TechnicalNote

affymetrix USB

Primer selection guidelines utilizing VeriQuest® qPCR for microarray validation

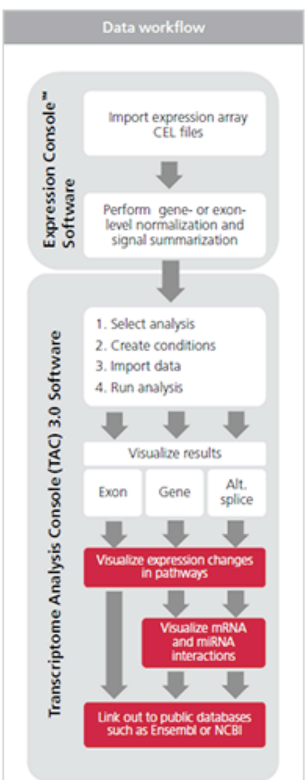
This tutorial provides examples of appropriate primer design strategies when using USB® VeriQuest qPCR master mixes to validate Affymetrix® microarrays. These guidelines generate optimal concordance of microarray and qPCR results when using **First-Strand cDNA Synthesis Kit (PN 75780)** followed by real-time PCR with **VeriQuest Probe (PN 75650)** or **VeriQuest SYBR® Green (PN 75600)** qPCR master mixes.

Note: Assays can be performed using either VeriQuest SYBR Green or Probe chemistries. SYBR assays are presented here to describe the most cost-effective solution.

Example 1: Validating GeneChip® Human Transcriptome Array (HTA) 2.0 data

GeneChip HTA 2.0 is the highest resolution microarray for whole transcriptome expression profiling of all transcript isoforms. A key advantage of HTA 2.0 is a simple, fast, and free data analysis package consisting of **Expression Console™ Software** and **Transcriptome Analysis Console (TAC) Software** that can be downloaded from the Affymetrix website (**Affymetrix.com**). Expression Console generated .chp files can be opened in TAC for analysis and easy interpretation of differential gene-level expression as well as alternative splicing between samples assayed on HTA 2.0.

A simple, easy-to-use, yet powerful analysis tool. With TAC Software you will be able to



- Visualize expression changes in biological pathways
- Perform statistical tests for differential expression
- Identify differentially expressed genes and exons
- Identify alternative splicing events
- Visualize gene models with exon and junction signal
- Visualize fold changes of gene-level and miRNA
- Link directly to available public annotations

When combined with the comprehensive coverage of Affymetrix' high-density microarray solutions, TAC Software provides an easy-to-use, integrated solution to take you from raw data to biological results in just a few clicks.

Download information and instructions

Download the latest version of Transcriptome Analysis Console Software, version 3.0 (64-bit). This version is supported on Microsoft Windows®7 (SP 1) and Windows 8.1 Professional.

Download

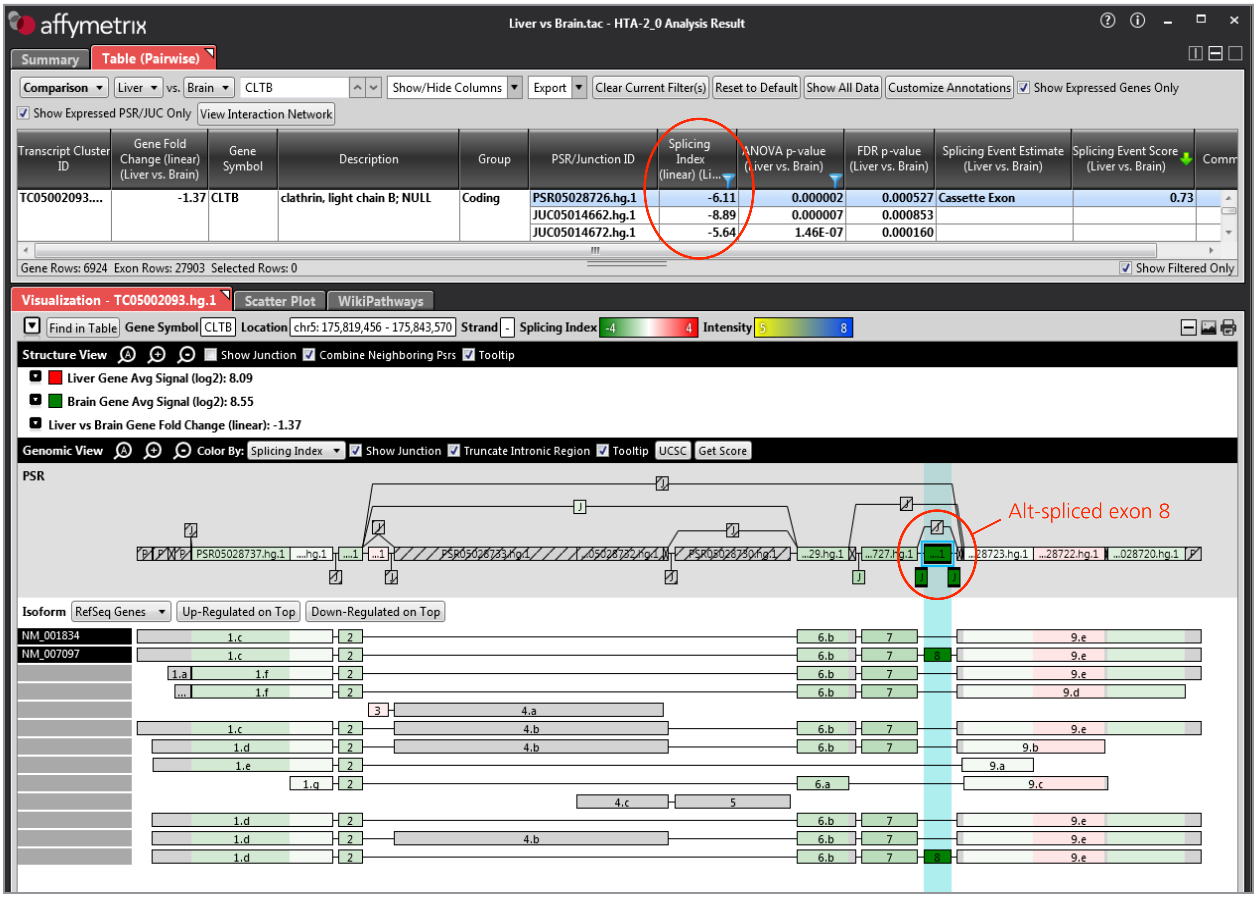
Additional analysis solutions from Affymetrix

Expression Console™ Software
Affymetrix Power Tools – developed internally at Affymetrix

