure

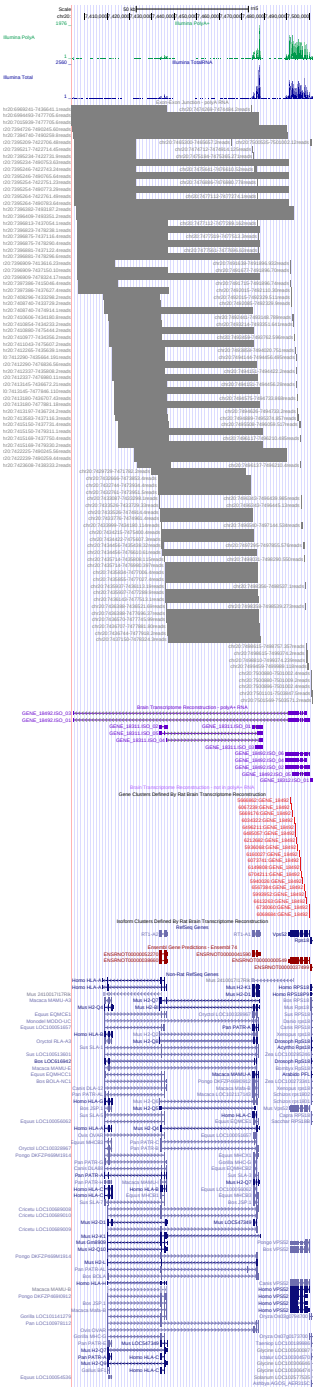


Figure 23.1: Reconstructed Transcripts of Vps52. The Illumina PolyA track (green) displays read coverage when including aligned reads from the polyA+-selected RNA from all six progenitor (SHR and BN-Lx) samples. The Illumina Total track (blue) displays read coverage when including aligned reads from the ribosomal RNA-depleted total RNA from all six progenitor (SHR and BN-Lx) samples. The Exon-Exon Junction - polyA+ RNA track (grey) indicates the regions of the genome that may be spliced out in the mature mRNA, i.e., exon-exon junction. The evidence for these splice site are reads in which a portion of the read aligns to the area upstream of the junction and another portion of the same read aligns to the area downstream of the junction. Only reads that matched perfectly to the region surrounding the junction were included in the graphic. The junctions were labeled by their location and the number of reads (including all samples) that cover the junction. The Brain Transcriptome Reconstruction tracks (purple) display the isoforms identified in the genome-guided de novo transcriptome reconstruction. The Brain Transcriptome Reconstruction - polyA+ track (dark purple) only contains isoforms from genes that were identified in the polyA+ RNA. The Brain Transcriptome Reconstruction - not in polyA+ RNA track (light purple) includes isoforms from genes that were ONLY identified in the ribosomal RNA-depleted total RNA. The Gene Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of the Affymetrix Exon Array probe sets that were included in designating gene clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the gene identifier generated during the transcriptome reconstruction. The probe sets are colored based on their correlation with each other across the RI as displayed in the following dendrograms. The Isoform Clusters Identified From Rat Brain Transcriptome Reconstruction track displays the location of probe sets that were included in defining isoform clusters. The probe set labels contains the probe set ID used by Affymetrix followed by the isoform identifier generated during the transcriptome reconstruction. The color-coding of probe sets is the same as the previous track. The remaining annotation tracks and the image were generated the UCSC Genome Browser (http://genome.ucsc.edu).

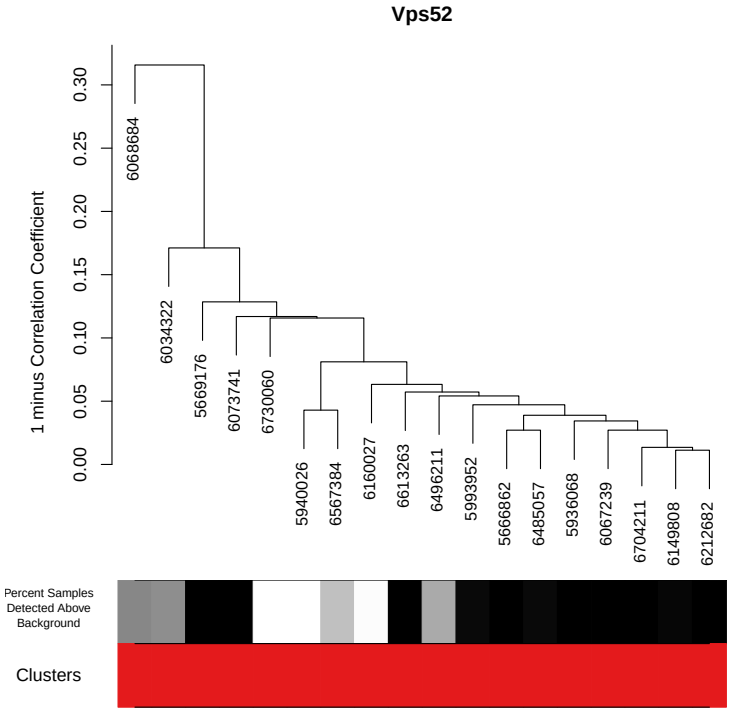


Figure 23.2: Correlation Structure - All Probe Sets Related To Vps52. The correlation structure among probe sets used to generate the gene-level expression estimates was determined using hierarchical clustering and 1 minus the Pearson correlation coefficient as the distance measure between probe sets and clusters. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The bar at the bottom of the graphic labeled Clusters indicates clusters of probe sets formed when cutting the dendrogram at a height of 0.6, i.e., correlation coefficient>0.40. Clusters are designated by color. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

