Supplemental Data

Memory T Cell RNA Rearrangement Programmed by Heterogeneous Nuclear Ribonucleoprotein hnRNPLL

Zuopeng Wu, Xinying Jia, Laura de la Cruz, Xun-Cheng Su, Bruz Marzolf, Pamela Troisch, Daniel Zak, Adam Hamilton, Belinda Whittle, Di Yu, Daniel Sheahan, Edward Bertram, Alan Aderem, Gottfried Otting, Christopher C. Goodnow, and Gerard F. Hoyne

Supplemental Notes on Group A Neural Spliced Genes

The following are notes on published references to Hnrpll-dependent exons that are alternatively spliced in the nervous system.

Mobp

Mol Cell Neurosci. 1999 Apr;13(4):229-36. Splicing pattern, transcript start distribution, and DNA sequence of the mouse gene (Mobp) encoding myelin-associated oligodendrocytic basic protein. McCallion AS, Stewart GJ, Montague P, Griffiths IR, Davies RW.
We have cloned the mouse gene Mobp, encoding the family of myelin-associated oligodendrocytic basic proteins (MOBP), to facilitate elucidation of its genomic organization and regulation. We report near complete sequence analysis of the Mobp gene (>11 kb), including complete sequence of all exons and their associated splice junctions. The Mobp gene comprises eight discrete exons and encompasses a genomic region in excess of 15 kb. We provide a definitive analysis of the alternative splicing events and exon usage required in the generation of the reported splice variants of Mobp transcripts. We identify sequences corresponding to the coding regions of all reported protein isoforms. Consequently, we demonstrate that sequence regions, predicted to encode unique portions of two putative protein isoforms in the rat (MOBP 71 and MOBP 99), are not fully conserved between the rat and the mouse: we predict that the mouse equivalents are two distinct polypeptides of 73 amino acids, MOBP73A and MOBP73B, respectively. We have analyzed sequence from 63 oligo-capped, cloned cDNA fragments and identify six transcription start points associated with the Mobp gene at postnatal day 26. This study provides the platform for a more detailed analysis of the function of the Mobp gene product and subsequent evaluation of its possible involvement in known neuropathies.

Myelin-associated oligodendrocytic basic protein (MOBP) is a newly recognized major CNS-specific myelin component, which is present throughout compact myelin. It has been suggested that MOBP may play a role in compacting or stabilizing the myelin sheath (Yamamoto *et al.,* 1994; Holz *et al.,* 1996; Montague *et al.,* 1997), possibly by binding the negatively charged acidic phospholipids of the cytoplasmic membrane.

Six splice variants have been identified which are predicted to encode five different protein isoforms of MOBP (Yamamoto *et al.,* 1994; Holz *et al.,* 1996). In the rat, these are predicted to be 69, 71, 81, 99, and 170 amino acids long, respectively, proline rich, and basic. Though all isoforms thus far identified (Yamamoto *et al.,* 1994; Holz *et al.,* 1996; Montague *et al.,* 1997) have amino acid residues 1–68 in common, they differ in the length and polarity of their respective C-terminal regions (Holz *et al.,* 1996; Montague

et al., 1997). These protein isoforms are encoded by a single gene, *Mobp.* Six different mRNA splice variants have been reported based on cDNA sequence data (Yamamoto *et al.,* 1994; Holz *et al.,* 1996). There is clear evidence of differential developmental occurrence of mRNA splice variants: two reported splice variants, Mobp 81A and 81B (Holz *et al.,* 1996; Montague *et al.,* 1997), have the same protein coding region, but different 38UTR sequences, and have been shown to localize to different parts of the oligodendrocyte (Montague *et al.,* 1998).

The 58UTR of the *Mobp* transcript is provided by exons 1 and 2. Exons downstream of exons 1 and 2 have been named exons 3–8, corresponding to their linear order in the mouse genome. Amino acid residues 1–68, present in all previously reported MOBP isoforms (Yamamoto *et al.,* 1994; Holz *et al.,* 1996; Montague *et al.,* 1997), were found to be encoded by exon 3. The complex relationship between exons 4 and 8, the six splice variants of the *Mobp* transcript reported in rat and mouse, and the proteins that they encode (Yamamoto *et al.,* 1994; Holz *et al.,* 1996; Montague *et al.,* 1997), is illustrated in Fig. 3 and Tables 1 and 2. Transcript Mobp 69 is produced from exons 3, 5, and 7a/b. Transcript Mobp 170, the only transcript present at an embryonic time-point, is produced from exons 3 and 4. Exon 4 encodes the 102 amino acids unique to the protein isoform MOBP 170 (Montague *et al.,* 1997). MOBP 170 is also known as rOPRP1 (Yamamoto *et al.,* 1994).

Glia. 2005 Apr 1;50(1):80-5. Characterization of the murine splice variant Mobp155: developmental CNS expression pattern and subcellular localization of epitope-tagged protein. Montague P, McCallion AS, Barrie JE, Edgar JM, McLaughlin M, Davies RW, Griffiths IR.
Members of the myelin-associated oligodendrocytic basic protein (MOBP) family constitute the third most abundant protein in CNS myelin. Although MOBP localizes to the major dense line (MDL) of CNS myelin, the function of the individual isoforms is unknown. Alternative splicing of pre-Mobp mRNA gives rise to six characterized splice variants in both the mouse and the rat. These splice variants share a common N-terminal encoded in Mobp exon 3 comprising 68 amino acids. The predicted protein isoforms differ in their C-termini. Sequence analysis of intron 3 revealed the presence of a putative initiation codon followed by an open reading frame (ORF) encoding 53 amino acids that extends in frame into Mobp exon 4 yielding a predicted MOBP isoform comprising 155 amino acids, designated MOBP155. This newly characterized isoform possessing a novel N-terminus shares a common C-terminus with MOBP170. Mobp170 message is detectable at low abundance throughout myelinogenesis. In contrast, the novel splice variant encoding MOBP155 is expressed at modest levels late in CNS development, coincident with the expression of the abundant splice variant, Mobp81A. Immunostaining of Cos7 cells transiently expressing an epitope-tagged MOBP155 suggested that most of the product was translocated to mitochondria. Although Mobp155 and Mobp170 encode a common predicted C-terminus they have different expression profiles and their products are targeted to mitochondria and the nucleus, respectively, in transiently transfected Cos7 cells.

Smn1: Survival motor neuron 1

A single gene in the mouse, duplicated in human to SMN1 and SMN2. SMN1 is deleted or mutated in Spinal Muscular Atrophy. SMN2 compensates poorly because it has a

single nucleotide substitution 6bp within exon 7 (exon 8 in the current assembly) that causes frequent silencing. This appears to be due to creating a better Exon Splicing Silencer sequence for binding and silencing by HnRNPA or G (Kashima… Manley JL 2007 Hum Mol Genetics 16, 3149-3159), although it may also be due to decreased activity of an exon splicing enhancer recognized by the SR protein SF2.

Variant splice products that skip exon 5 (exon 6 in the current assembly are also present in human, and this exon corresponds to the enhanced exon in thunder T cells.

Role of Smn1 is in snRNP assembly and splicing factor processing and assembly in nuclear organelles (Gems).

Gennaralli et al 1995 (Biochem Biophys Res Comm 213, 342-348) demonstrated that approximately 15% of SMN1 transcripts exclude exon 5, as well as a similar proportion that exclude this exon and exon 7. This reduces the coiled domain and potential myristoylation sequences.

Cell, Vol. 80, 155-165, January 13, 1995 Identification and Characterization of a Spinal Muscular Atrophy-Determining Gene Suzie Lefebvre,*'r Lydie B/irglen,*t Sophie Reboullet,* Olivier Clermont*, Philippe Burlet,* Louis Viollet,* Bernard Benichou,* Corinne Cruaud,~ Philippe Millasseau,§ Massimo Zeviani,*ll Denis Le Paslier,§ Jean Fr6zal,* Daniel Cohen,§ Jean Weissenbach,~ Arnold Munnich,* and Judith Melki*
Spinal muscular atrophy (SMA) is a common fatal autosomal recessive disorder characterized by degeneration of lower motor neurons, leading to progressive paralysis with muscular atrophy. The gene for SMA has been mapped to chromosome 5q13, where large-scale deletions have been reported. We describe here the inverted duplication of a 500 kb element in normal chromosomes and narrow the critical region to 140 kb within the telomeric region. This interval contains a 20 kb gene encoding a novel protein of 294 amino acids. An highly homologous gene is present in the centromeric element of 95% of controls. The telomeric gene is either lacking or interrupted in 226 of 229 patients, and patients retaining this gene (3 of 229) carry either a point mutation (Y272C) or short deletions in the consensus splice sites of introns 6 and 7. These data suggest that this gene, termed the survival motor neuron *(SMN)* gene, is an SMA-determining gene.

To demonstrate that the *SMN* gene *(rBCD541)* is indeed a determining gene for SMA, we searched for mutations in the three SMA patients whose *SMN* gene showed no rearrangements. PCR amplification of exon 7 and its flanking regions revealed, in addition to the normal-sized PCR product, a smaller fragment in patient SA (Figure 5a). Direct sequence analysis of the smaller fragment showed a 7 bp deletion in the 3' splice acceptor site of *SMN* intron 6 (Figure 6). Sequence analysis of the normal-sized PCR product detected the sequence specific to the centromeric exon 7 but not that of the *SMN* exon 7 (data not shown). These results demonstrated that patient SA carried an intronic deletion of the *SMN* gene on one mutant allele and lacked the *SMN* exon 7 on the other mutant allele. This intronic deletion was inherited from the mother, and, in contrast with that of her affected child, the allelic fragment was specific to the *SMN* gene, confirming that the unaffected mother is, indeed, heterozygous for the mutation (data not shown). In patient BI, a 4 bp deletion was found in the 5' consensus splice donor site of *SMN* intron 7 (Figure 6). Sequence analysis of the nondeleted amplification product recognized the centromeric exon 7, but not the *SMN* exon 7, suggesting that the other mutant chromosome lacked the *SMN* exon 7 (data not shown). These two mutations were absent in 246 control individuals. Such mutations have been found in other genes

and typically abolish or disrupt splicing (Bottema et al., 1990; Murru et al., 1991). Finally, patient HU was heterozygous for a point mutation at codon 272 (TAT--*TG'I'), changing a tyrosine into a cysteine in the protein (Figure 6). This mutation was absent in 100 normal chromosomes, ruling out rare polymorphisms.

Hum Mol Genet. 2007 Dec 15;16(24):3149-59. hnRNP A1 functions with specificity in repression of SMN2 exon 7 splicing. Kashima T, Rao N, David CJ, Manley JL.
Homozygous deletion or mutation of the survival of motor neuron 1 gene (SMN1) causes spinal muscular atrophy. SMN1 has been duplicated in humans to create SMN2, which produces a low level of functional SMN protein. However, most SMN2 transcripts lack exon 7, resulting in a non-functional protein. A single nucleotide difference near the 5' end of exon 7 largely accounts for SMN2 exon 7 skipping, an effect that has been attributed to loss of an exonic splicing enhancer (ESE) dependent on the SR protein splicing factor ASF/SF2 or to the creation of an exonic splicing silencer (ESS) element that functions by binding of the splicing repressor hnRNP A1. Our earlier experiments favored the latter mechanism and here we provide further evidence supporting the ESS model. We demonstrate that the striking effect of hnRNP A1 depletion on SMN2 exon 7 splicing is specific, as hnRNP A1 depletion has little or no effect on other inefficient splicing events tested, and ASF/SF2 depletion does not affect SMN1/2 splicing. By two different methods, we find a strong and specific interaction of hnRNPA1 with SMN2 exon 7 and only weak and equivalent interactions between ASF/SF2 and other SR proteins with the 5' ends of SMN1 and SMN2 exon 7. Finally, we describe two disease-related exon-skipping mutations that create hnRNP A1 binding sites, but show that splicing can be restored only modestly or not at all by hnRNP A1 depletion. Together our results provide strong support for the idea that SMN2 exon 7 splicing is repressed by an hnRNPA1-dependent ESS, but also indicate that creation of such elements is context-dependent.

Trends Mol Med. 2006 Mar;12(3):113-21. Spinal muscular atrophy: the RNP connection.
Eggert C, Chari A, Laggerbauer B, Fischer U.
Degenerated motor neurons in the spinal cord are the pathological hallmark of spinal muscular atrophy (SMA). SMA is caused by mutations in the ubiquitously expressed survival motor neuron 1 (SMN1) gene, which lead to reduced levels of functional SMN protein. Many different functions have been assigned to SMN, including assembly of ribonucleoproteins (RNPs), splicing, transcription and axonal mRNA transport. Recently, tissue from SMA patients and animal models has been used to determine which function of SMN is affected in SMA patients. A surprising picture has emerged: the impaired assembly of RNP subunits of the spliceosome seems to be responsible for SMA pathogenesis. Here, we present a model of how this defect might cause motor-neuron degeneration and consider potential therapies.

Kinesin family 21a: KIF21a

Confirmed alternative splice events at 3’end of transcript (also in human in exons 29-31), at the junction between the segment encoding the stalk and the seven WD40 domains predicted to determine cargo specificity.

J Cell Biol. 1999 May 3;145(3):469-79. Novel dendritic kinesin sorting identified by different process targeting of two related kinesins: KIF21A and KIF21B. Marszalek JR, Weiner JA, Farlow SJ, Chun J, Goldstein LS.
Neurons use kinesin and dynein microtubule-dependent motor proteins to transport essential cellular components along axonal and dendritic microtubules. In a search for new kinesin-like proteins, we identified two neuronally enriched mouse kinesins that provide insight into a unique intracellular kinesin targeting mechanism in neurons. KIF21A and KIF21B share colinear amino acid similarity to each other, but not to any previously identified kinesins outside of the motor domain. Each protein also contains a domain of seven WD-40 repeats, which may be involved in binding to cargoes. Despite the amino acid sequence similarity between KIF21A and KIF21B, these proteins localize differently to dendrites and axons. KIF21A protein is localized throughout neurons, while KIF21B protein is highly enriched in dendrites. The plus end-directed motor activity of KIF21B and its enrichment in dendrites indicate that models suggesting that minus end-directed motor activity is sufficient for dendrite specific motor localization are inadequate. We suggest that a novel kinesin sorting mechanism is used by neurons to localize KIF21B protein to dendrites since its mRNA is restricted to the cell body.

From Nature Genetics FEOM paper: Nat Genet. 2003 Dec;35(4):318-21. Heterozygous mutations of the kinesin KIF21A in congenital fibrosis of the extraocular muscles type 1 (CFEOM1). Yamada K, Andrews C, Chan WM, McKeown CA, Magli A, de Berardinis T, Loewenstein A, Lazar M, O'Keefe M, Letson R, London A, Ruttum M, Matsumoto N, Saito N, Morris L, Del Monte M, Johnson RH, Uyama E, Houtman WA, de Vries B, Carlow TJ, Hart BL, Krawiecki N, Shoffner J, Vogel MC, Katowitz J, Goldstein SM, Levin AV, Sener EC, Ozturk BT, Akarsu AN, Brodsky MC, Hanisch F, Cruse RP, Zubcov AA, Robb RM, Roggenkäemper P, Gottlob I, Kowal L, Battu R, Traboulsi EI, Franceschini P, Newlin A, Demer JL, Engle EC.
Congenital fibrosis of the extraocular muscles type 1 (CFEOM1; OMIM #135700) is an autosomal dominant strabismus disorder associated with defects of the oculomotor nerve. We show that individuals with CFEOM1 harbor heterozygous missense mutations in a kinesin motor protein encoded by KIF21A. We identified six different mutations in 44 of 45 probands. The primary mutational hotspots are in the stalk domain, highlighting an important new role for KIF21A and its stalk in the formation of the oculomotor axis.
…whose phenotypes map to *FEOM1* for mutations in transcripts from this region. After excluding several candidates, we examined two adjacent partial transcripts, *FLJ20052* and *KIAA1708*, that share similarity to mouse *Kif21a*6. Combining database sequence with that generated from human fetal brain cDNA by RT-PCR, we determined the cDNA and genomic organization of human *KIF21A*. The *KIF21A* open reading frame is 5,022 bp in length, comprises 38 exons with alternative splicing of exon 12 and exons 29–31, encompasses a genomic region of ~150 kb and is predicted to encode a protein of 1,674 amino acids. Kinesins and dyneins are molecular motors responsible for microtubule-dependent transport of cargo; in neurons, they are responsible for anterograde and retrograde axonal transport, respectively. KIF21A, KIF21B and KIF4 comprise one of the □14 classes of human kinesins9, and their predicted structures are similar to that of classical kinesin, with an N-terminal motor domain that interacts with the microtubule track, a central coiled-coil stalk and a C-terminal tail that interacts with the transported cargo. Mouse *Kif21a* was found to be enriched in adult neural tissues6. Notably, Kif4 transports membranous organelles containing L1, a cell adhesion molecule important to axonal formation and extension in young neurons10.

Glutamate receptor interacting protein 1: GRIP1

Confirmed two alternative promoters and first exons, Grip1a and Grip1b, conferring different N-terminal sequences with 2 and one Cysteine respectively. Only Grip1b appears to be palmitoylated. This is the start that is enhanced in thunder T cells.

Grip1 interacts also with Ephrin-B2.

Maya Yamazakia, Masahiro Fukayab, Manabu Abea, Kanju Ikenoa, Toshikazu Kakizakia, Masahiko Watanabeb, Kenji Sakimuraa,c,* (2001) Differential palmitoylation of two mouse glutamate receptor interacting protein 1 forms with different N-terminal sequences. *Neuroscience Letters* 304:81-84

Glutamate receptor interacting protein (GRIP) is a member of the PDZ domain-containing protein family that is localized in the postsynaptic density area. This protein has been reported to interact speci®cally with the C-termini of AMPA-selective glutamate receptor channel subunits, GluRa2 and GluRa3 through its PDZ domains. To clarify the physiological functions of GRIP, we cloned mouse GRIP1, and found that there are three sites for alternative splicing and two putative translational start codons by characterizing GRIP1 cDNA clones and reverse transcription-polymerase chain reaction products. Metabolic labeling of COS-7 cells expressing two N-terminal GRIP1 proteins demonstrated that these proteins differed in their pattern of palmitoylation. These findings suggested that the molecular diversity of GRIP1underlies the localization and functional heterogeneity of this protein.

Glutamate receptor channels mediate most of the fast excitatory synaptic transmission in the central nervous system. At excitatory synapses, ionotropic glutamate receptors such as a-amino-3-hydroxy-5-methyl-4-isoxazole propionic acid (AMPA)/kainate type and N-methyl-d-aspartate (NMDA) type are highly concentrated in the postsynaptic density (PSD). It has recently been proposed that protein±protein interactions of some cytosolic proteins with the C-termini of glutamate receptor subunits are involved in anchoring/clustering receptors at the PSD and coupling the receptors to cytoplasmic signaling molecules. The most common type of such interactions is mediated between PDZ domains of a binding protein and a short amino acid motif at the carboxyl end of the receptor. AMPA receptor subunits GluRa2 (GluR2) and GluRa3 (GluR3) (C-terminal sequence-ESVKI) bind specifically to PDZ proteins such as glutamate receptor interacting protein (GRIP) [5], AMPA receptor-binding protein (ABP) [15], and a protein interacting with C kinase 1 (PICK1) [4,19]. ABP resembles GRIP in its primary sequence and has been reported to be a shorter splice variant of ABP-L/GRIP2 [2,6,18]. Moreover, many of the scaffold proteins in synapses such as the PSD-95/SAP90 family have been reported to have alternative splicing variants [1,9,10,13,14]. The molecular diversity of these proteins may play important roles in synaptic targeting or signal transduction [1,9,10,13,14].

Erik I. Charych, Rongwen Li, David R. Serwanski, Xuejing Li, Celia P. Miralles, Noelia Pinal and Angel L. De Blas. Identification and characterization of two novel splice forms of GRIP1 in the rat brain. *Journal of Neurochemistry*, 2006, 97, 884–898.

The glutamate receptor interacting protein (GRIP) family comprises highly homologous multiple PDZ domain-containing proteins that are encoded by the GRIP1 (Dong et al. 1997, 1999) and GRIP2 (Bruckner et al. 1999; Dong et al. 1999; Wyszynski et al. 1999) genes. GRIP1 was originally reported as a 1112 amino acid polypeptide containing 7 PDZ domains (Dong et al. 1997). GRIP1 seems to play a role in the targeting and postsynaptic

localization of AMPA receptors (Bruckner et al. 1999; Dong et al. 1999; Li et al. 1999; Osten et al. 2000; Song and Huganir 2002) and in the activity-dependent synaptic reorganization of AMPA receptors during long term depression (Song and Huganir 2002; Bredt and Nicoll 2003). GRIP–/– mouse mutants have shown that GRIP1 is required for normal cell–matrix interactions during embryonic development (Bladt et al. 2002; Takamiya et al. 2004). GRIP1 also plays an important role in dendrite morphogenesis (Hoogenraad et al. 2005), and we have recently reported that GRIP1 is present at GABAergic synapses and in the intact brain (Charych et al. 2004a; Li et al. 2005a). In addition, GRIP1 has been shown to interact with the GABAAR interacting protein GABARAP via PDZ domains 4–6 (Kittler et al. 2004). GRIP1 exhibits a high degree of molecular heterogeneity resulting from alternative splicing (Fig. 1). Alternate 18 and 19 amino acid N-terminal peptides were shown in mouse to give rise to GRIP1a and GRIP1b, respectively, such that GRIP1b can be palmitoylated and GRIP1a cannot (Yamazaki et al. 2001). Consequently, GRIP1b is associated with membranes whereas GRIP1a is localized to the cytoplasm (Yamazaki et al. 2001). Alternative splicing of GRIP1 also gives rise to a 90 kDa short form of GRIP1, called DIP2 (DLX

Intersectin 1: Itsn1

Genomics. 1998 Nov 1;53(3):369-76. Two isoforms of a human intersectin (ITSN) protein are produced by brain-specific alternative splicing in a stop codon. Guipponi M, Scott HS, Chen H, Schebesta A, Rossier C, Antonarakis SE.
Using selected trapped exons with homology to specific protein domains, we identified a new full-length cDNA encoding a protein containing many motifs for protein-protein interactions. There are two major mRNA transcripts, a ubiquitously expressed mRNA of 5.3 kb and a brain-specific transcript of approximately 15 kb, encoding proteins of 1220 and 1721 amino acids, respectively. The stop codon of the ORF of the shorter transcript is split between adjacent exons. In brain tissues the last exon of the short transcript is skipped, and an alternative downstream exon, the first of several additional, is used to produce the 15-kb mRNA. The putative human protein is highly homologous to Xenopus intersectin (81% identical) and to Drosophila dynamin-associated protein, Dap160 (31% identical) and was termed intersectin (ITSN). Both human proteins contain five SH3 (Src homology 3) domains, two EH (Eps15 homology) domains, and an alpha-helix-forming region. The brain-specific long transcript encodes for three additional domains: a GEF (guanine-nucleotide exchange factors), a PH (pleckstrin homology), and a C2 domain. The Drosophila homologue is associated with dynamin, a protein family involved in the endocytic pathway and/or synaptic vesicle recycling. The structure of the human ITSN protein is consistent with its involvement in membrane-associated molecular trafficking and signal transduction pathways. The human ITSN gene has been mapped to 21q22. 1-q22.2 between markers D21S319 and D21S65, and its importance in Down syndrome and monogenic disorders is currently unknown.

J Biol Chem. 1999 May 28;274(22):15671-7. Splice variants of intersectin are components of the endocytic machinery in neurons and nonneuronal cells. Hussain NK, Yamabhai M, Ramjaun AR, Guy AM, Baranes D, O'Bryan JP, Der CJ, Kay BK, McPherson PS.
We recently identified and cloned intersectin, a protein containing two Eps15 homology (EH) domains and five Src homology 3 (SH3) domains. Using a newly developed intersectin antibody, we demonstrate that endogenous COS-7 cell intersectin localizes to

clathrin-coated pits, and transfection studies suggest that the EH domains may direct this localization. Through alternative splicing in a stop codon, a long form of intersectin is generated with a C-terminal extension containing Dbl homology (DH), pleckstrin homology (PH), and C2 domains. Western blots reveal that the long form of intersectin is expressed specifically in neurons, whereas the short isoform is expressed at lower levels in glia and other nonneuronal cells. Immunofluorescence analysis of cultured hippocampal neurons reveals that intersectin is found at the plasma membrane where it is co-localized with clathrin. Ibp2, a protein identified based on its interactions with the EH domains of intersectin, binds to clathrin through the N terminus of the heavy chain, suggesting a mechanism for the localization of intersectin at clathrin-coated pits. Ibp2 also binds to the clathrin adaptor AP2, and antibodies against intersectin co-immunoprecipitate clathrin, AP2, and dynamin from brain extracts. These data suggest that the long and short forms of intersectin are components of the endocytic machinery in neurons and nonneuronal cells.

From: *Genomics*. 2004 Jul;84(1):106-13. Alternative splicing of mammalian Intersectin 1: domain associations and tissue specificities. Tsyba L, Skrypkina I, Rynditch A, Nikolaienko O, Ferenets G, Fortna A, Gardiner K.
The Intersectin 1 (ITSN1) protein functions in clathrin-mediated endocytosis and in MAP kinase signaling. The complex domain structure comprises two EH and five SH3 domains in the short isoform, plus RhoGEF, pleckstrin, and putative calcium-interaction domains in the long isoform. Alternative splicing of exon 20, affecting the SH3A domain, has been shown in rat and that of exons 25 + 26, affecting the SH3C domain, has been shown in human and rat. Here we report 7 novel splice variants of the human and mouse ITSN1 genes and demonstrate conservation of alternative splicing affecting SH3A and SH3C in mouse. The novel variants encode transcripts with altered EH domain spacing and RhoGEF domain structure and possible targets of nonsense-mediated decay. Eight and 16 protein variants of the short and long ITSN1 isoforms, respectively, are predicted. These isoforms likely serve to modulate the many complex protein interactions and functions of ITSN1.

Intersectin 1 (ITSN1) is an evolutionarily conserved, multidomain protein that functions in clathrin-associated endocytosis and as a mediator in MAPK signaling pathways [1]. The major ITSN1 transcripts described in mammals are the short form, ITSN1-s, which is ubiquitously expressed, and the long form, ITSN1-l, which shows neuron-specific expression [2 and 3]. Both short and long forms of ITSN1 contain two Esp15 homology domains (EH1 and EH2) near the amino-terminus, followed by a coiled-coil domain enriched in lysine, leucine, glutamate, arginine, and glutamine (the KLERQ region) and five Src-homology 3 (SH3) domains (SH3A, B, C, D, and E). ITSN1-l contains a carboxy-terminal extension with a Dbl-homology (DH), or RhoGEF, domain, a pleckstrin homology (PH) domain, and a putative calcium-dependent (C2) interaction domain. A number of functions have been demonstrated for ITSN1 and several are modulated by or involve multiple domains of the protein. For example, expression in tissue culture cells of a cDNA containing the two EH domains activates the transcription factor Elk, but the level of this activation is lessened if the cDNA construct also encodes the KLERQ and SH3 domains, i.e., the complete ITSN1-s form [4]. The DH domain has guanine nucleotide exchange activity for Cdc42 [5]. cDNA constructs encoding only the DH or the DH + PH domains have greater activity than the complete ITSN1-l form, implying modulation by the EH, KLERQ, or SH3 domains. In addition, ITSN1-l activity for Cdc42 is enhanced by binding of N-WASP (neuronal Wiskott–Aldrich syndrome protein) to the

SH3 domains; binding efficiency is variable among domains with SH3A > E > C [5]. The ITSN1 SH3A, B, and D domains also bind CdGAP, a GTPase-activating protein that functions to inhibit Rac and Cdc42. Binding to ITSN1 inhibits CdGAP activity [6], thereby possibly increasing Cdc42 activity. The ITSN1 SH3A domain competes with Grb2 for binding to SOS (Son-of-Sevenless), a guanine nucleotide exchange factor for Ras. Overexpression of cDNA constructs encoding ITSN1 SH3 domains blocks Ras activation [7 and 8]. Additional ITSN1 protein interactions include several endocytosis-related proteins binding to the EH domains, SNAP25 (a protein involved in exocytosis) binding to the KLERQ region, the endocytosis protein dynamin and the phosphoinositol phosphatase synaptojanin binding to subsets of the SH3 domains, and phosphoinositol phosphates binding to the PH domain [9, 10, 11 and 12].

Using primers designed to amplify across exon 6 of the mouse Itsn cDNA, RT-PCR of mRNA from adult mouse tissues generated two products: a predominant band at 604 bp corresponding to the published Itsn messages and a second band at 493 bp (Fig. 2A). Cloning and sequencing of the novel PCR product from lung identified a deletion of 111 bp that truncated exon 6 without altering the open reading frame. This transcript, termed isoform 4, was detected in all mouse tissues tested; expression levels are relatively low, requiring hybridization for visualization. A human transcript corresponding to isoform 4 is detected in human fetal tissues (Fig. 2B) and is also found as a NEDO cDNA obtained from placenta (Accession No. AK027846).

Comparison of human and mouse genomic DNA with ITSN1 exon 6 sequence revealed that deletion of the 111 nucleotides reflects the use of a consensus splice site internal to exon 6, creating a new 5′ end. This truncation deletes 37 amino acids between the EH1 and the EH2 domains, decreasing the interdomain spacing from 105 amino acids to 68.

RT-PCR analysis of mouse Itsn using primers within exons 5 and 16 identified a product lacking exons 6–14 (isoform 3) (Fig. 2C), which encode the second EH domain and the part of the coiled-coil KLERQ region. This event shifts the open reading frame and introduces a stop codon truncating the protein at amino acid 133, after inclusion of an additional 15 residues (SANAWKTYSRETDTQ). This transcript is seen at low levels in brain, liver, and lung.

From: *Biochim Biophys Acta.* 2001 Oct 31;1521(1-3):1-11. The human intersectin genes and their spliced variants are differentially expressed. Pucharcos C, Casas C, Nadal M, Estivill X, de la Luna S.
Human intersectins (ITSN1 and ITSN2) are members of a conserved family of proteins involved in clathrin-mediated endocytosis. A short and a long isoform with different protein domain compositions have been described for both human intersectins. Here, we have resolved the exon/intron structure of the ITSN2 gene to explain the genomic origin of its alternatively spliced transcripts. Comparison of the two ITSN human genes shows a high level of conservation in their genomic organization, including the main alternative splicing events. An extensive tissue expression analysis of the two predominant transcripts as well as other minor variants shows that ITSN expression is under tissue and developmental controls. Their differential expression is made more evident when the expression of both intersectins is studied by in situ hybridization in mouse brain.

Clathrin-mediated endocytosis at the plasma membrane of eukaryotic cells is a specialized mechanism involved in the traffic of nutrients and receptor–ligand complexes

in all the cells. In neurons, this process also plays an essential role in the recycling of synaptic vesicles. Up to now, more than 20 proteins have been identified to participate in the molecular machinery of clathrin-mediated endocytosis (reviewed in [1, 2, 3 and 4]). These molecules interact among themselves through a limited set of protein domains to form intricate macromolecular complexes. The uptake process begins with the recruitment of the heterotetramer adapter protein 2 (AP-2) to the plasma membrane that drives the assembly of clathrin to the nascent coat pit. AP-2 binds to accessory proteins including Eps15 [5 and 6], epsin [7], amphiphysin heterodimers [8 and 9], auxilin [10] and AP180 [11], promoting the growth of clathrin-coated vesicles (CCVs). Lipid-modifying enzymes such as endophilins [12] and the phosphatase synaptojanin [13] are also involved in vesicle formation. Finally, it is thought that CCV fission is mediated by dynamin, although the molecular mechanism is still a matter of controversy [14 and 15]. Dynamin interacts with several SH3 (Src 3 homology) domain-containing proteins including amphiphysin I and II, endophilin, syndapin/PACSIN and intersectin (reviewed in [1 and 16. S.L. Schmid, M.A. McNiven and P. De Camilli. Curr. Opin. Cell Biol. 10 (1998), pp. 504–512.

Intersectin 1 is a protein involved in clathrin-mediated endocytosis which has been identified in humans [17 and 18], rat (named EHSH1 [19]), mouse (named Ese1 [20]), Xenopus laevis [21] and Drosophila melanogaster (named Dap160 [22]). In humans, intersectin 1 exists as two main isoforms, a short one composed of two EH (Eps15 homology) domains, a central coiled-coil region and five consecutive SH3 domains, and a long isoform with an extended carboxy-end region composed of a DH (Dbl homology), a PH (pleckstrin homology) and a C2 domain [17 and 18]. Intersectin has been proposed as a scaffolding protein due to its ability to bind several endocytic-related proteins. Thus, the EH domains bind epsins [20 and 21] and SCAMP1 [23], the coiled-coil region interacts with the central region of Eps15 and Eps15R [20], and with the exocytosis-related proteins SNAP23 and SNAP25 [19], and the SH3 domains interact with dynamin [19, 20 and 21] and synaptojanin [19 and 21]. Recently, intersectin 1 has been involved in signal transduction pathways since one of its SH3 domains has the ability to interact with the Ras exchange factor mSos1 [24], and it can regulate mitogenic signalling pathways [25 and 26], providing a link between endocytosis and cell signalling.

Intersectin 1 is highly similar to intersectin 2, described so far in humans and mouse [20 and 27]. Intersectin 2 has two main isoforms produced by alternative splicing giving a short and a long isoform with the same multimodular structure as intersectin 1. Both intersectin 1 and 2 show a subcellular distribution similar to other components of the endocytic machinery, and their overexpression blocks endocytosis [27]. Here we report the genomic organization of ITSN2, the characterization of some spliced variants of intersectin 1 and 2, their expression profiles and an extensive comparative expression study of both genes, including in situ hybridization analysis of mouse brains.

Ptprt = PTPRrho = LOC666183

BMC Genomics. 2004 Feb 11;5(1):14. Genomic structure and alternative splicing of murine R2B receptor protein tyrosine phosphatases (PTPkappa, mu, rho and PCP-2).
Besco J, Popesco MC, Davuluri RV, Frostholm A, Rotter A.
BACKGROUND: Four genes designated as PTPRK (PTPkappa), PTPRL/U (PCP-2), PTPRM (PTPmu) and PTPRT (PTPrho) code for a subfamily (type R2B) of receptor protein tyrosine phosphatases (RPTPs) uniquely characterized by the presence of an N-

terminal MAM domain. These transmembrane molecules have been implicated in homophilic cell adhesion. In the human, the PTPRK gene is located on chromosome 6, PTPRL/U on 1, PTPRM on 18 and PTPRT on 20. In the mouse, the four genes ptprk, ptprl, ptprm and ptprt are located in syntenic regions of chromosomes 10, 4, 17 and 2, respectively. RESULTS: The genomic organization of murine R2B RPTP genes is described. The four genes varied greatly in size ranging from approximately 64 kb to approximately 1 Mb, primarily due to proportional differences in intron lengths. Although there were also minor variations in exon length, the number of exons and the phases of exon/intron junctions were highly conserved. In situ hybridization with digoxigenin-labeled cRNA probes was used to localize each of the four R2B transcripts to specific cell types within the murine central nervous system. Phylogenetic analysis of complete sequences indicated that PTPrho and PTPmu were most closely related, followed by PTPkappa. The most distant family member was PCP-2. Alignment of RPTP polypeptide sequences predicted putative alternatively spliced exons. PCR experiments revealed that five of these exons were alternatively spliced, and that each of the four phosphatases incorporated them differently. The greatest variability in genomic organization and the majority of alternatively spliced exons were observed in the juxtamembrane domain, a region critical for the regulation of signal transduction. CONCLUSIONS: Comparison of the four R2B RPTP genes revealed virtually identical principles of genomic organization, despite great disparities in gene size due to variations in intron length. Although subtle differences in exon length were also observed, it is likely that functional differences among these genes arise from the specific combinations of exons generated by alternative splicing.

Common to all Type 2 RPTPs is an extracellular segment containing a combination of multiple fibronectin and immunoglobulin (Ig)-like domains, and a single transmembrane region. The intracellular region contains a membrane proximal juxtamembrane domain, followed by a catalytically active tyrosine phosphatase domain and a second inactive domain. Type 2 RPTPs have been further subdivided into two distinct classes (R2A and R2B). Genes in the R2B class are differentiated from the R2A class by an additional MAM (Meprin/ A5/PTP mu) domain at the N-terminus [5]. In addition to a putative role in signal transduction, R2B molecules have cell adhesive properties [6]. The four murine R2B genes (ptprk/PTPκ, ptprl/PCP-2, ptprm/PTPμ, and ptprt/PTPρ) are located on mouse chromosomes 10, 4, 17, and 2, respectively.

Three major regions were delineated, each with varying degrees of sequence identity: Exons 2–13 comprised the extracellular segment (MAM, Ig and four fibronectin (FN) type III domains), exon 14–18 (juxtamembrane region), and exons 19–32 (two phosphatase domains).
PCR experiments showed that five of the eight tested exons (14, 16, 17a, 20a, and 22a) were alternatively spliced. Exons 7 and 8 were present and exon 28a was absent in all R2B transcripts tested. All but one of the alternatively spliced exons (14) was located in the R2B intracellular segment. Exon 14 preceded the transmembrane region; exons 16 and 17a encoded intracellular juxtamembrane sequences, and the last two exons (20a, and 22a) encoded portions of the catalytically active, first phosphatase domain.

Each of the four R2B genes expressed in the brain used the five alternatively spliced exons in a different combination: In PTPρ transcripts, exon 17a and 20a were absent, and exons 14, 16, and 22a were alternatively spliced (Figure 11). In PTPμ transcripts, exons 14, 16, 20a and 22a were absent; exon 17a was present and not alternatively spliced. The alternative use of two 5' splice consensus sites resulted in the transcription of an

additional 58 bp of the intron between exons 13 and 15 (Figure 12). In PTPκ mRNA, exons 14 and 22a were absent, and exons 16, 17a and 20a were alternatively spliced (Figure 13).

The high frequency of alternatively spliced exons in the R2B juxtamembrane segment suggests that the region has highly specialized functions. The importance of alternatively spliced exons has been well documented for the closely related Type 2 RPTP, LAR, in which a small (27 bp) alternatively spliced exon (LASE-c) was identified in the fifth FN-III domain [31]. Subsequently, a 33 bp exon (LASE-a), was identified in the intracellular juxtamembrane region [32]. LASE-a, which was shown to be brain specific and developmentally regulated, was present in cell bodies of cultured granule cells, but was absent in neurites. Conversely, the LASE-c isoform was absent in cell bodies and present in neurites. Using in vitro ligand binding assays, the laminin-nidogen extracellular matrix complex was identified as a ligand for LAR, specifically interacting with the fifth FN-III domain [33]. When LAR bound the laminin-nidogen complex, cells formed long processes. Inclusion of the alternatively spliced 27 bp LASE-c exon disrupted this binding, causing changes in cell morphology. These studies imply a role for alternatively spliced exons in neurite extension through modification of cell adhesion.