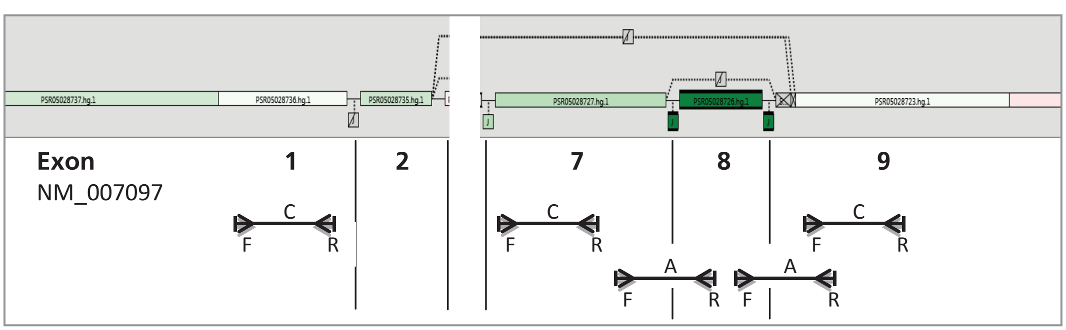
The following workflow demonstrates how to design VeriQuest qPCR assays for validating alternative splicing. HTA 2.0 data for Human Total RNA from liver and brain were subjected to Alternative Splicing Analysis in TAC to obtain a list of differentially expressed genes and alt-splicing events (liver vs. brain). To increase the confidence of the TAC analysis, a p value of 0.001 was used, which is more stringent than the default value of 0.05. Significant alt-splicing events were selected based on **Probe Selection Region (PSR)**

Splicing Index (SI) values. In this example, the Clathrin Light Chain B (CLTB) gene, which helps regulate receptor-mediated endocytosis, was selected based on its criteria of having a low **Gene Fold Change** = <2 or >-2 (**-1.37**) and a significant **PSR SI** value = >2 or <-2 (**-6.11**). The **Junction Probe (JUC) SI** values (**-8.89**, **-5.64**) provide further supporting evidence for differential splicing.

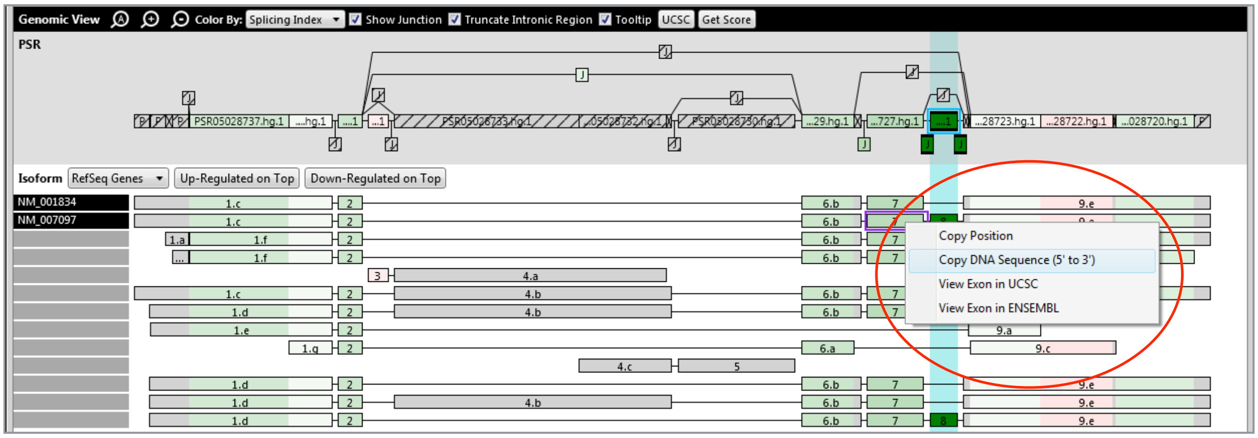
By clicking on the row containing CLTB in TAC, all known transcript isoforms in the database are displayed. In this gene structure view, significant alt-splicing events are darker in color and will highlight when clicking on the respective PSR as shown below (view filtered to show only the accession numbers for isoforms in RefSeq). In this example, CLTB exon 8 is characterized as a Cassette Exon that is skipped more often in liver versus brain.



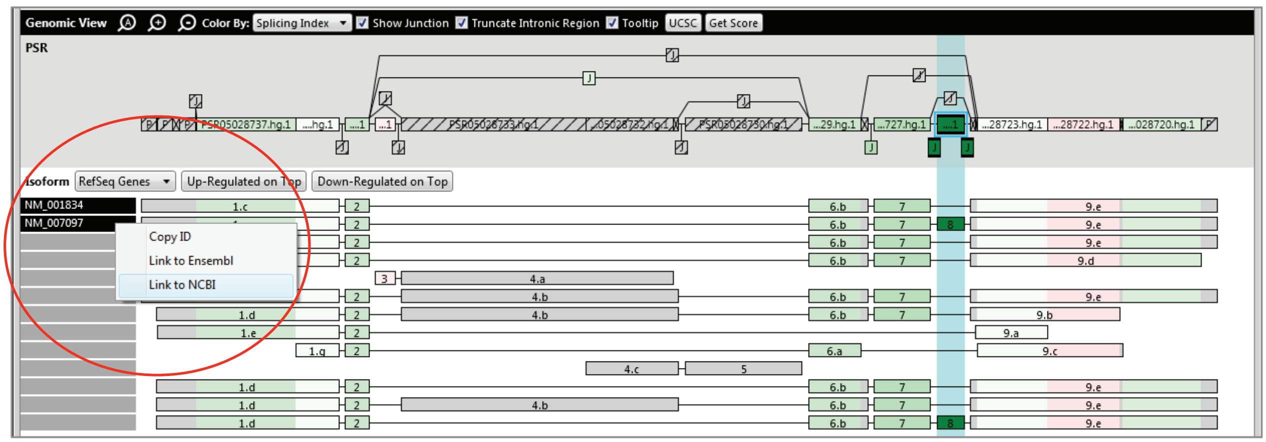
To validate the CLTB cassette exon alt-splicing event with VeriQuest qPCR, design multiple primer sets (2-4) within constitutive exons (exons that are expressed in both conditions or always "on") and multiple primer sets (2-4) within the alt-spliced exon. In this example, CLTB exons 1, 2, 6, 7, and 9 are constitutive and exon 8 is alternatively spliced. If the exon length is less than 70 bp, as for CLTB exon 8, design qPCR primers with one of the pair within the alt-spliced exon and the other within the flanking constitutive exon as shown below. Real-time PCR data from constitutive (C) amplicons provide gene-level expression information. Data from alt-spliced (A) amplicons are used to calculate qPCR Splicing Indexes. This strategy enables accurate calculation of Splicing Index from VeriQuest qPCR data (***Appendix A***) for optimal array concordance. The figure below shows the CLTB gene structure view from TAC (to scale) and approximate positions of qPCR amplicons (not to scale) designed to validate HTA 2.0 data.