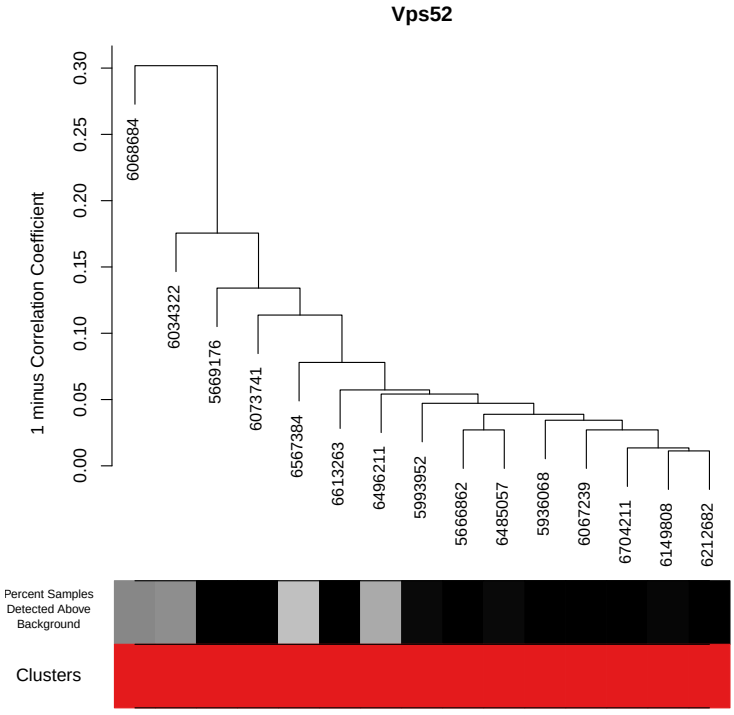


Figure 23.3: Correlation Structure - Probe Sets for Vps52 Detected Above Background. Using the same methods as the previous graphic, ONLY probe sets from the previous figure that were detected above background on the exon array were included in the hierarchical clustering. The bar underneath the dendrogram labeled Percent Samples Detected Above Background indicates using a grey scale the number of individual samples from the HXB/BXH RI panel whose expression for that probe set was detected above background (DABG p-value<0.0001). The scale ranges from solid black which indicates that the expression was above background in 100% of samples to white which indicates that the expression was above background in 0% of RI samples. The second bar at the bottom of the graphic labeled Clusters indicates what cluster each probe set belong to in the previous graphic that including ALL probe sets for the gene. The same colors are used in the UCSC genome browser version of the gene to label the positions of these probe sets.

Comments

All of the probe sets for this gene cluster were highly correlated.

23.2 Association with voluntary alcohol consumption

Comparison	Log2 Expression Difference	P-value
Selected Lines		
HAD1 vs. LAD1	-0.81	<0.0001
HAD2 vs. LAD2	-0.98	<0.0001
P vs. NP	-0.15	0.1637
AA vs. ANA	-0.50	0.0040
sP vs. sNP	0.16	0.1896
UChB vs. UChA	-0.27	<0.0001
meta-analysis	-	<0.0001
HXB/BXH Panel		
Correlation*	0.34	0.1269

Table 23.1: Association of Expression Levels with Alcohol Consumption - Vps52 (GENE_18492). The differences reported for the pairs of selected lines in the table are the differences in mean log base 2 transformed expression values between high and low drinking lines calculated using a mixed linear regression model with a random effect for breeding pair with an adaptive variance structure. In all cases, differences are calculated by subtracting the mean of the low drinking line from the high drinking line, i.e. a positive estimate indicates a positive association between RNA expression and drinking. For the HXB/BXH panel, the correlation coefficient for the panel is reported instead of a difference in expression. P-values were NOT adjusted for multiple comparisons.

23.3 Quantification of RNA-Seq expression

	SHR	BN-Lx
GENE_18492	11.86	7.48
GENE_18492.ISO_01	5.07	0.83
GENE_18492.ISO_02	2.16	3.21
GENE_18492.ISO_03	1.89	2.13
GENE_18492.ISO_04	1.88	1.30
GENE_18492.ISO_05	0.82	0.00
GENE_18492.ISO_06	0.04	0.00

Table 23.2: RNA-Seq - Vps52. Read Fragments Per Kilobase of Transcript Per Million Mapped Reads (FPKM) was reported for each isoform of the gene expressed in brain and for the gene (sum of expression across all isoforms) using read fragments generated from the ribosomal RNA-depleted total RNA. When only one isoform was detected, the isoform expression estimate and the gene expression estimate are equal. FPKM values were calculated using the cuffdiff function in the CuffLinks package.