BMC Genomics. 2001;2(1):1. Genomic organization and alternative splicing of the human and mouse RPTPrho genes. Besco JA, Frostholm A, Popesco MC, Burghes AH, Rotter A.
BACKGROUND: Receptor protein tyrosine phosphatase rho (RPTPrho, gene symbol PTPRT) is a member of the type IIB RPTP family. These transmembrane molecules have been linked to signal transduction, cell adhesion and neurite extension. The extracellular segment contains MAM, Ig-like and fibronectin type III domains, and the intracellular segment contains two phosphatase domains. The human RPTPrho gene is located on chromosome 20q12-13.1, and the mouse gene is located on a syntenic region of chromosome 2. RPTPrho expression is restricted to the central nervous system. RESULTS: The cloning of the mouse cDNA, identification of alternatively spliced exons, detection of an 8 kb 3'-UTR, and the genomic organization of human and mouse RPTPrho genes are described. The two genes are comprised of at least 33 exons. Both RPTPrho genes span over 1 Mbp and are the largest RPTP genes characterized. Exons encoding the extracellular segment through the intracellular juxtamembrane 'wedge' region are widely spaced, with introns ranging from 9.7 to 303.7 kb. In contrast, exons encoding the two phosphatase domains are more tightly clustered, with 15 exons spanning approximately 60 kb, and introns ranging in size from 0.6 kb to 13.1 kb. Phase 0 introns predominate in the intracellular, and phase 1 in the extracellular segment. CONCLUSIONS: We report the first genomic characterization of a RPTP type IIB gene. Alternatively spliced variants may result in different RPTPrho isoforms. Our findings suggest that RPTPrho extracellular and intracellular segments originated as separate modular proteins that fused into a single transmembrane molecule during a later evolutionary period.

Brain Res Mol Brain Res. 1998 Sep 18;60(1):1-12. LAR tyrosine phosphatase receptor: proximal membrane alternative splicing is coordinated with regional expression and intraneuronal localization. Honkaniemi J, Zhang JS, Yang T, Zhang C, Tisi MA, Longo FM.
Examination of null-mutant Drosophila and Leukocyte Common Antigen-Related (LAR)-deficient transgenic mice has demonstrated that the LAR protein tyrosine phosphatase (PTP) receptor promotes neurite outgrowth. In the absence of known ligands, the mechanisms by which LAR-type PTP receptors are regulated are unknown. We

hypothesized that an alternatively spliced eleven amino acid proximal membrane segment of LAR (LAR alternatively spliced element-a; LASE-a) contributes to regulation of LAR function. Human, rat and mouse LAR cDNA sequences demonstrated that the predicted eleven amino acid inserts in rat and mouse are identical and share nine of eleven residues with the human insert. LASE-a splicing led to the introduction of a Ser residue into LAR at a position analogous to Ser residues undergoing regulated phosphorylation in other PTPs. In-situ studies revealed increasingly region-specific expression of LASE-a containing LAR transcripts during postnatal development. RT-PCR analysis of cortical and hippocampal tissue confirmed that the proportion of LAR transcripts containing LASE-a decreases during development. Immunostaining of cultured PC12 cells, cerebellar granule neurons, dorsal root ganglia and sciatic nerve sections with antibody directed against the LASE-a insert demonstrated signal in cell bodies but little if any along neurites. In contrast, staining with antibody directed to a separate domain of LAR showed accumulation of LAR along neurites. The findings that LASE-a splicing is conserved across human, rat and mouse, that the LASE-a insert introduces a Ser at a site likely to be targeted for regulated phosphorylation and that developmentally regulated splicing is coordinated with specific regional and intraneuronal localization point to important novel potential mechanisms regulating LAR-type tyrosine phosphatase receptor function in the nervous system.

J Cell Biol. 1995 Feb;128(3):415-31. LAR tyrosine phosphatase receptor: alternative splicing is preferential to the nervous system, coordinated with cell growth and generates novel isoforms containing extensive CAG repeats. Zhang JS, Longo FM.
Receptor-linked tyrosine phosphatases regulate cell growth by dephosphorylating proteins involved in tyrosine kinase signal transduction. The leukocyte common antigen-related (LAR) tyrosine phosphatase receptor has sequence similarity to the neural cell adhesion molecule N-CAM and is located in a chromosomal region (1p32-33) frequently altered in neuroectodermal tumors. To understand the function of receptor-linked tyrosine phosphatases in neural development, we sought to identify LAR isoforms preferentially expressed in the nervous system and cellular processes regulating LAR alternative splicing. We report here the isolation of a series of rat LAR cDNA clones arising from complex combinatorial alternative splicing, not previously demonstrated for the tyrosine phosphatase-receptor gene family in general. Isoforms included: (a) deletions of the fourth, sixth and seventh fibronectin type III-like domains; (b) an alternatively spliced novel cassette exon in the fifth fibronectin type III-like domain; (c) two alternatively spliced novel cassette exons in the juxtamembrane region; (d) a retained intron in the extracellular region with in-frame stop codons predicting a secreted LAR isoform; and (e) an LAR transcript including an alternative 3' untranslated region containing multiple stretches of tandem CAG repeats up to 21 repeats in length. This number of repeats was in the range found in normal alleles of genes in which expansions of repeats are associated with neurodegenerative disease and the genetic phenomenon of anticipation. RT-PCR and Northern analysis demonstrated that LAR alternative splicing occurred preferentially in neuromuscular tissue in vivo and in neurons compared to astrocytes in vitro and was developmentally regulated. Alternative splicing was also regulated in PC12 cells by NGF, in 3T3 fibroblasts by cell confluence and in sciatic nerve and muscle

subsequent to nerve transection. Western blot analysis demonstrated that alternatively spliced cassette exons result in the presence of corresponding amino acid segments of LAR protein in vivo. These studies suggest specialized functions of LAR isoforms in the nervous system and support our hypothesis that LAR-like tyrosine phosphatase receptors play a role in neural development and regeneration.

Gabbr1

Cytogenet Cell Genet. 2001;92(1-2):116-21. The murine GABA(B) receptor 1: cDNA cloning, tissue distribution, structure of the Gabbr1 gene, and mapping to chromosome 17. Lamp K, Humeny A, Nikolic Z, Imai K, Adamski J, Schiebel K, Becker CM.
GABA (gamma-aminobutyric acid) is a major inhibitory neurotransmitter in the central nervous system (CNS) which activates both ionotropic (GABA(A)/GABA(C)) and metabotropic (GABA(B)) receptor systems. We identified two alternatively spliced cDNA variants of the murine GABA(B) receptor 1 that are predominantly expressed in the CNS. Deduced protein structures are highly homologous to the previously characterized rat and human receptors. Comparison of the genomic structures of mouse and human revealed that alternative splicing occurred at the same position, whereas the mouse gene has an additional 5' exon. Radiation hybrid mapping, combined with database searches, indicated that the GABA(B) receptor gene (Gabbr1) is located on mouse chromosome 17, adjacent to the marker D17Mit24 in a region homologous to human chromosome 6p21.3.

GABAB receptor

Neuron. 2006 May 18;50(4):589-601. Differential compartmentalization and distinct functions of GABAB receptor variants.Vigot R, Barbieri S, Bräuner-Osborne H, Turecek R, Shigemoto R, Zhang YP, Luján R, Jacobson LH, Biermann B, Fritschy JM, Vacher CM, Müller M, Sansig G, Guetg N, Cryan JF, Kaupmann K, Gassmann M, Oertner TG, Bettler B.
GABAB receptors are the G protein-coupled receptors for the main inhibitory neurotransmitter in the brain, gamma-aminobutyric acid (GABA). Molecular diversity in the GABAB system arises from the GABAB1a and GABAB1b subunit isoforms that solely differ in their ectodomains by a pair of sushi repeats that is unique to GABAB1a. Using a combined genetic, physiological, and morphological approach, we now demonstrate that GABAB1 isoforms localize to distinct synaptic sites and convey separate functions in vivo. At hippocampal CA3-to-CA1 synapses, GABAB1a assembles heteroreceptors inhibiting glutamate release, while predominantly GABAB1b mediates postsynaptic inhibition. Electron microscopy reveals a synaptic distribution of GABAB1 isoforms that agrees with the observed functional differences. Transfected CA3 neurons selectively express GABAB1a in distal axons, suggesting that the sushi repeats, a conserved protein interaction motif, specify heteroreceptor localization. The constitutive absence of GABAB1a but not GABAB1b results in impaired synaptic plasticity and hippocampus-dependent memory, emphasizing molecular differences in synaptic GABAB functions.

4. GABAB-receptor isoforms

When it became apparent that probably all GABAB receptors in the vertebrate brain are the sole products of the GABAB(1) and GABAB(2) genes, much attention focused on

subunit isoforms. Many in the field wondered whether isoforms encoded pharmacological differences and accounted for the heterogeneity observed with native GABAB receptors. Rapidly numerous GABAB isoforms were identified (recently reviewed in Ref. 29). A close inspection of GABAB gene structures indicates that not all of these splice variants are real and that some do not occur across different species. A summary of confirmed GABAB isoforms is shown in Figure 2. While in most laboratories mixing and matching of isoforms did not produce GABAB receptors with distinct functional and pharmacological properties, others reported differences that are, however, highly controversial (see sect. IIIA4A). GABAB isoforms may afford the means for a differential subcellular targeting and/or coupling to distinct intracellular signaling pathways. To some extent a coupling to different effector systems could mimic a differential pharmacology and explain some of the differences that were observed with native GABAB receptors.

The distinct ECD in GABAB(1b) results from the presence of an alternative transcription initiation site within the GABAB(1a) intron upstream of exon 6, thereby extending exon 6 at its 5'-end

GABAB(1e) encodes the extracellular ligand-binding domain of the GABAB(1a) subunit. GABAB(1e) is generated by skipping of exon 15 and is detected both in rats and humans. While the GABAB(1e) transcript is a minor component in the CNS, it is very prominent in peripheral tissues. GABAB(1e) is secreted into the culture medium when expressed in transfected mammalian cells. Additionally, it also forms stable heteromeric complexes with GABAB(2) at the plasma membrane.

GABAB(1g) is characterized by an insertion of 124 bp between exon 4 and 5, which generates a frameshift. GABAB(1g) encodes a COOH-terminally truncated polypeptide of 239 amino acid residues of unknown function.

Cadps2

From BMC Neurosci. 2007; 8: 25. Alternative splicing variations in mouse CAPS2: differential expression and functional properties of splicing variants. Tetsushi Sadakata,1 Miwa Washida,1 and Teiichi Furuichi
In this study, we defined 31 exons in the mouse CAPS2 gene and identified six alternative splicing variants, CAPS2a-f. CAPS2a is an isoform lacking exons 22 and 25, which encode part of the Munc13-1-homologous domain (MHD). CAPS2b lacks exon 25. CAPS2c lacks exons 11 and 22. CAPS2d, 2e, and 2f have C-terminal deletions from exon 14, exon 12, and exon 5, respectively. On the other hand, a mouse counterpart of CAPS2Δexon3 was not detected in the mouse tissues tested. CAPS2b was expressed exclusively in the brain, and the other isoforms were highly expressed in the brain, but also in some non-neural tissues. In the brain, all isoforms showed predominant expression patterns in the cerebellum. In the developing cerebellum, CAPS2b showed an up-regulated expression pattern, whereas the other isoforms exhibited transiently peaked expression patterns. CAPS2 proteins were mostly recovered in soluble fractions, but some were present in membrane fractions, except for CAPS2c and 2f, both of which lack the PH domain, suggesting that the PH domain is important for membrane association. In contrast to CAPS2a and 2b, CAPS2c showed slightly decreased BDNF-releasing activity, which is likely due to the Cterminal truncation of the PH domain in CAPS2c.

The Ca2+-dependent activator protein for secretion (CAPS/CADPS) family consists of two members, CAPS1/CADPS1 [1-3] and CAPS2/CADPS2 [4-6]. CAPS1 was shown to play a role in the ATP-dependent priming step of Ca2+-triggered dense-core vesicle (DCV) exocytosis by binding to both phosphatidyl inositol 4,5-bisphosphate and DCVs in response to Ca2+ increase [7,8]; it also interacts with the dopamine receptor [9]. A study of CAPS1-deficient mice suggested that CAPS1 is involved in catecholamine loading of DCVs in embryonic chromaffin cells [10]. On the other hand, CAPS2 is associated with secretory vesicles containing the neurotrophins BDNF and NT-3 in the parallel fiber terminals of cerebellar granule cells, and is involved in the release of BDNF and NT-3 [6]. CAPS2 is widely localized across various brain regions, and shows overlapping distribution patterns with BDNF in many areas [11]. Our recent studies of CAPS2-deficient mice indicated that CAPS2 plays pivotal roles in BDNF release, cellular phenotypes (e.g., neuronal survival and differentiation, synapse structure and function), and behavioral phenotypes (e.g., water-maze spatial learning, anxiety, circadian rhythm, maternal behavior) [12,13].

The expression of an aberrant splicing variant of CAPS2 in some patients with autism has also been reported [12]. The autistic isoform, CAPS2[Delta]exon3, is missing exon 3, which encodes the dynactin 1-interacting domain (DID) [12]. Exogenously expressed CAPS2[Delta]exon3 failed to localize to axons but accumulated in somato-dendritic areas in cultured cerebellar and cortical neurons, suggesting a possible disturbance of local CAPS2-mediated neurotrophin release [12]. In this study, we further explore splicing variations in CAPS2 and identified five additional splice isoforms. We characterized the expression patterns and neurotrophin release properties of these isoforms.

The full-length CAPS2 protein consists of the following functional domains, as shown in Figure 2: a dynactin 1-interacting domain (DID), a C2 domain, a pleckstrin homology (PH) domain and a Munc13-1-homologous domain (MHD). CAPS2b lacks only exon 25, which encodes a part of the MHD. CAPS2a lacks two exons, 22 and 25, which encode part of the MHD sequence. CAPS2c lacks two exons, exon 11 encoding a C-terminal part of the PH domain, and exon 22. In contrast to these three long isoforms (CAPS2a-c), CAPS2d, 2e and 2f have large C-terminal large truncations from exon 14, 12, and 5, respectively; thus, these isoforms have neither the MHD nor the C-tail domain, which mediates the association of CAPS with DCVs [7]. These three short isoforms (CAPS2d-f) have an additional unique exon at their 3' end. The shortest isoform, CAPS2f, has only the first four authentic exons and none of the functional domains identified so far.

All isoforms (CAPS2a-f) have exon 3, which contains the domain interacting with p150Glued [12]. The mouse CAPS2 exon 3 coding sequence is identical to the human CAPS2 exon 3 coding sequence, which is absent in some autistic patients. As shown in Figure 3A, no exon 3 skipping was observed during the development of the cerebellum, from E18 to P56, provided there were sufficient PCR cycles. Moreover, no exon 3 skipping was observed in whole brain, thymus, lung, heart, liver, spleen, kidney, or testis from P7 or P21 mice (Figure 3A). These results suggest that exon 3 skipping does not normally occur in mouse CAPS2.

Otof

Am J Hum Genet. 2000 September; 67(3): 591–600. OTOF encodes multiple long and short isoforms: genetic evidence that the long ones underlie recessive deafness DFNB9.

Yasunaga S, Grati M, Chardenoux S, Smith TN, Friedman TB, Lalwani AK, Wilcox ER, Petit C.
Deafness is one of the most frequent hereditary disorders in humans. The vast majority of genetic prelingual hearing loss is inherited on the autosomal recessive mode (DFNB forms) (for review, see work by Petit [1996]). To date, 26 DFNB loci have been mapped (see Web site for Nonsyndromic Hearing Impairment Autosomal Recessive Loci) and six genes have been identified (Kelsell et al. 1997; Liu et al. 1997; Weil et al. 1997; Li et al. 1998; Wang et al. 1998; Mustapha et al. 1999; Yasunaga et al. 1999). Using a positional cloning strategy combined to a candidate gene approach, we have identified the OTOF gene, encoding otoferlin, as responsible for DFNB9 (MIM 601071) (Yasunaga et al. 1999). The 5-kb OTOF cDNA from a total fetus polyA+ mRNA preparation, encodes a putative 1,230-aa membrane-anchored cytosolic protein containing three predicted C2 domains (Yasunaga et al. 1999) (GenBank AF107403). A single nonsense mutation has been detected in the four analyzed Lebanese families. This mutation is predicted to truncate the 500 C-terminal amino acids of the protein (Yasunaga et al. 1999). OTOF is a member of a mammalian gene family related to the Caenorhabditis elegans gene fer-1 (Achanzar and Ward 1997). This family also comprises DYSF (Bashir et al. 1998; Liu et al. 1998) and MYOF (Davis et al. 2000), encoding dysferlin (2080 aa) and myoferlin (2061 aa), respectively. Both dysferlin and myoferlin are membrane-anchored cytosolic proteins that contain six predicted C2 domains, four of which are expected to bind Ca2+ (Rizo and Südhof 1998). Subsequent detection of a 7-kb otoferlin mRNA by northern blot analysis prompted us to characterize this transcript which had not been detected in our initial study.

Interestingly, several alternatively spliced long forms of otoferlin transcripts were detected in both mice and humans. Firstly, in the murine brain and cochlea, cDNAs were found that resulted from the skipping of exon 6 (fig. 4A). Exon 6 encodes 15 aa located in the N-terminal part of the first inter-C2 domain region (aa 169–184 in the murine cochlear form [GenBank AF183183]). RT-PCR experiments on human brain RNA also revealed a single product that resulted from the skipping of exon 6 (data not shown). Secondly, RT-PCR analysis of the cDNA sequence corresponding to exon 31 in the murine brain and cochlea, revealed two different products resulting from the alternative use of an acceptor splice site within exon 31 (table 2) (fig. 4B). The same results were obtained in the human brain (GenBank AF183186). By contrast, in the human heart, kidney, and total fetus, only the shorter PCR product corresponding to the use of the internal splice site was detected (data not shown). The additional 20 aa encoded by the 5′ part of this exon belong to the fourth inter C2 domain region (aa position 1244–1264 in the murine brain sequence GenBank AF183184, and aa position 1245–1265 in the human brain sequence [GenBank AF183185] [fig. 1]).

Cobl

Cell. 2007 Oct 19;131(2):337-50. Cordon-bleu is an actin nucleation factor and controls neuronal morphology. Ahuja R, Pinyol R, Reichenbach N, Custer L, Klingensmith J, Kessels MM, Qualmann B.
Despite the wealth of different actin structures formed, only two actin nucleation factors are well established in vertebrates: the Arp2/3 complex and formins. Here, we describe a further nucleator, cordon-bleu (Cobl). Cobl is a brain-enriched protein using three Wiskott-Aldrich syndrome protein homology 2 (WH2) domains for actin binding. Cobl promotes nonbundled, unbranched filaments. Filament formation relies on barbed-end

growth and requires all three Cobl WH2 domains and the extended linker L2. We suggest that the nucleation power of Cobl is based on the assembly of three actin monomers in cross-filament orientation. Cobl localizes to sites of high actin dynamics and modulates cell morphology. In neurons, induction of both neurites and neurite branching is dramatically increased by Cobl expression—effects that critically depend on Cobl's actin nucleation ability. Correspondingly, Cobl depletion results in decreased dendritic arborization. Thus, Cobl is an actin nucleator controlling neuronal morphology and development.

Dev Biol. 2003 Oct 1;262(1):16-31. Cordon-bleu is a conserved gene involved in neural tube formation. Carroll EA, Gerrelli D, Gasca S, Berg E, Beier DR, Copp AJ, Klingensmith J.
The axial midline is an important source of patterning and morphogenesis cues in the vertebrate embryo. The midline derives from a small group of cells in the gastrulating embryo, known as "the organizer" in recognition of its ability to organize an entire body plan. The mammalian organizer, the node, gives rise to axial midline structures: the notochord, dorsal foregut, and part of the floor plate of the neural tube. Only some of the genes that direct midline development are known. In this study, we present the complete coding sequence for a novel gene, cordon-bleu (cobl), expressed specifically in the node and its derivatives until organogenesis stages. The deduced sequence does not resemble any gene of known function. However, cobl is widely conserved: apparent orthologs and paralogs are found in many vertebrate species, with several sequence domains of high conservation but unknown function. We find that chicken cordon-bleu is similarly expressed in the node and its derivatives, suggesting functional conservation. We also report the sequence and nonoverlapping expression of a related mouse gene, Coblr1. Finally, we show that cobl interacts with the neurulation gene Vangl2 to facilitate midbrain neural tube closure, demonstrating roles for both cobl and Vangl2 in midbrain neurulation.

Although the sequence of cobl is novel, we have identified homologues in several vertebrate species by database searches. None represents a gene product associated with known functions. The human clone KIAA0633 shows homology with mouse Cobl throughout its length, with an overall sequence identity of 67% and similarity of 86%, reflecting conserved amino acid substitutions (Fig. 2A). The greatest conservation is found at the N-terminal third of the proteins, the region containing the three repeated “KRAP” motifs, with identity of 84% and similarity of 94%. Considerable homology is also found in the last 200 residues, including the 3 WH2 domains (74% identical, 92% conserved). This probable human cobl orthologue was found in an effort to sequence large cDNAs from adult human brain [Ishikawa et al., 1998], but fragments of it have also been sequenced in cDNAs from whole fetus (8–9 weeks), and normal adult lung, liver, and pancreas, tissues where mouse cobl is also expressed.

Human genome sequencing projects localize human cobl at chromosomal position 7p12. The genomic DNA encoding KIAA0633 resides in two sequenced PAC clones in this region ([The Sanger Centre and The Washington University Genome Sequencing Center, 1998]; Accession Nos. AC005535 and AC004414). Although cobl is located near the imprinted gene Grb10 in both human and mouse, neither cobl homolog is imprinted [Hitchins et al., 2002]; data not shown).

Partial cDNAs characterized as expressed sequence tags (ESTs) from a variety of

vertebrate species have also shown homology to cobl. An alignment of several homologous sequences is shown in Fig. 2B, with putative cobl orthologues identified in the human, chicken, and Xenopus genomes. The most significant homology is found in the N-terminal third of the protein, with sequence similarity of 37% for all, excluding an alternatively spliced exon, found only in mouse and human cDNAs to date.

Mice that are homozygous for the coblC101 gene trap do not have any phenotype, in either an outbred or inbred background [Gasca et al., 1995]. We have demonstrated here that wild-type message is still transcribed in coblC101 homozygous mice. The explanation for the paradoxical presence of wild-type message in the absence of wild-type alleles lies in the nature of the gene trap vector used in the gene trap screen that identified cobl [Gasca et al., 1995]. In some portion of coblC101 mRNAs, the gene trap is spliced out, presumably via utilization of a splice acceptor other than that in the vector. Northern blots from homozygous ES cells suggest that wild-type message occurs at significant levels, though less abundant than the fusion transcript [Gasca et al., 1995]. Splicing around gene trap insertions has been observed previously (e.g., [Voss et al., 1998]. In a coblC101/coblC101 animal, the transcripts from the cobl locus are a mix of wild-type and cobl-lacZ fusion messages. Thus, the level of normal cobl transcript is lower in coblC101/coblC101 mice than in heterozygous or wild-type mice, and likely results in lower levels of Cobl protein being made. At least for developmentally regulated genes, reduced expression of wild-type transcript from an allele often makes it a hypomorph [Meyers et al., 1998 and Nagy et al., 1998]. In the case of cobl, it appears that the level of protein resulting from the coblC101 allele is enough to maintain function in the absence of other interacting mutations. However, the new phenotype generated when coblC101 is homozygous in combination with Lp demonstrates that coblC101 is not wild-type in function. Furthermore, there is no phenotype associated with decreased Lp when coblC101 is heterozygous. This suggests that coblC101 is not a gain-of-function allele that acts only in the presence of decreased Vang12 activity. Taken together, these data indicate that coblC101 is a hypomorphic allele. Most likely, it is weakly hypomorphic, since defects have only been seen in the presence of decreased Vang12 and since considerable wild-type message remains in coblC101 homozygotes. The consequences of a null allele remain undetermined at present.

Phenotypic analysis of animals carrying both coblC101 and loop-tail (Vang12Lp) mutant alleles reveals that these genes have a role in neural tube closure in the midbrain. This phenotype is quite evident by E9.5, at a time when cobl is expressed in the floor plate but not more dorsally. Despite the presence of cobl expression in the dorsal neural tube by E11.5, then, this domain does not account for the neural tube closure defect.

SORBS2 = ArgBP2

Arg and c-Abl represent the mammalian members of the Abelson family of non-receptor protein-tyrosine kinases. They interact with the Arg/Abl binding proteins via the SH3 domains present in the carboxy end of the latter group of proteins. This gene encodes the sorbin and SH3 domain containing 2 protein. It has three C-terminal SH3 domains and an N-terminal sorbin homology (SoHo) domain that interacts with lipid raft proteins. The subcellular localization of this protein in epithelial and cardiac muscle cells suggests that it functions as an adapter protein to assemble signaling complexes in stress fibers, and that it is a potential link between Abl family kinases and the actin cytoskeleton.

Alternative splicing results in multiple transcript variants encoding different isoforms.

J Biol Chem. 1997 Jul 11;272(28):17542-50. ArgBP2, a multiple Src homology 3 domain-containing, Arg/Abl-interacting protein, is phosphorylated in v-Abl-transformed cells and localized in stress fibers and cardiocyte Z-disks. Wang B, Golemis EA, Kruh GD.
Arg and c-Abl represent the mammalian members of the Abelson family of protein-tyrosine kinases. A novel Arg/Abl-binding protein, ArgBP2, was isolated using a segment of the Arg COOH-terminal domain as bait in the yeast two-hybrid system. ArgBP2 contains three COOH-terminal Src homology 3 domains, a serine/threonine-rich domain, and several potential Abl phosphorylation sites. ArgBP2 associates with and is a substrate of Arg and v-Abl, and is phosphorylated on tyrosine in v-Abl-transformed cells. ArgBP2 is widely expressed in human tissues and extremely abundant in heart. In epithelial cells ArgBP2 is located in stress fibers and the nucleus, similar to the reported localization of c-Abl. In cardiac muscle cells ArgBP2 is located in the Z-disks of sarcomeres. These observations suggest that ArgBP2 functions as an adapter protein to assemble signaling complexes in stress fibers, and that ArgBP2 is a potential link between Abl family kinases and the actin cytoskeleton. In addition, the localization of ArgBP2 to Z-disks suggests that ArgBP2 may influence the contractile or elastic properties of cardiac sarcomeres and that the Z-disk is a target of signal transduction cascades.
In the present study we describe ArgBP2, the second Arg-binding protein that we isolated using a segment of the Arg COOH-terminal domain as bait in the yeast two-hybrid system. ArgBP2 contains three COOH-terminal SH3 domains, several potential phosphorylation sites for c-Abl, and a serine/threonine-rich region. We show that ArgBP2 associates with Arg and v-Abl in living cells and that the endogenous ArgBP2 protein is phosphorylated on tyrosine residues in v-Abl-transformed NIH3T3 cells. ArgBP2 is widely expressed in human tissues and extremely abundant in heart. In epithelial cells ArgBP2 is located in stress fibers and the nucleus, similar to the reported localization of c-Abl. In cardiac muscle cells ArgBP2 is localized in the Z-disks of sarcomeres. Thus, this study suggests that ArgBP2 functions as an adapter protein to assemble signaling complexes in stress fibers, and that ArgBP2 is a potential link between Abl family kinases and the actin cytoskeleton. In addition, ArgBP2 is the first known component of the Z-disk that associates with signaling cascades. The localization of ArgBP2 in Z-disks thus suggests that ArgBP2 may function to influence the contractile or elastic properties of cardiac sarcomeres, and that the Z-disk may be target of signal transduction cascades.

The largest ArgBP2 insert contained a 250-amino acid open reading frame, followed by a stop codon and 250 base pairs of 39-untranslated sequences. Additional cDNA clones were isolated from a human heart cDNA library by plaque hybridization using an ArgBP2 cDNA insert as the radiolabeled probe. Several cDNA clones were characterized and nucleotide sequence analysis revealed that they represented two distinct classes of ArgBP2 cDNAs (Fig. 1). Analysis of one class of cDNAs (five clones), designated ArgBP2A, indicated that they spanned a total of 5.6 kb. The ArgBP2A sequence encoded a predicted protein of 666 amino acids bordered by in-frame stop codons, indicating that it represented a full-length open reading frame. A second class of cDNAs (two clones), designated ArgBP2B, encoded a partial open reading frame of 618 amino acids followed by a stop codon and 257 base pairs of 39-untranslated sequence. The partial ArgBP2B predicted protein, which began at the 27th codon of ArgBP2A, was identical to the latter predicted protein over most of its sequences, but differed at its midportion and COOH terminus. To obtain additional 59 sequences of ArgBP2B, a reverse transcription PCR

approach was employed using a primer located 59 of the upstream in-frame stop codon of ArgBP2A and two nested primers located in the unique ArgBP2B midregion. Nucleotide sequence analysis of the PCR product indicated that the aminoterminal sequence of ArgBP2B was identical to that of ArgBP2A.

J Biol Chem. 2005 Jun 3;280(22):21483-90. Epub 2005 Mar 22. ArgBP2gamma interacts with Akt and p21-activated kinase-1 and promotes cell survival. Yuan ZQ, Kim D, Kaneko S, Sussman M, Bokoch GM, Kruh GD, Nicosia SV, Testa JR, Cheng JQ.
Akt/protein kinase B is a major cell survival pathway through phosphorylation of proapoptotic proteins Bad and Bax and of additional apoptotic pathways linked to Forkhead proteins glycogen synthase kinase-3beta and ASK1. To further explore the mechanism by which Akt regulates cell survival, we identified an Akt interaction protein by yeast two-hybrid screening. It is highly homologous to ARG-binding protein 2 (ArgBP2) with splicing exon 8 of the coding region of the ArgBP2. As two splicing isoforms (ArgBP2alpha and -beta) of ArgBP2 have been identified (Wang, B., Golemis, E. A., and Kruh, G. D. (1997) J. Biol. Chem. 272, 17542-17550), it was named ArgBP2gamma. ArgBP2gamma contains four Akt phosphorylation consensus sites, a SoHo motif, and three Src homology (SH) 3 domains and binds to C-terminal proline-rich motifs of Akt through its first and second SH3 domains. It also interacts with p21-activated protein kinase (PAK1) via its first and third SH3 domains, indicating the SH3 domains of ArgBP2gamma as docking sites for Akt and PAK1. Akt phosphorylates ArgBP2gamma in vitro and in vivo. Expression of ArgBP2gamma induces PAK1 activity and overrides apoptosis induced by ectopic expression of Bad or DNA damage. Nonphosphorylatable ArgBP2gamma-4A and SH3 domain-truncated mutant ArgBP2gamma inhibit Akt-induced PAK1 activation and reduce Akt and PAK1 phosphorylation of Bad and antiapoptotic function. These data indicate that ArgBP2gamma is a physiological substrate of Akt, functions as an adaptor for Akt and PAK1, and plays a role in Akt/PAK1 cell survival pathway.

J Cell Sci. 2004 May 15;117(Pt 12):2557-68. Epub 2004 May 5.Recruitment of Pyk2 and Cbl to lipid rafts mediates signals important for actin reorganization in growing neurites. Haglund K, Ivankovic-Dikic I, Shimokawa N, Kruh GD, Dikic I.
Protein tyrosine kinase Pyk2 and multifunctional adaptor protein Cbl are implicated in the regulation of the cytoskeleton in several cell types. We report that Pyk2 and Cbl form a signaling complex that is translocated to lipid rafts and is enriched in growth cones of differentiating PC12 cells following growth factor stimulation. We found that Pyk2 and Cbl interacted with the adaptor protein ArgBP2, which also bound to flotillin-1, a component of lipid raft microdomains. These interactions contributed to recruitment of the Pyk2/Cbl complex to lipid raft compartments. In addition, Pyk2, Cbl and ArgBP2 were found co-localized with actin in axons and growth cones of differentiated PC12 cells. Moreover, co-expression of Pyk2, ArgBP2 and Cbl facilitated growth factor-induced formation of lamellipodia at the tip of neurites. Formation of these growth cone lamellipodia was dependent on intact lipid rafts and the Cbl-associated effectors Crk and phosphatidylinositol 3 (PI 3)-kinase. Our results indicate that recruitment of Pyk2/Cbl complexes to lipid rafts participates in growth factor-induced regulation of the actin cytoskeleton in growing neurites.

Mlx

Oncogene. 2000 Jul 6;19(29):3266-77. Mlx, a new Max-like bHLHZip family member: the center stage of a novel transcription factors regulatory pathway? Meroni G, Cairo S, Merla G, Messali S, Brent R, Ballabio A, Reymond A.
The Myc proto-oncogene family members have been identified as the cellular homologs of the transforming oncogene of avian retroviruses. They encode central regulators of mammalian cell proliferation and apoptosis, and they associate with the bHLHZip protein Max to bind specific DNA sequences and regulate the expression of genes important for cell cycle progression. The other family members, Mad1, Mxi1, Mad3, Mad4 and Rox (Mnt) antagonize their activities. The Mads and Rox compete with Myc in heterodimerizing with Max and in binding to the same specific target sequences. These Mads:Max and Rox:Max dimers repress transcription through binding to the mSIN3 corepressor protein and by tethering histone deacetylase-containing complexes to the DNA. In a screen for Rox interactors we isolated Mlx, a bHLHZip protein previously identified in a screen for Mad1 interactors. In the present work we extend the known dimerization partners of Mlx by demonstrating its ability to interact with Rox. Moreover, we show that contrary to previous reports Mlx is able to homodimerize and to bind E-box sequences at low concentration levels. The possible role of Mlx in an emerging regulatory pathway and acting parallel to the Max driven network is discussed.
The basic region recognizes a canonical CANNTG DNA-binding sequence called E-box (Blackwell et al., 1993). The HLH and Zip motifs participate in protein dimerization, a prerequisite for DNA-binding. Max is a widely expressed and stable bHLHZip protein that forms DNA-binding heterodimers with the Myc and Mad proteins (Ayer et al., 1993; Blackwood and Eisenman, 1991; Hurlin et al., 1995b; Prendergast et al., 1991; Zervos et al., 1993). Although Myc-Max and Mad-Max heterodimers bind the same E-box-related DNA sequences, they elicit different transcriptional responses. Reporter genes under the control of E-box elements are activated by Myc-Max, while Mad-Max represses their transcription (Amati et al., 1992; Ayer et al., 1993; Hurlin et al., 1995b; Kretzner et al., 1992). Thus it appears that Myc activates, while Mad represses, genes involved in promoting cellular proliferation (reviewed in Cole and McMahon, 1999; Dang, 1999; Peters and Taparowsky, 1998).

This screening allowed us to identify four new Rox-interactors. In this report we describe the cloning and functional characterization of one of these interactors, a new mammalian bHLHZip protein (see Figure 1a). Due to its sequence, size, overall similarity and function features we named the gene BigMax (see below).

From the two-hybrid screening, we recovered four clones corresponding to two independent and overlapping BigMax cDNAs (clones 630, 640, 692 and 694, see Figure 1a). These cDNAs represent the human homolog of murine transcription factor like protein 4 (TCFL4) (Bjerknes and Cheng, 1996). TCFL4 was cloned from the mouse 17q21 syntenic region where BRCA1 was mapped (Friedman et al., 1994, 1995). However, these clones lack 162 bp of exon 1 and constitute an alternatively spliced form of the human TCFL4 (isoform BigMax-beta, see Figure 1d). A human clone corresponding to complete murine TCFL4 was also isolated through a database search (isoform BigMax-gamma, see Figure 1d) (IMAGE clone 488171). Moreover, a third isoform lacking exon 3 and 162 bp in exon 1 (isoform BigMax-alpha, see Figure 1d) was identified in dbEST (IMAGE clones 488652 and 502642). Nested PCR analysis on cDNA libraries (serum starved WI-38 cells, fetal brain, HeLa and keratinocyte cells) excluded the possibility that the two newly identified splice forms resulted from cloning artifacts. During the preparation of this manuscript, BigMax-beta was independently isolated and

named Mlx by Donald E Ayer and collaborators in a screen for new Mad1 interactors (Billin et al., 1999). For the sake of clarity and to avoid confusion, we decided to adopt their nomenclature and name the BigMax-alpha, -beta and -gamma isoforms, Mlx-alpha, -beta and -gamma, respectively.

The entire Mlx gene spans approximately 7 kb and is contained in a fully sequenced cosmid (GenBank accession number U34879, (Zhao et al., 1996)). It consists of eight exons which can be alternatively spliced to encode the different isoforms (see Figure 1a,d). Interestingly, the two shorter isoforms (-alpha and -beta) use a cryptic AGgcaagc donor site in exon 1A.

We sequenced the three alternative Mlx-alpha, -beta and -gamma isoforms and deposited them in GenBank (accession numbers AF213666, AF213667 and AF213668, respectively). The putative initiation codon properly fulfills Kozak's criteria (Kozak, 1984), but no in frame 5'UTR STOP codon was identified in any of the six independent clones we isolated. We also cloned and sequenced the murine Mlx alternatively spliced forms to confirm their existence in rodents (mMlx-alpha, -beta and -gamma, GenBank accession numbers AF213670, AF213671 and AF213672, respectively).

To test whether Mlx isoforms are tissue specific, a qualitative study of the presence or absence of the three Mlx isoforms was performed by PCR on the Express-Check panel of 62 tissue-specific cDNA libraries (ATCC). We found the three Mlx-alpha, -beta and -gamma isoforms throughout the panel (data not shown). These data show that Mlx isoforms, like the Max isoforms, are not tissue specific. The semi-quantitative PCR conditions used in these experiments allowed us to conclude that in a large number of tissues, Mlx-alpha and -gamma are more abundant than the rarer Mlx-beta, whereas this ratio is inverted in favor of isoform beta in the basal ganglia, liver, placenta and pancreatic islets (data not shown). These results demonstrate that while the isoforms are not tissue specific, their ratio probably is, as previously shown for some of the Max protein isoforms (Makela et al., 1992).

Alternative splice forms may induce subtle differences in DNA binding as in the case of Max1 versus Max2 (Prochownik and VanAntwerp, 1993), but no obvious changes in E-box binding were observed with Mlx-alpha, -beta and -gamma. These results confirm the ability of Mlx to readily homodimerize and bind DNA even at low concentration.

The two shorter isoforms, Mlx-alpha and -beta, are predominantly cytoplasmic and only partially nuclear, while the Mlx-gamma isoform shows a predominantly nuclear localization. The longer isoform differs from both of the two shorter isoforms by the presence of exon 1B, which is 54 residues longer than exon 1A (see Figure 1a,b,c). Two sequences GRAGRARARRGAGRR and PACAKV, resembling a bipartite nuclear localization signal (NLS) and a portion of the c-Myc NLS (PAAKRVKLD), are present in this exon (reviewed in Nigg, 1997).

Osbpl1a

Mol Biol Cell. 2003 Mar;14(3):903-15. The two variants of oxysterol binding protein-related protein-1 display different tissue expression patterns, have different intracellular localization, and are functionally distinct. Johansson M, Bocher V, Lehto M, Chinetti G, Kuismanen E, Ehnholm C, Staels B, Olkkonen VM.