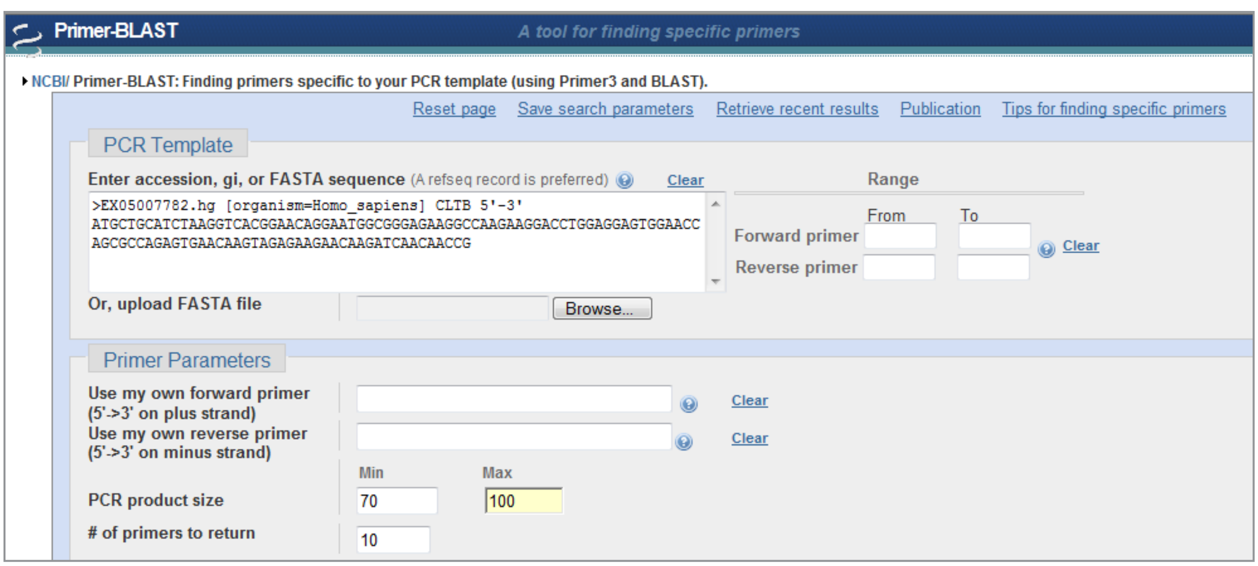
Designing primers for exons greater than 70 bp (optimal amplicon length 70-100 bp)
Designing qPCR primers to interrogate specific exons is simplified with TAC. By right clicking on an exon of interest (CLTB constitutive exon 7 below), and selecting **Copy DNA Sequence (5′ to 3′)**, the sequence can be copied and pasted into **NCBI Primer-BLAST** for efficient design of qPCR primers for use with VeriQuest SYBR Green qPCR Master Mix.



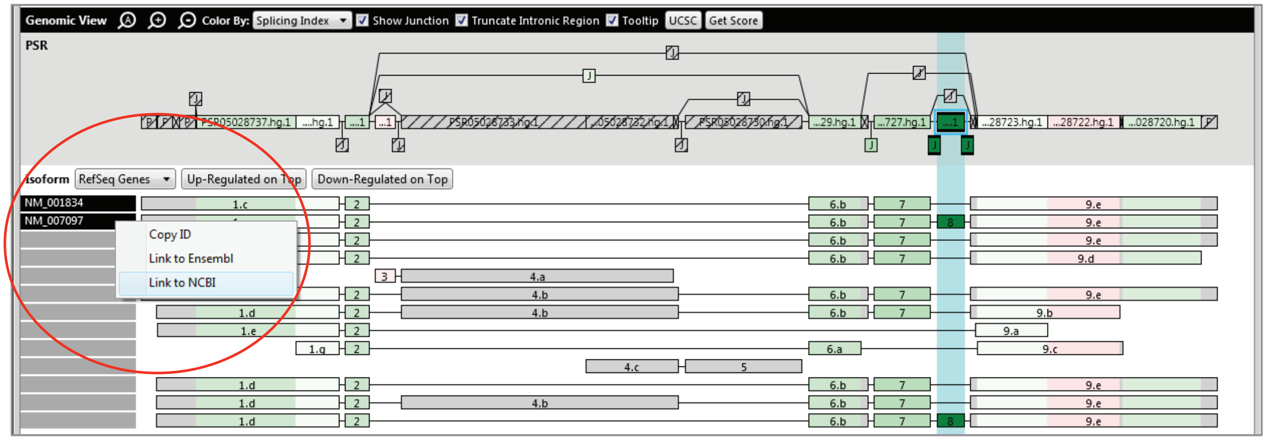
NCBI can be opened from TAC by right clicking on the RefSeq accession number and selecting **Link to NCBI**:



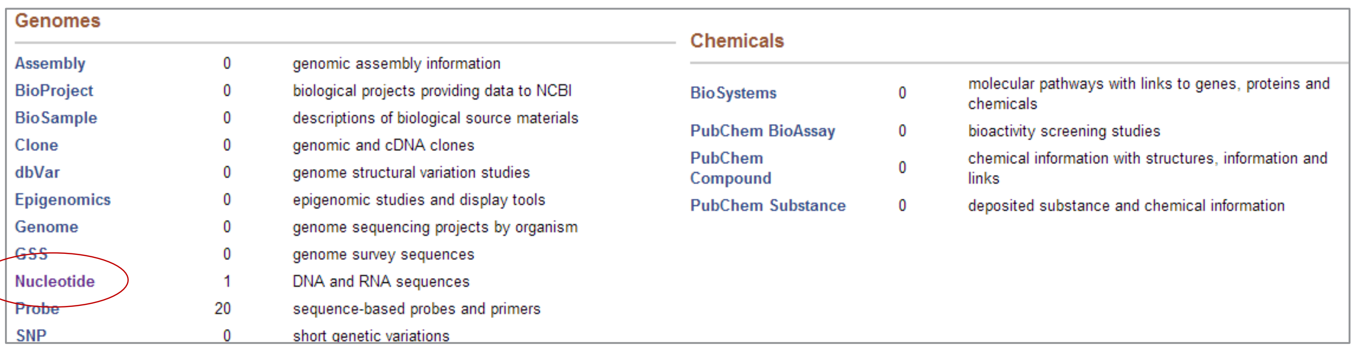
Paste the exon sequence into **NCBI Primer-BLAST**, set Max PCR product size to 100, and click **Get Primers**:



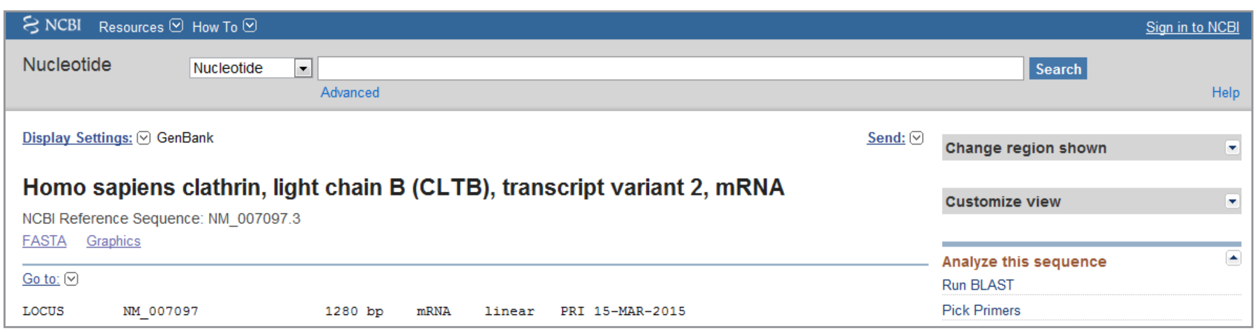
Designing primers for exons less than 70 bp (optimal amplicon length 70-100 bp)
To interrogate exons less than 70 bp, as is the case for CLTB alt-spliced exon 8 (54 bp), design qPCR primers spanning exon-exon junctions by right clicking on the RefSeq accession number that contains the alt-spliced exon 8 (in this case NM_007097) and selecting **Link to NCBI**:



Click the **Nucleotide** link under Genomes:



This will open the NCBI Reference Sequence for CLTB:



The exon sequences can be viewed towards the bottom of the page and start/end positions determined.

```
                     /inference="alignment:Splign:1.39.8"
     exon            670..723
                     /gene="CLTB"
                     /gene_synonym="LCB"
                     /inference="alignment:Splign:1.39.8"
     exon            724..1214
                     /gene="CLTB"
                     /gene_synonym="LCB"
                     /inference="alignment:Splign:1.39.8"
     STS             872..1111
                     /gene="CLTB"
                     /gene_synonym="LCB"
                     /standard_name="RH98811"
                     /db_xref="UniSTS:84078"
     STS             896..1170
                     /gene="CLTB"
                     /gene_synonym="LCB"
                     /standard_name="WI-7674"
                     /db_xref="UniSTS:58080"
     STS             968..1161
                     /gene="CLTB"
                     /gene_synonym="LCB"
                     /standard_name="RH69093"
                     /db_xref="UniSTS:58244"
     STS             987..1133
                     /gene="CLTB"
                     /gene_synonym="LCB"
                     /standard_name="G54012"
                     /db_xref="UniSTS:109394"
     polyA_site      1214
                     /gene="CLTB"
                     /gene_synonym="LCB"
ORIGIN
        1 ccgctcctcc cgccgcggcg cccgcaccgc agtgacagcc agccgggccc ggtggcggag
       61 aggaagtgcg gtccgcgcca agcccgtccc cgccgacgcc ggctccccgc ggctcgggtg
      121 acagcgtcgc ggccgccgga cgcagcgcgg ggcaggcgcg ggcagagccg agcgcagcgg
      181 aggctccggc ggaggcgcgg ggaaaatggc tgatgacttt ggcttcttct cgtcgtcgga
      241 gagcggtgcc ccggaggcgg cggaggagga cccggcggcc gccttcctgg cccagcagga
      301 gagcgagatt gcaggcatag agaacgacga gggcttcggg gcacctgccg gcagccatgc
      361 ggcccccgcg cagccgggcc ccacgagtgg ggctggttct gaggacatgg ggaccacagt
      421 caatggagat gtgtttcagg aggccaacgg tcctgctgat ggctacgcag ccattgccca
      481 ggctgacagg ctgacccagg agcctgagag catccgcaag tggcgagagg agcagaggaa
      541 acggctgcaa gagctggatg ctgcatctaa ggtcacggaa caggaatggc gggagaaggc
      601 caagaaggac ctggaggagt ggaaccagcg ccagagtgaa caagtagaga agaacaagat
      661 caacaaccgg atcgctgaca aagcattcta ccagcagcca gatgctgata tcatcggcta
      721 cgtggcatcc gaggaggctt tcgtgaagga atccaaggag gagaccccag gcacagagtg
      781 ggagaaggtg gcccagctat gtgacttcaa ccccaagagc agcaagcagt gcaaagatgt
      841 gtcccgcctg cgctcggtgc tcatgtccct gaagcagacg ccactgtccc gctaggtgcc
      901 tgctaggtgc atggccacag agcatgggct gggcctgggc acaggaggag cagctgcttt
      961 ggtcggggtg gagactcgca gcagctgcta cccacagcct attccactcc tccccatctc
     1021 caggcgctgg gagggggccc ctcaccccat cacgcctcgc tccctcctgg ccctctggtc
     1081 cagcccctca cgcctcctct cagtctactc aattgtgact gtccctcctg atgtattttt
     1141 tttcttggct taaagggtgt gttgttgact ctttttacac ttatttatta tcattctcac
```

CLTB NM_007097

exon	start	end	length
1	1	392	392
2	393	439	47
6	440	557	118
7	558	669	112
8*	**670**	**723**	**54**
9	724	1214	491

* CLTB exon 8 alt-spliced as determined by TAC

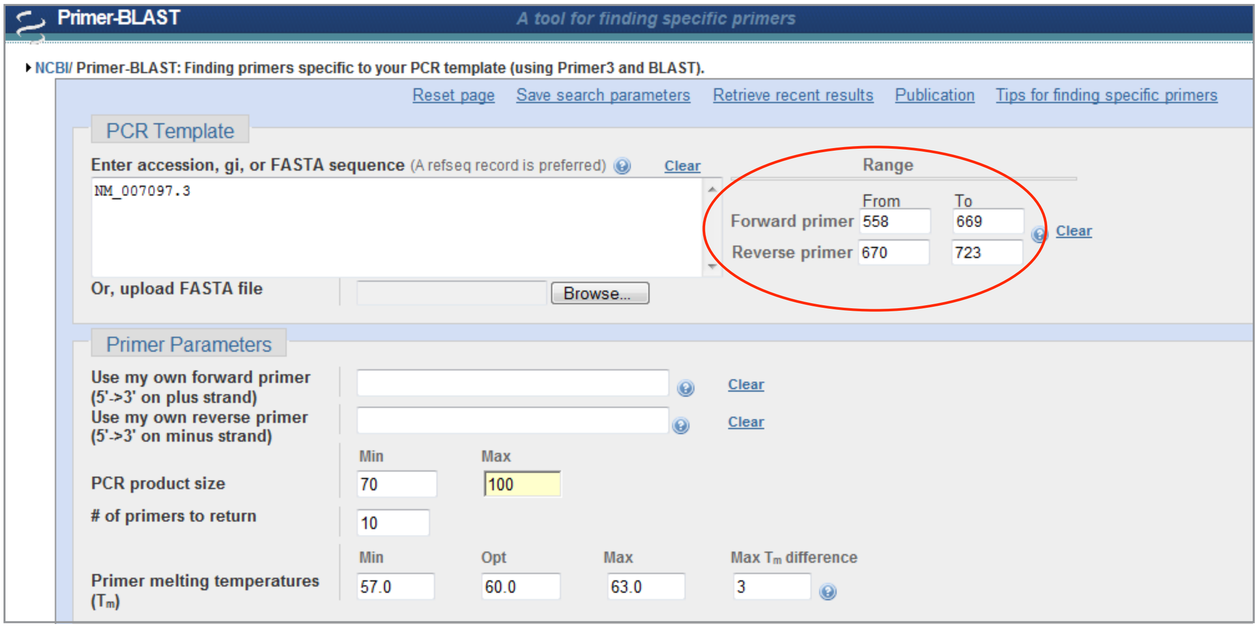
CLTB exon 8 sequence selected in NCBI. This sequence can be cross-referenced with TAC for confirmation of the correct alt-spliced exon.

By clicking **Pick Primers** under Analyze this sequence, qPCR primer sets can be designed using **NCBI Primer-BLAST**.