Doc. S3. Functional Analysis

Phospholipase Cδ (Plcδ4) [1] is one of the top 10 differentially expressed genes across the high and low drinking rats of the selected lines, as well as being a member of the Indian Red 4 module. The delta isoforms of phospholipase C are dependent on calcium and membrane bound phosphatidylinositol (PIP_2) (not Gq) for their activity and produce two intracellular messengers, DAG and IP_3. The metabolism of membrane-bound PIP_2 by PLCδ4 is also a component of cell membrane dynamics which promote vesicular fusion and exocytosis in non-excitable tissues [2-4]. Phospholipase Cδ4 is highly expressed in brain [5] and certain of the PLCδ isoforms play an essential role in responding to calcium influx through the ARC channels and the activation of PLCδ acts as an "amplifier" [1] of cellular calcium signals through activation of IP_3-sensitive ryanodine receptors on the endoplasmic reticulum. Although recently, store-operated channel control of intracellular calcium has also been identified in "excitable" cells of the CNS (neurons), earlier thinking and experimentation localized the store-operated calcium channels (CRAC) and ARC channels to non-excitable cells such as astrocytes and microglia [6,7], in which cytosolic calcium transients can control a multitude of effectors ranging from kinases and proteases to gliotransmission. Another mechanism for controlling cytosolic calcium levels in non-excitable tissue as well as neurons [6] is through receptor gated, calcium permeable, channels. One such channel, highly permeable to calcium, is the P2X4 purinergic receptor [8]. The P2x4 receptor, which is activated by ATP, is particularly important in activation of microglia and astrocytes in brain and controlling the secretion of inflammatory cytokines, as well as GDNF, and glutamate and GABA from these cells [9]. The expression of the P2x4 receptor was negatively correlated with alcohol consumption, thus linking lower alcohol consumption to higher expression of an element leading to higher cytosolic calcium levels. Cytosolic calcium levels also impact the function of another member of the Indian Red 4 module important for neuronal, astrocyte and microglial function and glial proliferation. Ift81 is a critical

component of cilium formation and cilium elongation in astrocytes, neurons and other mammalian cell types [10]. The “primary cilium” in mammalian cells is usually a non-mobile element surrounded by the plasma membrane (cilium membrane). The cilium is positioned to sense physical and biochemical extracellular signals [11,12], to couple these signals to proliferation in cells capable of proliferation (e.g., glia), and to participate in endocytotic and exocytotic events [12]. One of the potent mechanisms controlling the length of a cilium (inhibiting growth) is cytosolic calcium released from the endoplasmic reticulum or mitochondria or entering through plasma membrane channels impacting the function of Ift81 [13]. In neurons of the rat and mouse CNS, cilia have concentrated levels of somatostatin, 5-HT_6 and dopamine D1 receptors [14-16] and cilia have been referred to as an extrasynaptic signaling device [17]. The cilium in the CNS may also act as a nutrient sensor [18] and ciliary malfunction has been shown to control consumatory behavior [19].

The regulation of the localization of the P2x4 receptors to the plasma membrane occurs through rapid delivery and endocytosis of these receptors, a process shown to occur in neurons and glia [20]. In the process of endocytosis, the P2x4 receptors are incorporated into late endosomes for transport to lysosomes or the Golgi for recycling [20]. The fusion of the endosomes carrying P2x4 receptors with the Golgi apparatus and/or lysosomes has recently been shown to be mediated by the Golgi associated retrograde protein (GARP) complex [21]. A critical component of this complex is Vps52, which is another of the gene products identified as one of the top 10 gene products differentially expressed across the six selectively bred lines of rats.

Further “functional” connections within the Indian Red 4 module are evident between elements linked to control of cytosolic and mitochondrial calcium and mitochondrial enzymes and scaffolding molecules derived from *Maats1* and *Coq5* genes. Coenzyme Q5 (Coq5) is an enzyme component in the synthesis of reduced ubiquinone (QH_2) [22] which is critical as a catalytic component of mitochondrial complexes 1 (NADH-ubiquinone oxidoreductase).

Complex 1, by catalyzing the transfer of electrons from NADH to ubiquinone, generates a protein gradient which is utilized to synthesize ATP from ADP [23]. Complex 1 is composed of multiple protein subunits, many of which can be phosphorylated by PKA present in the mitochondria [24]. cAMP for activation of PKA in the mitochondria is generated by a soluble dicarbonate activatable adenylyl cyclase in the mitochondrial matrix [25]. The phosphorylation of Complex 1 subunits (e.g., NDUFS4 protein) promotes Complex 1 activity, increasing the production of ATP and simultaneously reducing the mitochondrial production of ROS [22]. The PKA in mitochondria is similar to the PKA associated with the plasma membrane of neurons and glial cells in brain. The PKA holoenzyme is immobilized and kept in its inactive form by its association of its regulatory subunit and the association of the regulatory subunit with scaffolding proteins. cAMP interacts with the regulatory subunit and releases the catalytic subunit to perform its phosphorylation reactions. In mitochondria, the Maats1 (AAT-1) protein (Indian Red 4) interacts with AKAP to form a scaffold for the PKA regulatory subunit [26,27].

Recently Han and Ayer [28] demonstrated that adenosine can regulate the transcription of *Txnip* (Indian Red 4 module member). Txnip, thioredoxin interacting protein, binds to thioredoxin and inhibits its disulfide reductase activity, and in this way negatively regulates thioredoxin's actions on molecules affected by ROS and levels of ROS *per se* [29]. The regulation of thioredoxin is only one of the actions of Txnip. Txnip also plays important roles in glucose uptake into glia and modulates the production and release of inflammatory cytokines [30].

Txnip's role in cellular glucose metabolism is through an anapleurotic mechanism which involves the transcription factor Mondo A. Activation of MondoA by glucose leads to the increased transcription of Txnip which in turn, in a feedback fashion, reduces glucose uptake by the glucose transporter, Glut1, [31] and reduces cellular energy production. Glutamine, which can be produced from glutamate in astrocytes and other glial cells by glutamine synthetase localized to glia (astrocytes) [32] is, on the other hand, a blocker of the transcriptional activity of

MondoA and can reduce cellular levels of Txnip, and increase glucose uptake. This process defines a metabolic "check point" through a Txnip-modulated balance between glutamine and glucose availability and utilization of these substrates for energy production in astrocytes.