Oxysterol binding protein (OSBP) homologs comprise a family of 12 proteins in humans (Jaworski et al., 2001; Lehto et al., 2001). Two variants of OSBP-related protein (ORP) 1 have been identified: a short one that consists of the carboxy-terminal ligand binding domain only (ORP1S, 437 aa) and a longer N-terminally extended form (ORP1L, 950 aa) encompassing three ankyrin repeats and a pleckstrin homology domain (PHD). We now report that the two mRNAs show marked differences in tissue expression. ORP1S predominates in skeletal muscle and heart, whereas ORP1L is the most abundant form in brain and lung. On differentiation of primary human monocytes into macrophages, both ORP1S and ORP1L mRNAs were induced, the up-regulation of ORP1L being >100-fold. The intracellular localization of the two ORP1 variants was found to be different. Whereas ORP1S is largely cytosolic, the ORP1L variant localizes to late endosomes. A significant amount of ORP1S but only little ORP1L was found in the nucleus. The ORP1L ankyrin repeat region (aa 1-237) was found to localize to late endosomes such as the full-length protein. This localization was even more pronounced for a fragment that additionally includes the PHD (aa 1-408). The amino-terminal region of ORP1L consisting of the ankyrin repeat and PHDs is therefore likely to be responsible for the targeting of ORP1L to late endosomes. Interestingly, overexpression of ORP1L was found to enhance the LXRalpha-mediated transactivation of a reporter gene, whereas ORP1S failed to influence this process. The results suggest that the two forms of ORP1 are functionally distinct and that ORP1L is involved in control of cellular lipid metabolism.

Two mRNA variants of the ORP1 gene have recently been reported (Lehto et al., 2001 blue right-pointing triangle; Xu et al., 2001 blue right-pointing triangle). One of these, denoted hereafter as ORP1S (ORP1 “short”), encodes a predicted 50-kDa protein with 437 aa residues that consists of a ligand binding domain only (Xu et al., 2001 blue right-pointing triangle). The other, ORP1L (ORP1 “long”), encodes a 108-kDa protein with 950 aa. This protein structurally resembles OSBP, which has a PHD (Lemmon and Ferguson, 2000 blue right-pointing triangle) near its amino terminus. In addition, ORP1L has three ankyrin repeats (Sedgwick and Smerdon, 1999 blue right-pointing triangle) situated near the N terminus (Lehto et al., 2001 blue right-pointing triangle). The ORP1L transcript consists of sequences from 28 exons (Lehto et al., 2001 blue right-pointing triangle). The ORP1S mRNA is apparently generated through the use of an alternative first exon, named exon 1B, which in the genome is located between exons 15 and 16 (Figure 1). The two transcripts have been suggested to arise from the use of two different promoters (Jaworski et al., 2001 blue right-pointing triangle).

Real-time PCR analysis revealed that ORP1L expression was induced 100- to 160-fold during monocyte-macrophage differentiation. The expression of ORP1S was also up-regulated, but to a lesser extent (3- to 5-fold).

Pctk1

Gene. 1996 Oct 17;176(1-2):243-7. Expression of alternatively spliced PCTAIRE-1 mRNA in PC12 cells and neonatal rat brain. Gao CY, Chauthaiwale VM, Rampalli AM, Zelenka PS.
A rat PCTAIRE-1 cDNA clone was isolated by immunoscreening of a PC12 cDNA library, followed by 5' RACE (rapid amplification of cDNA ends) to determine the 5' end. The rat PCTAIRE-1 cDNA sequence is 96% identical to mouse PCTAIRE-1 and contains an alternatively spliced exon of 131 bp near the 5' end. Although a mouse cDNA

containing this exon has been reported, examination of several mouse cell lines provided no evidence for expression of the corresponding mRNA (Okuda et al., 1992). In contrast, reverse transcription and polymerase chain reaction (RT/PCR) across this region using RNA from proliferating, differentiated, and apoptotic PC12 cells demonstrated that alternatively spliced forms of PCTAIRE-1 mRNA with and without this exon are expressed. Both forms of PCTAIRE-1 mRNA are also expressed in vivo in neonatal rat brain, although other tissues examined contained only the form lacking the alternatively spliced exon. In the absence of the alternatively spliced exon PCTAIRE-1 mRNA contains an open reading frame of 1488 bp, corresponding to a 55-kDa protein that is 97% identical to mouse PCTAIRE-1 protein. When the alternatively spliced exon is present (*actually an extension of exon 3 due to an upstream splice acceptor*), this open reading frame is terminated by a stop codon and a second open reading frame is initiated, predicting a second PCTAIRE-1 protein of 52 kDa. The two predicted PCTAIRE-1 proteins are identical downstream of the splice site, but share no homology at their N-terminal ends.

The rat PCTAIRE-1 cDNA sequence obtained by 5' RACE also contained a 131-bp sequence (453-583), which is not present in the human sequence (Meyerson et al., 1992). Since the 3' end of this sequence corresponds exactly to an intron/exon boundary of the human PCTAIRE-I gene (Okuda et al., 1994), it apparently represents an alternatively spliced exon. Although a corresponding sequence was previously observed in mouse PCTAIRE-1 cDNA, it could not be detected in mRNA of several cell lines tested (Okuda et al., 1992). To determine whether two, alternatively spliced forms of PCTAIRE-1 mRNA are expressed in PC12 cell mRNA, we performed RT/PCR using primers flanking this region and sequenced the PCR products. The results demonstrated that PCTAIRE-1 mRNAs with and without this sequence are expressed in PC12 cells (Fig. 3A). No change in the relative abundance of these forms was associated with differentiation or apoptosis (Fig. 3A). To determine whether these two PCTAIRE-1 mRNAs are also expressed in rat tissues in vivo, the same RT/PCR assay was performed using RNA isolated from brain, kidney, heart, and lung of 6-day-old neonatal rats and from adult rat lens fibers. A PCTAIRE-1 mRNA containing the alternatively spliced exon was observed only in the neonatal rat brain (Fig. 3B). It is not clear at the present whether both poly A addition sites (Fig. 2) are used for each of the alternatively spliced forms of mRNA detected by RT/PCR (Fig. 3).

Translation of the two PCTAIRE-1 mRNA sequences predicts the existence of two distinct PCTAIRE-1 proteins, which differ at the N-terminus. A 1488-bp open reading frame begins at nt 251 in the mRNA lacking the alternatively spliced exon and corresponds to a 55 kDa protein, which shares 97% identity with mouse PCTAIRE-1. Among the conserved features in the protein sequence are the ATP binding site (172-177), the PCTAIRE domain (205-211), the kinase domain (286-291), and three stretches of basic residues at 5-8 (KKIK), 107-110 (RKIS) and 147-150 (RRLR) in the N-terminal end, which resemble the consensus nuclear localization signal KR/KXR/K (Chelsky et al., 1989). In the PCTAIRE-1 mRNA containing the alternatively spliced exon, the open reading frame initiated at 251 is interrupted by a termination codon at 479 (Fig. 1). Consequently, the longest open reading in this mRNA extends from 487 to 1869 and corresponds to a 52-kDa protein. Amino acids 33-461 of this protein are identical to amino acids 68-496 of the 55-kDa protein, but the two predicted amino acid sequences bear no similarity at their N-terminal ends. The unique N-terminus of the 52 kDa form of PCTAIRE-1 contains four tyrosine residues and a possible site for phosphorylation by proline directed kinases (TP) (Nigg, 1993).

J. Biol. Chem., Vol. 281, Issue 15, 9852-9858, April 14, 2006. Pctaire1 phosphorylates N-ethylmaleimide-sensitive fusion protein: implications in the regulation of its hexamerization and exocytosis. Liu Y, Cheng K, Gong K, Fu AK, Ip NY.
Originally identified as a Cdc2-like kinase (1, 2), Pctaire1 is ubiquitously expressed in mammalian tissues and is particularly abundant in terminally differentiated cells such as postmitotic neurons (1–3). In contrast to classical Cdks,2 which depend on cyclins for activation, Pctaire1 is not activated by cyclins. We have reported previously that Pctaire1 interacts with the Cdk5 activator p35, and phosphorylation by Cdk5/p35 enhances the kinase activity of Pctaire1 (4). An alternative regulatory mechanism to inhibit the activity of Pctaire1 through protein kinase A has also been proposed (5). Similar to Cdk5, Pctaire1 is not involved in the regulation of cell cycle progression (1, 3). To date, the precise functions of Pctaire1 remain elusive. Recent studies indicate that Pctaire1 regulates neurite outgrowth in Neuro-2A cells, whereby overexpression of the kinase-dead (KD) mutant of Pctaire1 leads to enhanced neurite outgrowth (5).

Pctaire1 activity can be regulated by Cdk5/p35, which is known to play a pivotal role in neurotransmission through the phosphorylation of various synaptic proteins (33–36). Although Cdk5 does not directly interact with or phosphorylate NSF in vitro (data not shown), Cdk5/p35 associates with Pctaire1 and regulates its activity. Our findings in the present study reveal that Pctaire1 not only interacts with NSF in adult rat brain but also with a number of synaptic proteins including components of the SNARE complex. It would therefore be interesting to examine whether Pctaire1 and Cdk5 exist as components of a large protein complex and together play an important role in phosphorylating a repertoire of protein substrates to ensure proper transmitter release from synaptic vesicles. The present study reveals a potential novel role of Pctaire1 phosphorylation in modulating NSF-mediated exocytosis. Recent studies of NSF function in regulated secretion suggest that NSF serves to prime synaptic vesicles and rearrange the SNARE complex for Ca2+-triggered fusion after docking. This rearrangement is required to maintain the readily releasable pool of synaptic vesicles (37–39), whereas regulation of NSF phosphorylation by Pctaire1 might modulate the regulated secretion. Further studies on Pctaire1 functions in neurotransmitter release might shed light on its roles in neurons during development.

Journal of Cell Science 115, 3479-3490 (2002)
Messenger RNAs encoding the 45-58 kDa PCTAIRE family of protein kinase subunits have been identified in frogs, mice and humans, and consist of two or three isoforms. The mRNAs for PCTAIRE-2 and 3 are found almost exclusively in the brain, whereas PCTAIRE-1 is located in both brain and testis of adult mice and rats (Hirose et al., 1997Go; Rhee and Wolgemuth, 1995Go). Unlike CDK1, which has only 10 residues before the Walker A motif, the PCTAIRE family have 127-198 residues in N-terminal extensions (Meyerson et al., 1992Go; Okuda et al., 1992Go). No cyclin partner has been identified for any of the PCTAIREs, and these extensions may play the role of an attached cyclin for these kinases, although they do not display obvious homology to known cyclin sequences. Although expression of the PCTAIREs is restricted to the brain and testis, our attention was aroused by the finding that high levels of expression occurred in a wide range of transformed and immortalised cell lines of epithelial origin (Charrasse et al., 1999Go; Hirose et al., 1997Go; Rhee and Wolgemuth, 1995Go). There is great uncertainty about the function of the PCTAIREs, some authors suggesting a role in the cell cycle (Charrasse et al., 1999Go), others in the differentiation of neurons or testicular

cells (Besset et al., 1999Go; Rhee and Wolgemuth, 1995Go). Equally, the natural substrate for this family of protein kinases is not known, but some investigators found that PCTAIRE-1, and PCTAIRE-2 could phosphorylate histone H1 (Hirose et al., 1997Go; Le Bouffant et al., 1998Go), whereas others only detected MBP phosphorylation (Besset et al., 1999Go; Rhee and Wolgemuth, 1995Go). The levels of kinase activity, compared to CDK1 or protein kinase A (PKA), are extremely low. This makes it difficult to be sure that the observed kinase activity is due to PCTAIRE, rather than traces of contaminating protein kinases.

Dync1i2

Cytoplasmic dynein is the major microtubule minus end–directed motor protein of the cell; it is involved in many essential cellular processes such as membrane trafficking and mitosis. The cytoplasmic dynein complex is resolved on SDS–polyacrylamide gels into subunit polypeptides of ~530 (dynein heavy chains), ~74 (intermediate chains), ~53–59 (light intermediate chains), and ~10–14 kD (light chains). Since mammalian cytoplasmic dynein was first identified and characterized (Paschal et al., 1987), further understanding of the subunit complexity of cytoplasmic dyneins has emerged (for review see Vallee et al., 2004; Pfister et al., 2005). Only a single heavy chain gene has been identified for the initially described form of cytoplasmic dynein, but two intermediate chain and two light intermediate chain genes have been found (Mikami et al., 1993; Zhang et al., 1993; Gill et al., 1994; Hughes et al., 1995; Vaughan and Vallee, 1995). Three light chain families have been identified, which appear to be shared among some, but not all, cytoplasmic and axonemal dynein complexes (King et al., 1996a,b, 1998; Bowman et al., 1999; Wilson et al., 2001).

Science. 2003 May 2;300(5620):808-12. Mutations in dynein link motor neuron degeneration to defects in retrograde transport. Hafezparast M, Klocke R, Ruhrberg C, Marquardt A, Ahmad-Annuar A, Bowen S, Lalli G, Witherden AS, Hummerich H, Nicholson S, Morgan PJ, Oozageer R, Priestley JV, Averill S, King VR, Ball S, Peters J, Toda T, Yamamoto A, Hiraoka Y, Augustin M, Korthaus D, Wattler S, Wabnitz P, Dickneite C, Lampel S, Boehme F, Peraus G, Popp A, Rudelius M, Schlegel J, Fuchs H, Hrabe de Angelis M, Schiavo G, Shima DT, Russ AP, Stumm G, Martin JE, Fisher EM.
Degenerative disorders of motor neurons include a range of progressive fatal diseases such as amyotrophic lateral sclerosis (ALS), spinal-bulbar muscular atrophy (SBMA), and spinal muscular atrophy (SMA). Although the causative genetic alterations are known for some cases, the molecular basis of many SMA and SBMA-like syndromes and most ALS cases is unknown. Here we show that missense point mutations in the cytoplasmic dynein heavy chain result in progressive motor neuron degeneration in heterozygous mice, and in homozygotes this is accompanied by the formation of Lewy-like inclusion bodies, thus resembling key features of human pathology. These mutations exclusively perturb neuron-specific functions of dynein.

Genomics. 1999 Feb 1;55(3):257-67. Cloning and characterization of two cytoplasmic dynein intermediate chain genes in mouse and human. Crackower MA, Sinasac DS, Xia J, Motoyama J, Prochazka M, Rommens JM, Scherer SW, Tsui LC.
Three of the eight isolated cDNA clones characterized did not include nucleotides 411–461 of the consensus sequence, which would translate into an open reading frame of 628 amino acids, suggesting the presence of alternative splicing (Fig. 3A). For discussion we

will refer to the larger transcript as *DNCI1A* and the smaller alternative transcript as *DNCI1C*. It appeared that the alternatively spliced *DNCI1C* transcript arose as a result of a cryptic splice-acceptor site located within exon 4 (Figs. 2 and 3A).

we chose to identify the equivalent mouse gene by low-stringency hybridization screening of a mouse embryonic cDNA library and a mouse adult brain cDNA library. The *Dnci1* gene was identified from the adult brain cDNA library, and DNA sequence analysis revealed a composite cDNA sequence of 2586 bp with a putative open reading frame of 628 amino acids (Fig. 3B). This isoform was most similar to *DNCI1C* from human and thus named accordingly. *Dnci1c* was found to be 97 and 98% identical at the amino acid level to the human and rat orthologues, respectively. Sequence analysis also identified a second isoform of *Dnci1* (tentatively named *Dnci1e*) in which amino acids 106–125 of *Dnci1c* were absent (Fig. 3A). Alternative splicing at this same position had also been reported for the rat orthologue (Paschal *et al.,* 1992). To determine whether additional splice variants similar to those found in rat and human are also present in mice, RT-PCR was performed with adult mouse brain RNA where the region of alternative splicing was amplified. The presence of four mouse brain *Dnci1* isoforms was identified, with varying contribution (Fig. 4A). Sequence analyses of these bands showed that the predominant species was isoform *Dnci1c.* In addition, *Dnci1e* was present. As expected, a splice variant homologous to *Dnci1a* was found, but in a relatively small amount (Figs. 3A and 4A); however, isoform *Dnci1b* was not present in mouse brain. The fourth isoform present from mouse brain, *Dnci1d,* was found to be a novel isoform not previously described (Figs. 3A and 4A). Alternative splicing has been reported for both *Dnci1* and *Dnci2* in rat (Paschal *et al.,* 1992; Vaughan and Vallee, 1995).

In this study we have identified several alternatively spliced isoforms for both human and mouse *DNCI1.* The regions of alternative splicing found in both mouse and human were highly similar to that reported for rat. On the other hand, novel isoforms not reported in rat were found in human, mouse, or both. The differences in identified isoforms likely reflects the methodology employed or the tissue source analyzed. Nevertheless, rat isoform *Dncib* was not found in human or mouse even following RT-PCR analysis. In this case RT-PCR was performed on mouse brain RNA, the same tissue used for rat gene identification (Paschal *et al.,* 1992). Thus, it would seem that *Dnci1b* may be specific to the rat brain. The alternative splicing in human *DNCI1* is probably the result of a cryptic splice site within exon 4. The same splice isoform was identified in mouse and is likely a result of a similar cryptic splice site. In addition, based on the human genomic structure of *DNCI1,* it is likely that mouse *Dnci1d* is the result of the utilization of a second cryptic splice-acceptor site in mouse exon 4. It would be necessary, however, to determine the genomic structure of mouse *Dnci1* to confirm this.

Figure and Table Titles and Legends

Figure S1. Increased Expression of CD45RA and RC on Diverse Leukocytes in thunder Mice

Figure S2. A Cell-Autonomous Role for hnRNPLL Controlling Peripheral T Cell Longevity

Figure S3. Normal Proliferation and Differentiation of *thunder* T Cells

Figure S4. List of Genes in Thunder Chromosomal Interval

Figure S5. *Hnrpll* Expression Pattern in Different Lymphocyte Subsets and Tissues

Figure S6. NMR Spectra of HnRNPLL RRM1 Domain
(A-D) ^{15}N-HSQC spectra of wildtype and mutant RRM1 in absence (A, B) and presence of ARS RNA (C, D). (E) ^{15}N-HSQC spectra of wild-type RRM1 in presence of mutant ARS RNA. (F) Rosetta all-atom energy of all structures calculated versus the C^a root mean square deviation (r.m.s.d.) from the structure with the lowest energy. (G) Stereoviews of lowest-energy structure from CSRosetta superimposed on 1WEX and 1SJQ structures, and primary sequence comparison of RRM1 domains of hnRNPLL and PTB1.

Figure S7. Chemical-Shift Changes of Backbone Amide Protons versus Amino Acid Sequence

Figure S8. Isothermal Calorimetry Data

Figure S9. Group A Gene Views in Alphabetical Order

Figure S10. Group B Gene Views in Alphabetical Order

Figure S11. Group C Gene Views in Alphabetical Order

Figure S12. Group D Gene Views in Alphabetical Order

Figure S13. Group E Gene Views in Alphabetical Order

Figure S14. Group F Gene Views in Alphabetical Order

Figure S15. Gene Views for CD44 and Stat5a

Table S1. False Discovery Rates for Specific Cell-Type Comparisons

Table S2. Data for Group A Genes

Table S3. Data for Group B Genes

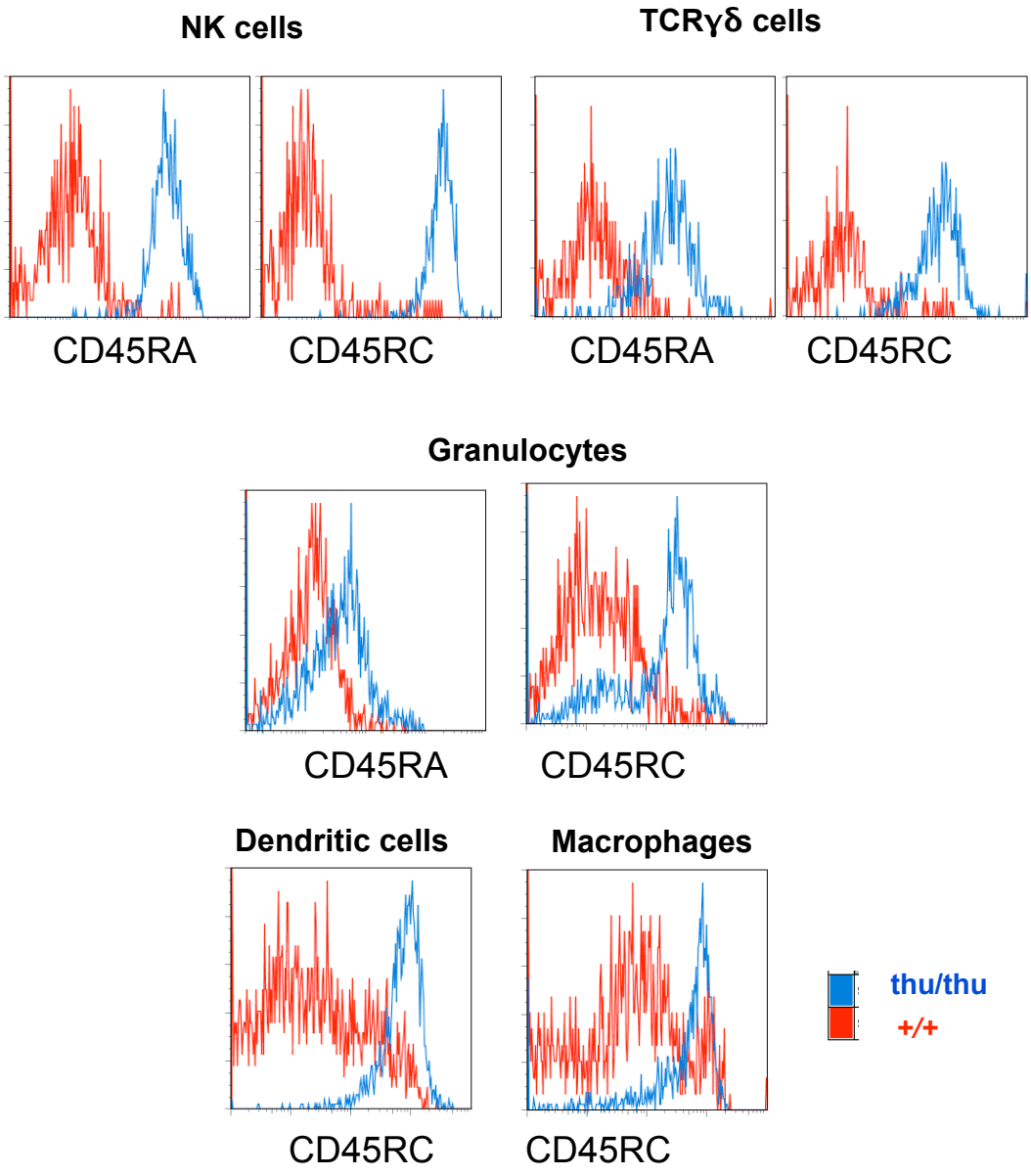
Table S4. Data for Group C Genes

Table S5. Data for Group D Genes

Table S6. Data for Group E Genes

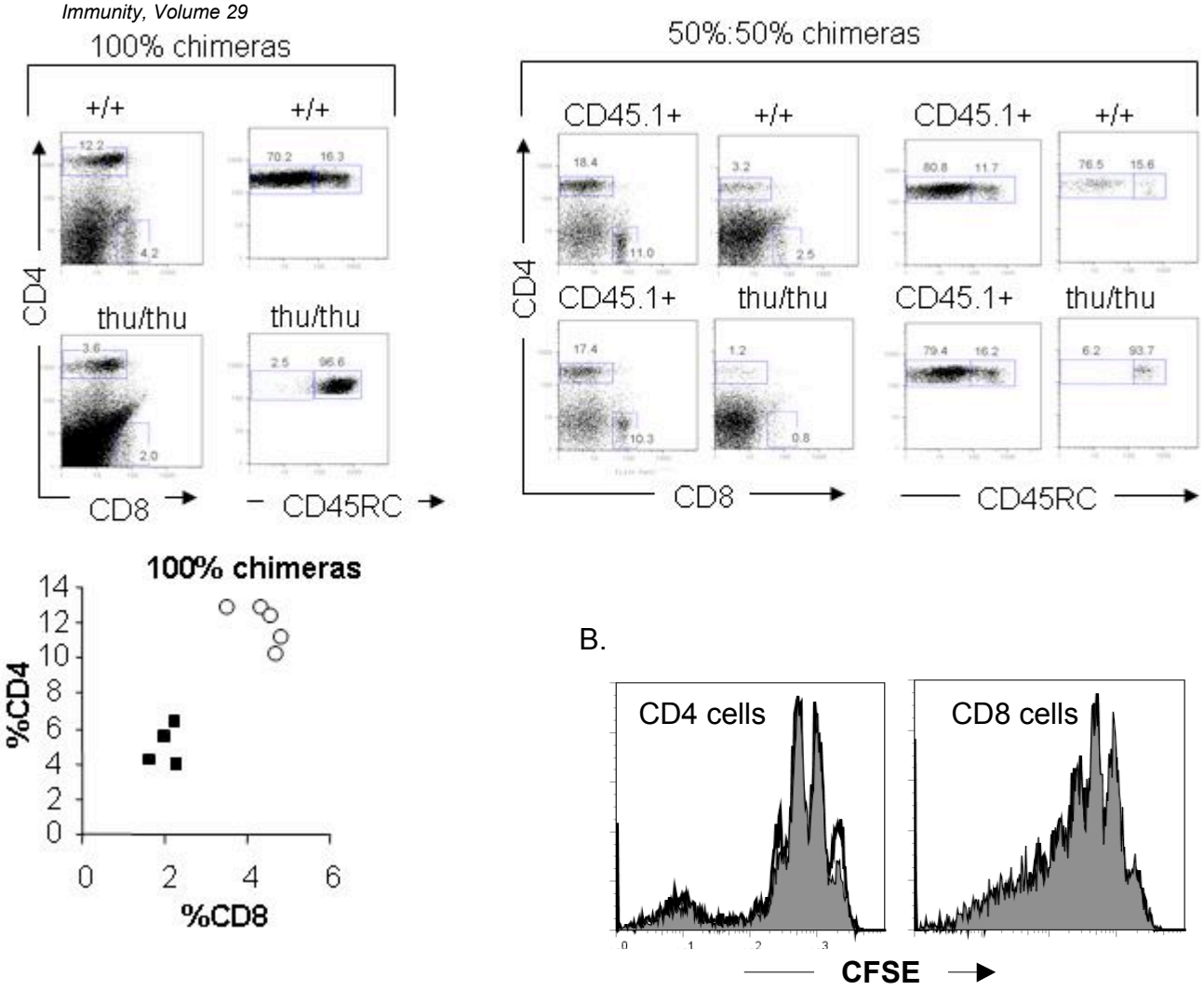
Table S7. Data for Group F Genes

Table S8. Data for Highest-Ranked Genes in Wild-Type Memory CD4 and CD8 Cells



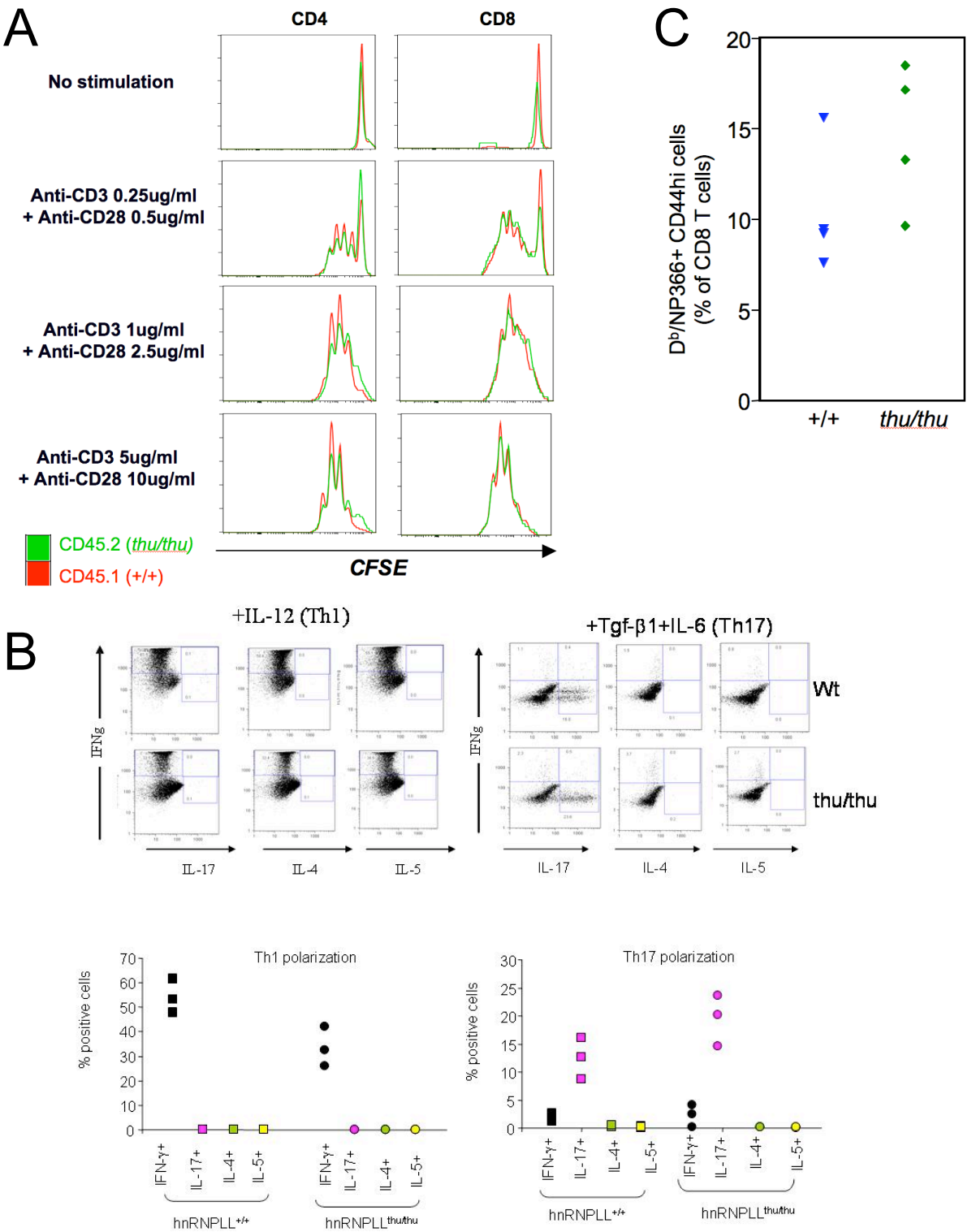
Supplementary Figure 1. The *thunder* mutation disrupts silencing of CD45RA and RC in numerous cell types.

Representative histograms comparing CD45RA and CD45RC expression on NK and TCRγδ+ cells in the thymus, and on Gr-1+ granulocytes, CD11c+ dendritic cells, and CD11b+ macrophages in the spleen. (red lines = +/+ cells, blue line=*thu/thu* cells)



Supplementary Figure 2. A cell autonomous role for hnRNPLL is controlling peripheral T cell longevity. A. Bone marrow chimeras were prepared by injecting 100% wild type CD45.2+ cells or 100% *thu/thu* T depleted bone marrow cells into an irradiated CD45.1+ recipient. (left hand panel). Alternatively mixed bone marrow chimeras were prepared with 50:50 mixture of CD45.1+ wild type and CD45.2 wild type bone marrow cells, or a 50:50 mixture of CD45.1+ and CD45.2+ *thu/thu* bone marrow cells. Mice were analysed 12 weeks after the transfer and stained for CD4, CD8 CD45.1, CD45.2 and CD45RC. Representative FACS dot plots are shown for the 100% and 50:50 mixed chimeras. Wild type CD45.2+ cells (open circles) and thu/thu CD45.2+ (black squares).

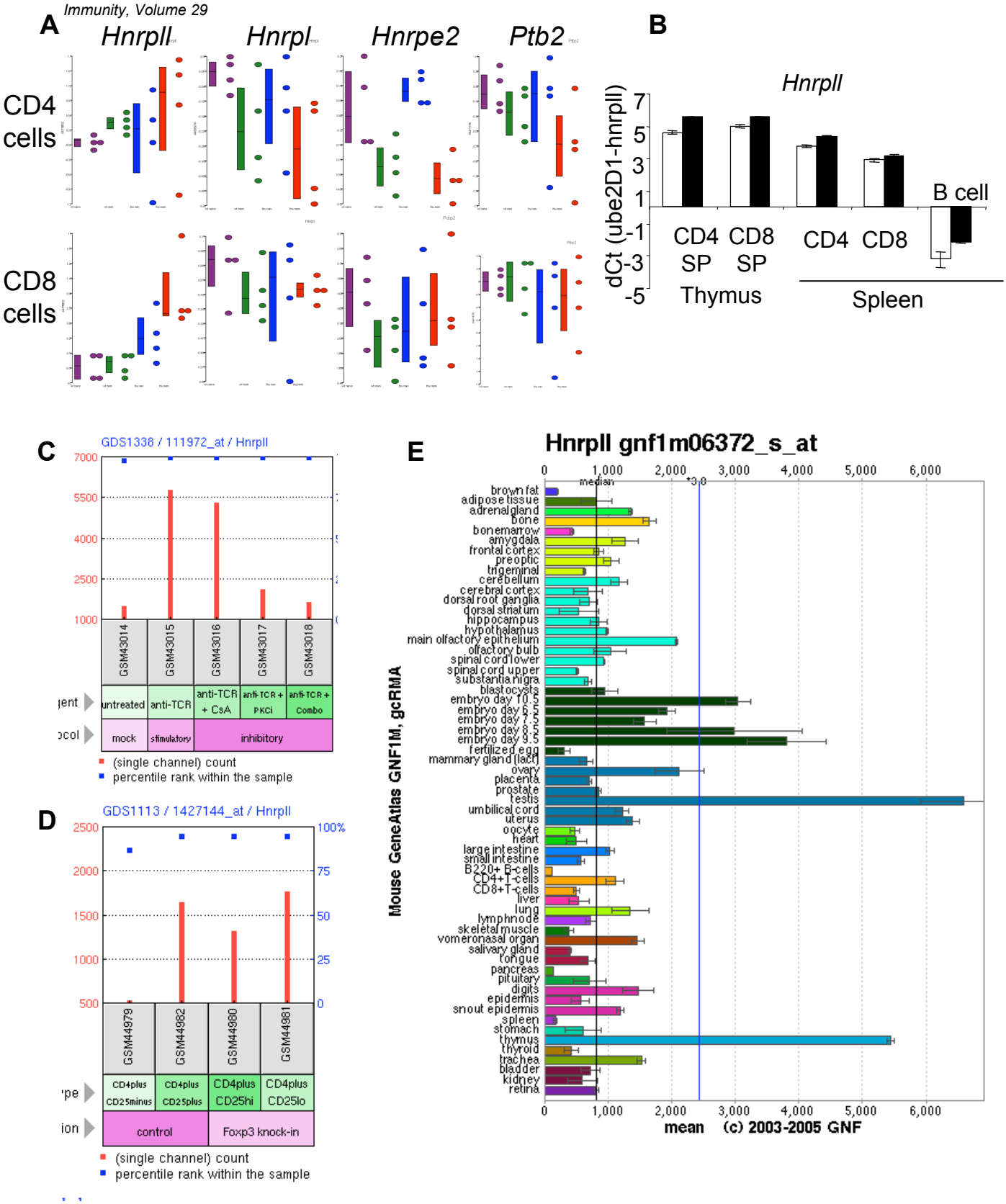
B. Representative FACS histograms of CFSE staining of donor CD4+ and CD8+ T cells from irradiated C57BŁ6J mice injected with a 50:50 mix of CFSE labelled wild type CD45.1+ and CD45.2 *thu/thu* thymocytes. Data shows an over lay of wild type cells (line) and *thu/thu* cells (filled histogram).



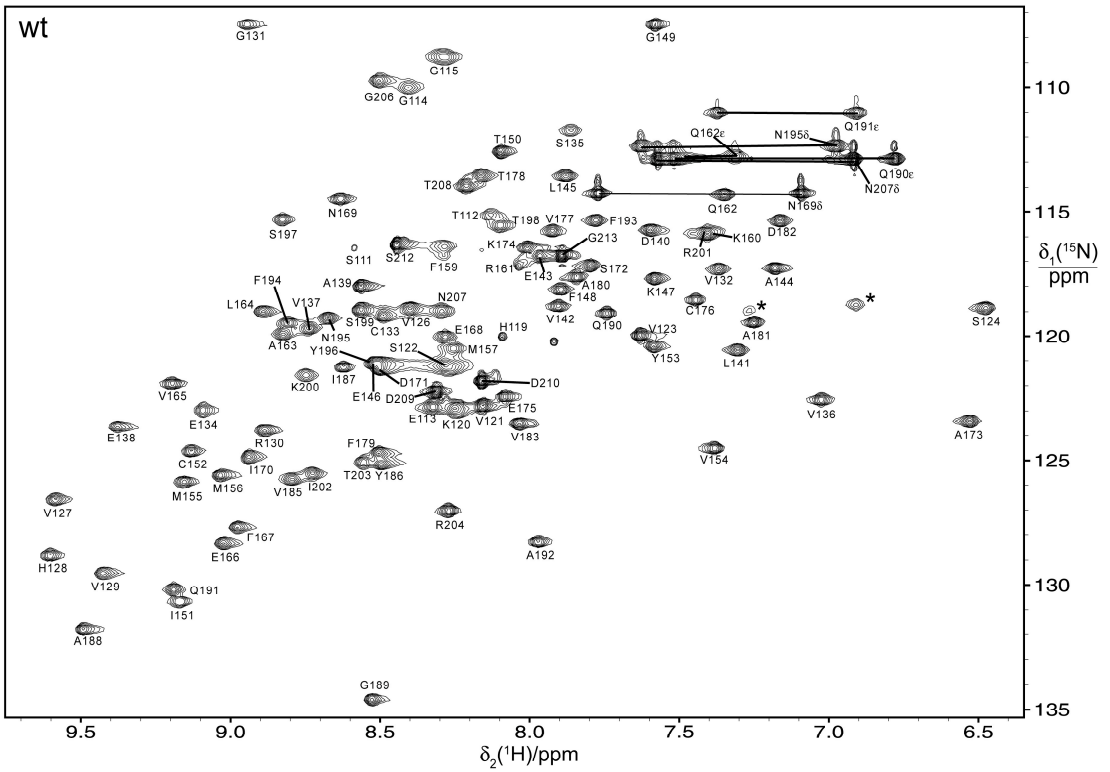
Supp Figure 3. Normal proliferation and differentiation of *thunder* T cells. A. Wild-type spleen cells bearing the CD45.1 marker were mixed in equal numbers with *thu/thu* spleen cells bearing the CD45.2 marker, labelled with CFSE, and the mixtures cultured in medium containing the indicated concentrations of anti-CD3 and anti-CD28 antibody and cell division compared between mutant and wildtype T cells on day 4 as shown on the overlay plots. **B.** Spleen cells from individual wild-type or *thunder* mice were cultured for 7 days with anti-CD3 & anti-CD28 antibodies plus 1ug/ml IL-12 or I ng/ml TGFβ & 10ng/ml IL-6. The cells were then restimulated for 5 hr with PMA and ionomycin in the presence of Golgi-stop, harvested, fixed & permeabilized, stained for intracellular cytokines, and analyzed by flow cytometry as shown. **C.** Wild-type or *thunder* mice were intraperitoneally injected with influenza viruses HKx31 (N3N2) on day 0 and PR8 (H1N1) on day 21. The CD8 T cell response to the dominant cross-reactive nucleoprotein (NP) epitope was enumerated in individual mice by flow cytometric staining with influenza NP366/D^b tetramers in blood collected on day 28.