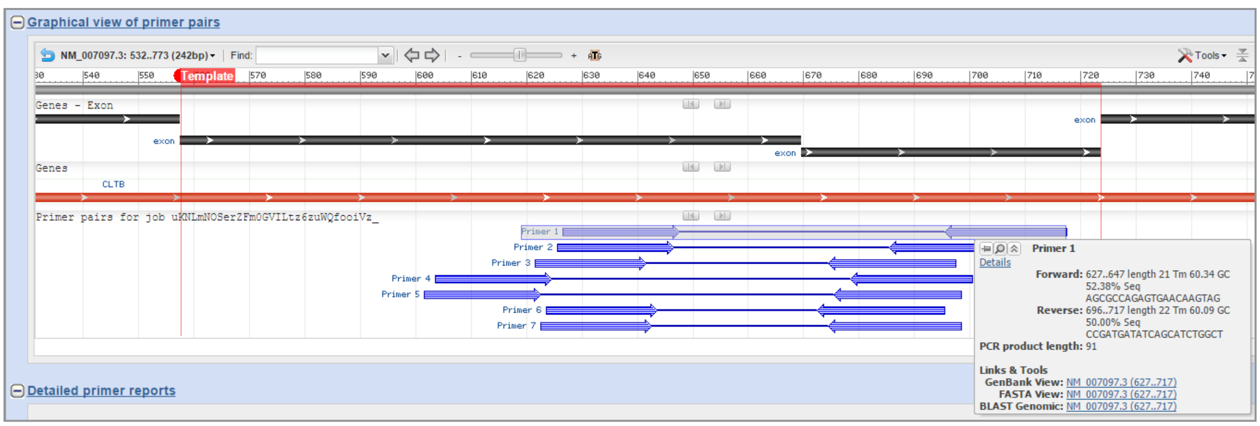
```
NCBI  Resources  How To                                              Sign in to NCBI
Nucleotide    Nucleotide [v] [                                  ] Search
              Advanced                                                        Help

Display Settings: GenBank                        Send:     Change region shown
                                                           Customize view
Homo sapiens clathrin, light chain B (CLTB), transcript variant 2, mRNA
NCBI Reference Sequence: NM_007097.3                       Analyze this sequence
FASTA  Graphics                                            Run BLAST
                                                           Pick Primers
Go to:                                                     Highlight Sequence Features
                                                           Find in this Sequence
LOCUS       NM_007097           1280 bp    mRNA    linear   PRI 15-MAR-2015
DEFINITION  Homo sapiens clathrin, light chain B (CLTB), transcript variant 2,
            mRNA.                                          Articles about the CLTB gene
ACCESSION   NM_007097                                      Genome-wide shRNA screening identifies host
VERSION     NM_007097.3  GI:365906278
KEYWORDS    RefSeq.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
```

To design qPCR primers that span exon-exon junctions with one of the pair within the exon of interest (in this case CLTB alt-spliced exon 8), enter the start/end positions of exon 7 (558/669) in the **Forward primer From/To** fields and the start/end positions of exon 8 (670/723) in the **Reverse primer From/To** fields. Set Max PCR product size to 100, and click **Get Primers**.

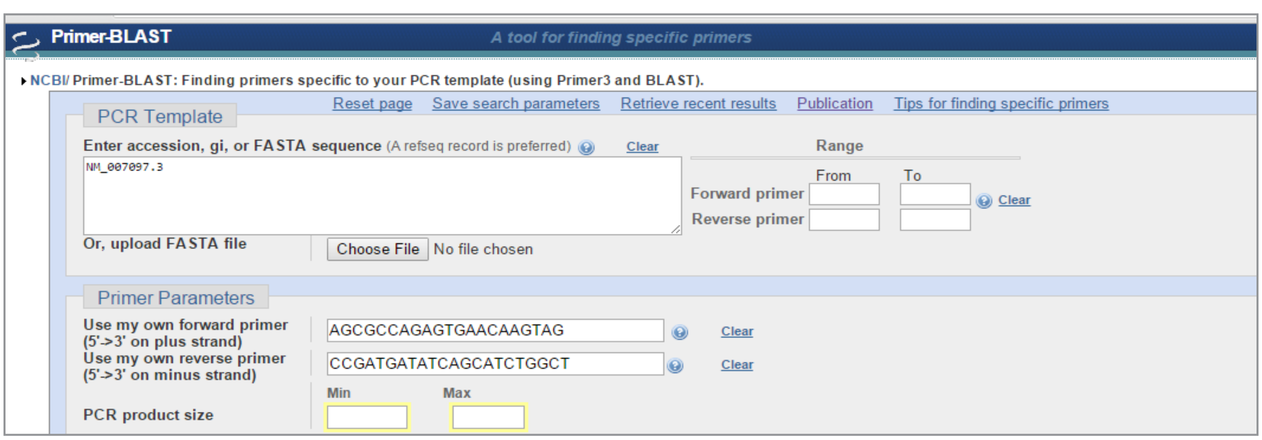


Select a qPCR primer pair (optimal amplicon length 70-100 bp) for interrogating the CLTB alt-spliced exon 8.