Txnip transcription is also controlled by the PERK and IRE1 pathways which are responsive to ROS and endoplasmic reticulum stress (unfolded protein response [33]). *Txnip* transcription is induced by ER stress and ROS production [34] but ROS also promotes the dissociation of Txnip from thioredoxin, thus promoting Txnip interaction with other proteins, while releasing thioredoxin from inhibition. One such Txnip interaction is with the proteins of the NLRP3 inflammasome (which promotes the activation of the inflammasome (i.e., Caspase 1 activation, and IL-1β and IL-18 production) [35]. Although the components of the NLRP3 inflammasome have been noted in neurons [36], much of the innate immunity response in the CNS is mediated by the function of microglia and astrocytes [37]. As already mentioned, hyperglycemia can drive the increased transcription of *Txnip*, and lead to activation of the NLRP3 inflammasome, which has been referred to a "sensor for metabolic danger" [38]. Another avenue for activation of the inflammasome is extracellular ATP acting through P2X4 receptors [39]. This ATP can arise from autocrine release from microglia and astrocytes [40-42] or can be released with activated neurotransmission from neurons, particularly GABA and glutamatergic neurons [43,44]. The ATP can interact with purginergic receptors on glia and activation of P2X7 and P2X4 receptors on microglia and astrocytes can in turn lead to inflammasome activation (P2X7 by reducing intracellular K^+ levels and P2X4 by allowing Ca^{++} entry) [35,45]. As already mentioned, the Indian Red 4 module also contains a regulator of glial cytosolic calcium levels (*Plcδ4)*. ATP is also a substrate for oligoadenylate synthase (Oas1a), which generates 2'-5'-oligoadenylates that are mandatory activators of RNAse L [46]. Oas1a activity is inhibited by Oas1b, the product of one of our candidate transcripts, and recent

evidence indicates that RNAse L activation is an important component of the innate immune response [47-49].

The ATP and glutamate used for gliotransmission are generated from de novo synthesis of these compounds in glia. Glutamate synthesis is initiated by the enzyme pyruvate carboxylase which in brain, is primarily localized in astrocytes [50,51] and the generation of ATP is through the TCA cycle and mitochondrial oxidative phosphorylation. Thus, the TCA cycle and mitochondrial oxidative phosphorylation generate not only the ATP to satisfy cellular energy needs, but also the ATP to be released as a gliotransmitter. Mitochondrial oxidative phosphorylation, in addition to producing ATP, is a major control of cellular NAD levels and cellular redox state. NADH produced by glycolysis and the TCA cycle is converted to NAD in concert with the generation of ATP, and the NAD/NADH ratio in cells constitutes a control for a multitude of dehydrogenases as well as being a factor in the control of ROS.

When we examined the list of the top 10 genes which arose from the meta-analysis of the six bidirectionally selected lines (six high drinking lines-six low drinking lines) we found that two of the gene products already discussed above (Oas1b and Plcδ4) were part of this list and several of the other gene products were functionally related to the members of the Indian Red4 module.

The Ctss gene product is the lysosomal cysteine protease S (Cathepsin S). Cathepsin S has been shown to be released from activated microglia and participate in the cleavage of the transmembrane chemokine fractalkine which is expressed on neuronal membranes [52]. The release of Cathepsin S from microglia is mediated in part by the purinergic P2X7 receptors [53] and the release of Cathepsin S and its cleavage of fractalkine on neurons generates a peptide ligand for the CX3CR1 receptor expressed predominantly on microglia [54]. Thus, communication between glia $\rightarrow$ neuron $\rightarrow$ glia is established by this system [54]. Under normal conditions the Cathepsin S, fractalkine, CX3CR1 system can be neuroprotective [54,55] and in

fact participates positively in adult neurogenesis [56]. There is some evidence that the CX3CR1 receptor may also be expressed on some neurons in the CNS and participate in processes related to regulation of neurotransmission and synaptic plasticity [54,57,58]. However, activation of microglia in the process of inflammation leads to neuronal damage via the Cathepsin S, fractalkine, CX3CR1 system [54,59].

Data derived from the meta-analysis of differentially expressed genes across the six lines of rats selectively bred for differences in alcohol consumption, also identified Tmem14a as a candidate gene whose expression was associated with the drinking phenotype. This was the second example of a member of the TMEM family being identified through our analysis. Tmem116, a member of both the Indian Red 4 and Aquamarine 1 modules, lacked much mention in the literature, and thus its functional significance within our group of identified gene products could not be ascertained. Tmem14a has, however, recently been linked to a role in maintaining the membrane potential in mitochondria [60]. The maintenance of such potentials is a critical component for continued function of oxidative phosphorylation in mitochondria and prevention of the permeabilization of the mitochondrial outer membrane and suppression of pro-apoptotic protein release from mitochondria [60]. The activity of Tmem14a in stabilizing mitochondrial function has been postulated to be countered by ROS activation of the Bcl-2 family proteins and thus Tmem14a has been referred to as a suppressor of Bax-mediated apoptosis [60]. The other gene products which were found through the selected lines meta-analysis to be significantly differentially expressed and associated with differences in drinking behavior between the members of the six high/low drinking lines were Fbxo 45 (a ubiquitin ligase), Fibrulin 1, Tgm2 (transglutaminase 2), Nxph1 (neurexin1) and the Cd74 antigen. Interestingly, this group of candidates are located at the synapse between neurons or are part of the extracellular matrix important in cell-cell contact or cell-cell functional interactions. *Fbxo45* is selectively expressed in the central nervous system, and is an interaction partner of NOS in some tissues [61]. iNOS is a glial form of nitric oxide synthetase, and production of nitric oxide