	Gene ID	CDS (bp)	Transcripts	Description
1	Arl6ip2	2842	2	ADP-ribosylation factor-like 6 interacting protein 2 (nucleotide binding)
2	UNKNOWN		1	
3	UNKNOWN		1	
4	**Hnrpll**	3002	1	**heterogeneous nuclear ribonucleoprotein L-like**
5	Galm	2245	1	galactose mutarotase
6	**Sfrs7**	**2272**	**3**	**splicing factor, arginine/serine-rich 7**
7	4930560E09Rik		1	PREDICTED: hypothetical protein LOC67737
8	**Gemin6**	**1086**	**1**	**gem (nuclear organelle) associated protein 6**
9	AW494914	4101	3	expressed sequence AW494914 (ATP dependent RNA helicase)
10	Mopt	647	2	protein containing single MORN motif in testis
11	Gm941	1207	1	gene model 941, (NCBI)
12	Q8BSD7_MOUSE	324	1	12 days embryo embryonic body between diaphragm region and neck cDNA,
13	Sos1	4747	1	Son of sevenless homolog 1 (Drosophila)
14	XP_355031.3	984	4	PREDICTED: similar to cyclin-dependent kinase-like 4
15	Map4k3	3382	2	mitogen-activated protein kinase kinase kinase kinase 3
16	UNKNOWN			
17	UNKNOWN	248	1	
18	2810417M05Rik	1690	1	**RIKEN** cDNA 2810417M05 gene
19	**Thumpd2**	**1848**	**2**	**THUMP domain containing 2**
20	NM_178927.2			expressed sequence AV344025 (AV344025), mRNA
21	Slc8a1	2823	4	solute carrier family 8 (sodium/calcium exchanger), member 1

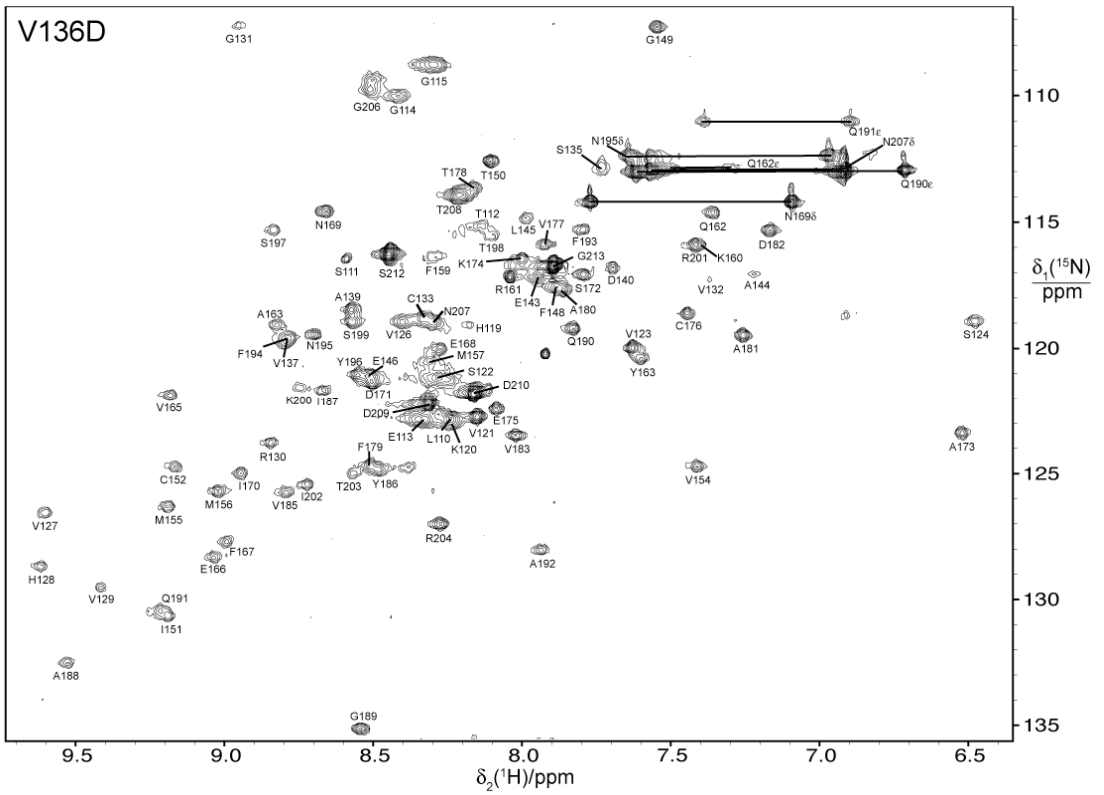
Supplementary Figure 4 **The thunder locus located on distal Chromosome 17.** 21 transcription units were identified in the minimal thunder interval defined by D17Mit42 (79.2Mbp) and the SNP 17-0802 (81.3Mbp). Genes shown in bold text have been implicated in RNA binding. Sfrs7 is a splicing factor and hnRNPLL is a member of the hnRNP protein family which are implicated in alternative splicing of mRNA.



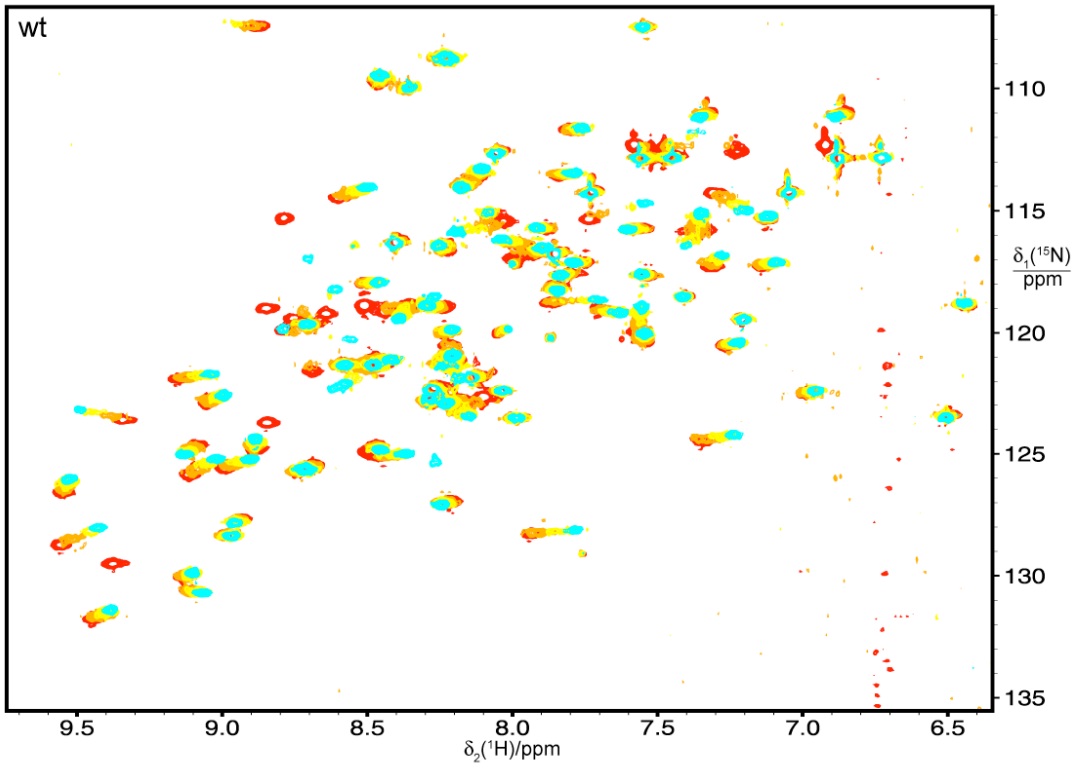
Supplementary Figure 5. *hnRNPLL* gene expression pattern.
a, Relative mRNA abundance as determined from exon array data for various hnRNP mRNAs in CD4+ (top panels) and CD8+ (lower panels) T cells isolated from wild type naïve (purple circles), and memory (green circles) T cells and thunder naïve (blue circles) and memory (red circles) T cells. Each circle represents an independent sample for each genotype. **b.** Relative *Hnrpll* mRNA in T cells versus B cells measured by realtime PCR. Open bars, +/+; solid bars, *thu/thu*. **c.** Relative *Hnrpll* mRNA in unstimulated and stimulated AE.7 T cells (from GEO deposited array data GDS1338 of Safford M et al. 2005 *Nat Immunol* 6:472-80). **d.** Relative *Hnrpll* mRNA in CD4+25+ versus CD4+25- spleen T cells (from GEO deposited array data GDS1113 of Fontenot JD et al 2005 *Immunity* 22:329-41). **e.** The expression of hnRNPLL gene in multiple tissues as determined by GNF Symatlas (Su AI et al. 2004. *Proc Natl Acad Sci U S A* 101:6062-7).



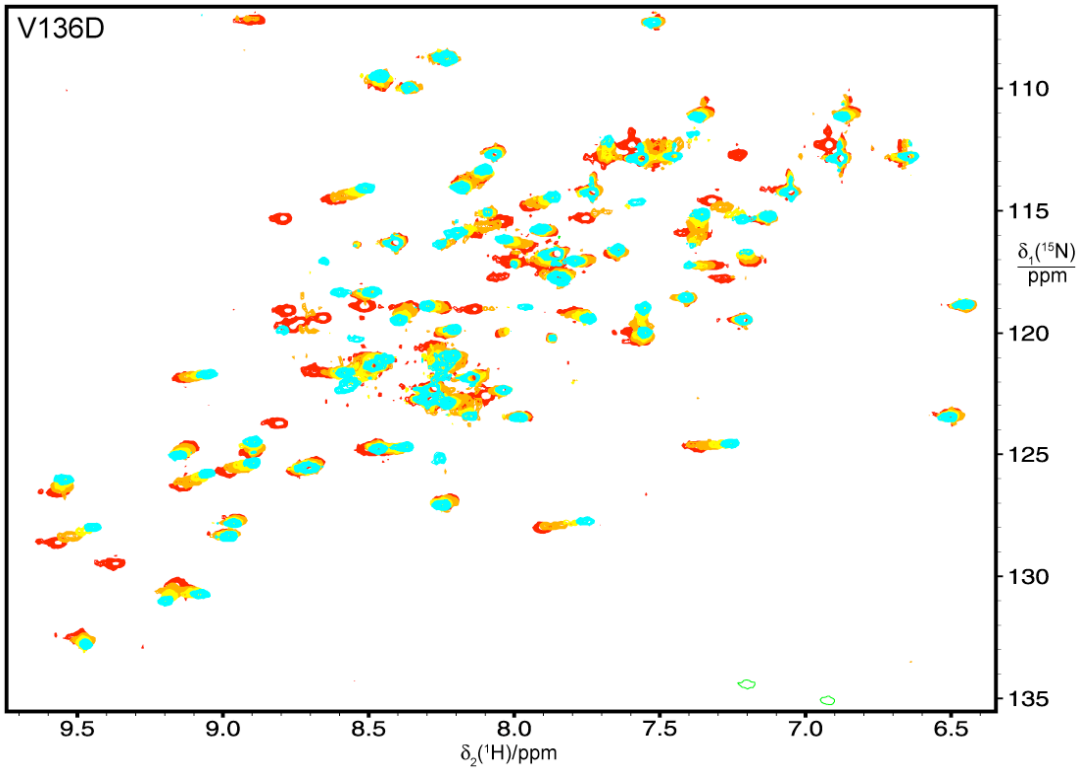
Supp Fig 6A. ^{15}N-HSQC spectrum of a 0.8 mM solution of the wild-type RRM1 domain in NMR buffer at 25 ºC. The cross-peaks of the backbone amides are assigned with the one-letter amino acid code and the sequence number of the residues. Horizontal lines connect pairs of cross-peaks from side-chain amides of Asn and Gln. Folded peaks from arginine side chains are labelled with a star. The amide cross-peak of Gly189 is folded from $\delta_1(^{15}N)$ = 101.3 ppm.



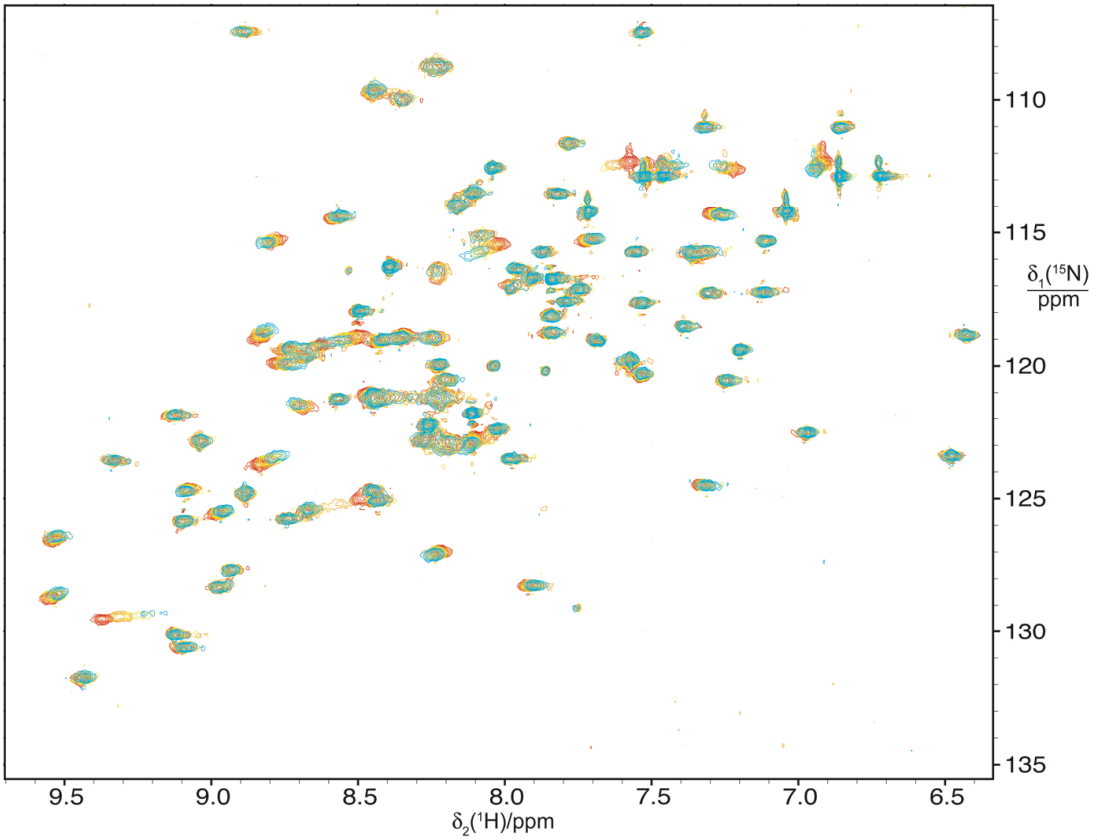
Supp Fig 6B. Same as Supp Fig 6A, but for the V136D mutant.



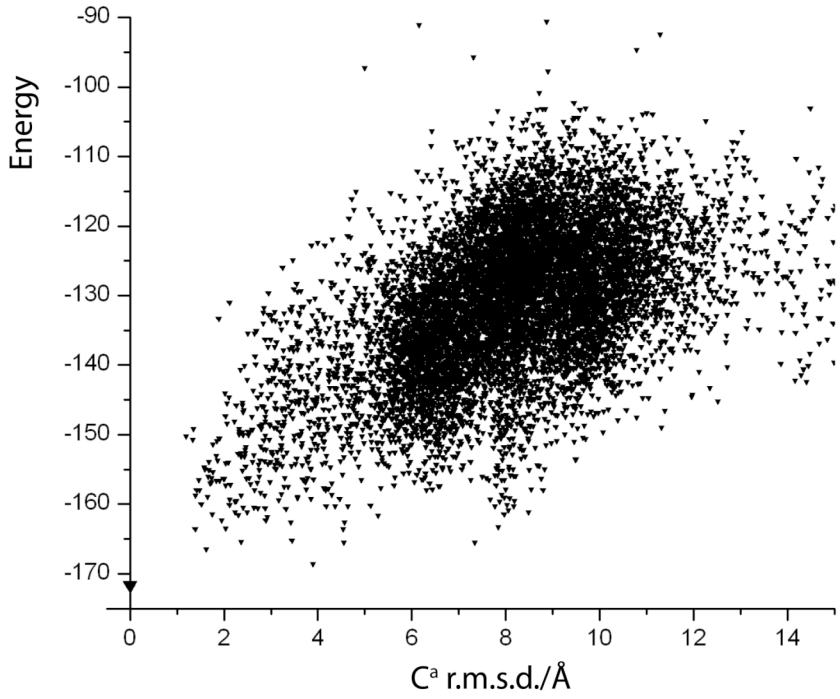
Supp Fig 6C. Superimposition of four ^{15}N-HSQC spectra of uniformly ^{15}N-labeled wild-type RRM1 in the presence of different amounts of RNA. The nucleotide sequence was 5'-CCUUACCUGCACGCA-3'. The spectra shown correspond to [protein]:[RNA] ratios of 1:0 (red), 1:0.3 (orange), 1:0.7 (yellow), and 1:1.5 (cyan). The gradual change in the peak positions confirms that the exchange between bound and free RNA is fast on the chemical shift time scale (milliseconds). The changes in chemical shifts as a function of titration ratio are plotted in Fig. 4C and D for selected peaks (see Supp Figs 6A and 6B for the resonance assignments).



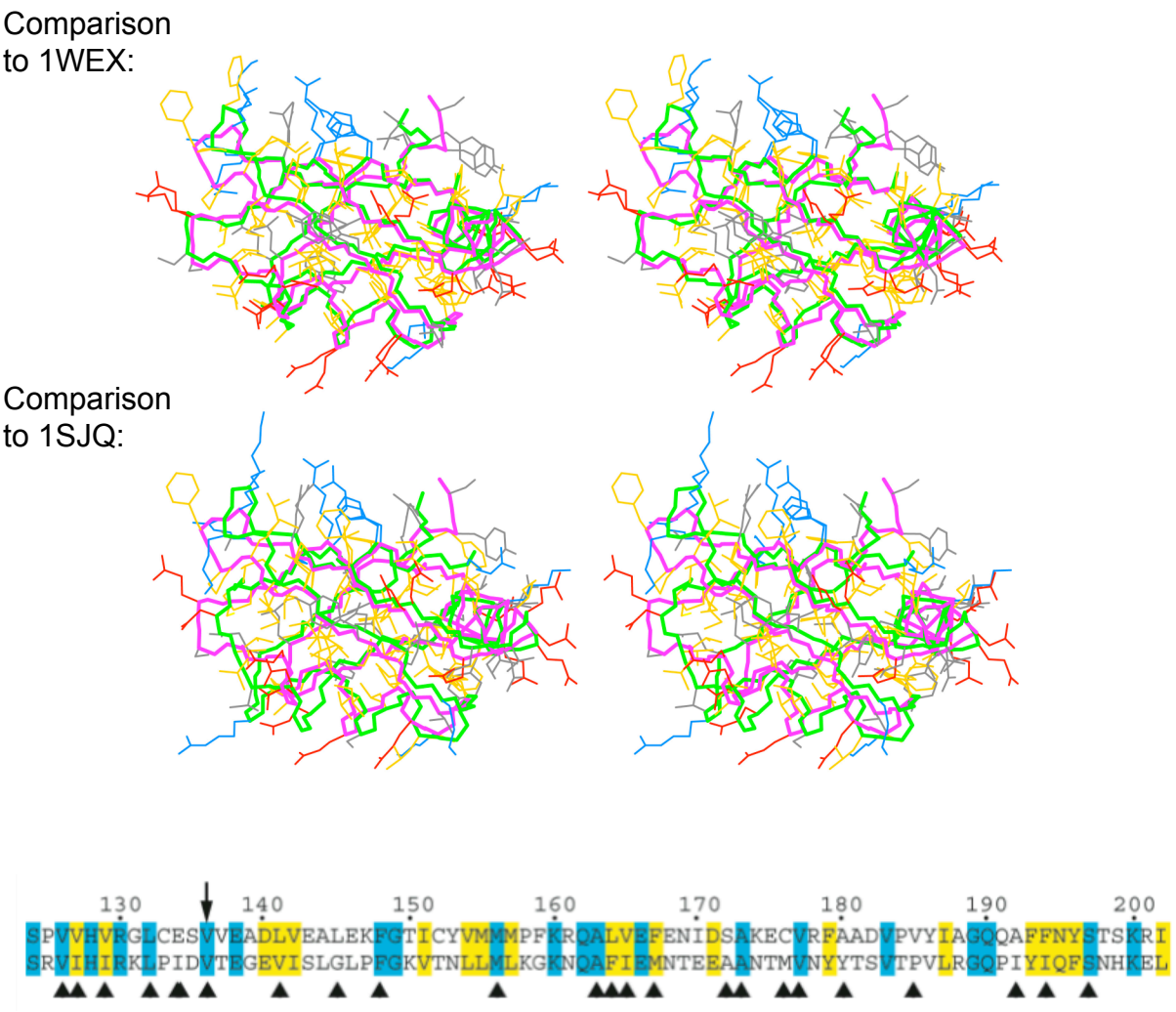
Supp Fig 6D. Same as Supp Fig 6C, but for the V136D mutant.



Supp Fig 6E. Same as Supp Fig 6C, except using the ARS mutant RNA oligonucleotide 5'-CCUUACCUGGACGUA-3'. The smaller shift changes and the greatly reduced number of shifted peaks suggesting that this RNA oligomer does not bind specifically. The dissociation constant derived from the shifted peaks was ~5 times greater than for the RNA oligomer of Supp Fig 6C.



Supp Fig 6F. Rosetta all-atom energy of all structures calculated versus the C^{α} root mean square deviation (r.m.s.d.) from the structure with the lowest energy. The structure with the lowest energy (residues 124-202) is shown in Figure 4A and B. All residues in this structure are in the allowed region of the Ramachandran plot, except for three residues which are in the additional allowed region. The correctness of the fold was verified by nuclear Overhauser effects (NOEs) that had not been used in the structure calculation. In particular, NOEs observed between the methyl groups of V136 and the side-chains of Leu134 and Leu143 confirmed that the side-chain of V136 is buried in the structure.

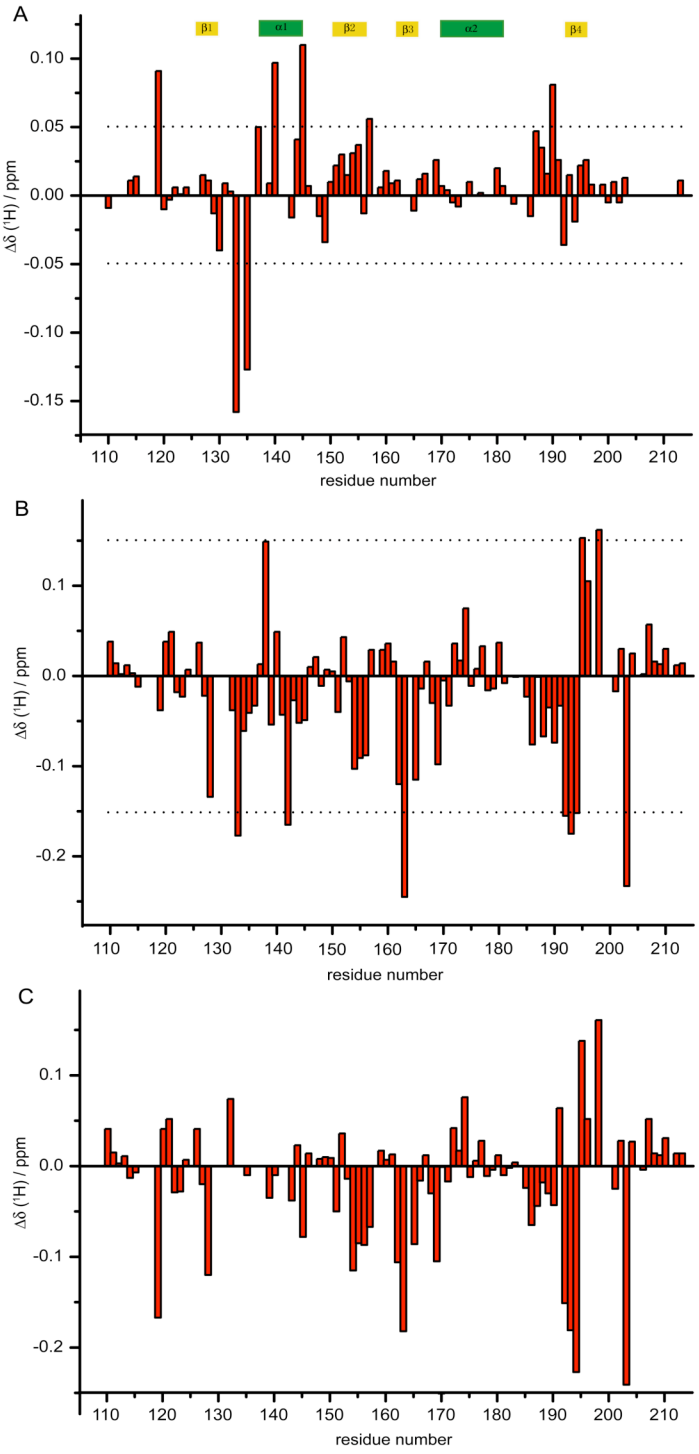


Supp Fig 6G.

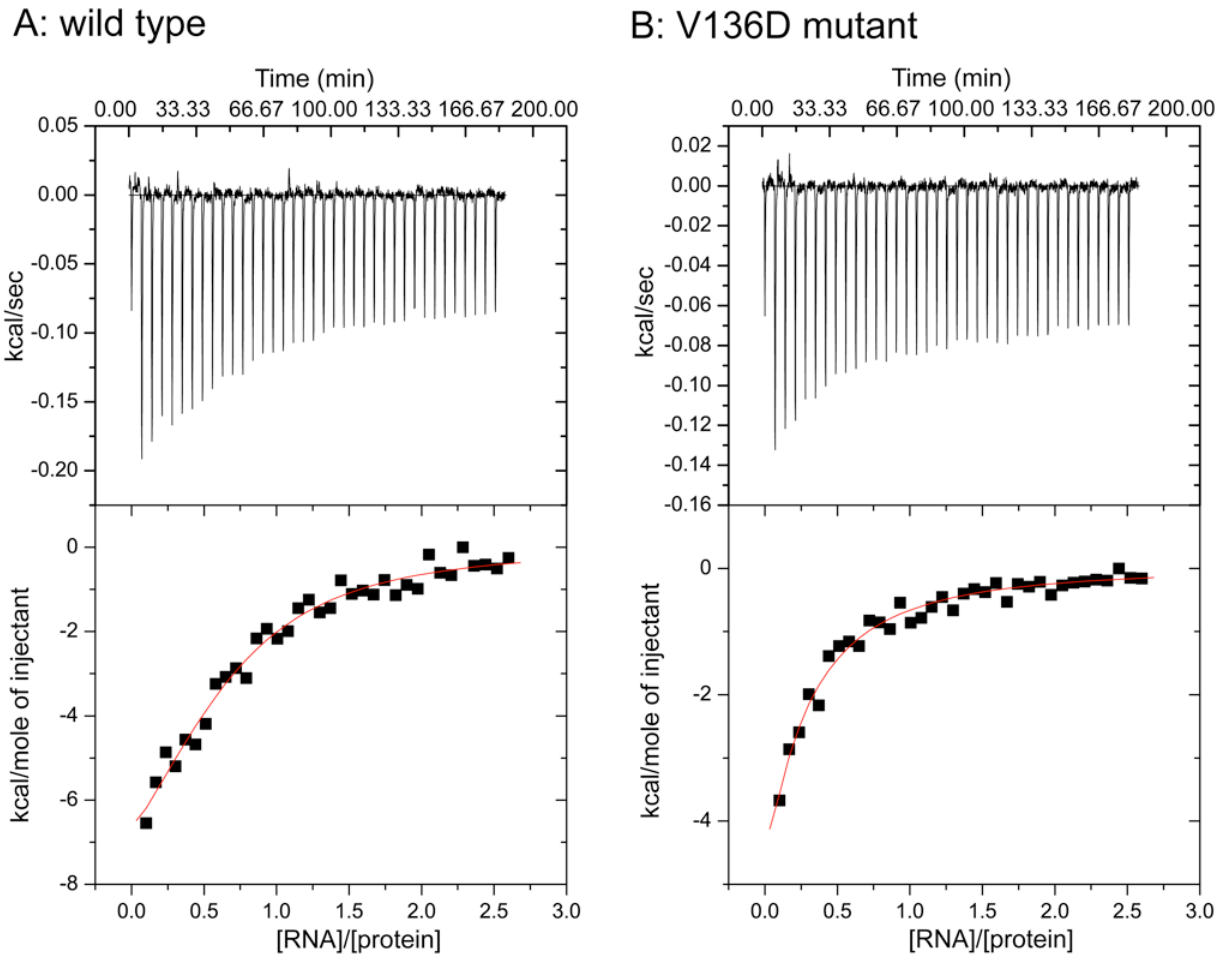
(Top) Stereoview of a heavy-atom representation of the lowest-energy structure from CSRosetta (magenta backbone) superimposed on the first conformer of the structure determined by the RIKEN Structural Genomics/Proteomics Initiative (PDB ID: 1WEX; these coordinates have been deposited without supporting information and we cannot assess their reliability). The average pairwise backbone and heavy atom r.m.s.d.s are1.0 and 2.0 Å, respectively, comparing residues 126-196 of the CSRosetta structure with all conformers of the 1WEX structure.

(Middle) Same as top, but comparing the CSRosetta structure with the first conformer of the PTB1 RRM1 domain structure (PDB ID: 1SJQ). The average pairwise backbone r.m.s.d. is 1.9 Å, comparing residues 126-196 of the CSRosetta structure with all conformers of the 1SJQ structure.

(Bottom) Amino acid sequence alignment of the hnRNPLL RRM1 domain at the top and the sequence of the PTB1 RRM1 domain below. Identical and similar residues are highlighted in cyan and yellow, respectively. The residue numbering of the hnRNPLL RRM1 domain is indicated. An arrow identifies the site of the thunder mutation. The location of buried side-chains with less than 11% solvent exposure is indicated for the PTB1 RRM1 domain by triangles below the sequence. Structural similarity between the hnRNPLL and PTB1 RRM1 domains can be inferred from the high sequence conservation (52% sequence similarity), the functional conservation as RNA-binding domains, the conservation of characteristic residues for RNA-binding (e.g., His128 and Gln162), and the fact that almost all residues with buried side-chains in the PTB1 RRM1 structure are hydrophobic in both proteins.

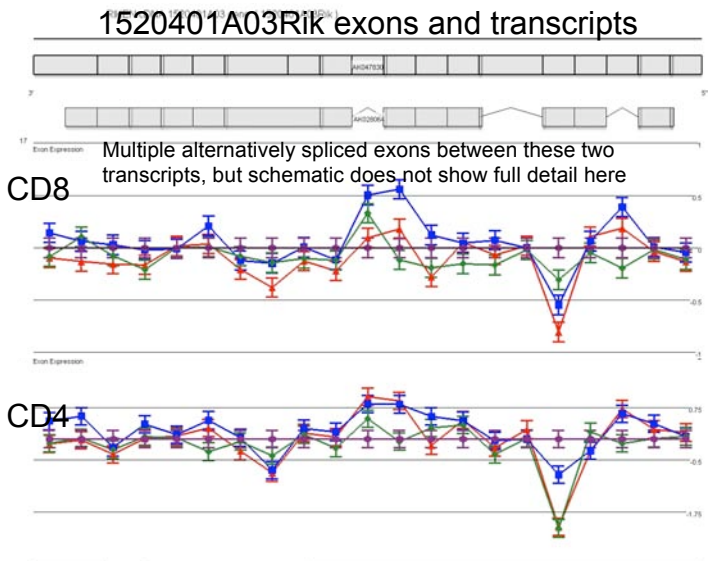


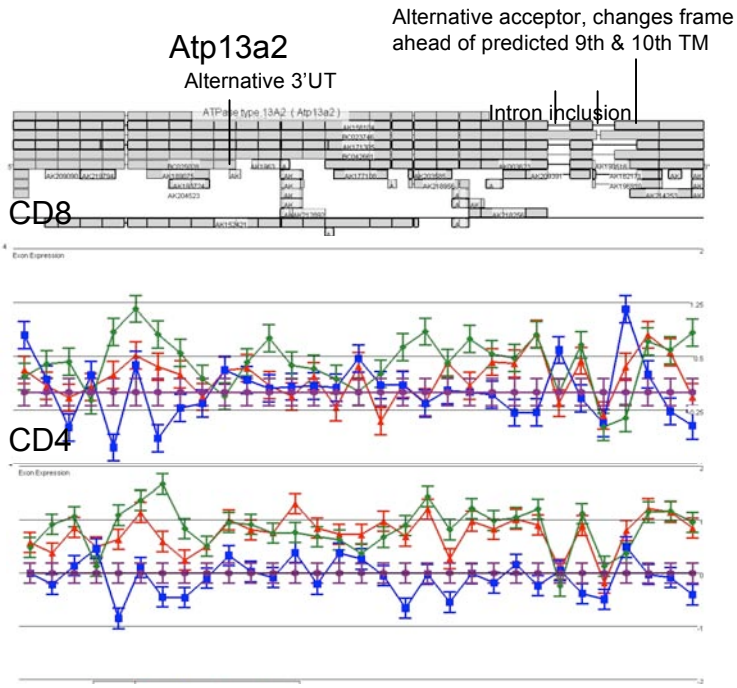
Supp Figure 7. Chemical shift changes of backbone amide protons versus amino-acid sequence. The chemical shifts are compared to those of the wild-type RRM1 domain. All amide protons of the wild-type protein (residues 110-213) were assigned, except for those of residues 116-118. The regular secondary structure elements are indicated at the top. The dotted lines indicate the cutoffs used to highlight amide protons with significant chemical shift changes in Figs. 3A and B. (A) Chemical shifts of the V136D mutant minus those of the wild-type protein. As for the wild-type protein, residues 116-118 could not be assigned. In addition, the ^{15}N-HSQC cross-peaks of Glu134, Val136, Glu138, Leu141, Val142, Lys147 and Leu164 were broadened beyond detection. (B) Chemical shifts of the wild-type RRM1 domain in the presence of the 15-mer 5'-CCUUACCUGCACGCA-3' (RNA:protein ratio of 2:1) minus those observed in the absence of 15-mer. In addition to the unassigned amides of residues 116-118, the ^{15}N-HSQC cross-peaks of Val129, Arg130, Gly131, Leu164, Ser197, Ser199 and Lys200 were broadened beyond detection in the presence of 15-mer. The cross-peaks of the side-chain amides of Gln162 and Asn195 also disappeared in the presence of 15-mer. (C) Chemical shifts of the RRM1 V136D mutant in the presence minus those in the absence of the 15-mer (RNA:protein ratio of 2:1). In addition to the unassigned amides of residues 116-118, 134, 136, 138, 141, 142, 147 and 164, the ^{15}N-HSQC cross-peaks of Val129, Arg130, Gly131, Val137, Ser197, Ser199, and Lys200 were broadened beyond detection in the presence of 15-mer. As in the case of the wild-type protein, also the cross-peaks of the side-chain amides of Gln162 and Asn195 disappeared in the presence of 15-mer. The similarity in chemical shift changes induced by the RNA in the wild-type and mutant RRM1 domain indicates a similar RNA binding mode.

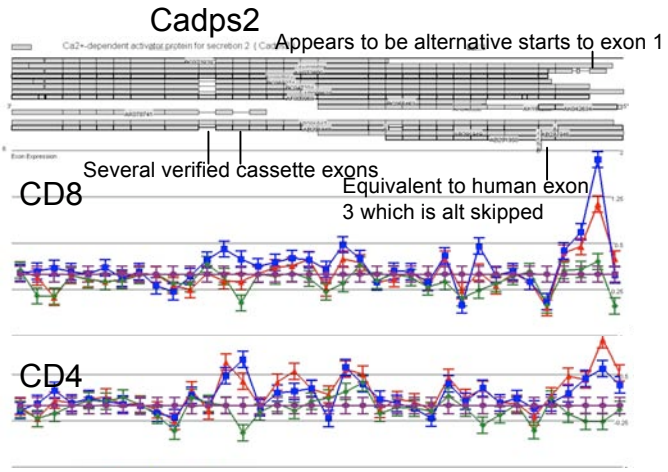


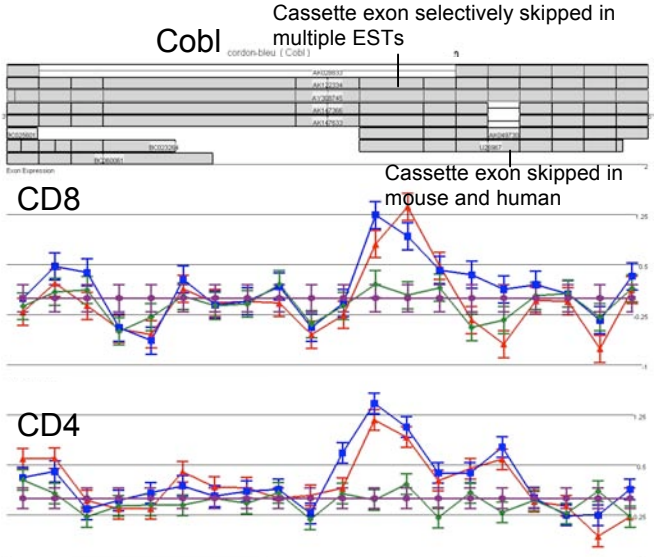
Supp Fig 8. ITC data recorded for titrations of wild-type (A) and mutant (B) RRM1 domains with the RNA oligomer 5'-CCUUACCUGCACGCA-3' at 15 ºC. Titration of the RNA into buffer generated significant heat effects which were subtracted from the titration data of the RRM1 domains in order to assess the binding affinity of the protein domains. The data of Fig. 3C and D were obtained by individual integration of each peak in order to minimize contributions from the baseline noise.