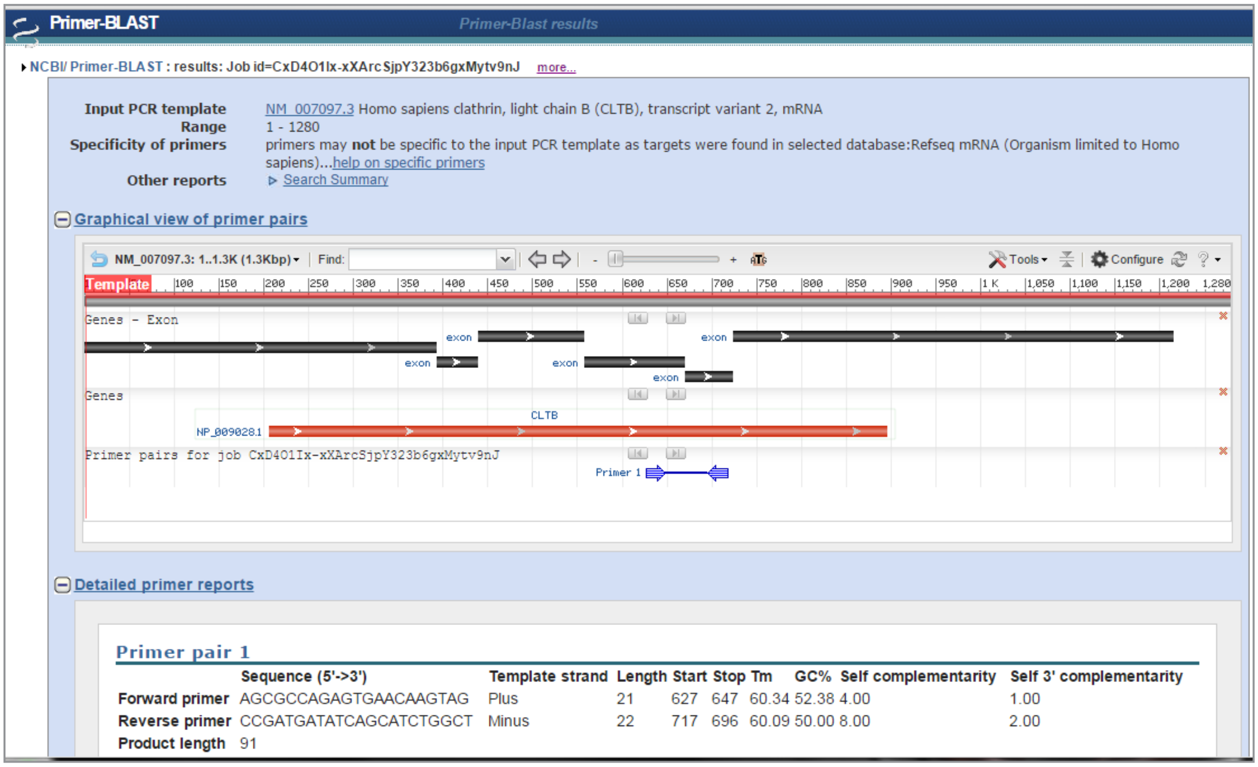


Checking primers for specificity and SNPs

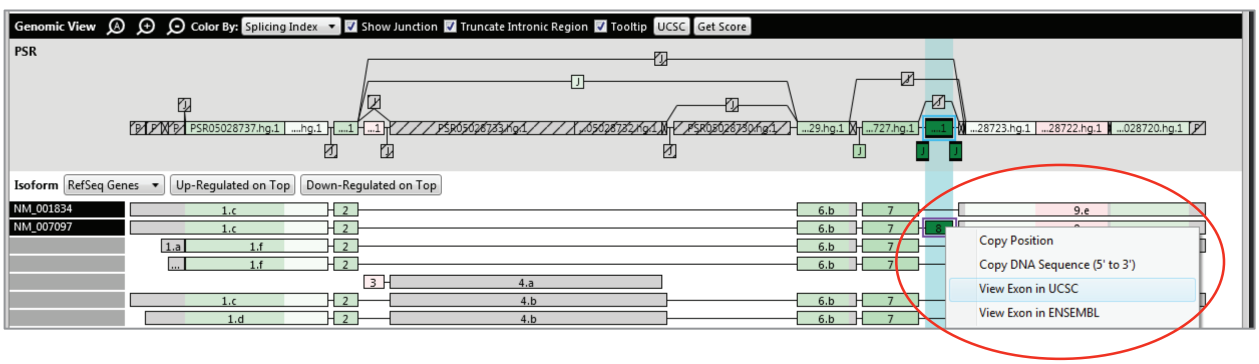
Primer pairs can be checked for specificity in **NCBI Primer-BLAST** by inputting the forward and reverse sequences in the **Use my own primer** fields and clicking **Get Primers**:



qPCR primer pairs that are unique for a specific target should only give products on the target of interest (in this case CLTB exon 7/8) of the appropriate amplicon length (in this case 91).



UCSC In-Silico PCR can also be used to check qPCR primer pairs for specificity and SNP positions. Common SNPs within primer annealing sites should be avoided. The UCSC database can be opened from TAC by right clicking on any exon and selecting **View Exon** in UCSC.



When primer pairs span exon-exon junctions, a longer amplicon may be amplified due to genomic DNA contamination and will show as a higher molecular weight band on agarose gels. This contamination will interfere with SYBR-based qPCR analysis and should be minimized by using isolation kits with a DNase step (*e.g.* **PrepEase RNA Spin Kit, PN 78766**) or by DNase treatment of isolated RNA using **rDNase I, RNase-Free (PN 78411)**. Genomic DNA contamination is more difficult to detect when primer pairs are within an exon boundary, and the inclusion of a no reverse transcriptase control is recommended in those situations.

Fold change from real-time PCR data is calculated using the $\Delta\Delta C_t$ method relative to an endogenous reference gene (*e.g.* ACTB). Gene-level fold change is calculated using the fold change of constitutive exons. Splicing Index from qPCR data is calculated for alt-spliced amplicons by normalizing their fold change to the average fold change of constitutive amplicons. In effect, Splicing Index is a fold-change measure of the alt-splice event. See ***Appendix A*** for detailed calculation and table of recommended reference genes and primers in ***Appendix B***.

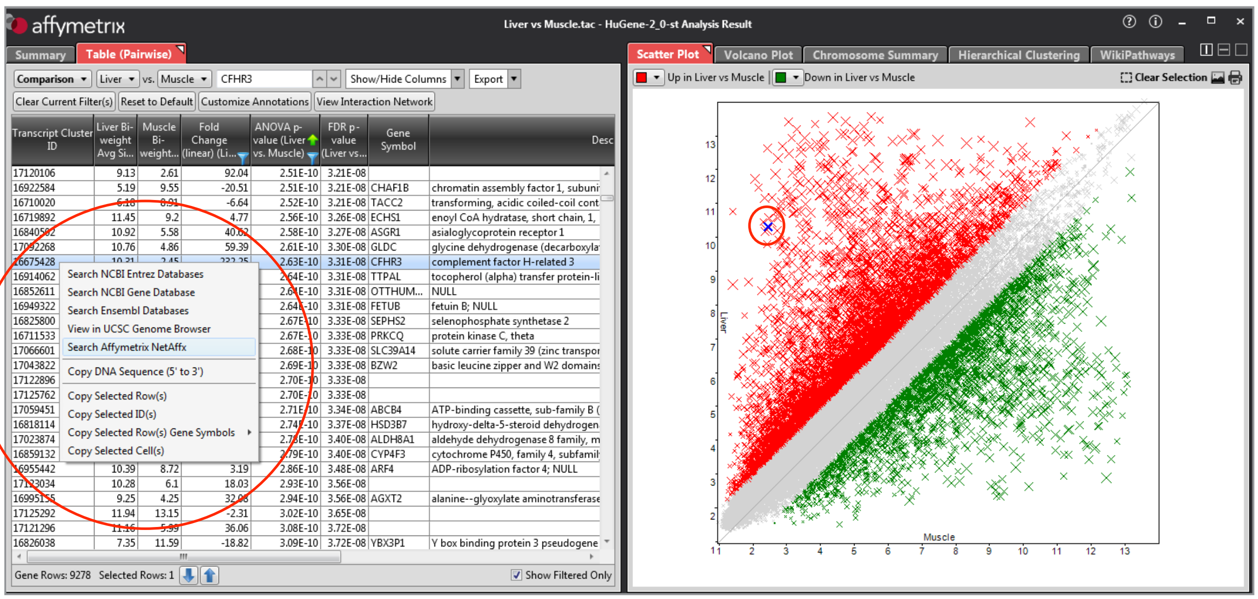
Validation of Affymetrix GeneChip HTA 2.0 alternative splicing data by VeriQuest SYBR Green qPCR assays may have differences in the magnitude of the splicing indexes. This is normal due to the inherent differences in the dynamic range and noise of each platform. In an internal study conducted by Affymetrix scientists, 16 separate differential alt-splicing events selected from HTA 2.0 data were confirmed by VeriQuest qPCR with 100%-directional concordance (see **White Paper**).

Example 2: Validating GeneChip Human Gene (HuGene) 2.0 ST Array

When using VeriQuest qPCR Master Mixes to validate Affymetrix 3' IVT (*e.g.* Human Genome U133 Plus 2.0) and Gene (*e.g.* Human, Mouse, Rat Gene 1.x or 2.x ST) microarrays, qPCR primers should be designed against constitutive exons using a workflow similar to ***Example 1***. While these arrays are not annotated to the extent of HTA 2.0 in **TAC**, **Expression Console** generated files can be analyzed in TAC Software for easy interpretation of differential expression between samples.