by this isoform in astrocytes has been shown to influence glutamatergic neuro-transmission [62]. Fbxo45 protein was found to be colocalized with VGlut1 and PSD95, also indicating involvement with glutamatergic transmission [63]. One of the functions of Fbxo45 has been shown to be important in the negative control of glutamatergic vesicle fusion with the synaptic membrane in a depolarization independent manner. Thus, a knock-down of *Fbxo45* in brain leads to an increase in mEPSC, and over-expression of *Fbxo45* diminishes mEPSCs [63]. Fbxo45 interacts with the Unc13-1 vesicular protein and leads to the ubiquitination and degradation of Unc-1 [63]. Unc-1 has recently been shown to be a central player in the "superpriming" of synaptic vesicles, enhancing sensitivity to calcium and making them "competent" for release of neurotransmitter [64].

Fibrulin 1 is a small extracellular matrix protein that plays a role in adhesion, migration, proliferation and differentiation in peripheral tissues [65]. Fibrulins, and particularly fibrulin 1, have a particular affinity for fibronectin-based fibrils, and bind to Type III repeats of fibronectin [66,67]. Our finding of *Fibrulin 1* expression in brain can indicate that fibrulin/fibronectin interactions take place in the CNS. The role of fibronectin in the CNS impacts significantly on microglial activation and function. Microglia must migrate through a dense extracellular matrix for their role in inflammatory processes and matrix remodeling, and microglia accomplish their migratory role by being able to degrade fibronectin [68]. Fibronectin also functions to stimulate the expression of P2X4 receptors on the surface of microglia during microglia activation. Fibronectin exposure of glial cells in culture leads to increased transcription of P2X4 mRNA and increased delivery to the cell surface [69,70]. Another of our candidates, transglutaminase 2, is also located in the extracellular space (ECS) as well as in various intracellular compartments. It is a calcium-dependent enzyme catalyzing deamidation, transamidation and protein cross-linking [71]. Transglutaminase also possesses GTPase activity which allows it to act as a signal transduction element for several cell surface receptors and couple the receptors to activation of PLCδ and increase in intracellular IP3 levels. In the ECS transglutaminase enables the

generation of stable covalent protein heterocomplexes, particulary between fibronectin and other ECS proteins [71]. Transglutaminase is expressed in astrocytes, microglia and neurons and one of its functions in brain is the control of mitochondrial integrity and energy production. The deletion of *Tgm2* causes dysregulation of respiratory complex 1 and 2 and reduction of ATP production, but increased mitrochondrial membrane potential and reduction in Bax recruitment and thus, reduction in apoptosis [72].

Nxph1 codes for neurexophylin 1α [73]. Neuroxophylin is expressed primarily in brain and its expression is limited to what have been described as inhibitory interneurons [73]. Neuroxophylin is rapidly N-glycosylated in many cell types, and proteolytic processing of the N-glycosylated pre-protein generates the mature form, which is neurexin 1α. This reaction only takes place in neurons [74]. Neurexin 1α is released into the extracellular space and partners with a number of synaptic proteins, and such interactions generate specificity in the actions of neurexin as an organizer of particular types of synapses. For instance, calsyntenin-3 interacts with neurexin-1α to generate the development of GABAergic and glutamatergic synapses even in adult brain tissue [74]. Just like neurexophylin, the expression of calsyntenin-3 seems to be primarily confined to interneurons [74]. The differentiation induced by neurexin-1α and its binding partners in postsynaptic membranes which evolve into GABA synapses includes the recruitment and clustering of gephyrin and the gamma 2 subunit of the GABA-A receptor [75]. It also should be noted that the synaptic development promoted by neurexin-1α and its partnering proteins are critical in the process of new network connections between glutamatergic and GABAergic axons as part of the process of neural plasticity [76].

The last of the annotated candidate gene products identified through our analysis is Cd74. Cd74 is a membrane protein also known as MHC class II invariant chain protein in cells of the immune system [77]. Cd74 is expressed primarily in brain and it can interact with the chemokine receptor CXCR4 to generate a functional complex that can bind the chemokine-like

inflammatory factor, MIF (macrophage migration inhibitory factor) [78]. MIF is expressed in astrocytes and microglia and can function intracellularly as a thioprotein-oxidoreductase similar to thioredoxin or can be released and function as autocrine innate immunity regulator [79]. Through interaction with the Cd74-CXCR2/4 complex, MIF counteracts the effects of glucocorticoids and can increase TNF-α, IL-8 and IL-1β [79]. One of the actions of MIF on glial cells may be an activation of glucose metabolism. Although not yet demonstrated in brain, MIF in peripheral tissues can elevate the function of phosphofructokinase and AMPK effects increasing glucose uptake and increasing lactate production. Production and release of lactate from astrocytes can fuel neuronal energy production [79].