Group A

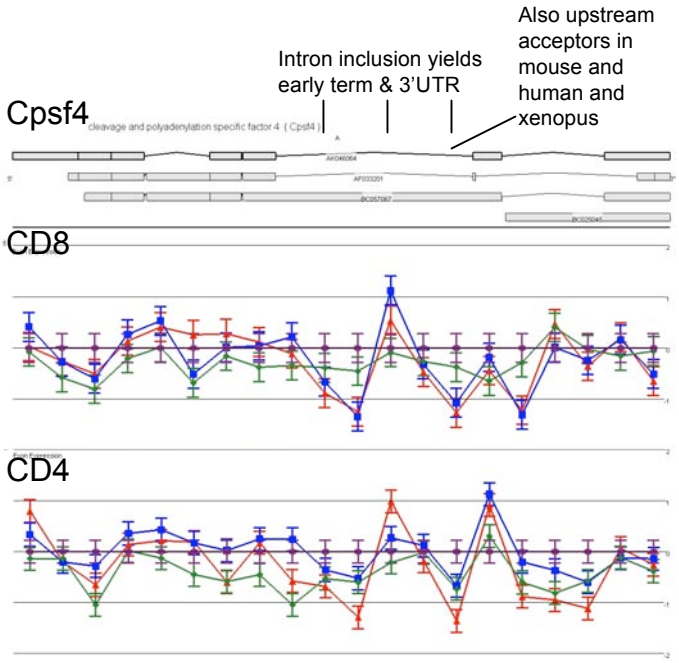

1520401A03Rik exons and transcripts
Multiple alternatively spliced exons between these two transcripts, but schematic does not show full detail here
CD8
CD4

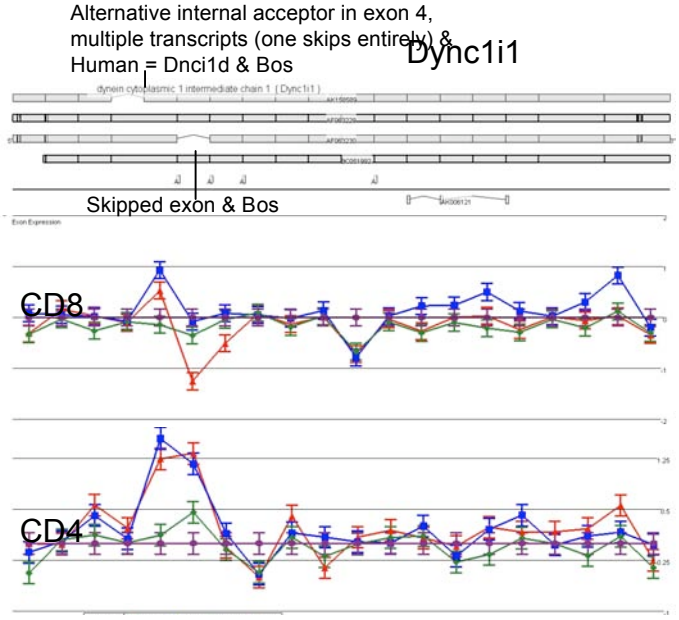

Atp13a2
Alternative 3'UT
Alternative acceptor, changes frame ahead of predicted 9th & 10th TM
Intron inclusion
CD8
CD4

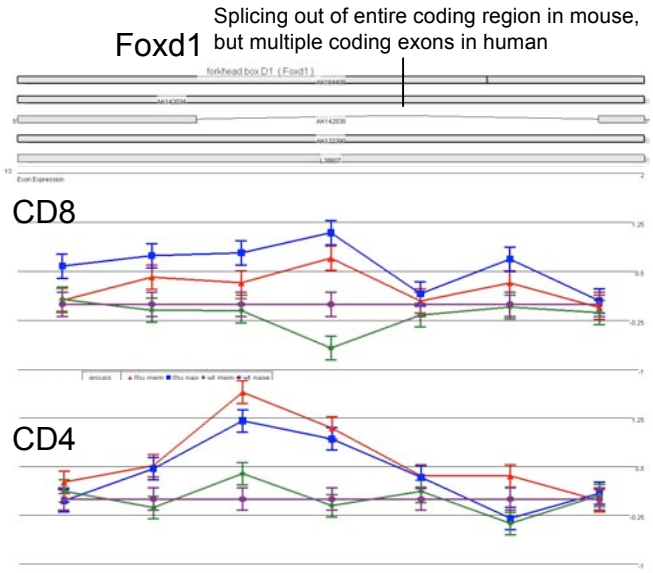

Cadps2
Appears to be alternative starts to exon 1
Several verified cassette exons
Equivalent to human exon 3 which is alt skipped
CD8
CD4

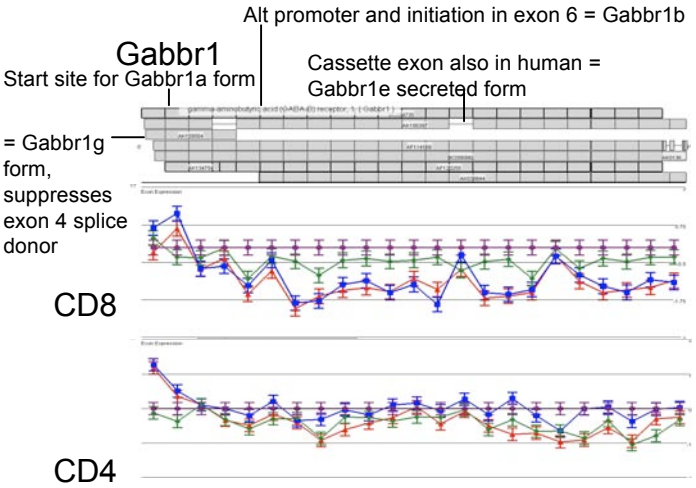

Cobl
Cassette exon selectively skipped in multiple ESTs
Cassette exon skipped in mouse and human
CD8
CD4

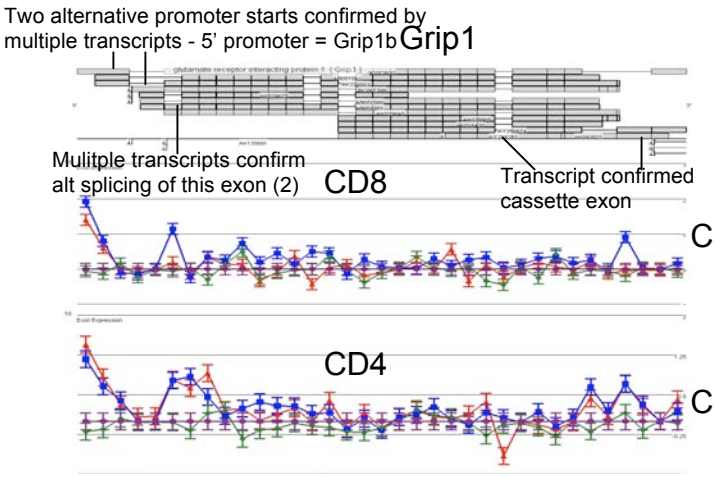
Group A

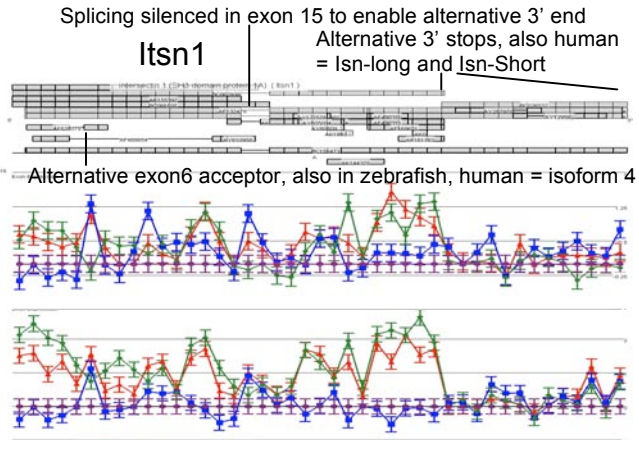

Also upstream acceptors in mouse and human and xenopus
Intron inclusion yields early term & 3'UTR
Cpsf4
CD8
CD4


Alternative internal acceptor in exon 4, multiple transcripts (one skips entirely) & Human = Dnci1d & Bos
Dync1i1
Skipped exon & Bos
CD8
CD4

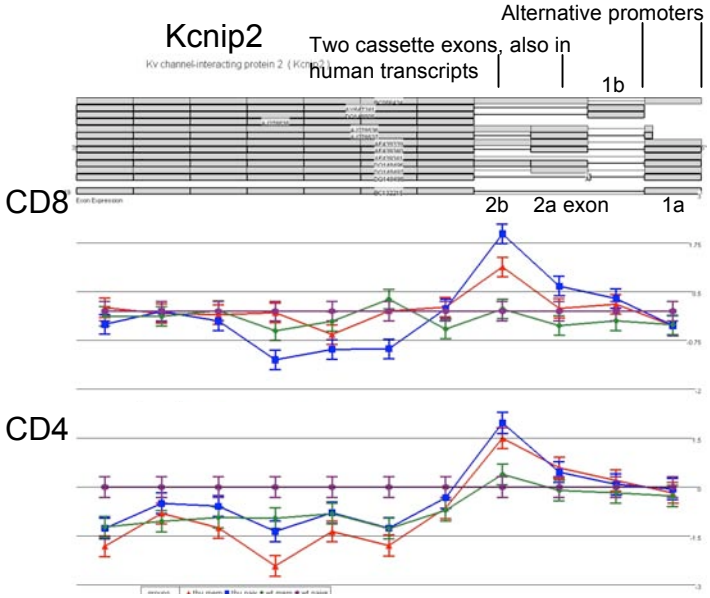

Splicing out of entire coding region in mouse, but multiple coding exons in human
Foxd1
CD8
CD4

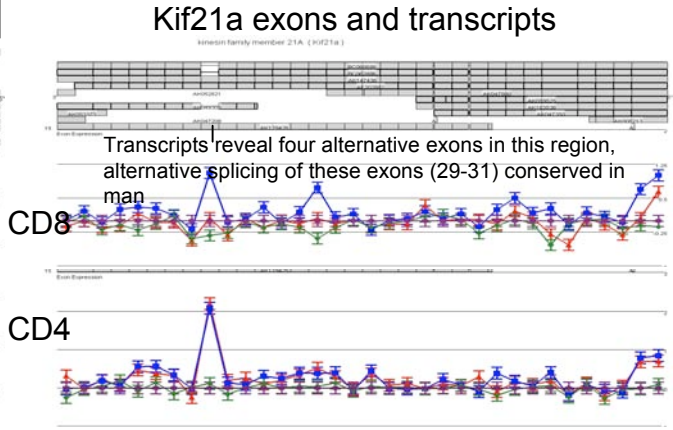

Alt promoter and initiation in exon 6 = Gabbr1b
Gabbr1
Start site for Gabbr1a form
Cassette exon also in human = Gabbr1e secreted form
= Gabbr1g form, suppresses exon 4 splice donor
CD8
CD4

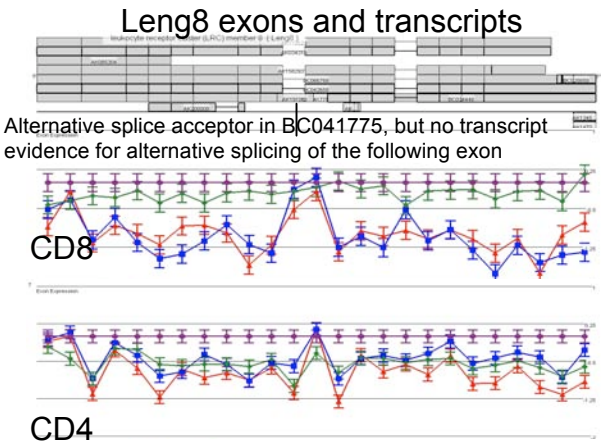

Two alternative promoter starts confirmed by multiple transcripts - 5' promoter = Grip1b
Grip1
Mulitple transcripts confirm alt splicing of this exon (2)
CD8
Transcript confirmed cassette exon
CD4

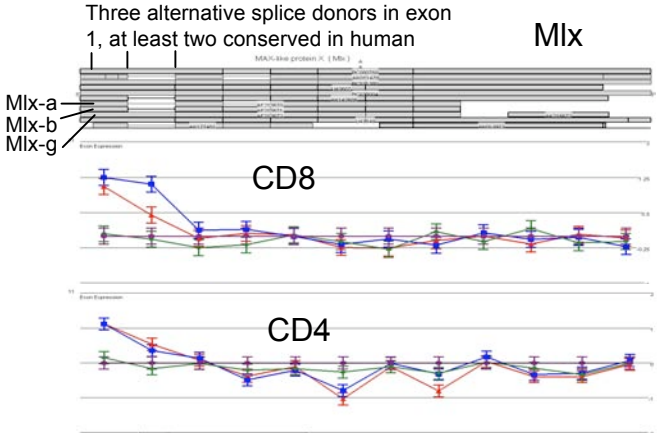

Splicing silenced in exon 15 to enable alternative 3' end
Itsn1
Alternative 3' stops, also human = Isn-long and Isn-Short
Alternative exon6 acceptor, also in zebrafish, human = isoform 4
CD8
CD4

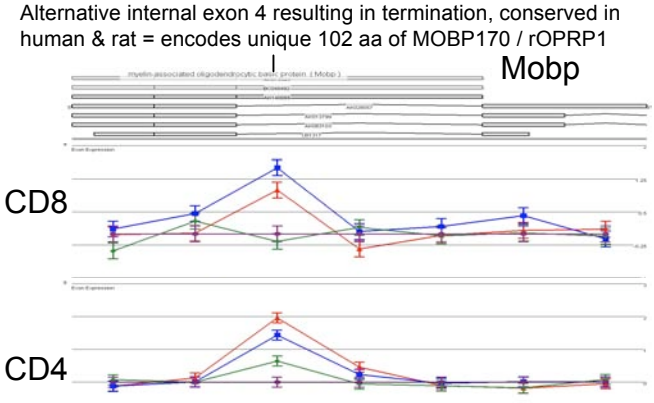
Group A

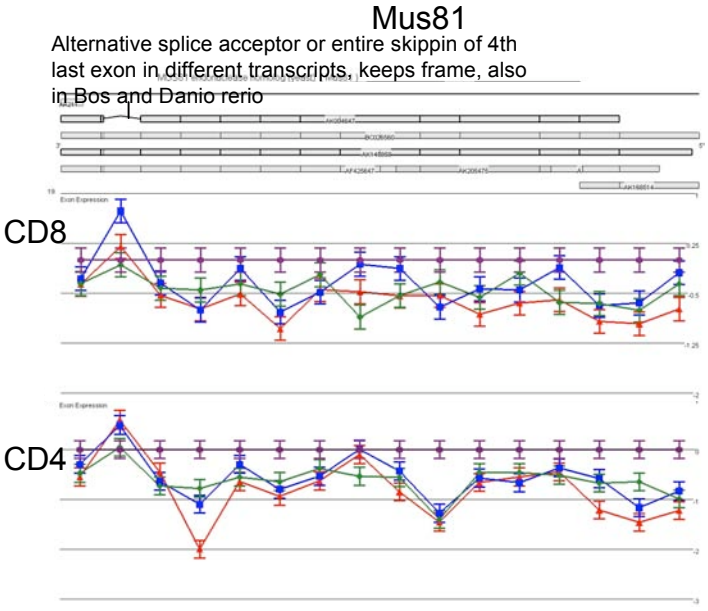

Kcnip2
Alternative promoters
Two cassette exons, also in human transcripts
1b
2b
2a exon
1a
CD8
CD4


Kif21a exons and transcripts
Transcripts reveal four alternative exons in this region, alternative splicing of these exons (29-31) conserved in man
CD8
CD4


Leng8 exons and transcripts
Alternative splice acceptor in BC041775, but no transcript evidence for alternative splicing of the following exon
CD8
CD4

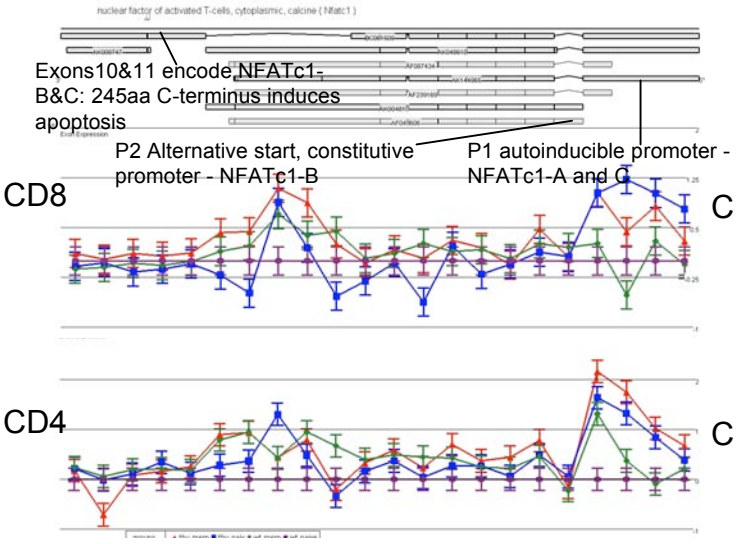

Three alternative splice donors in exon 1, at least two conserved in human
Mlx
Mlx-a
Mlx-b
Mlx-g
CD8
CD4


Alternative internal exon 4 resulting in termination, conserved in human & rat = encodes unique 102 aa of MOBP170 / rOPRP1
Mobp
CD8
CD4

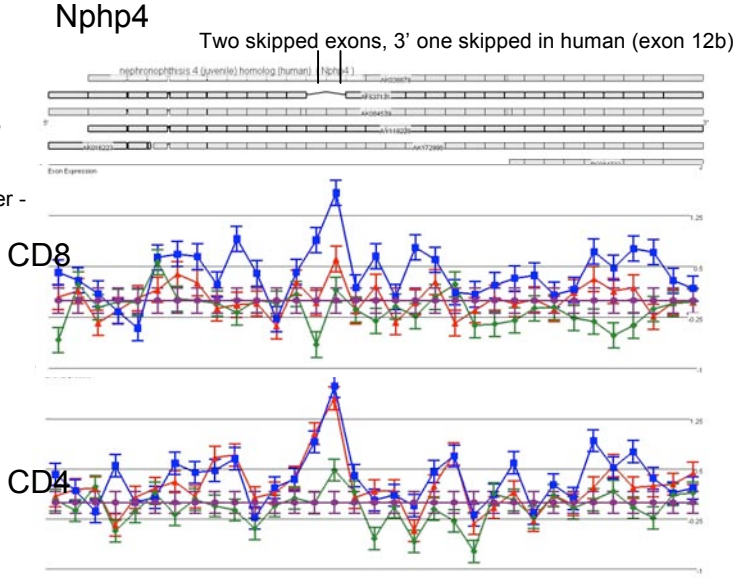

Mus81
Alternative splice acceptor or entire skippin of 4th last exon in different transcripts, keeps frame, also in Bos and Danio rerio
CD8
CD4

Group A

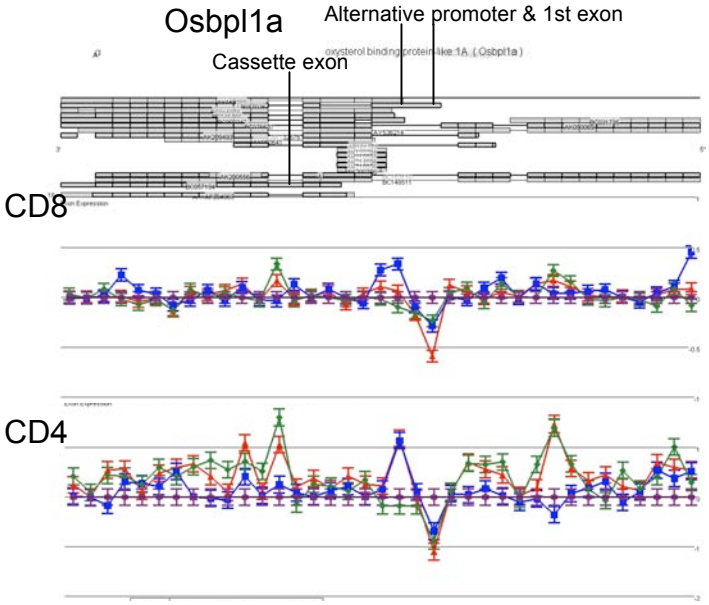
Nfatc1



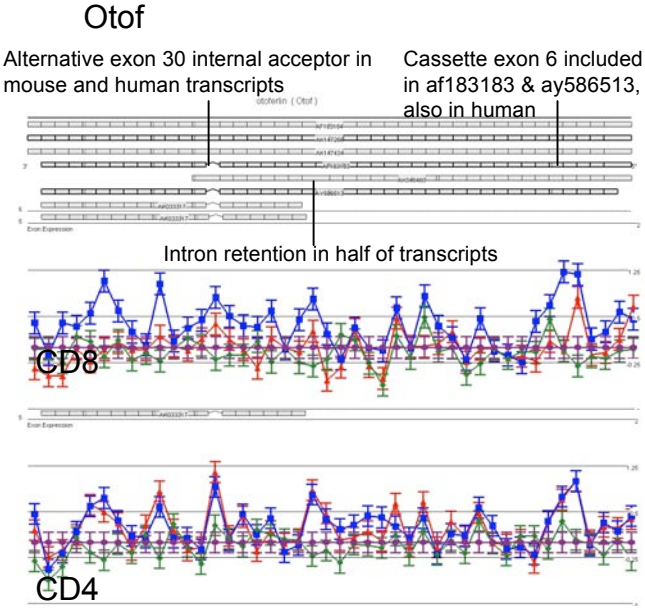
Nphp4



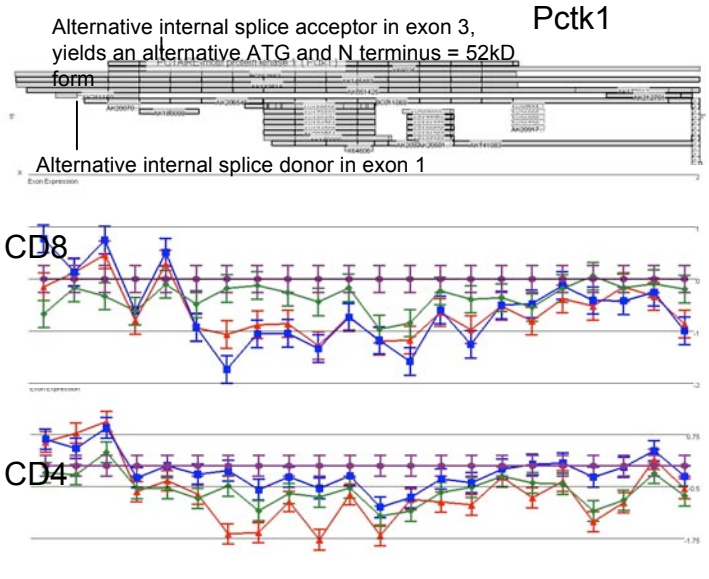
Osbpl1a



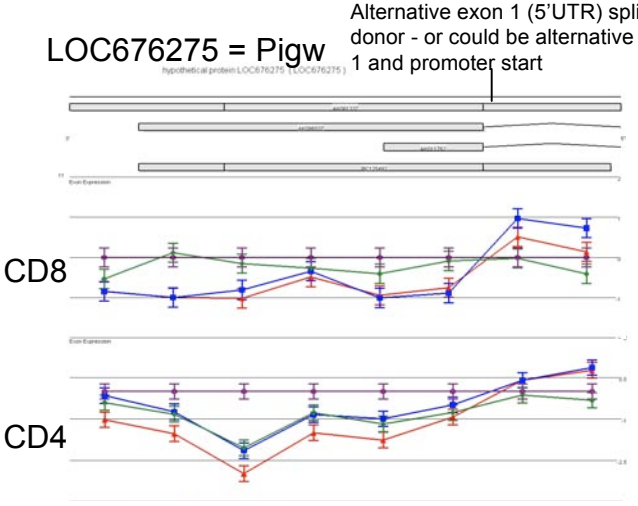
Otof



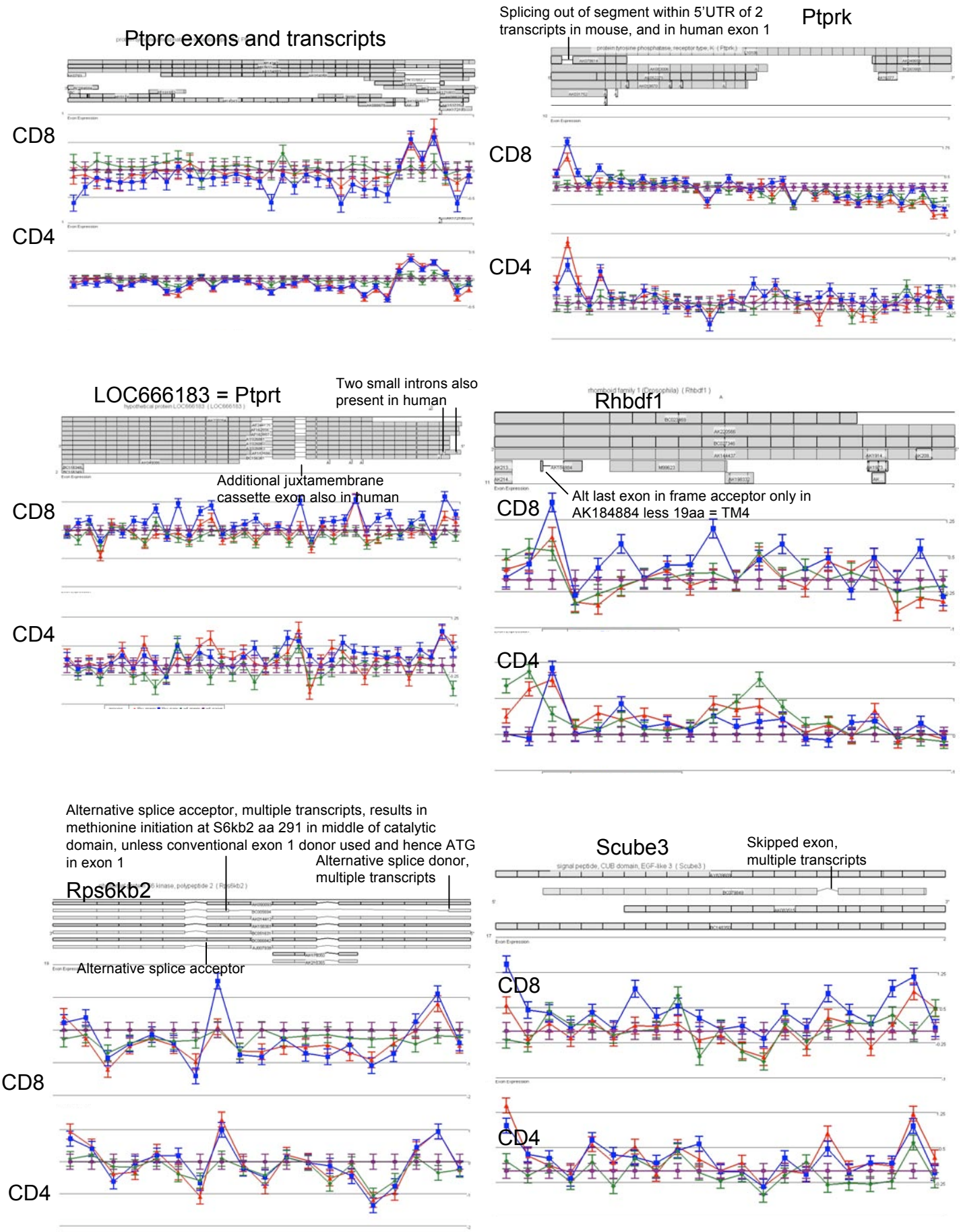
Pctk1



LOC676275 = Pigw



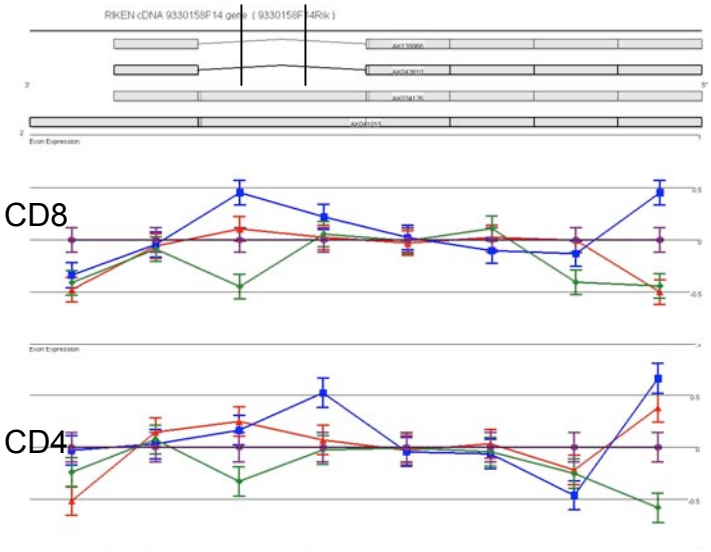
Group A


Ptprc exons and transcripts
CD8
CD4
Splicing out of segment within 5'UTR of 2 transcripts in mouse, and in human exon 1
Ptprk
CD8
CD4
LOC666183 = Ptprt
Two small introns also present in human
Additional juxtamembrane cassette exon also in human
CD8
CD4
Rhbdf1
Alt last exon in frame acceptor only in AK184884 less 19aa = TM4
CD8
CD4
Alternative splice acceptor, multiple transcripts, results in methionine initiation at S6kb2 aa 291 in middle of catalytic domain, unless conventional exon 1 donor used and hence ATG in exon 1
Alternative splice donor, multiple transcripts
Rps6kb2
Alternative splice acceptor
CD8
CD4
Scube3
Skipped exon, multiple transcripts
CD8
CD4

Group A

9330158F14Rik = Homo SLC38A11

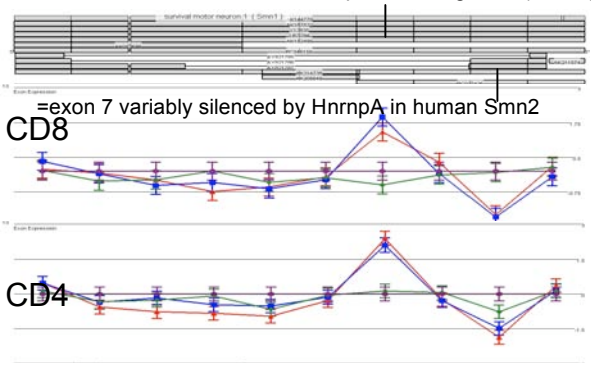
Two alternative exons yield early termination and 3'UTR



Smn1

This exon is selectively absent from a subset of human transcripts encoding Smn (exon5)

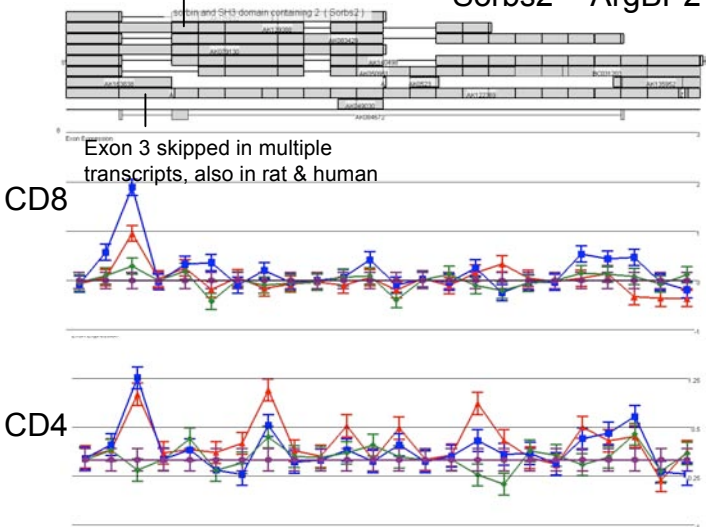
=exon 7 variably silenced by HnrnpA in human Smn2



Sorbs2 = ArgBP2

Exon 4 skipped in multiple transcripts

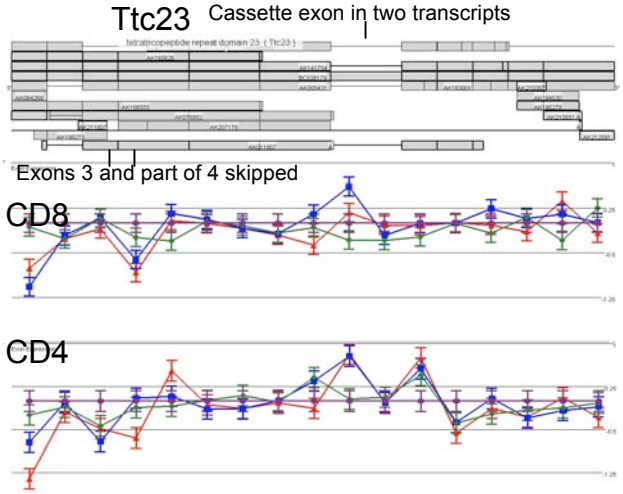
Exon 3 skipped in multiple transcripts, also in rat & human