In this example, HuGene 2.0 data from Human liver and muscle total RNA were analyzed by TAC Software to obtain a list of differentially expressed genes (see figure below). The Complement Factor H-Related 3 (CFHR3) gene, which encodes a protein that binds heparin and may be involved in complement regulation, was selected based on its significant gene-level fold change between samples (liver vs. muscle) and constitutive regions of significant length for interrogation by VeriQuest qPCR. The highlighted CFHR3 gene's differential expression is shown as a blue '**x**' in the displayed scatter plot (higher expression in liver vs. muscle). The **Transcript Cluster ID** (16675428) can be queried in **NetAffx™ Analysis Center**, which is accessible on the Affymetrix website (**Affymetrix.com**), for appropriate primer design strategy.

Right click the Transcript Cluster ID for CFHR3 and select **Search Affymetrix NetAffx**:

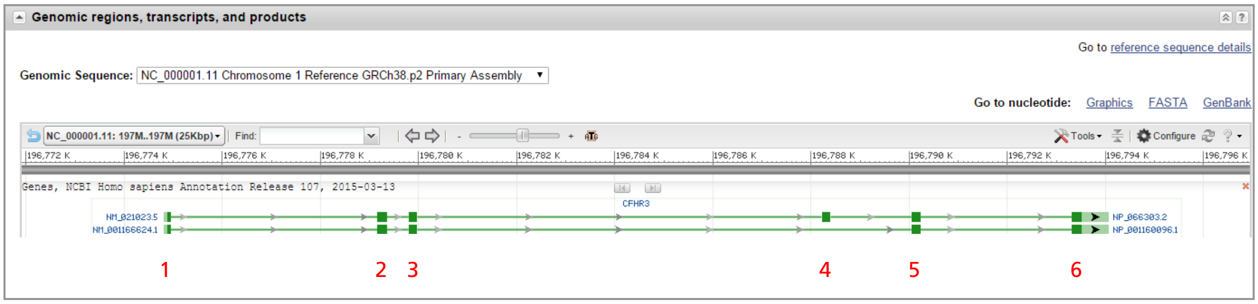


Click the **Entrez ID (10878)** under Transcripts Detected By 16675428 to link out to **NCBI** for the gene structure view:

Transcripts Detected By 16675428

Full Length	Gene Symbol (Entrez ID)	Description Pathways	Assign Score	Coverage
BC070259 GenBank	CFHR3 (10878)	Homo sapiens complement factor H-related 3, mRNA (cDNA clone IMAGE:4716670), complete cds.	100 (8/8)	67 (8/12)
ESTs	Gene Symbol (Entrez ID)	Description Pathways	Assign Score	Coverage
Predicted	Gene Symbol (Entrez ID)	Description Pathways	Assign Score	Coverage

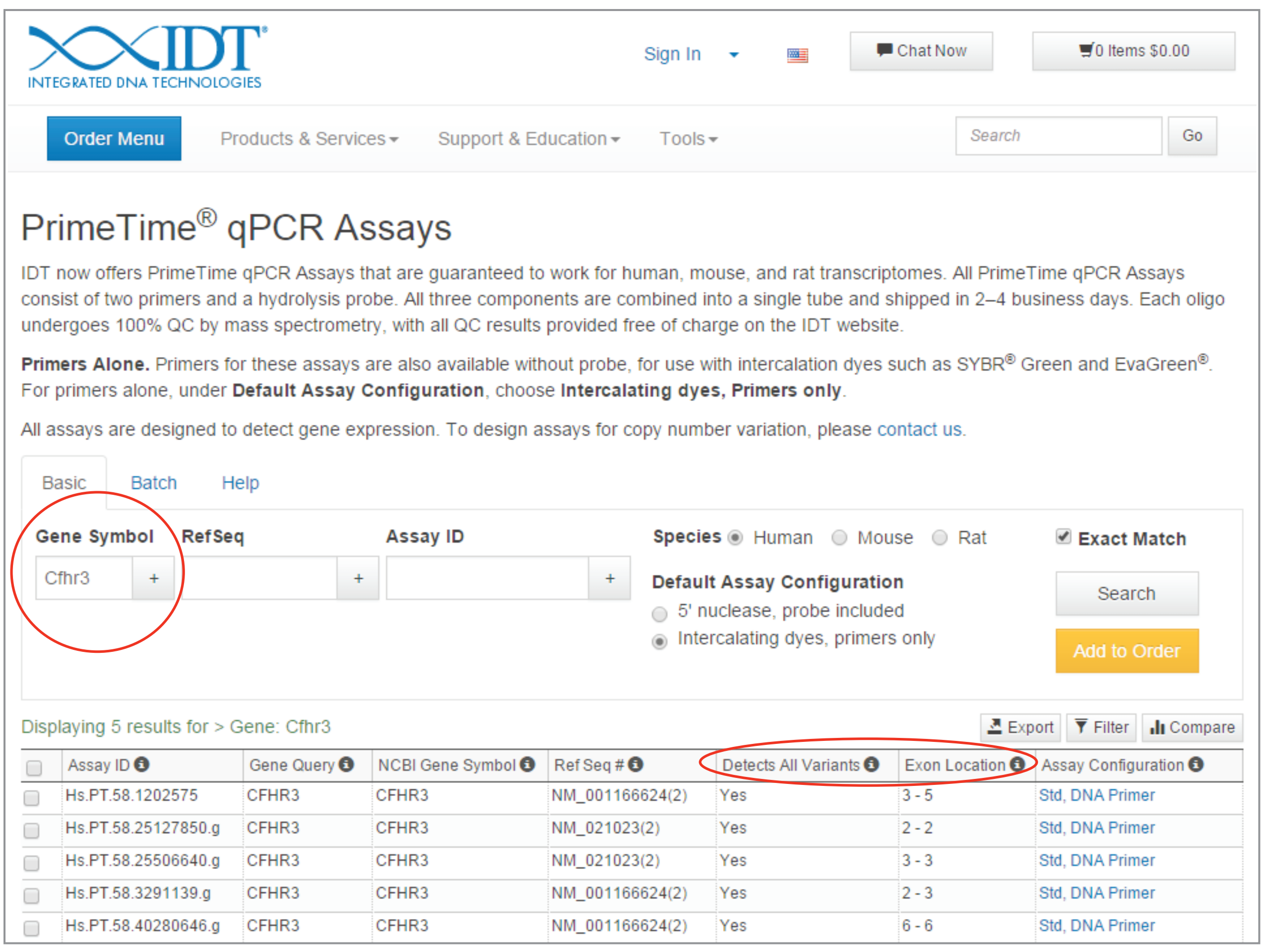
From the NCBI gene structure view of CFHR3, constitutive exons (*e.g.* CFHR3 exons 1-3, 5,6) are present in all variants where potentially alt-spliced exons (*e.g.* CFHR3 exon 4) are not. Use constitutive exon sequences for designing qPCR primers for gene-level validation in VeriQuest qPCR assays.



Note: Gene structures are subject to change with discovery of new variants. Non-concordance between assays due to undefined alt-splicing events can be resolved with a 2nd qPCR assay with primer pairs in a different constitutive region.