References

1. Thompson JL, Shuttleworth TJ (2011) Orai channel-dependent activation of phospholipase C-delta: a novel mechanism for the effects of calcium entry on calcium oscillations. The Journal of physiology 589: 5057-5069.
2. Ochocka AM, Pawelczyk T (2003) Isozymes delta of phosphoinositide-specific phospholipase C and their role in signal transduction in the cell. Acta biochimica Polonica 50: 1097-1110.
3. Ammar MR, Kassas N, Chasserot-Golaz S, Bader MF, Vitale N (2013) Lipids in Regulated Exocytosis: What are They Doing? Frontiers in endocrinology 4: 125.
4. Vitale N (2010) Synthesis of fusogenic lipids through activation of phospholipase D1 by GTPases and the kinase RSK2 is required for calcium-regulated exocytosis in neuroendocrine cells. Biochemical Society transactions 38: 167-171.
5. Pawelczyk T (1999) Isozymes delta of phosphoinositide-specific phospholipase C. Acta biochimica Polonica 46: 91-98.
6. Brawek B, Garaschuk O (2013) Microglial calcium signaling in the adult, aged and diseased brain. Cell calcium 53: 159-169.
7. Ohana L, Newell EW, Stanley EF, Schlichter LC (2009) The Ca2+ release-activated Ca2+ current (I(CRAC)) mediates store-operated Ca2+ entry in rat microglia. Channels 3: 129-139.
8. Coddou C, Yan Z, Obsil T, Huidobro-Toro JP, Stojilkovic SS (2011) Activation and regulation of purinergic P2X receptor channels. Pharmacological reviews 63: 641-683.
9. Del Puerto A, Wandosell F, Garrido JJ (2013) Neuronal and glial purinergic receptors functions in neuron development and brain disease. Frontiers in cellular neuroscience 7: 197.
10. Bhogaraju S, Cajanek L, Fort C, Blisnick T, Weber K, et al. (2013) Molecular basis of tubulin transport within the cilium by IFT74 and IFT81. Science 341: 1009-1012.
11. Ishikawa H, Marshall WF (2011) Ciliogenesis: building the cell's antenna. Nature reviews Molecular cell biology 12: 222-234.
12. Moser JJ, Fritzler MJ, Rattner JB (2009) Primary ciliogenesis defects are associated with human astrocytoma/glioblastoma cells. BMC cancer 9: 448.

13. Besschetnova TY, Kolpakova-Hart E, Guan Y, Zhou J, Olsen BR, et al. (2010) Identification of signaling pathways regulating primary cilium length and flow-mediated adaptation. Current biology : CB 20: 182-187.
14. Handel M, Schulz S, Stanarius A, Schreff M, Erdtmann-Vourliotis M, et al. (1999) Selective targeting of somatostatin receptor 3 to neuronal cilia. Neuroscience 89: 909-926.
15. Brailov I, Bancila M, Brisorgueil MJ, Miquel MC, Hamon M, et al. (2000) Localization of 5-HT(6) receptors at the plasma membrane of neuronal cilia in the rat brain. Brain research 872: 271-275.
16. Marley A, von Zastrow M (2010) DISC1 regulates primary cilia that display specific dopamine receptors. PloS one 5: e10902.
17. Whitfield JF (2004) The neuronal primary cilium--an extrasynaptic signaling device. Cellular signalling 16: 763-767.
18. Yoshimura K, Kawate T, Takeda S (2011) Signaling through the primary cilium affects glial cell survival under a stressed environment. Glia 59: 333-344.
19. Berbari NF, Pasek RC, Malarkey EB, Yazdi SM, McNair AD, et al. (2013) Leptin resistance is a secondary consequence of the obesity in ciliopathy mutant mice. Proceedings of the National Academy of Sciences of the United States of America 110: 7796-7801.
20. Robinson LE, Murrell-Lagnado RD (2013) The trafficking and targeting of P2X receptors. Frontiers in cellular neuroscience 7: 233.
21. Bonifacino JS, Hierro A (2011) Transport according to GARP: receiving retrograde cargo at the trans-Golgi network. Trends in cell biology 21: 159-167.
22. Tran UC, Clarke CF (2007) Endogenous synthesis of coenzyme Q in eukaryotes. Mitochondrion 7 Suppl: S62-71.
23. Papa S, Rasmo DD, Technikova-Dobrova Z, Panelli D, Signorile A, et al. (2012) Respiratory chain complex I, a main regulatory target of the cAMP/PKA pathway is defective in different human diseases. FEBS letters 586: 568-577.
24. Sardanelli AM, Signorile A, Nuzzi R, Rasmo DD, Technikova-Dobrova Z, et al. (2006) Occurrence of A-kinase anchor protein and associated cAMP-dependent protein kinase in the inner compartment of mammalian mitochondria. FEBS letters 580: 5690-5696.
25. Yuzefovych L, Wilson G, Rachek I (2010) Different effects of oleate vs. palmitate on mitochondrial function, apoptosis, and insulin signaling in L6 skeletal muscle cells: role of oxidative stress. Am J Physiol Endochondrial Metab 299: E1096-E10105.
26. Ilouz R, Bubis J, Wu J, Yim YY, Deal MS, et al. (2012) Localization and quaternary structure of the PKA RIbeta holoenzyme. Proceedings of the National Academy of Sciences of the United States of America 109: 12443-12448.
27. Yukitake H, Furusawa M, Taira T, Iguchi-Ariga SM, Ariga H (2002) AAT-1, a novel testis-specific AMY-1-binding protein, forms a quaternary complex with AMY-1, A-kinase anchor protein 84, and a regulatory subunit of cAMP-dependent protein kinase and is phosphorylated by its kinase. The Journal of biological chemistry 277: 45480-45492.
28. Han KS, Ayer DE (2013) MondoA senses adenine nucleotides: transcriptional induction of thioredoxin-interacting protein. The Biochemical journal 453: 209-218.
29. Spindel ON, World C, Berk BC (2012) Thioredoxin interacting protein: redox dependent and independent regulatory mechanisms. Antioxidants & redox signaling 16: 587-596.
30. Watanabe R, Nakamura H, Masutani H, Yodoi J (2010) Anti-oxidative, anti-cancer and anti-inflammatory actions by thioredoxin 1 and thioredoxin-binding protein-2. Pharmacology & therapeutics 127: 261-270.
31. Kaadige MR, Looper RE, Kamalanaadhan S, Ayer DE (2009) Glutamine-dependent anapleurosis dictates glucose uptake and cell growth by regulating MondoA transcriptional activity. Proceedings of the National Academy of Sciences of the United States of America 106: 14878-14883.