Ttc23

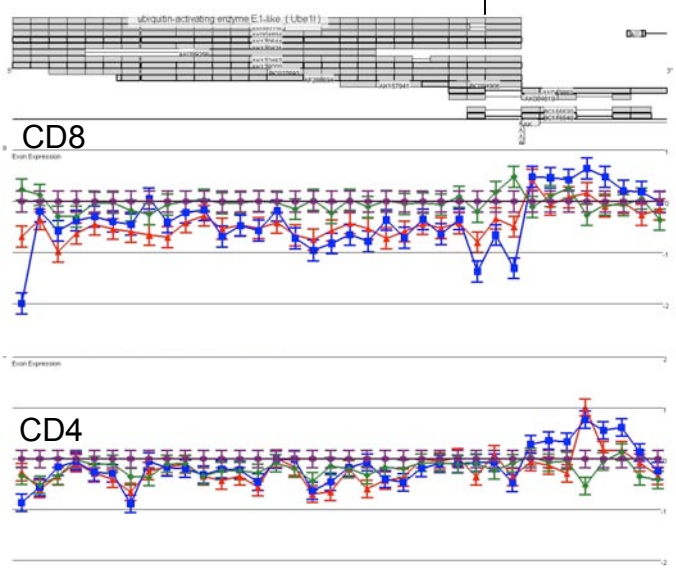
Cassette exon in two transcripts

Exons 3 and part of 4 skipped



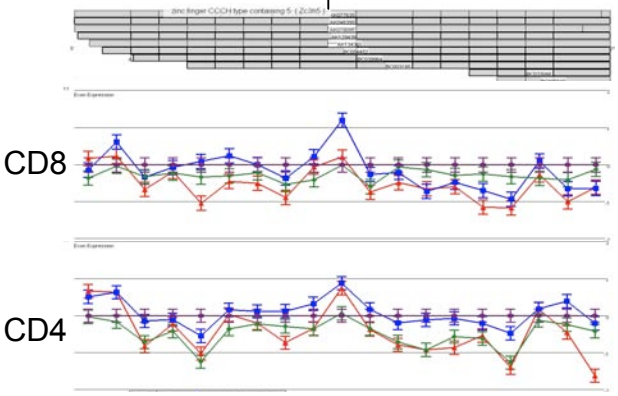
Ube1l

Alternative splicing to give alternative 3'UTRs



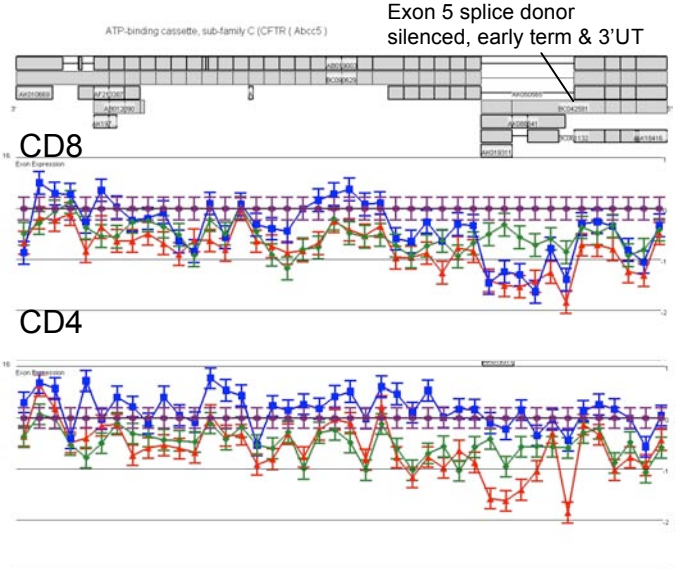
Zc3h5 = Unk

Alternative splice donor within exon 8, in frame deletion of 13aa

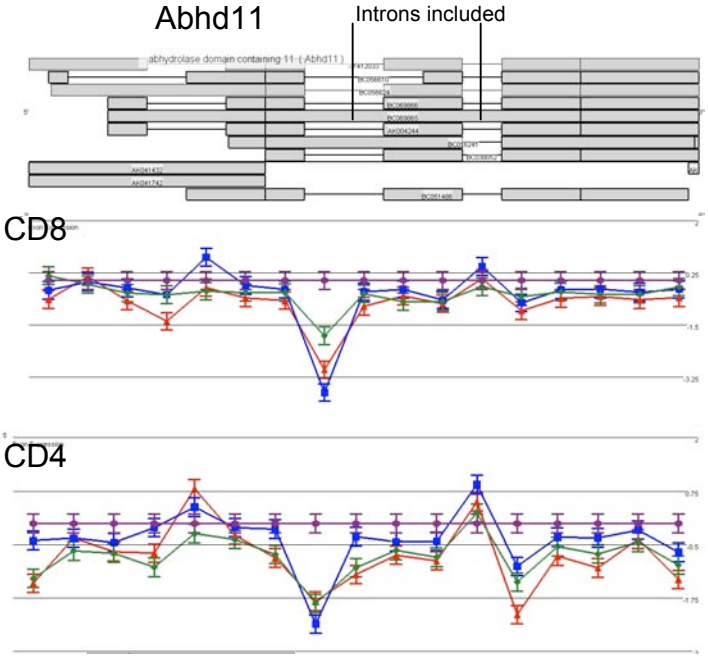


Group B

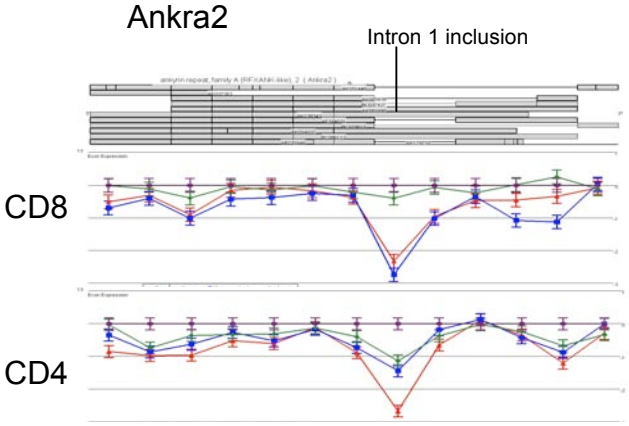
Abcc5



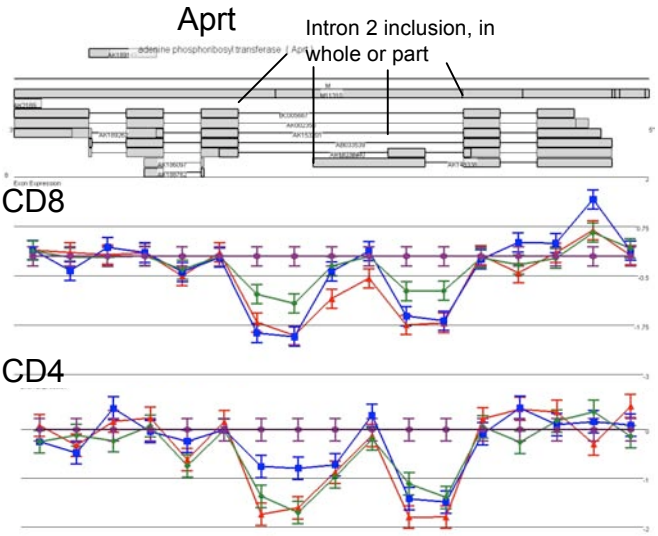
Abhd11



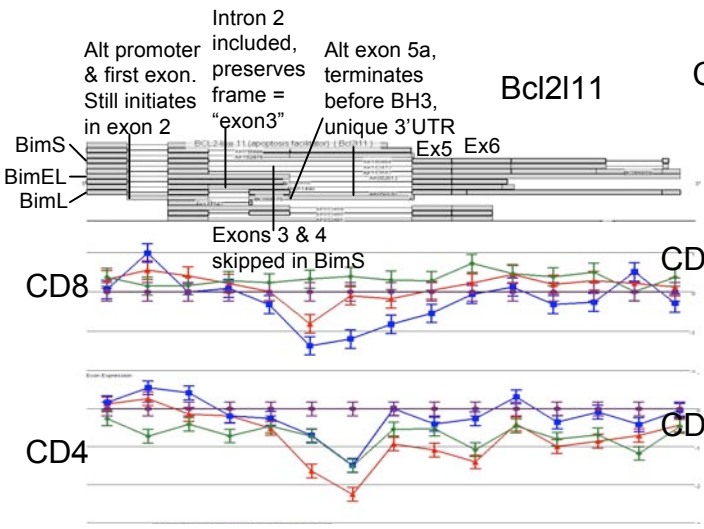
Ankra2



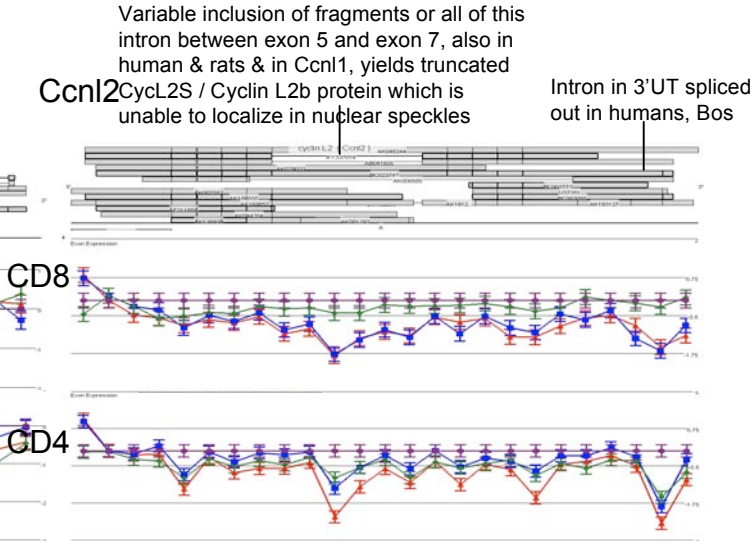
Aprt



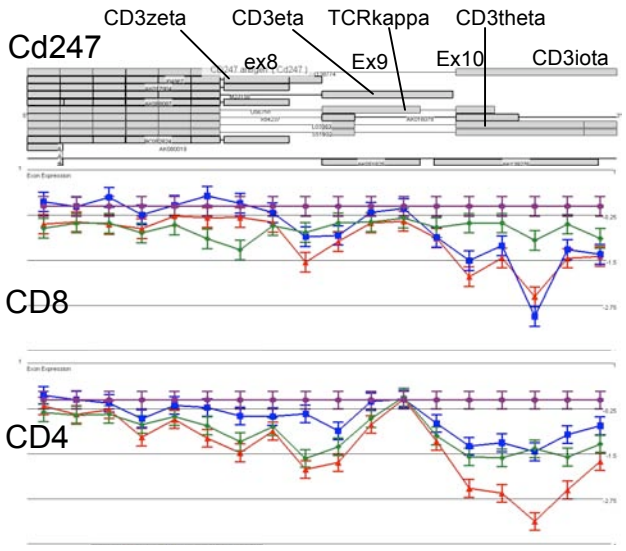
Bcl2l11

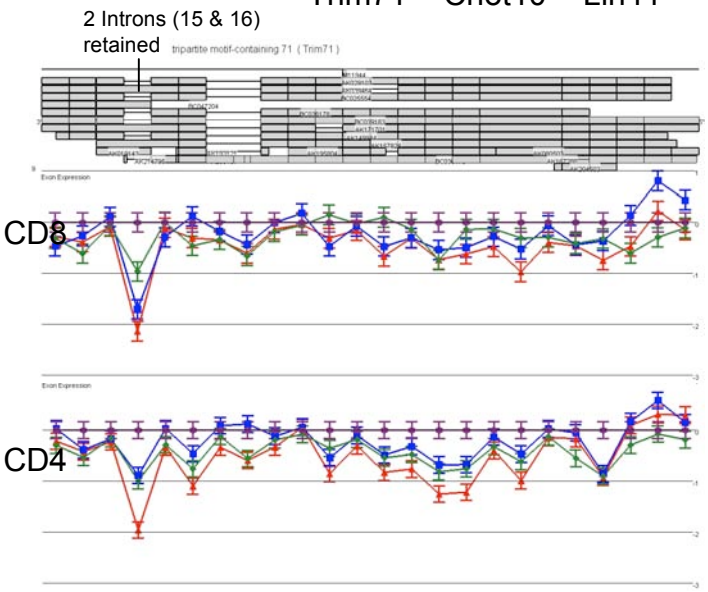


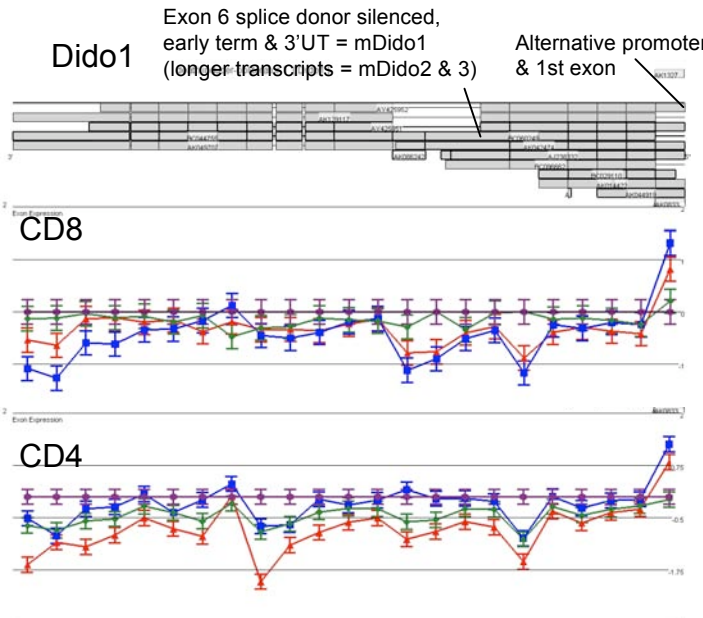
Ccnl2

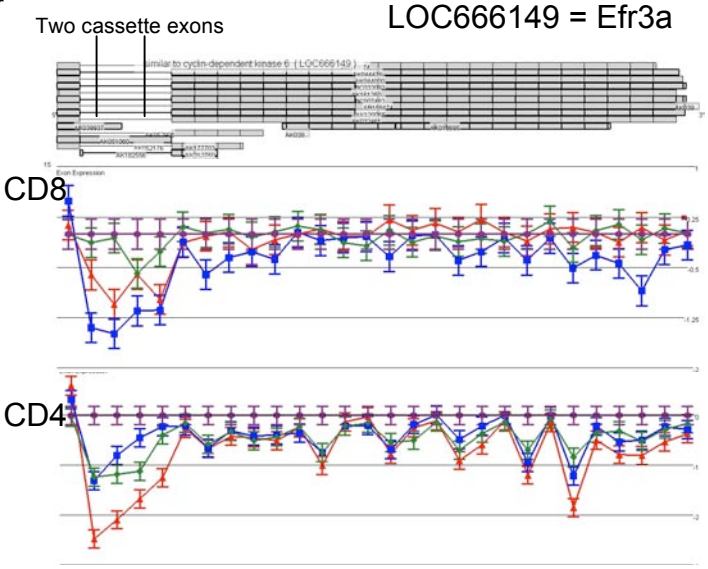


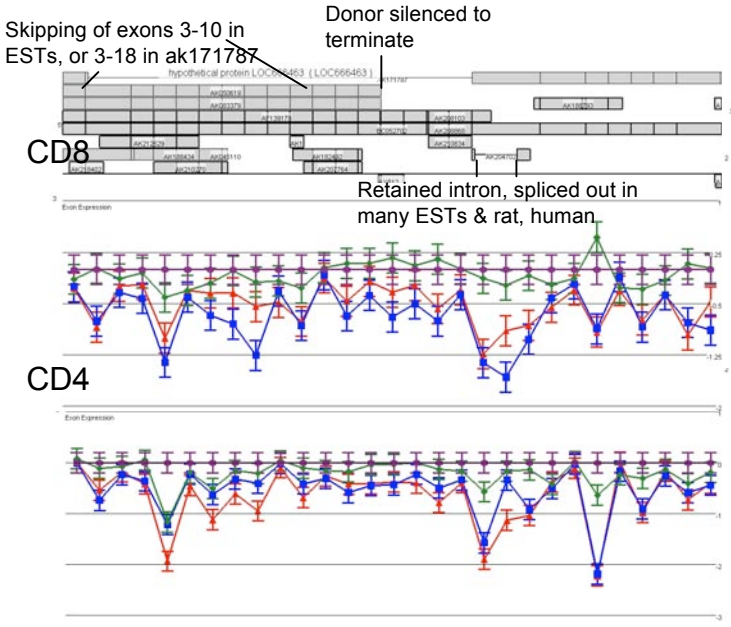
Group B

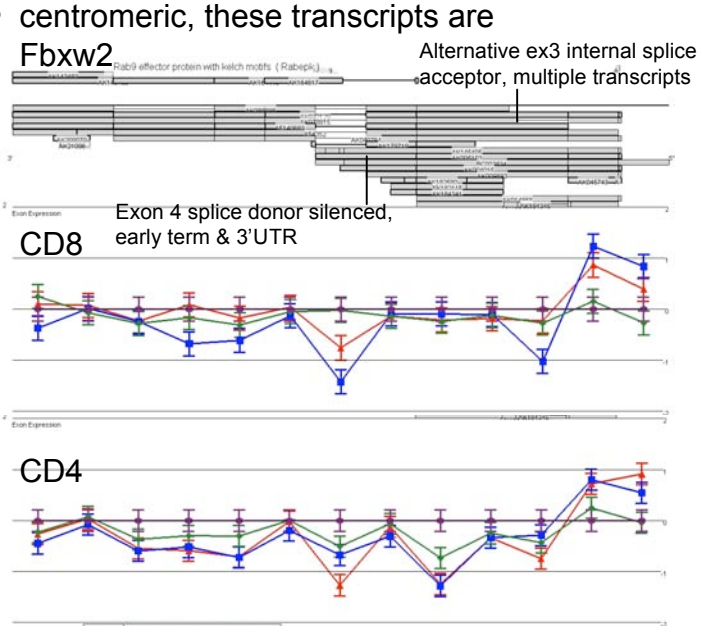

CD3zeta
CD3eta
TCRkappa
CD3theta
Cd247
ex8
Ex9
Ex10
CD3iota
CD8
CD4


Trim71 = Cnot10 = Lin41
2 Introns (15 & 16) retained
CD8
CD4


Dido1
Exon 6 splice donor silenced, early term & 3'UT = mDido1 (longer transcripts = mDido2 & 3)
Alternative promoter & 1st exon
CD8
CD4

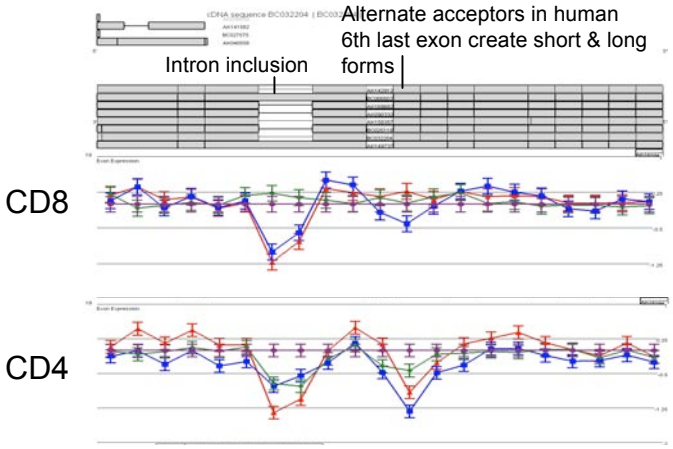

LOC666149 = Efr3a
Two cassette exons
CD8
CD4


Loc666463= FAM48a = P38IP
Skipping of exons 3-10 in ESTs, or 3-18 in ak171787
Donor silenced to terminate
Retained intron, spliced out in many ESTs & rat, human
CD8
CD4

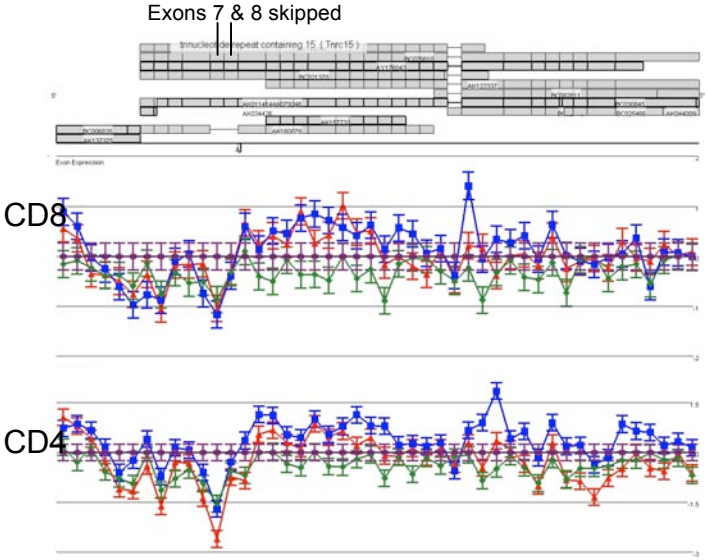

Rabepk (actually this is centromeric, these transcripts are Fbxw2
Alternative ex3 internal splice acceptor, multiple transcripts
Exon 4 splice donor silenced, early term & 3'UTR
CD8
CD4

Group B

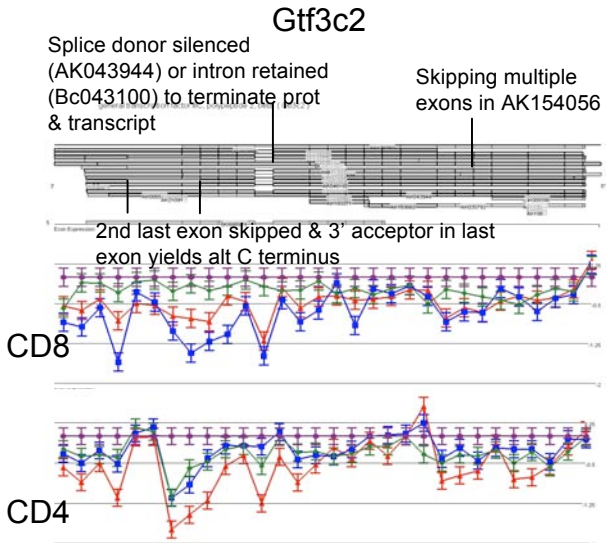
BC032204 = FERMT3



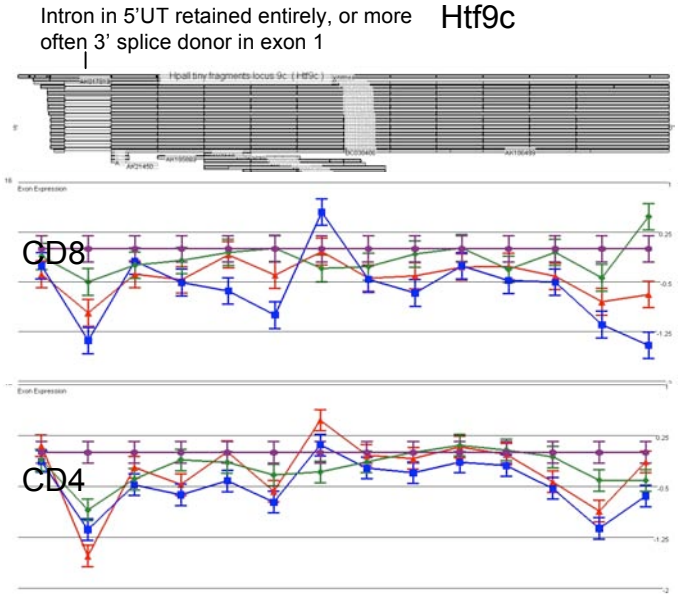
Tnrc15 = Gigyf2



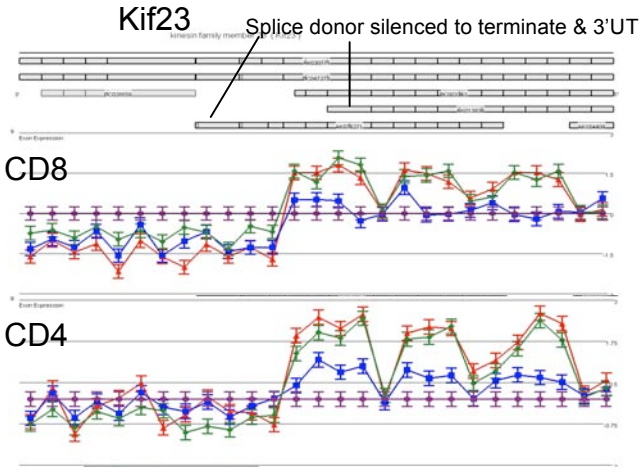
Gtf3c2



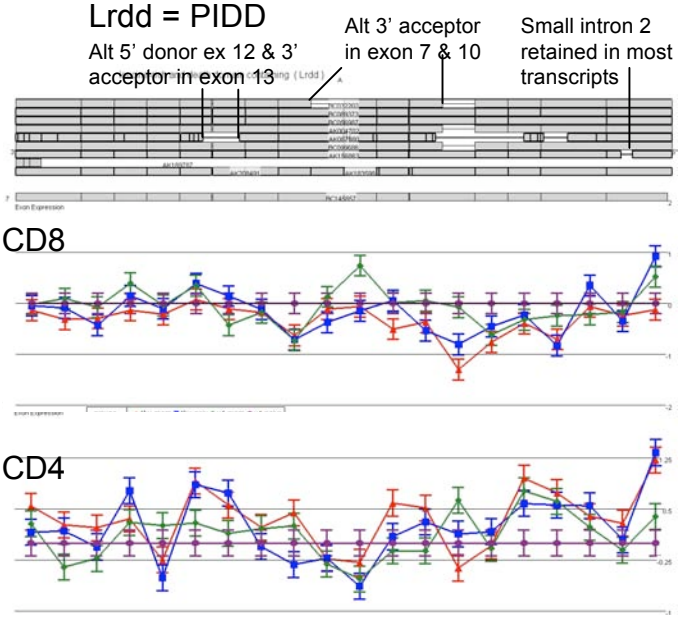
Htf9c



Kif23

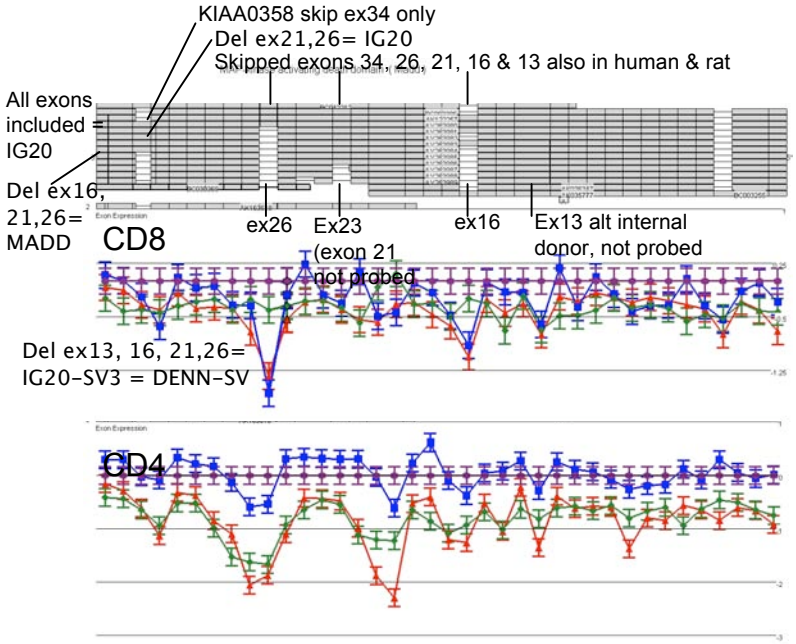


Lrdd = PIDD

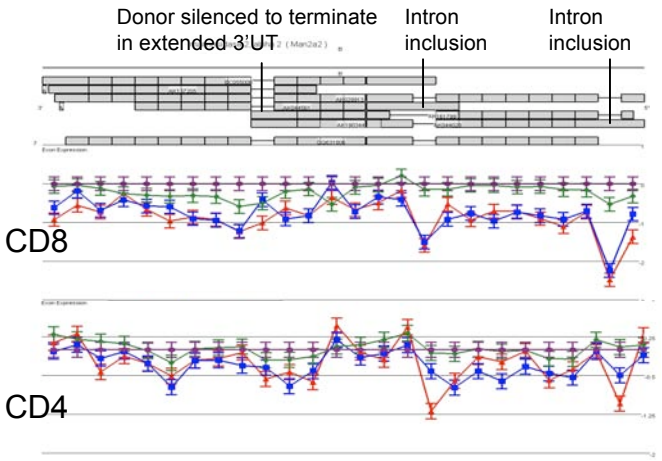


Group B

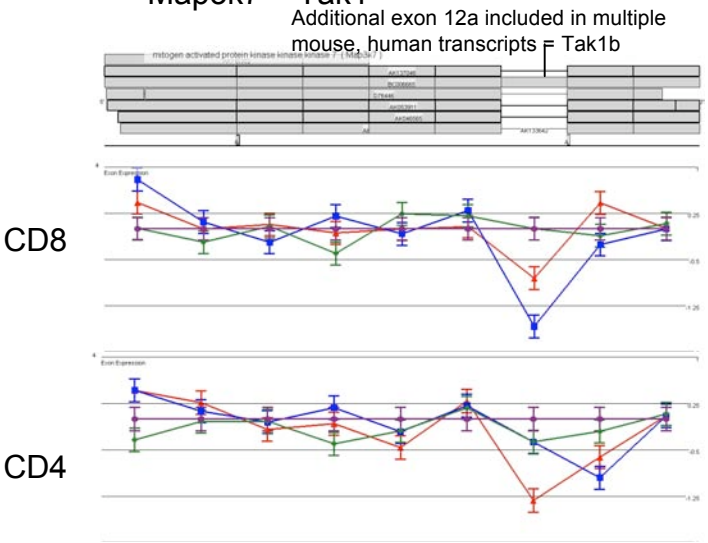
Madd



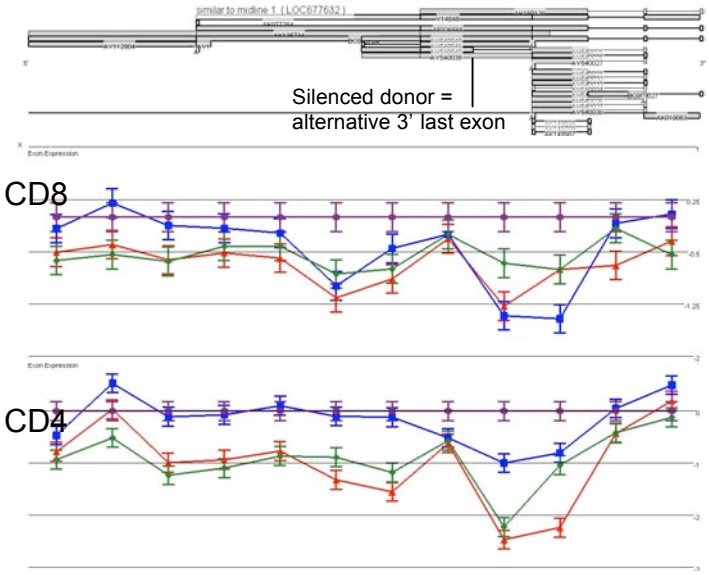
Man2a2



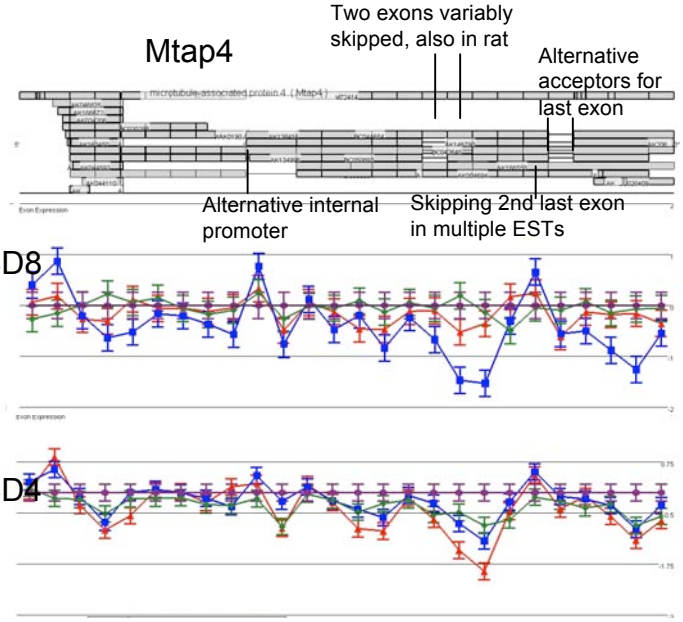
Map3k7 = Tak1



LOC667632 = Mid1



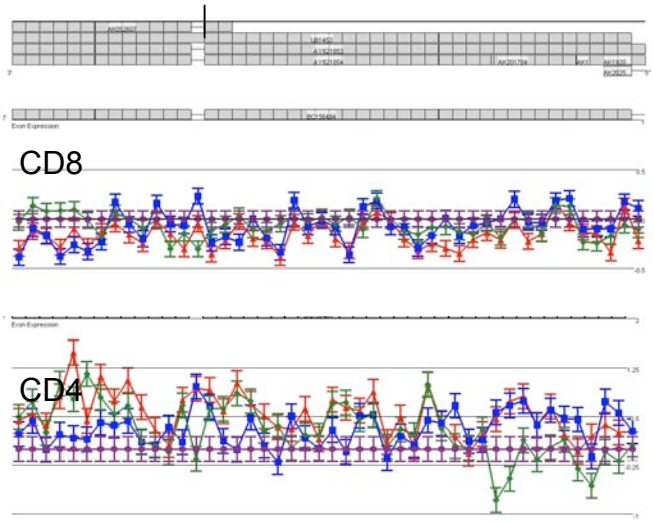
Mtap4



Group B

Myo7a

Alternative ex34 splice acceptor in intron, also in human

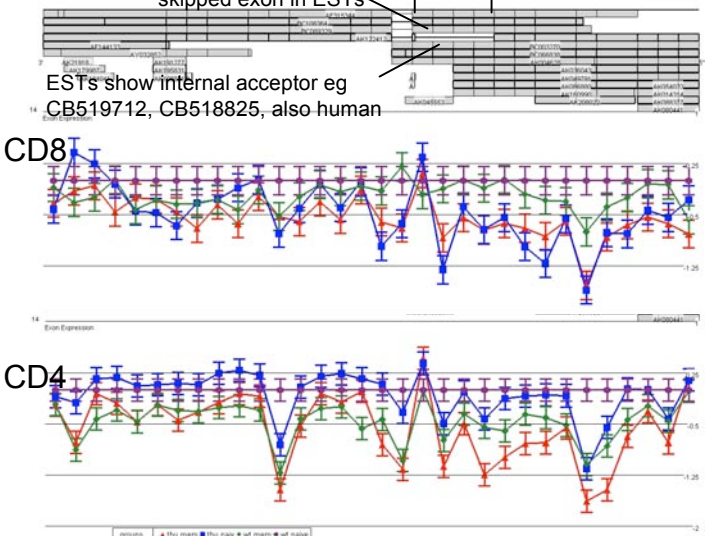


Nisch

Internal splice donor and acceptor and 2 variably skipped exons

skipped exon in ESTs

ESTs show internal acceptor eg CB519712, CB518825, also human



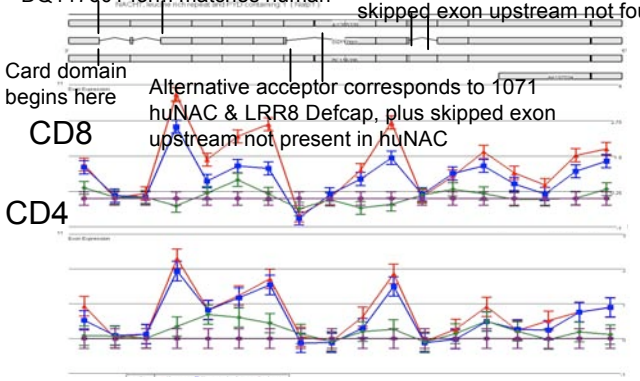
Nlrp1 = Nalp1

Alternative donors and acceptors between 2nd last and last exons,in frame, DQ117601 form matches human

Alternative acceptor corresponds to 901 huNAC in LRR4 & LRR5 in huDEFCAP, plus skipped exon upstream not found in huNAC

Card domain begins here

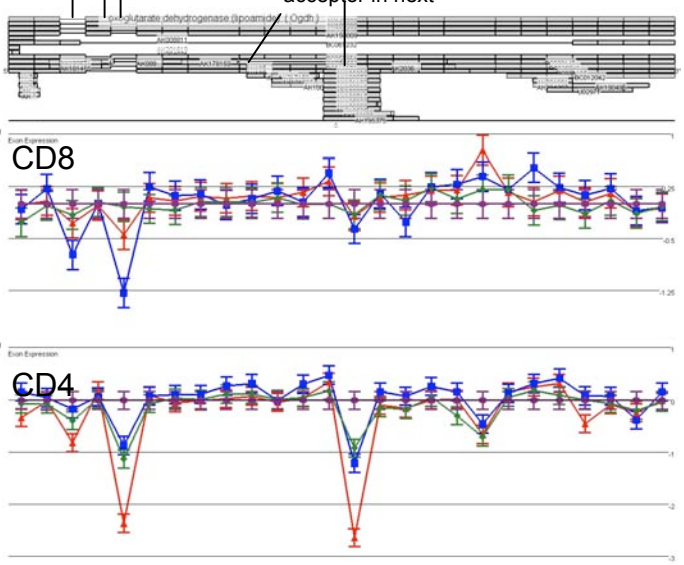
Alternative acceptor corresponds to 1071 huNAC & LRR8 Defcap, plus skipped exon upstream not present in huNAC



Ogdh

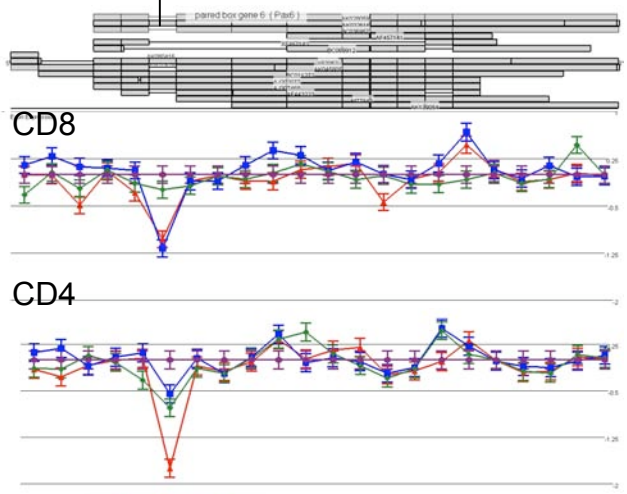
Alternative cassette exons

3 skipped exons + internal acceptor in next



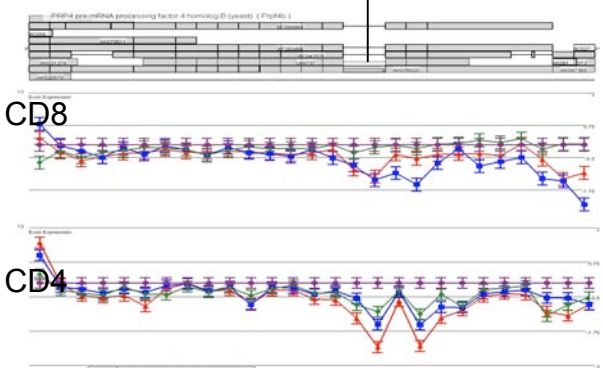
Pax6

5a exon encoding insert in paired domain variably skipped, also in man, rat, fish, quail etc



Prpf4b

Additional exon in BC003769

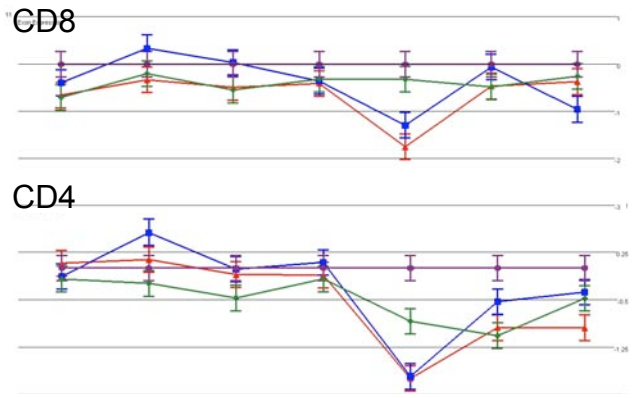


Group B

Rpain

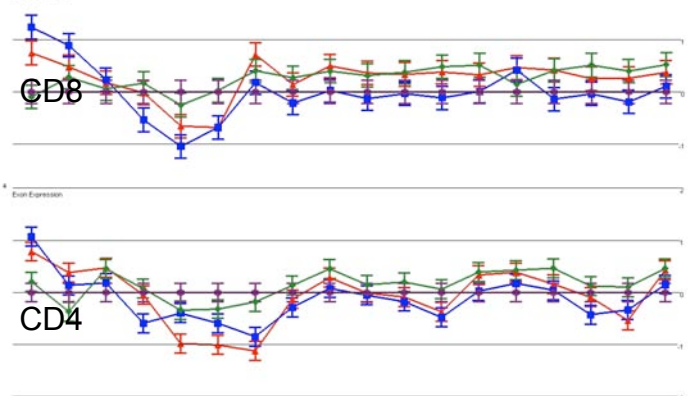
Skipped exon 4 = hRIPdelta1;
skip 4, 5 and 6= hRIPdelta3
also in human

Exon 5 Exon 7



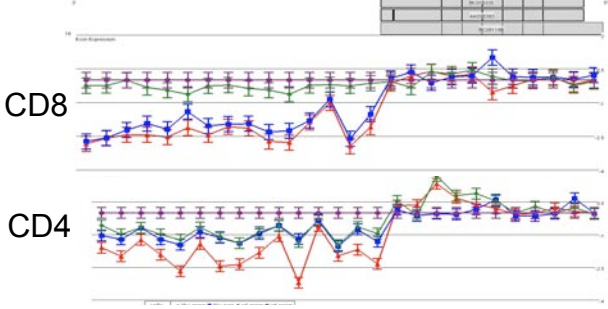
Rragc

Ex 2 Splice donor silenced, 3'
donor instead, early term



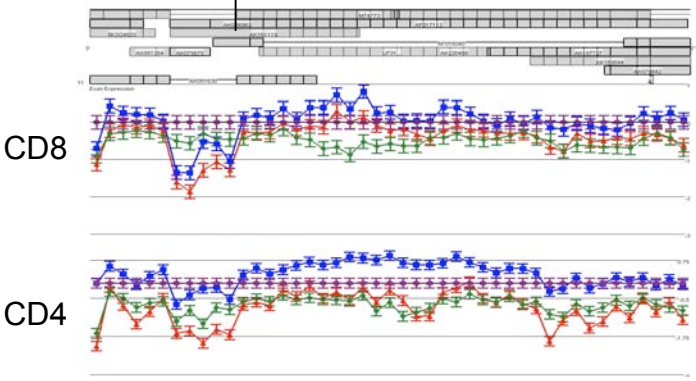
Slc15a2

Exon 7 splice donor suppressed, yielding
early 3'UT and termination



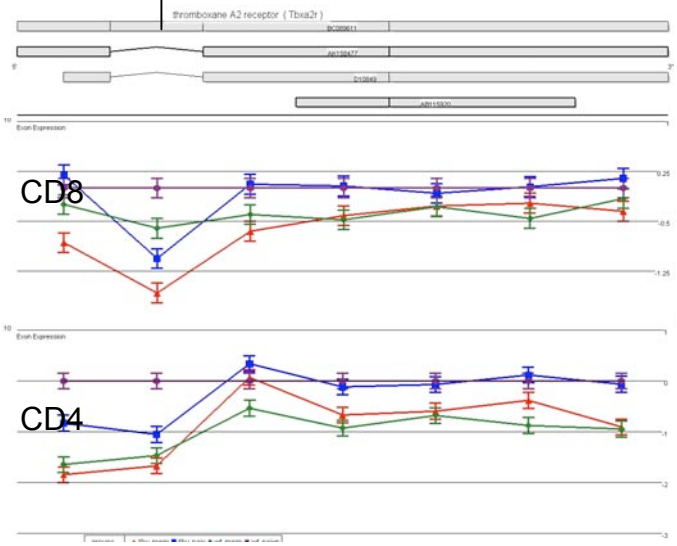
Spnb2

Donor silenced, early term & 3'UT = short form of Spectrin
beta 2 lacking PH domain and fodraxin binding site



Tbxa2r

Additional exon in 5'UT

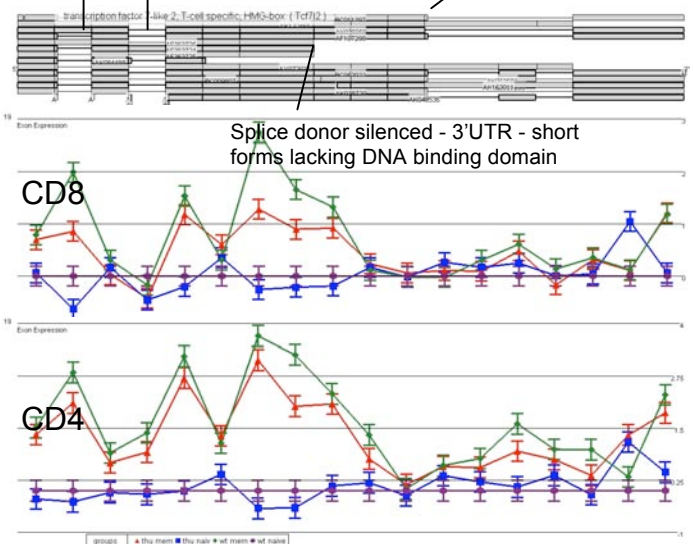


Tcf7l2 = TCF-4

Skipped exons 2 (1a) & 4
(2a) , also in human

Donor silenced,
early term & short
3'UT

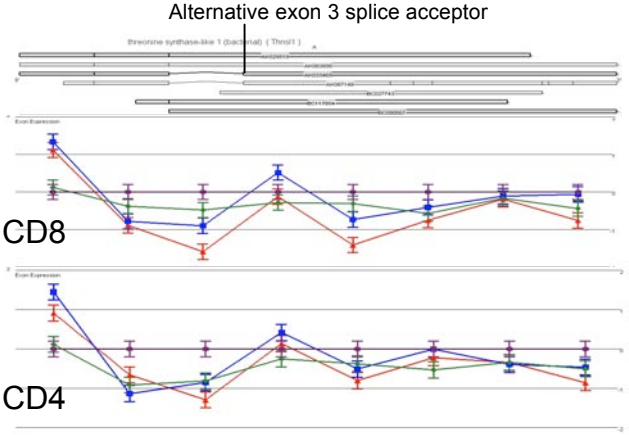
Splice donor silenced - 3'UTR - short
forms lacking DNA binding domain