The RefSeq accession number (NM_021023) can then be queried on the **NCBI** website to determine sequences of constitutive exons for qPCR primer design (similar to the workflow in ***Example 1***) for validation with VeriQuest SYBR Green qPCR Master Mixes.

As an alternative to primer design using **NCBI Primer-BLAST** and **UCSC In-Silico PCR,** both SYBR- and Probe-based pre-designed qPCR assays from **IDT** can be used for gene-level validation of Affymetrix expression arrays. By simply inputting the **Gene Symbol**, **Species**, and **Assay Configuration** preference (in this case CFHR3, Human, Intercalating dyes, primers only) and clicking **Submit**, available assays for **Detecting All Variants** can be purchased. Notice there are no assays available for exon 4 which is a non-constitutive exon and should be avoided in VeriQuest qPCR validation assays.



Fold change from real-time PCR data is calculated using the $\Delta\Delta C_t$ method relative to an endogenous reference gene (*e.g.* ACTB). Gene-level fold change is the fold change of constitutive exons.

Appendix A

The Splicing Index for HTA data represents the ratio of the exon intensities in condition 1 versus condition 2 after normalization to their respective gene intensities in each sample. A Splicing Index value of 0 indicates that the Probe Selection Region (PSR) is present at equal levels in both conditions. A positive Splicing Index value implies elevated inclusion, and a negative value suggests increased PSR skipping in condition 1 versus condition 2.

This is an example of fold change and Splicing Index calculation from VeriQuest qPCR data. See ***Example 1*** above for CLTB gene structure view and primer design.

														Fold Change		Splicing Index	
Gene	Amplicon		C_t per triplicates						Average C_t		Average ΔC_t (ACTB-CLTB)		$\Delta\Delta C_t$	$2^\wedge\Delta\Delta C_t$	$2^\wedge\Delta\Delta C_t$ or $-1/2^\wedge\Delta\Delta C_t$	FC(A)/AvgFC(C)	SI
			Liver			Brain			L	B	L	B	Liver-Brain	Liver vs Brain		Liver vs Brain	
CLTB	constitutive (C)	exon 1-1	31.3	31.2	31.3	29.9	29.8	29.8	31.3	29.8	-4.96	-3.95	-1.01	0.50	-2.0		
		exon 7-7	29.9	29.8	30.0	28.8	28.9	28.9	29.9	28.9	-3.56	-3.00	-0.56	0.68	-1.5		
		exon 9-9	31.5	31.2	31.3	29.9	29.9	29.9	31.3	29.9	-5.03	-4.02	-1.01	0.50	-2.0		
	alt-spliced (A)	exon 7-8	35.0	35.0	35.0	29.1	29.0	29.1	35.0	29.0	-8.70	-3.14	-5.55	0.021		0.038	-26
		exon 8-9	35.0	35.0	35.0	28.9	28.9	28.9	35.0	28.9	-8.70	-2.99	-5.70	0.019		0.034	-29
ACTB	constitutive	exon 3-4	26.3	26.3	26.3	25.9	25.9	25.9	26.3	25.9							

Fold change and Splicing Index from qPCR data are calculated using the $\Delta\Delta C_t$ method relative to an endogenous reference gene (ACTB in this example). A raw C_t of 35 is used as the limit of detection: C_t values are set at 35 for any replicates with C_t values not determined or >35.

Calculate average ΔC_t as the average C_t (ACTB) - average C_t (CLTB) for both conditions (liver and brain).
Calculate $\Delta\Delta C_t$ as ΔC_t (liver) - ΔC_t (brain).
Calculate $2^{\Delta\Delta C_t}$.

Gene-level fold change is reported as the fold change of constitutive exons (C).

Fold change = $2^{\Delta\Delta C_t}$ if $2^{\Delta\Delta C_t} \geq 1$
Fold change = $-1/2^{\Delta\Delta C_t}$ if $2^{\Delta\Delta C_t} < 1$

For the constitutive exon 1, the calculated $2^{\Delta\Delta C_t}$ value is less than 1 (0.50), indicating slightly decreased gene expression in liver versus brain. This is finally reported as -1/0.50 = -2.0, as a negative number.

Splicing Index from qPCR data is calculated for alt-spliced amplicons (A) by normalizing their fold change to the average fold change of constitutive exons.

Calculate FC (A)/average FC (C)
For amplicon spanning exons 7-8:
0.021/average (0.50, 0.68, 0.50) = 0.038

Splicing Index = FC (A)/average FC (C) if FC (A)/average FC (C) ≥1
Splicing Index = -1/(FC (A)/average FC (C)) if FC (A)/average FC (C) <1

For amplicon spanning exons 7-8, the calculated FC (A)/average FC (C) value is less than 1 (0.038), indicating decreased exon 8 inclusion in liver. This is finally reported as -1/0.038 = -26, as a negative number.

Appendix B

Real-time PCR data is normalized to a reference gene, an mRNA that is stably expressed under the experimental conditions employed. It is recommended to screen multiple reference genes to determine which expression levels fluctuate the least. Commonly used endogenous reference genes with possible forward and reverse primer sequences are listed in the table.

Gene ID	Description	Primer sequences (5′ to 3′)	
18S	18S ribosomal RNA	F	CGAAGACGATCAGATACCGT
		R	GGTCATGGGAATAACGCCG
28S	28S ribosomal RNA	F	TCGTCCGACCTGGGTATAG
		R	GCTATCCTGAGGGAAACTTCG
ACTB	actin, beta	F	AACCGCGAGAAGATGACCCAGAT
		R	TAGCACAGCCTGGATAGCAACGTA
GAPDH	glyceraldehyde-3-phosphate dehydrogenase	F	ATCGTGGAAGGACTCATGACCACA
		R	TAGAGGCAGGGATGATGTTCTGGA
HPRT1	hypoxanthine phosphoribosyltransferase 1	F	ATGGACAGGACTGAACGTCTTGCT
		R	TTGAGCACACAGAGGGCTACAATG
POLR2A	polymerase (RNA) II (DNA directed) polypeptide A	F	GCATTGACTTGCGTTTCCAC
		R	CCATCACACATGTGCCGTTC
PPIA	peptidylprolyl isomerase A (cyclophilin A)	F	GTCTCCTTTGAGCTGTTTGC
		R	AAGCAGGAACCCTTATAACC
TBP	TATA box binding protein	F	ACCAGGTGATGCCCTTCTGTAAGT
		R	ATGAGCAACTCACAGTCACGCT

Affymetrix, Inc.
USB® Products
26111 Miles Road
Cleveland, Ohio 44128
Tel: 800-321-9322 | 216-765-5000
Fax: 800-535-0898 | 216-464-5075
USBcustomerserv@affymetrix.com
usb.affymetrix.com

Affymetrix UK Ltd.
USB® Products
Voyager, Mercury Park,
Wycombe Lane, Wooburn Green,
High Wycombe HP10 0HH, UK
Tel: +44 (0)1628 55 2600
Fax: +44 (0)1628 55 2675
USBcustomerserveurope@affymetrix.com

Affymetrix Pte Ltd.
USB® Products
7 Gul Circle, #2M-01
Keppel Logistics Building
Singapore 629563
Tel: +65 63957310
Fax: +65 63957300
USBcustomerserv@affymetrix.com

 Rev 02/08/16
P/N USB05284 Rev. 1