32. Norenberg MD, Martinez-Hernandez A (1979) Fine structural localization of glutamine synthetase in astrocytes of rat brain. Brain research 161: 303-310.
33. Zhang K, Kaufman RJ (2008) From endoplasmic-reticulum stress to the inflammatory response. Nature 454: 455-462.
34. Oslowski CM, Hara T, O'Sullivan-Murphy B, Kanekura K, Lu S, et al. (2012) Thioredoxin-interacting protein mediates ER stress-induced beta cell death through initiation of the inflammasome. Cell metabolism 16: 265-273.
35. Tschopp J, Schroder K (2010) NLRP3 inflammasome activation: The convergence of multiple signalling pathways on ROS production? Nature reviews Immunology 10: 210-215.
36. Zou J, Crews FT (2012) Inflammasome-IL-1beta Signaling Mediates Ethanol Inhibition of Hippocampal Neurogenesis. Frontiers in neuroscience 6: 77.
37. Ransohoff RM, Brown MA (2012) Innate immunity in the central nervous system. The Journal of clinical investigation 122: 1164-1171.
38. Schroder K, Zhou R, Tschopp J (2010) The NLRP3 inflammasome: a sensor for metabolic danger? Science 327: 296-300.
39. Chen K, Zhang J, Zhang W, Zhang J, Yang J, et al. (2013) ATP-P2X4 signaling mediates NLRP3 inflammasome activation: a novel pathway of diabetic nephropathy. The international journal of biochemistry & cell biology 45: 932-943.
40. Domercq M, Vazquez-Villoldo N, Matute C (2013) Neurotransmitter signaling in the pathophysiology of microglia. Frontiers in cellular neuroscience 7: 49.
41. Sershen H (2012) Astrocyte origin of activity-dependent release of ATP and glutamate in hippocampal slices: real-time measurement utilizing microelectrode biosensors. British journal of pharmacology 167: 1000-1002.
42. Heinrich A, Ando RD, Turi G, Rozsa B, Sperlagh B (2012) K+ depolarization evokes ATP, adenosine and glutamate release from glia in rat hippocampus: a microelectrode biosensor study. British journal of pharmacology 167: 1003-1020.
43. Pankratov Y, Lalo U, Verkhratsky A, North RA (2006) Vesicular release of ATP at central synapses. Pflugers Archiv : European journal of physiology 452: 589-597.
44. Larsson M, Sawada K, Morland C, Hiasa M, Ormel L, et al. (2012) Functional and anatomical identification of a vesicular transporter mediating neuronal ATP release. Cerebral cortex 22: 1203-1214.
45. de Rivero Vaccari JP, Bastien D, Yurcisin G, Pineau I, Dietrich WD, et al. (2012) P2X4 receptors influence inflammasome activation after spinal cord injury. The Journal of neuroscience : the official journal of the Society for Neuroscience 32: 3058-3066.
46. Dong B, Silverman RH (1995) 2-5A-dependent RNase molecules dimerize during activation by 2-5A. The Journal of biological chemistry 270: 4133-4137.
47. Boo KH, Yang JS (2010) Intrinsic cellular defenses against virus infection by antiviral type I interferon. Yonsei Med J 51: 9-17.
48. Jha BK, Polyakova I, Kessler P, Dong B, Dickerman B, et al. (2011) Inhibition of RNase L and RNA-dependent protein kinase (PKR) by sunitinib impairs antiviral innate immunity. The Journal of biological chemistry 286: 26319-26326.
49. Banerjee S, Chakrabarti A, Jha BK, Weiss SR, Silverman RH (2014) Cell-Type-Specific Effects of RNase L on Viral Induction of Beta Interferon. MBio 5.
50. Zorec R, Araque A, Carmignoto G, Haydon PG, Verkhratsky A, et al. (2012) Astroglial excitability and gliotransmission: an appraisal of Ca2+ as a signalling route. ASN neuro 4.
51. Jitrapakdee S, Vidal-Puig A, Wallace JC (2006) Anaplerotic roles of pyruvate carboxylase in mammalian tissues. Cellular and molecular life sciences : CMLS 63: 843-854.
52. Clark AK, Malcangio M (2012) Microglial signalling mechanisms: Cathepsin S and Fractalkine. Experimental neurology 234: 283-292.