Group B

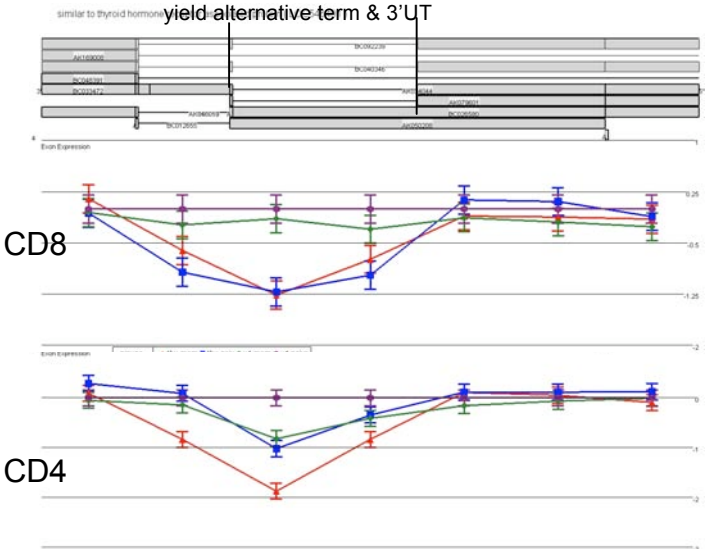
Thnsl1

Alternative exon 3 splice acceptor



LOC546401 = Thrap3 = Trap150

Splice donors silenced to yield alternative term & 3'UT



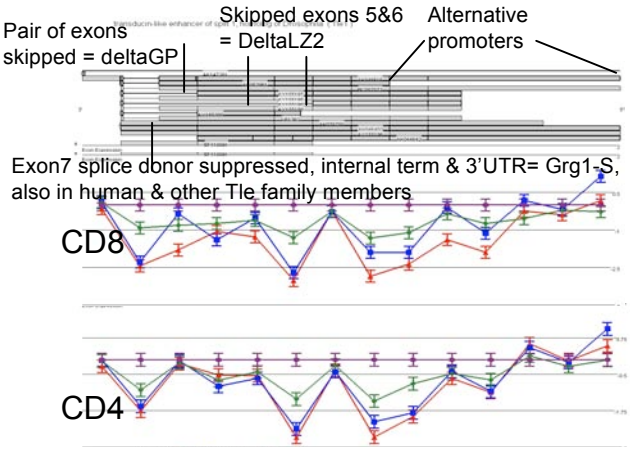
Tle1 = Grg1

Pair of exons skipped = deltaGP

Skipped exons 5&6 = DeltaLZ2

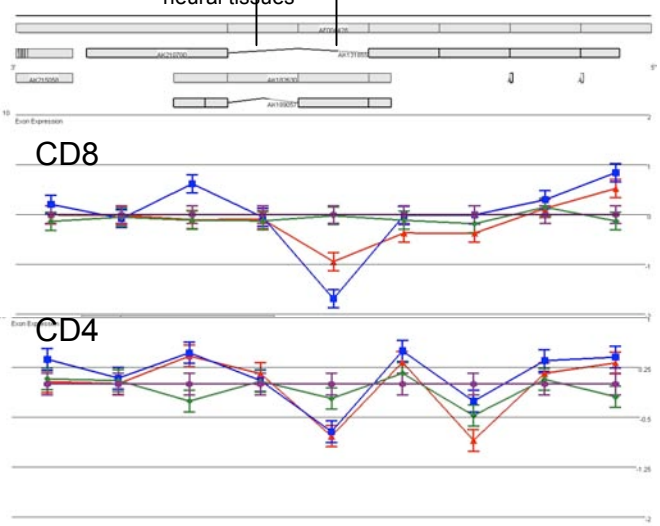
Alternative promoters

Exon7 splice donor suppressed, internal term & 3'UTR= Grg1-S, also in human & other Tle family members



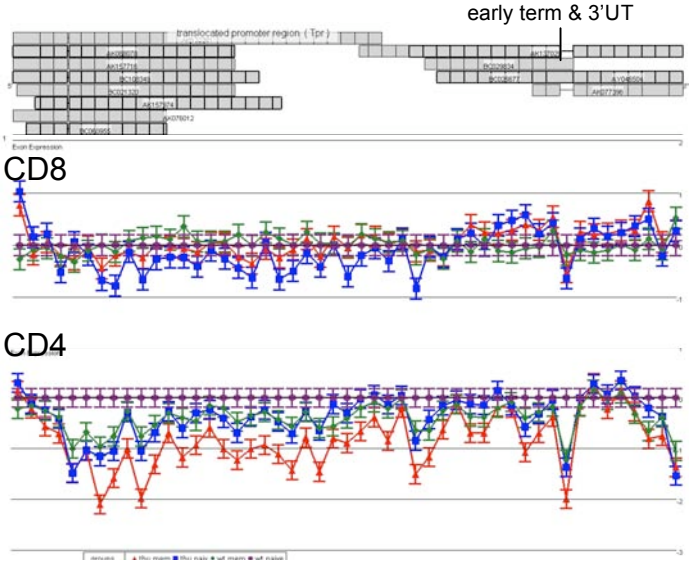
Tpd52l1 = hD53

Exons 5 & 6 skipped, also in human neural tissues



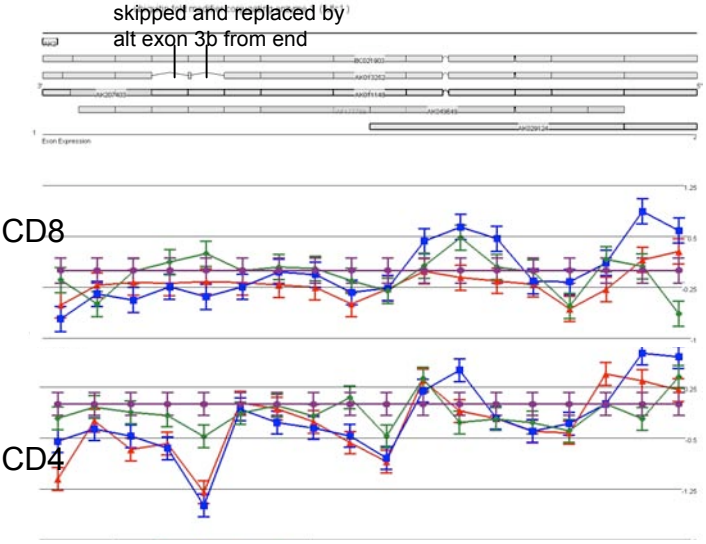
Tpr

Splice donor silenced, early term & 3'UT

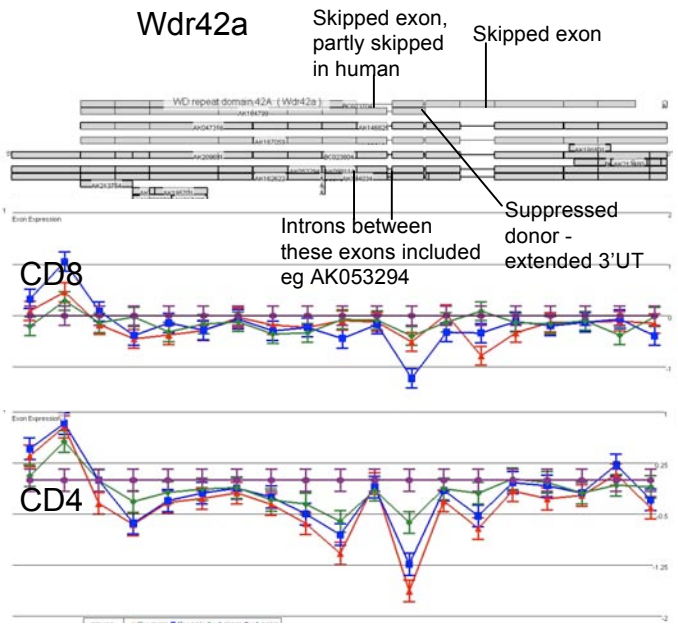


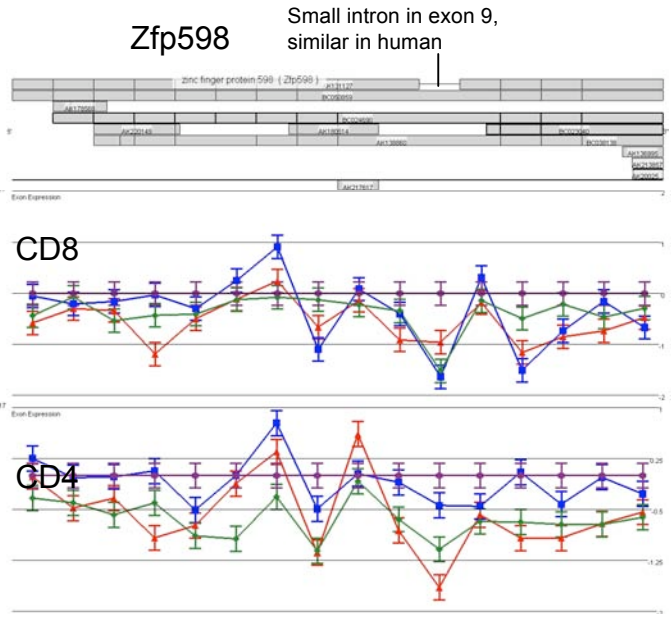
Ufc1 = Usp21

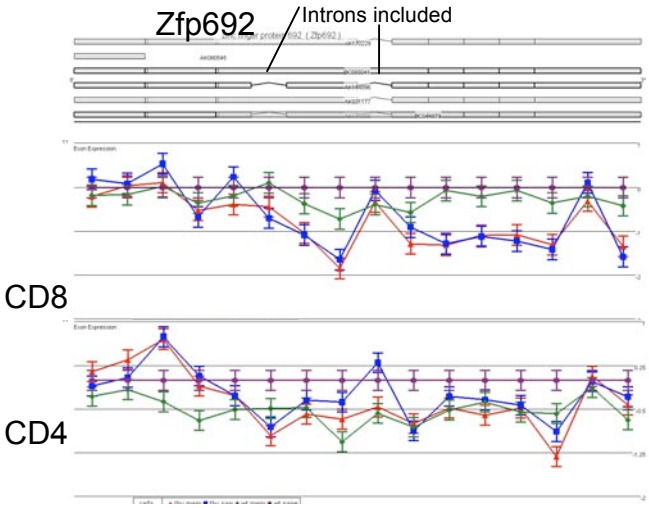
Exon 3 & 4 from end skipped and replaced by alt exon 3b from end



Group B

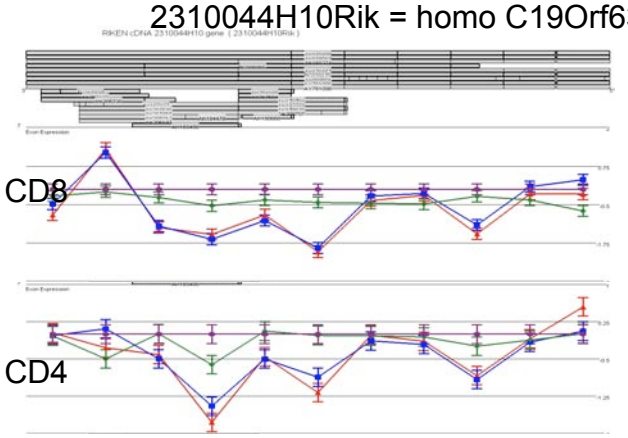

Wdr42a
Skipped exon, partly skipped in human
Skipped exon
Introns between these exons included eg AK053294
Suppressed donor - extended 3'UT
CD8
CD4