53. Clark AK, Wodarski R, Guida F, Sasso O, Malcangio M (2010) Cathepsin S release from primary cultured microglia is regulated by the P2X7 receptor. Glia 58: 1710-1726.
54. Sheridan GK, Murphy KJ (2013) Neuron-glia crosstalk in health and disease: fractalkine and CX3CR1 take centre stage. Open biology 3: 130181.
55. Suzuki M, El-Hage N, Zou S, Hahn YK, Sorrell ME, et al. (2011) Fractalkine/CX3CL1 protects striatal neurons from synergistic morphine and HIV-1 Tat-induced dendritic losses and death. Molecular neurodegeneration 6: 78.
56. Gemma C, Bachstetter AD (2013) The role of microglia in adult hippocampal neurogenesis. Frontiers in cellular neuroscience 7: 229.
57. Rogers JT, Morganti JM, Bachstetter AD, Hudson CE, Peters MM, et al. (2011) CX3CR1 deficiency leads to impairment of hippocampal cognitive function and synaptic plasticity. The Journal of neuroscience : the official journal of the Society for Neuroscience 31: 16241-16250.
58. Parkhurst CN, Yang G, Ninan I, Savas JN, Yates JR, 3rd, et al. (2013) Microglia promote learning-dependent synapse formation through brain-derived neurotrophic factor. Cell 155: 1596-1609.
59. Tang Z, Gan Y, Liu Q, Yin JX, Shi J, et al. (2014) CX3CR1 deficiency suppresses activation and neurotoxicity of microglia/macrophage in experimental ischemic stroke. Journal of neuroinflammation 11: 26.
60. Woo IS, Jin H, Kang ES, Kim HJ, Lee JH, et al. (2011) TMEM14A inhibits N-(4-hydroxyphenyl)retinamide-induced apoptosis through the stabilization of mitochondrial membrane potential. Cancer letters 309: 190-198.
61. Foster MW, Thompson JW, Forrester MT, Sha Y, McMahon TJ, et al. (2013) Proteomic analysis of the NOS2 interactome in human airway epithelial cells. Nitric oxide : biology and chemistry / official journal of the Nitric Oxide Society 34: 37-46.
62. Amitai Y (2010) Physiologic role for "inducible" nitric oxide synthase: a new form of astrocytic-neuronal interface. Glia 58: 1775-1781.
63. Tada H, Okano HJ, Takagi H, Shibata S, Yao I, et al. (2010) Fbxo45, a novel ubiquitin ligase, regulates synaptic activity. The Journal of biological chemistry 285: 3840-3849.
64. Lee JS, Ho WK, Neher E, Lee SH (2013) Superpriming of synaptic vesicles after their recruitment to the readily releasable pool. Proceedings of the National Academy of Sciences of the United States of America 110: 15079-15084.
65. de Vega S, Iwamoto T, Yamada Y (2009) Fibulins: multiple roles in matrix structures and tissue functions. Cellular and molecular life sciences : CMLS 66: 1890-1902.
66. Balbona K, Tran H, Godyna S, Ingham KC, Strickland DK, et al. (1992) Fibulin binds to itself and to the carboxyl-terminal heparin-binding region of fibronectin. The Journal of biological chemistry 267: 20120-20125.
67. Tran H, VanDusen WJ, Argraves WS (1997) The self-association and fibronectin-binding sites of fibulin-1 map to calcium-binding epidermal growth factor-like domains. The Journal of biological chemistry 272: 22600-22606.
68. Vincent C, Siddiqui TA, Schlichter LC (2012) Podosomes in migrating microglia: components and matrix degradation. Journal of neuroinflammation 9: 190.
69. Tsuda M, Toyomitsu E, Kometani M, Tozaki-Saitoh H, Inoue K (2009) Mechanisms underlying fibronectin-induced up-regulation of P2X4R expression in microglia: distinct roles of PI3K-Akt and MEK-ERK signalling pathways. Journal of cellular and molecular medicine 13: 3251-3259.
70. Nasu-Tada K, Koizumi S, Tsuda M, Kunifusa E, Inoue K (2006) Possible involvement of increase in spinal fibronectin following peripheral nerve injury in upregulation of microglial P2X4, a key molecule for mechanical allodynia. Glia 53: 769-775.
71. Nurminskaya MV, Belkin AM (2012) Cellular functions of tissue transglutaminase. International review of cell and molecular biology 294: 1-97.

72. Malorni W, Farrace MG, Rodolfo C, Piacentini M (2008) Type 2 transglutaminase in neurodegenerative diseases: the mitochondrial connection. Current pharmaceutical design 14: 278-288.
73. Petrenko AG, Ullrich B, Missler M, Krasnoperov V, Rosahl TW, et al. (1996) Structure and evolution of neurexophilin. The Journal of neuroscience : the official journal of the Society for Neuroscience 16: 4360-4369.
74. Missler M, Sudhof TC (1998) Neurexophilins form a conserved family of neuropeptide-like glycoproteins. The Journal of neuroscience : the official journal of the Society for Neuroscience 18: 3630-3638.
75. Pettem KL, Yokomaku D, Luo L, Linhoff MW, Prasad T, et al. (2013) The specific alpha-neurexin interactor calsyntenin-3 promotes excitatory and inhibitory synapse development. Neuron 80: 113-128.
76. Frias CP, Wierenga CJ (2013) Activity-dependent adaptations in inhibitory axons. Frontiers in cellular neuroscience 7: 219.
77. Stern LJ, Potolicchio I, Santambrogio L (2006) MHC class II compartment subtypes: structure and function. Current opinion in immunology 18: 64-69.
78. Schwartz V, Lue H, Kraemer S, Korbiel J, Krohn R, et al. (2009) A functional heteromeric MIF receptor formed by CD74 and CXCR4. FEBS letters 583: 2749-2757.
79. Savaskan NE, Fingerle-Rowson G, Buchfelder M, Eyupoglu IY (2012) Brain miffed by macrophage migration inhibitory factor. International journal of cell biology 2012: 139573.