Zfp598
Small intron in exon 9, similar in human
CD8
CD4

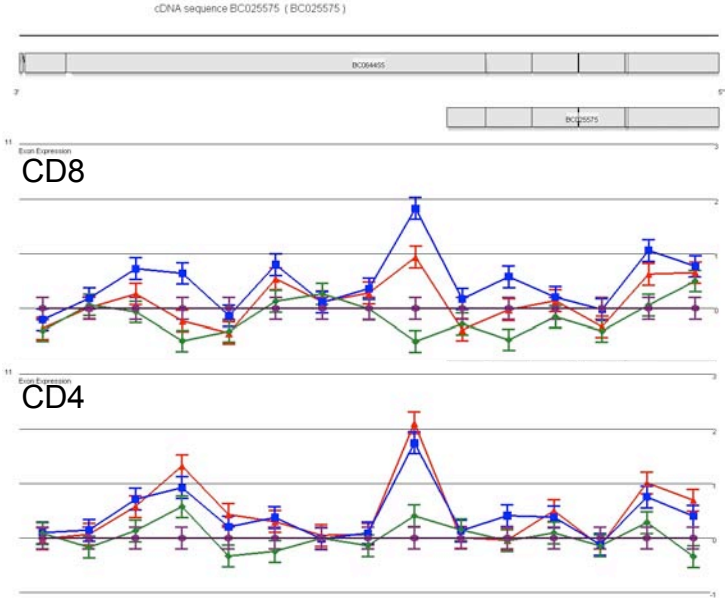

Zfp692
Introns included
CD8
CD4

Group C Genes

2310044H10Rik = homo C19Orf63

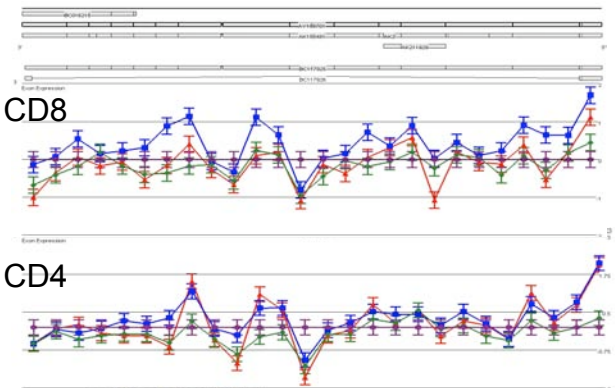


BC025575 =homo MCC34829 hypothetical protein

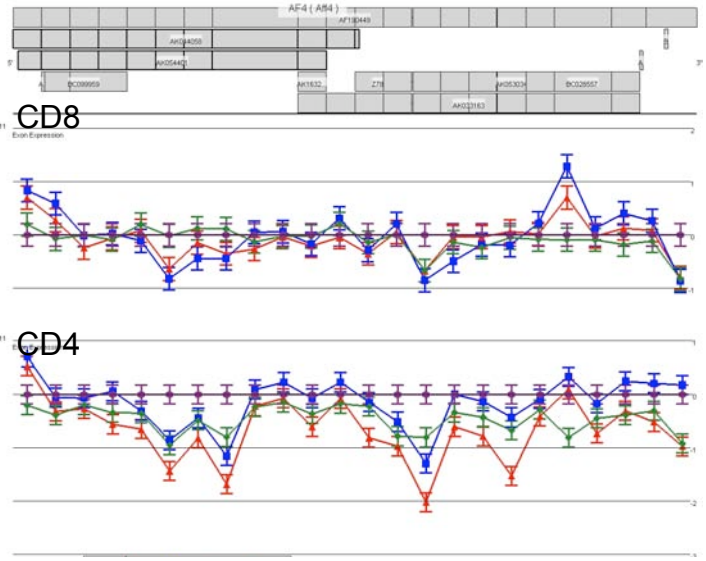


Group C Genes

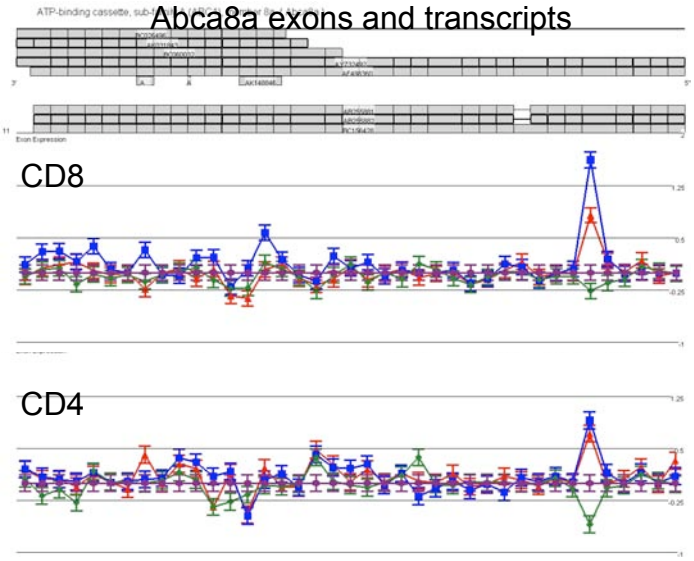
Adamtsl4 exons and transcripts



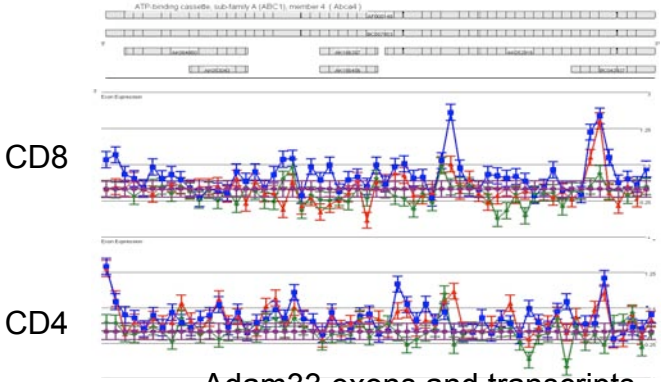
Aff4 exons and transcripts



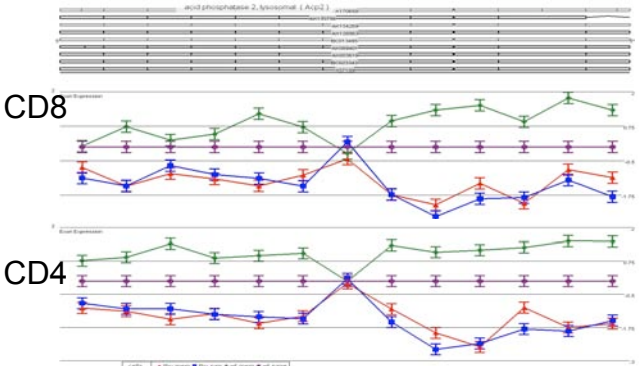
Abca8a exons and transcripts



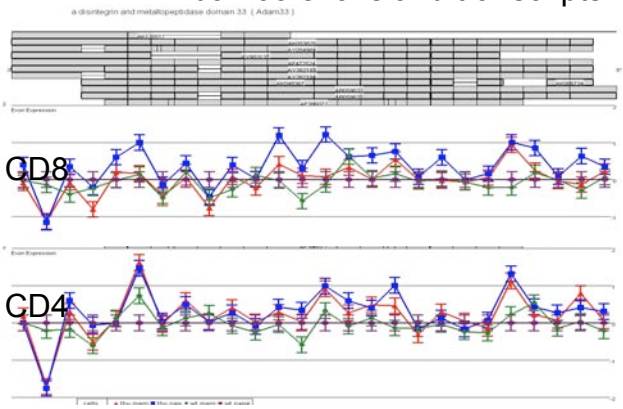
Abca4 exons and transcripts



Acp2 exons and transcripts

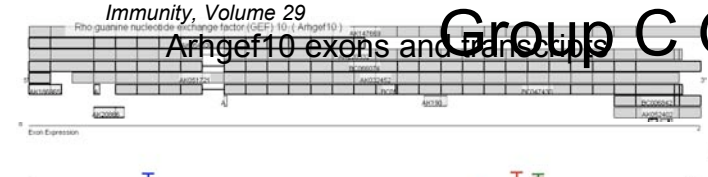


Adam33 exons and transcripts

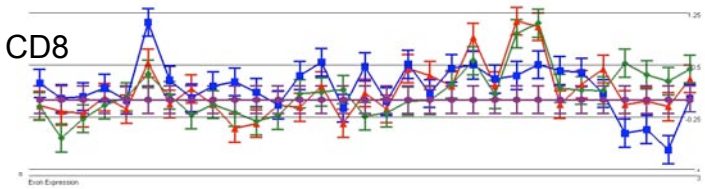


Group C Genes

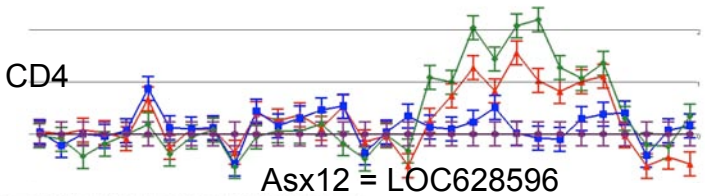
Arhgef10 exons and transcripts



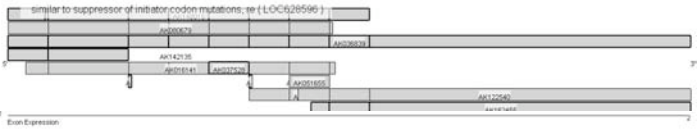
CD8



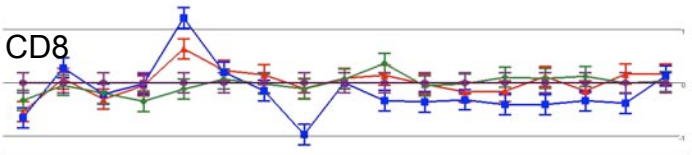
CD4

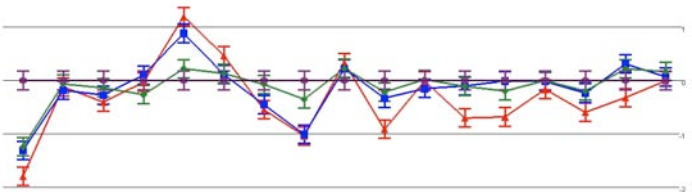


Asx12 = LOC628596



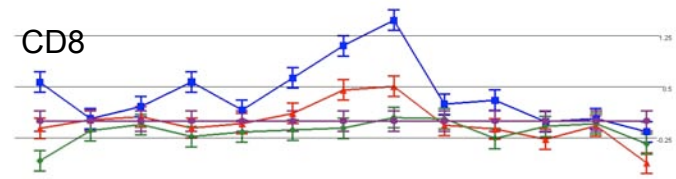
CD8



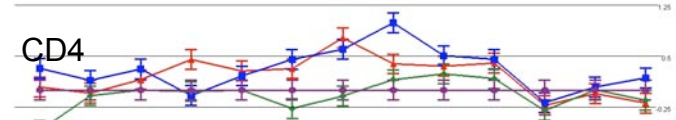


LOC669789 = Brunol6

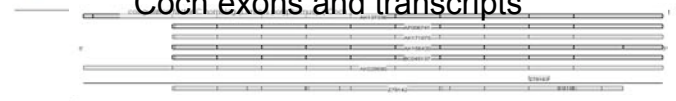
CD8



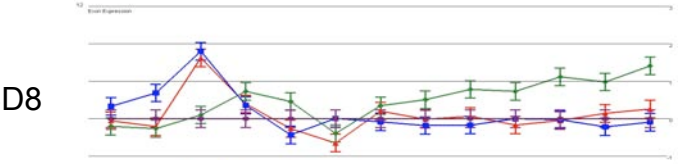
CD4



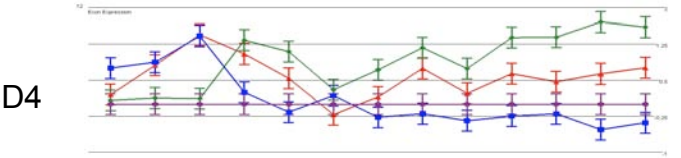
Coch exons and transcripts



CD8

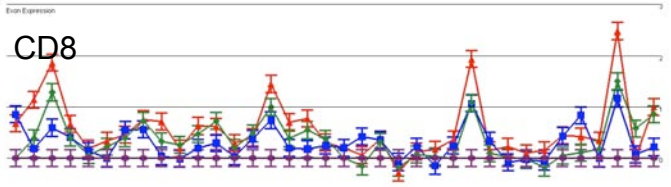


CD4

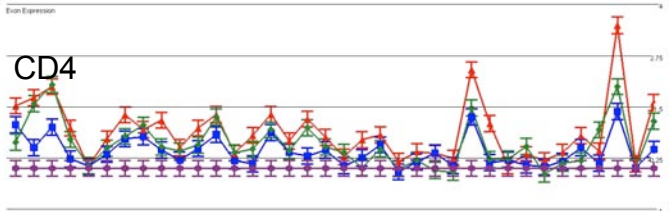


Aspm exons and transcripts

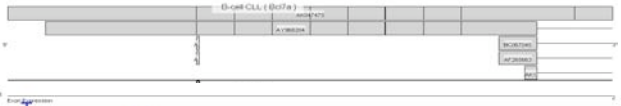
CD8



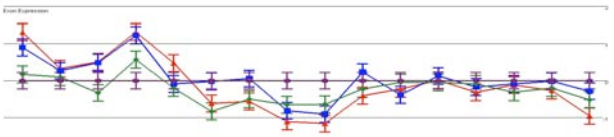
CD4



Bcl7a exons and transcripts



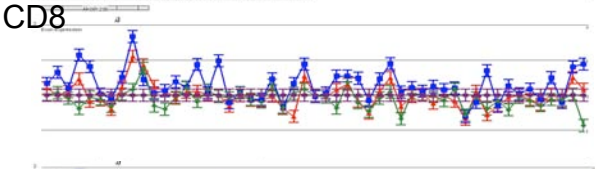
CD8



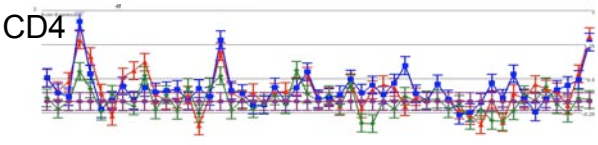
CD4

Cacna1b exons and transcripts

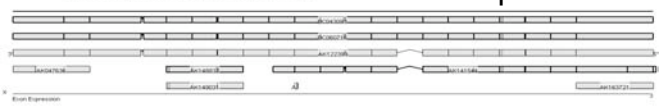
CD8



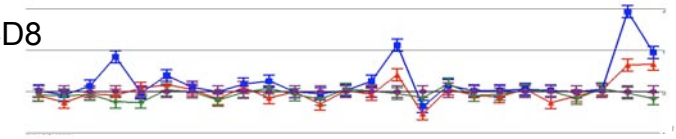
CD4



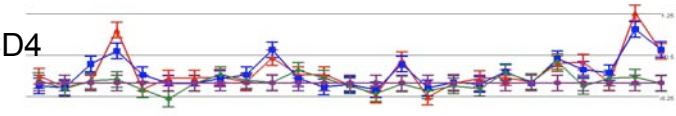
Cnksr2 exons and transcripts



CD8

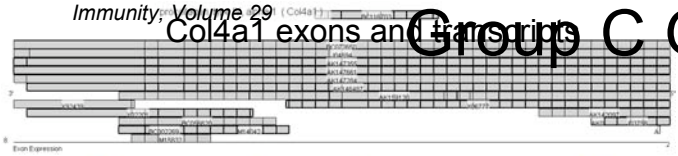


CD4

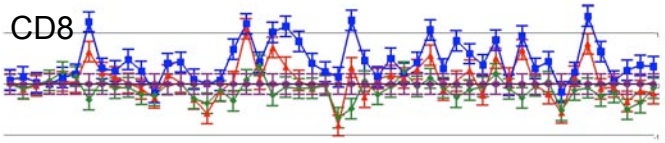


Group C Genes

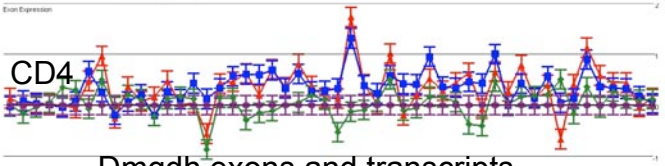
Col4a1 exons and transcripts



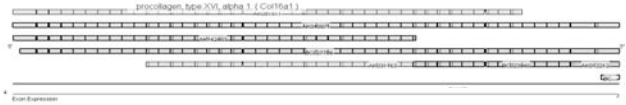
CD8



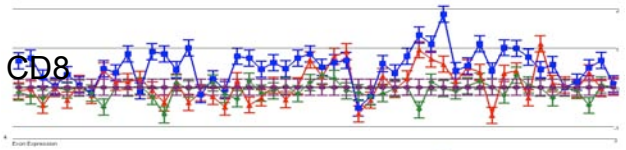
CD4



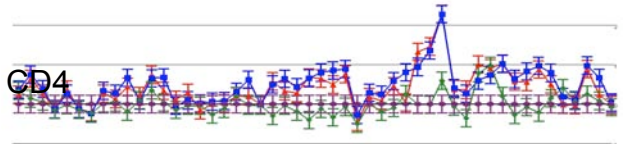
Col16a1 exons and transcripts



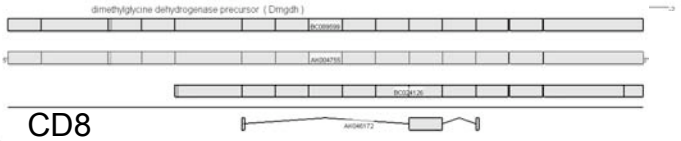
CD8



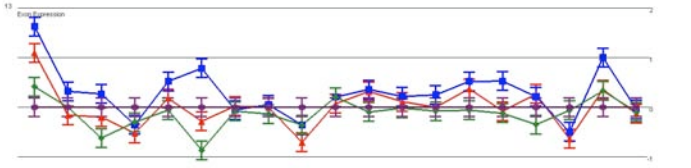
CD4



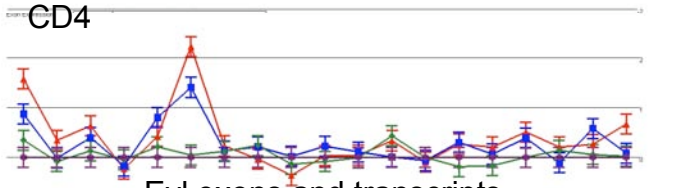
Dmgdh exons and transcripts



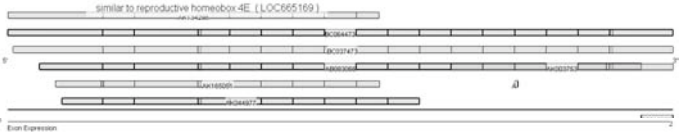
CD8



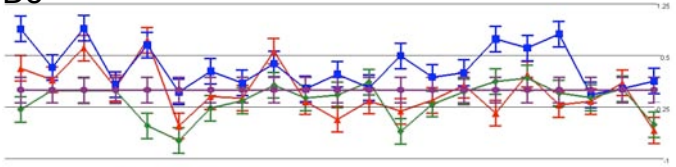
CD4



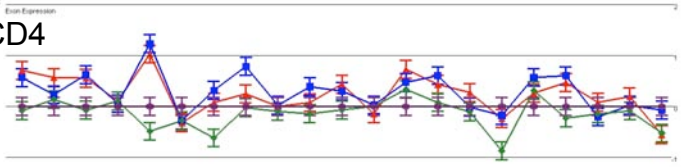
Evc2 = LOC665169



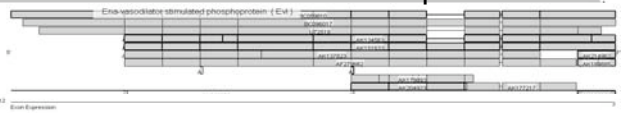
CD8



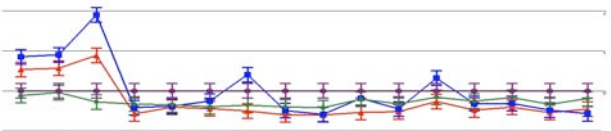
CD4



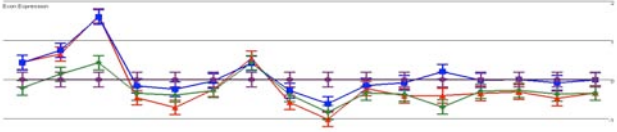
Evl exons and transcripts



CD8

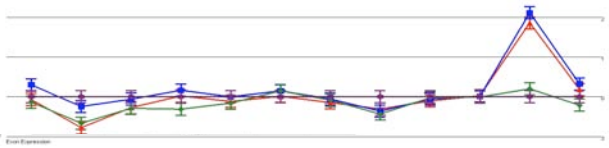


CD4

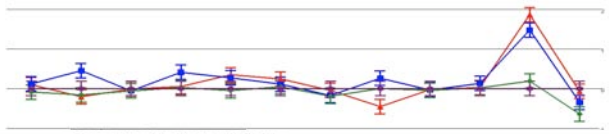


Fgb exons and transcripts

CD8

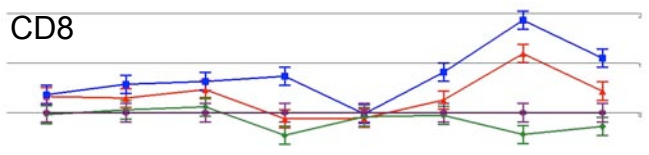


CD4

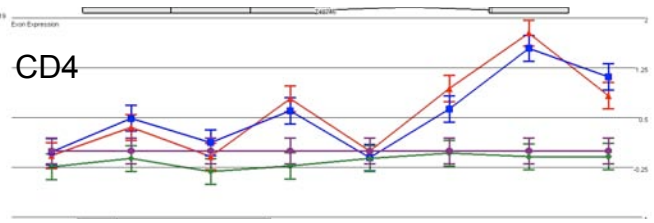


Fgf8 exons and transcripts

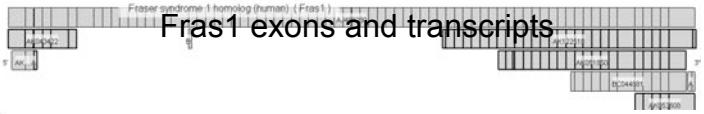
CD8



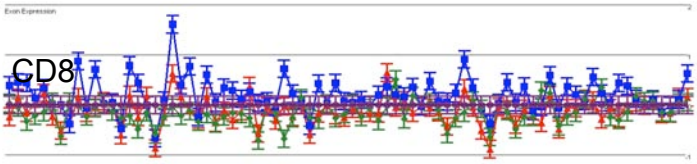
CD4



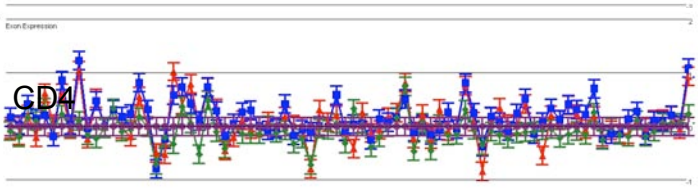
Fras1 exons and transcripts



CD8

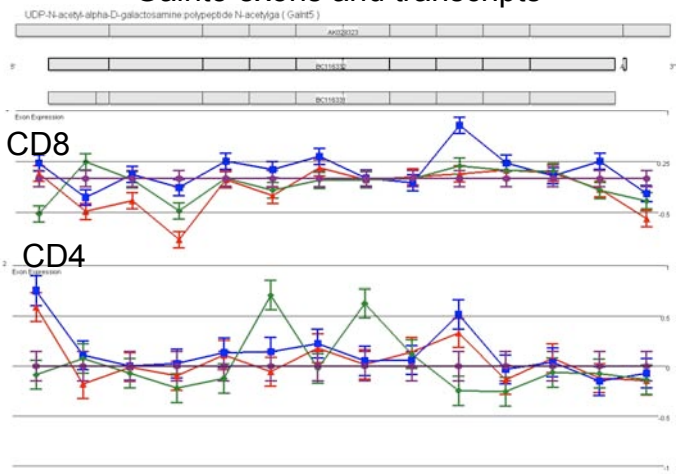


CD4

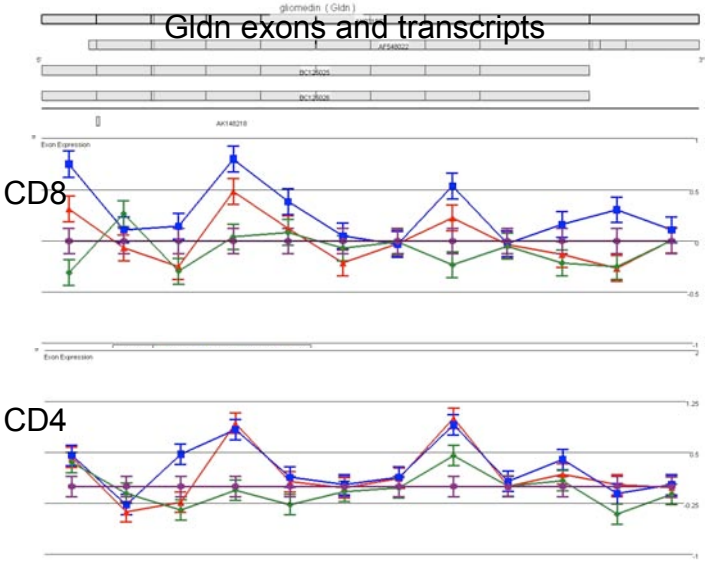


Group C Genes

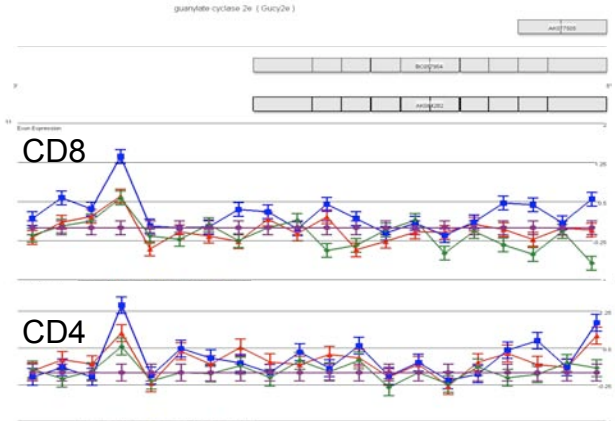
Galnt5 exons and transcripts



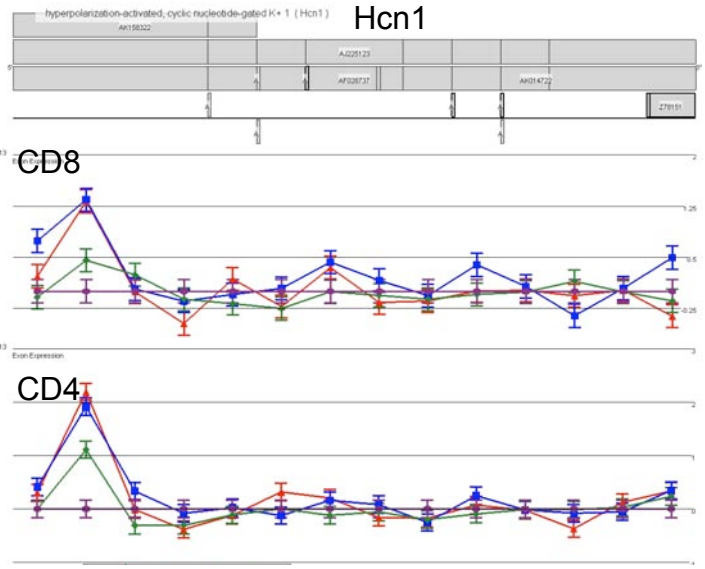
Gldn exons and transcripts



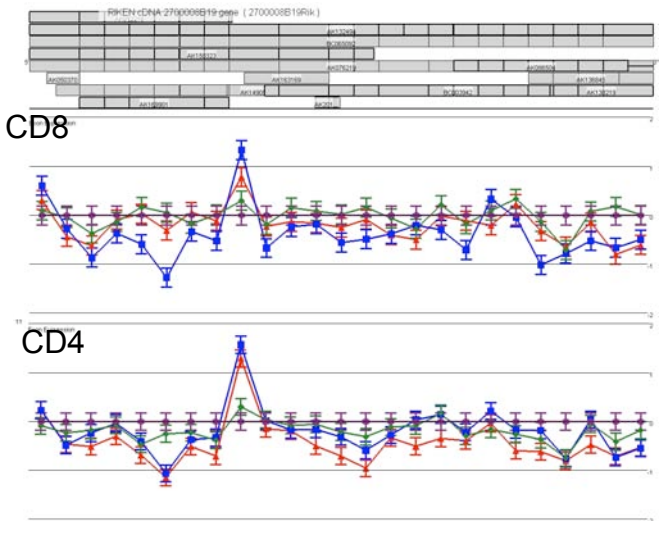
Gucy2e exons and transcripts



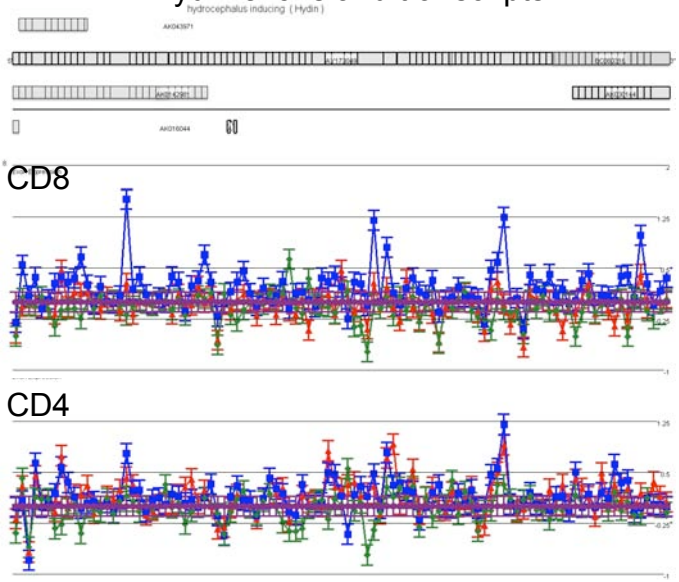
Hcn1



Heatr6 = 2700008B19

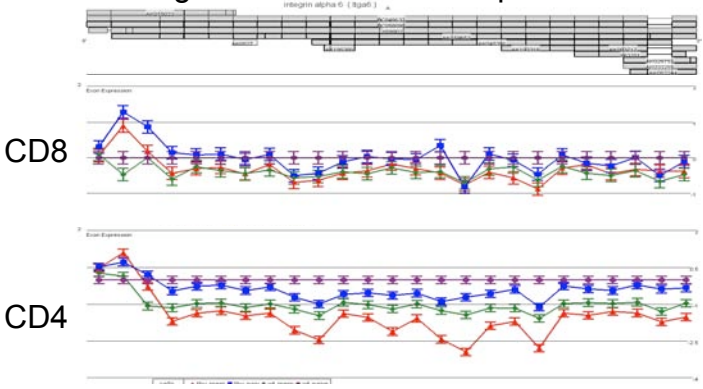


Hydin exons and transcripts

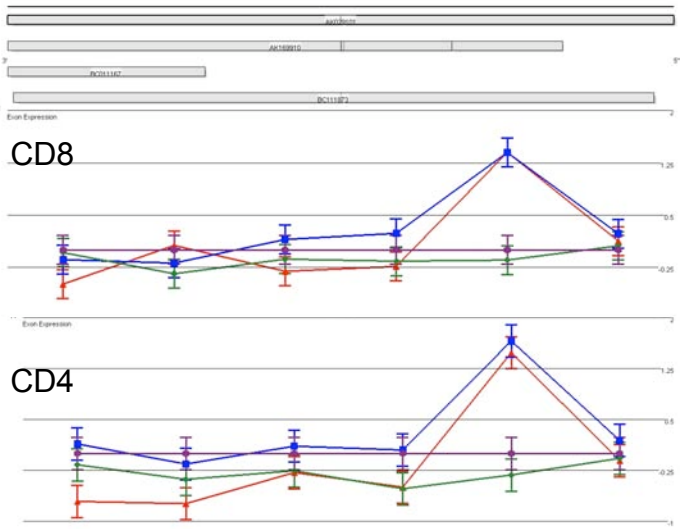


Group C Genes

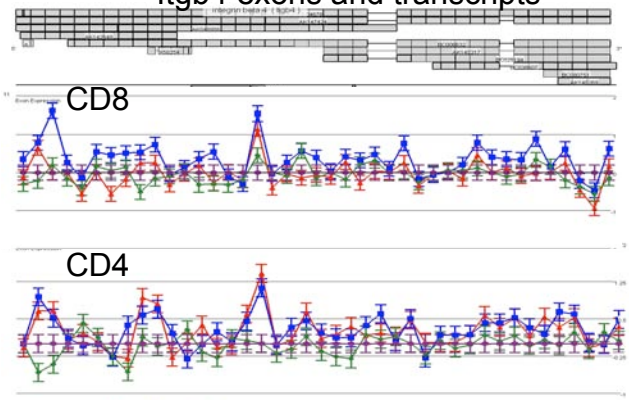
Itga6 exons and transcripts



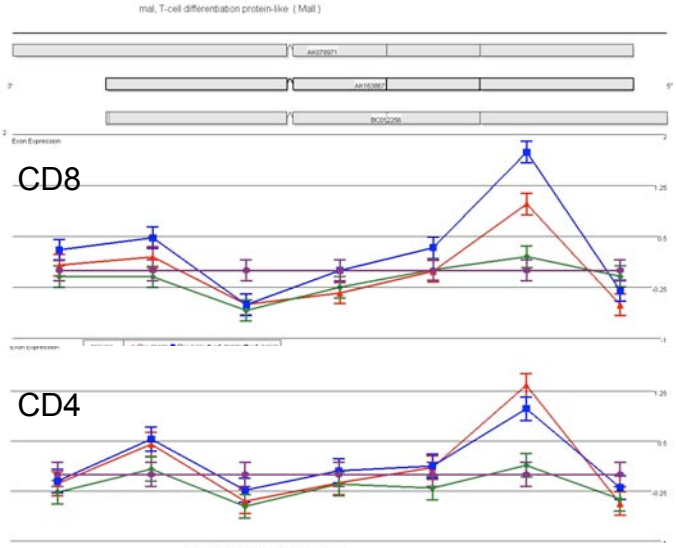
KLhl11 exons and transcripts



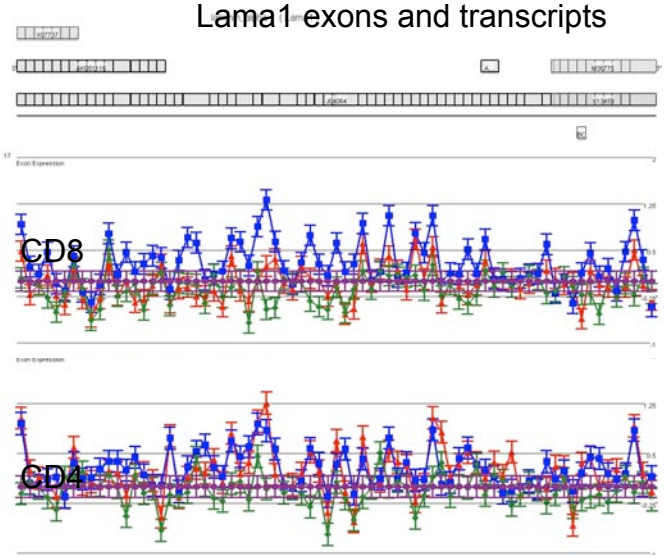
Itgb4 exons and transcripts



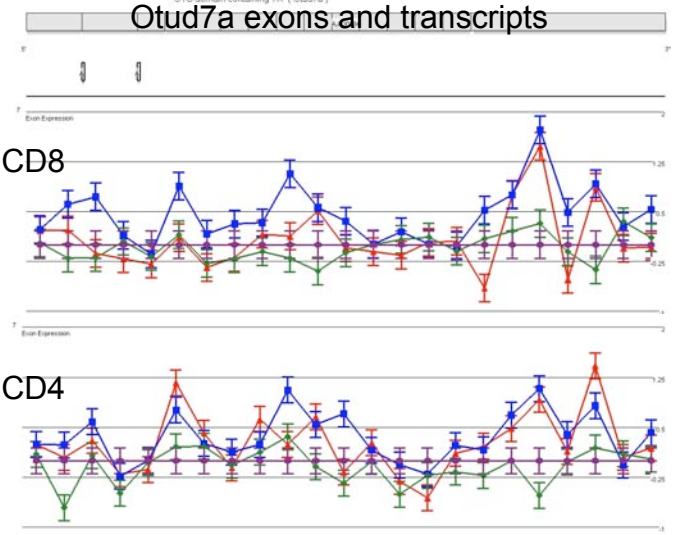
Mall exons and transcripts



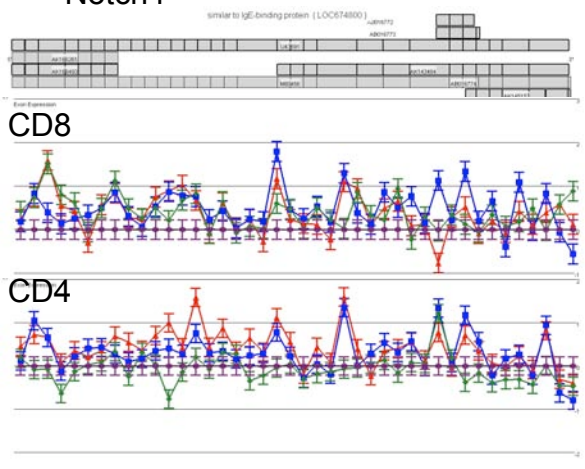
Lama1 exons and transcripts



Otud7a exons and transcripts

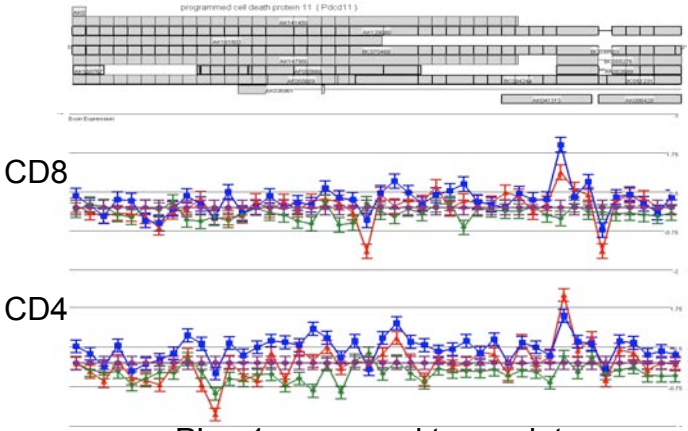


Notch4

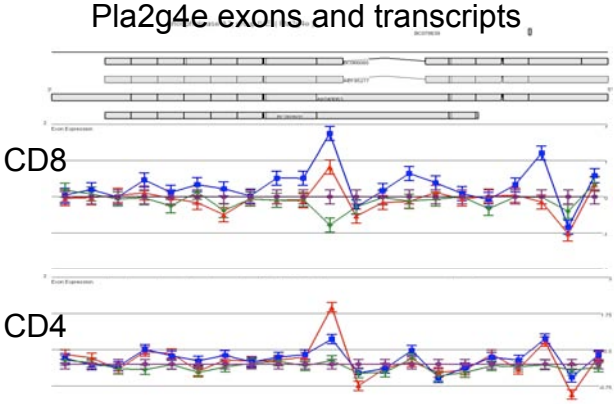


Group C Genes

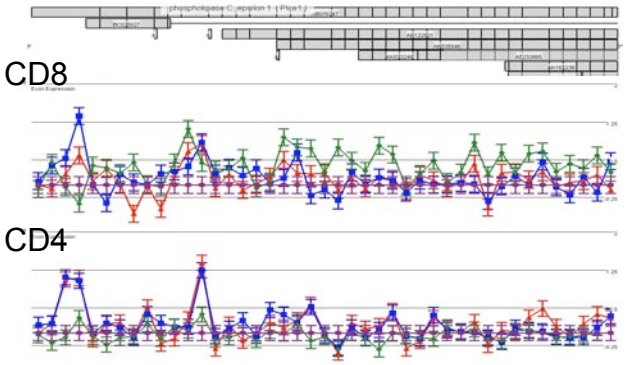
Pdcd11 exons and transcripts



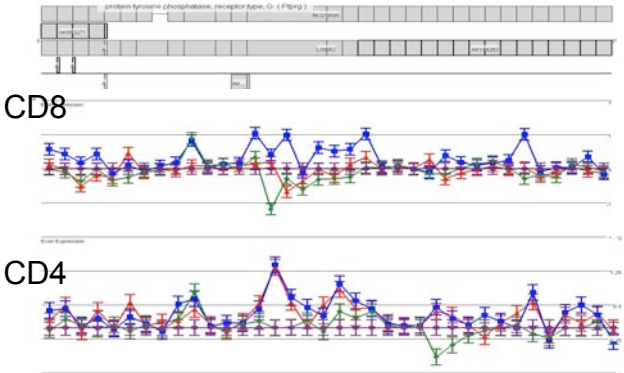
Pla2g4e exons and transcripts



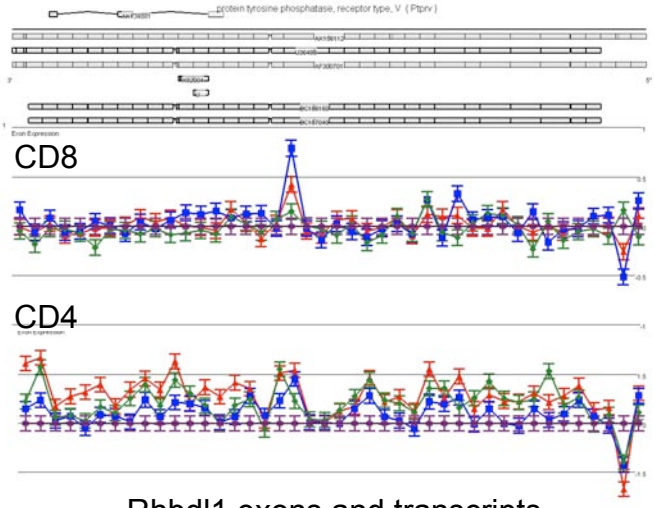
PLce1 exons and transcripts



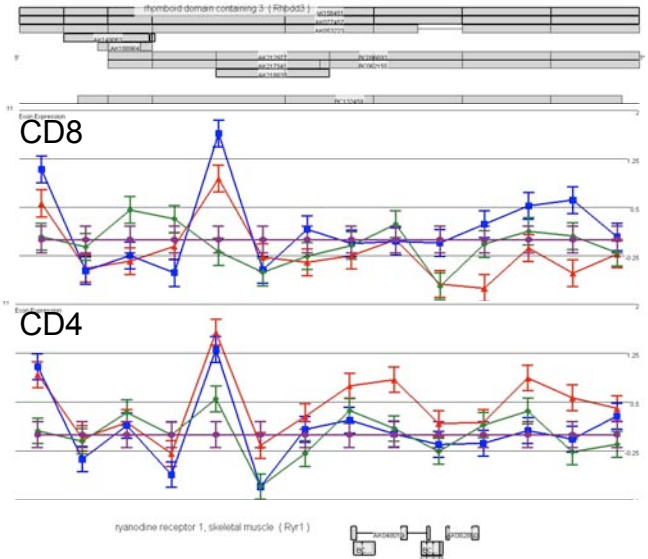
Ptprg exons and transcripts



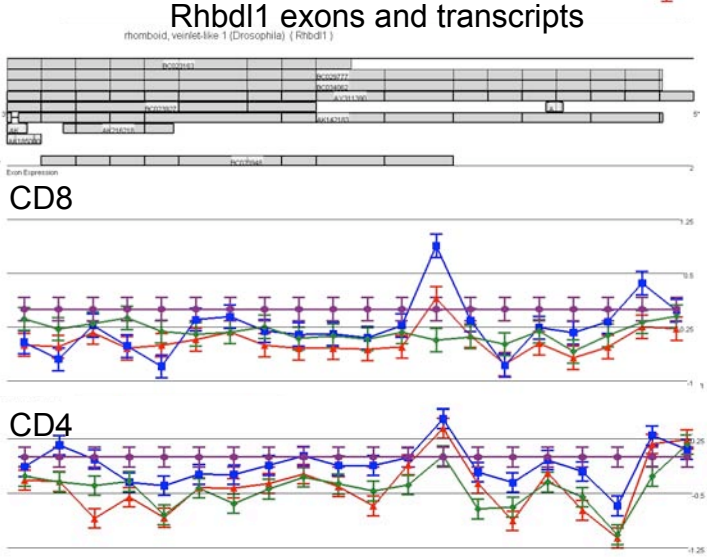
Ptprv exons and transcripts



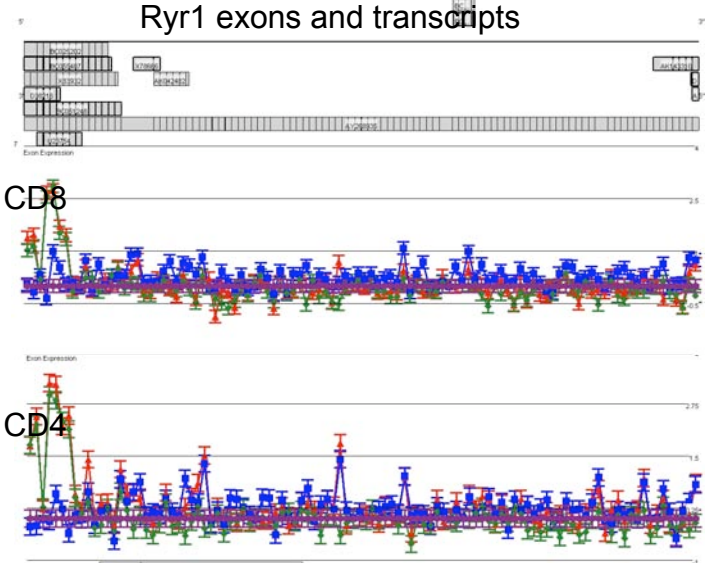
Rhbdd3 exons and transcripts



Rhbdl1 exons and transcripts

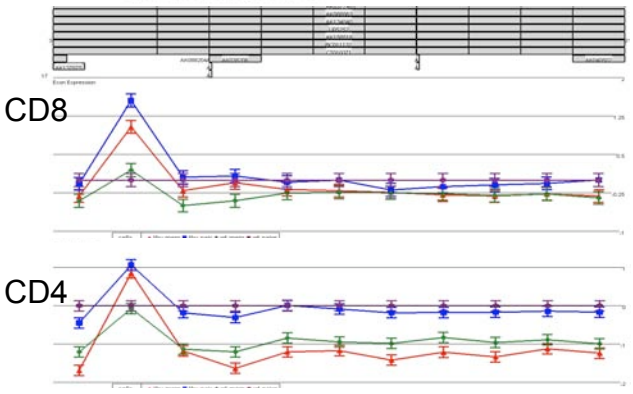


Ryr1 exons and transcripts

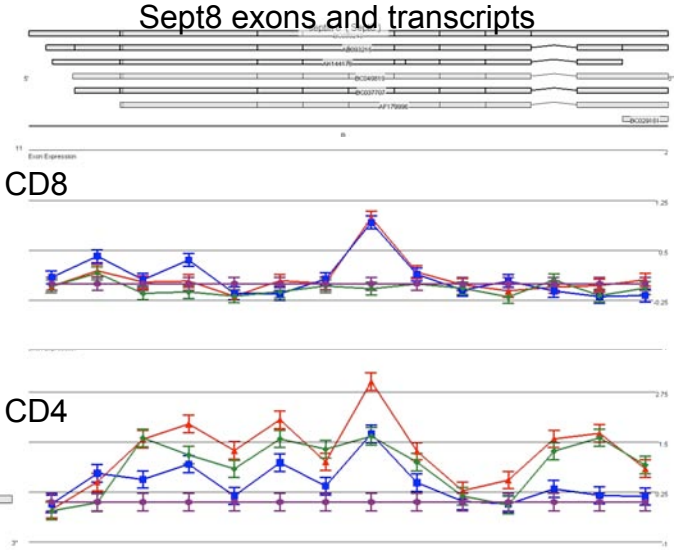


Group C Genes

Satb1 exons and transcripts

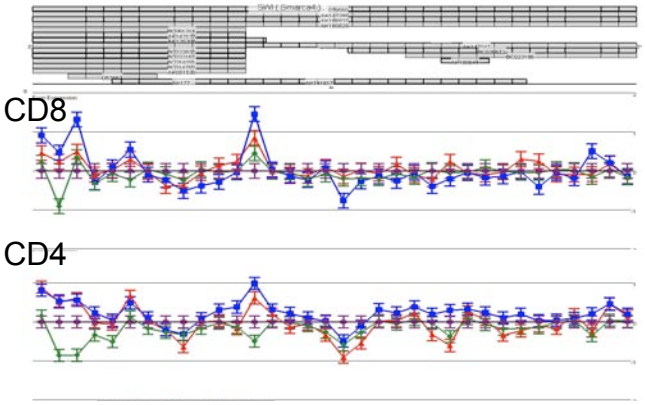


Sept8 exons and transcripts

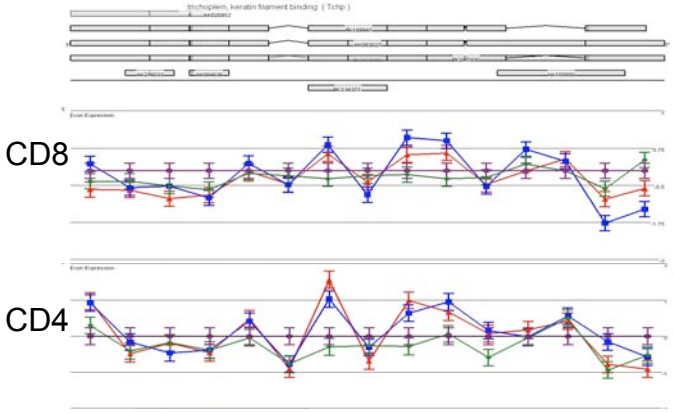


Slc6a13 exons and transcripts

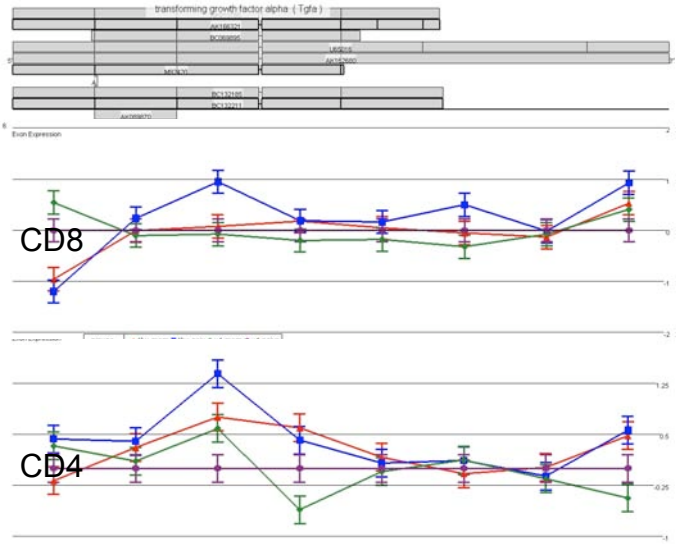
Smarca4 exons and transcripts



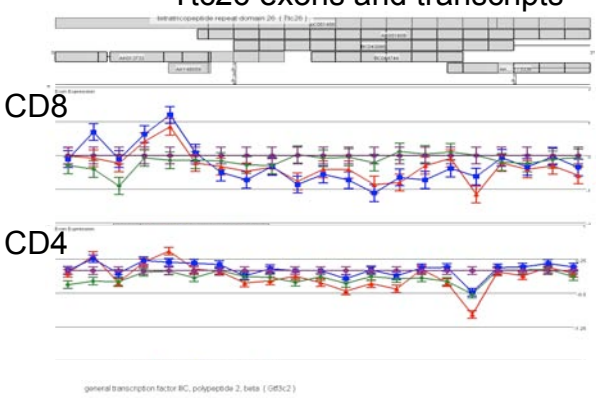
Tchp exons and transcripts



Tgfa exons and transcripts

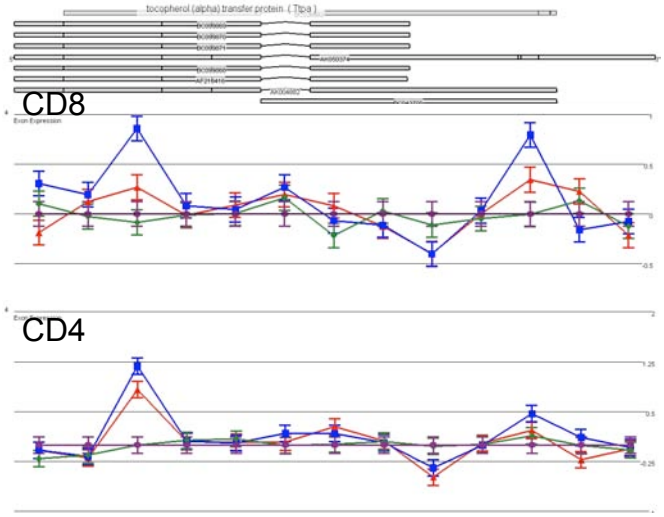


Ttc26 exons and transcripts

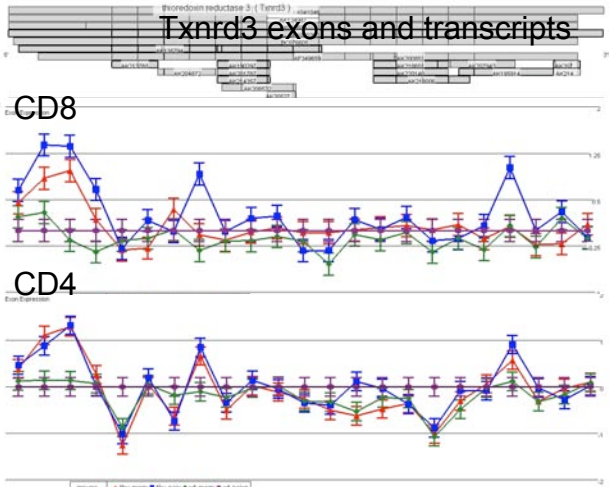


Group C Genes

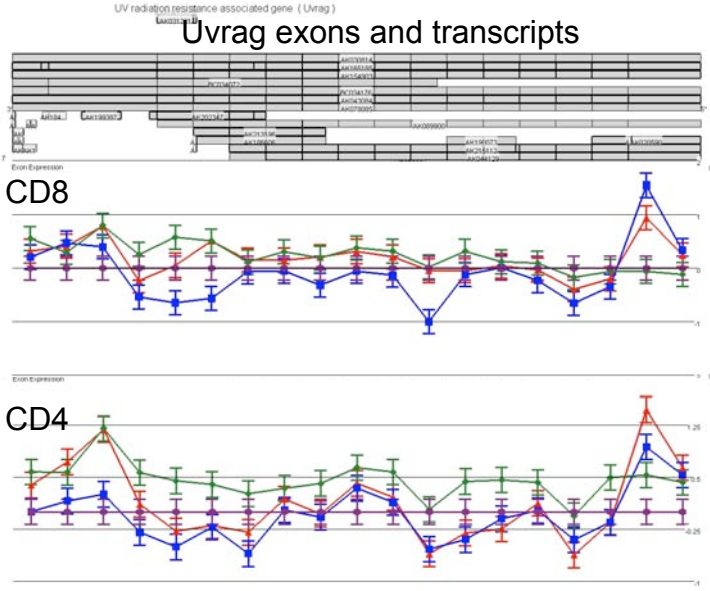
Ttpa exons and transcripts



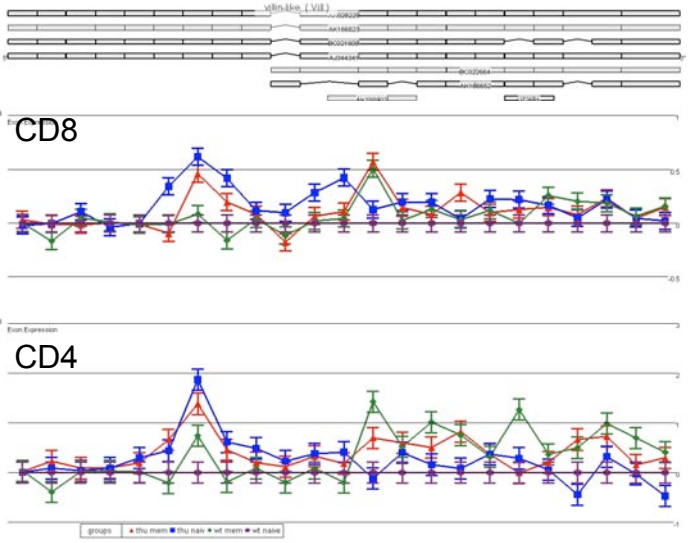
Txnrd3 exons and transcripts



Uvrag exons and transcripts

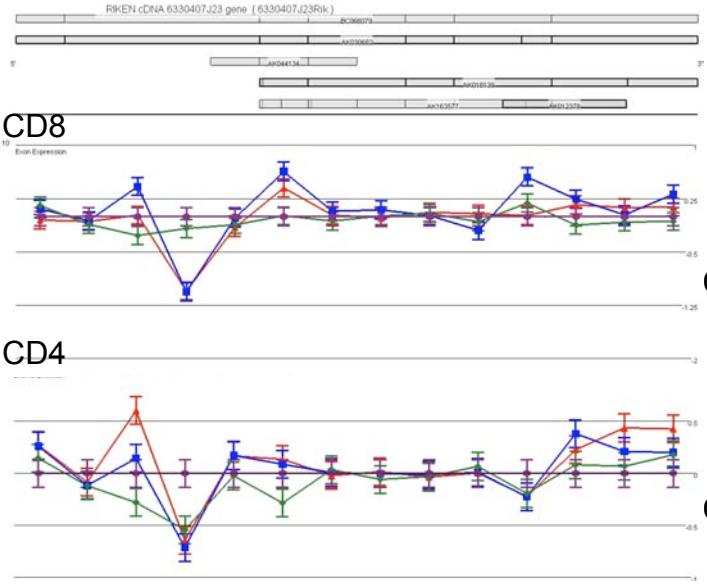


Vill exons and transcripts

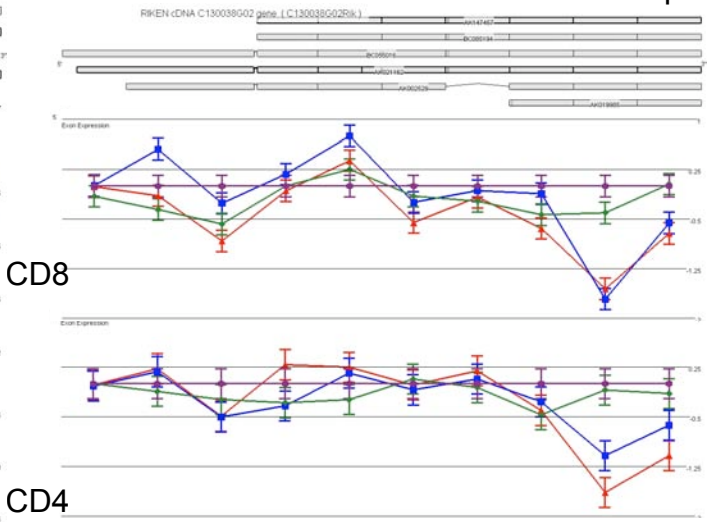


Group D Genes

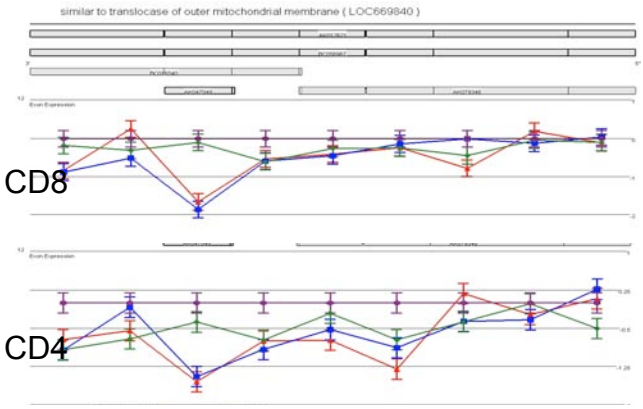
6330407J23Rik



C130038G02Rik exons and transcripts



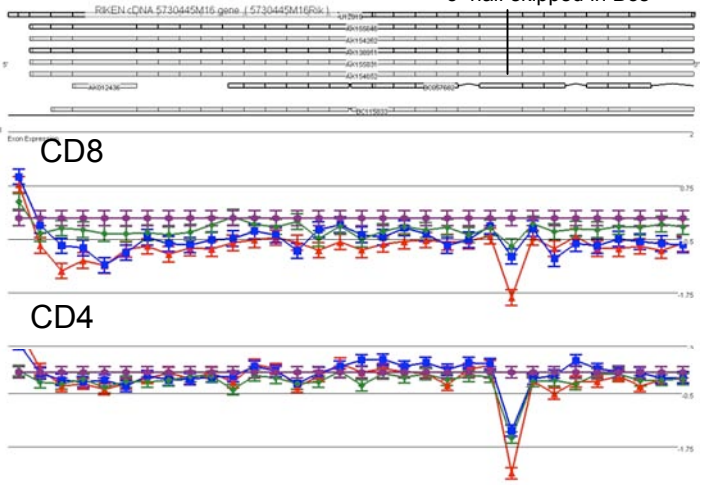
LOC669840 = 2700097O09Rik



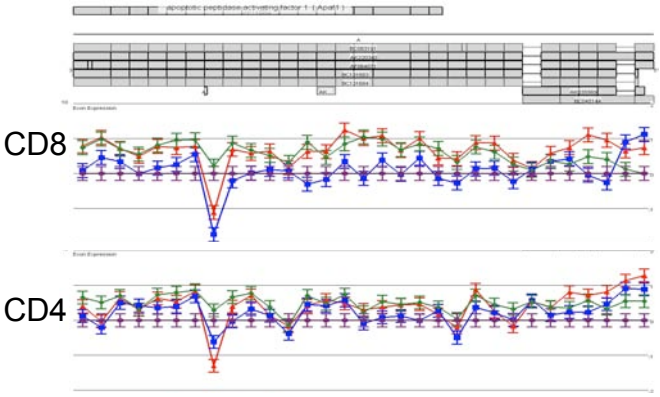
Group D Genes

5730445M16 = Adcy7

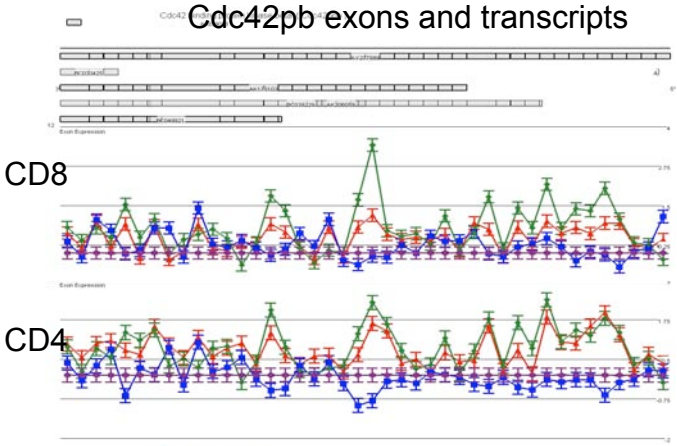
This exon skipped in Homo & 5' half skipped in Bos



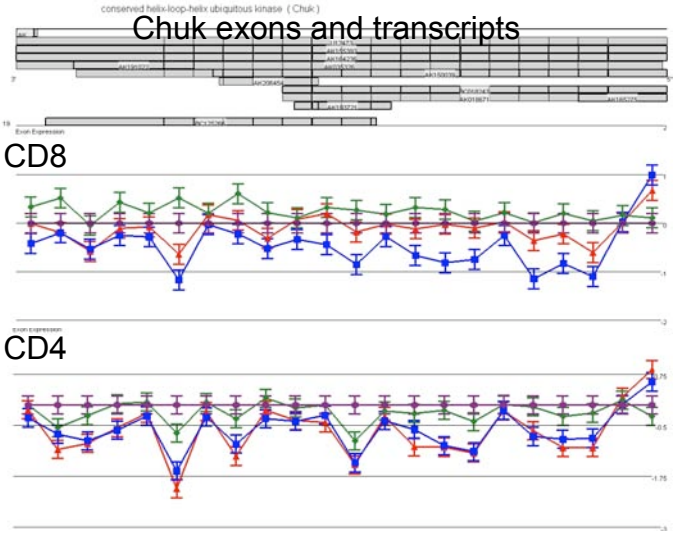
Apaf1 exons and transcripts



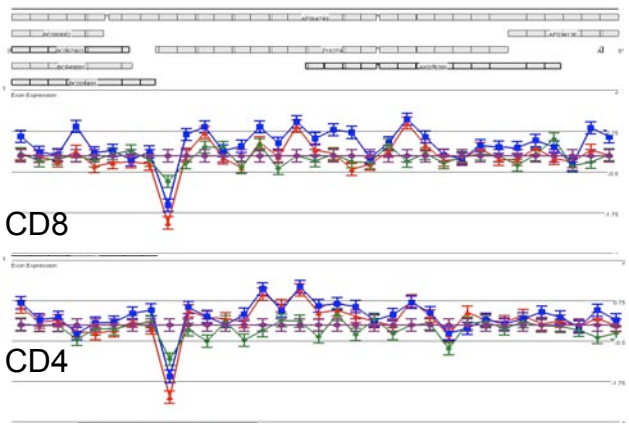
Cdc42pb exons and transcripts



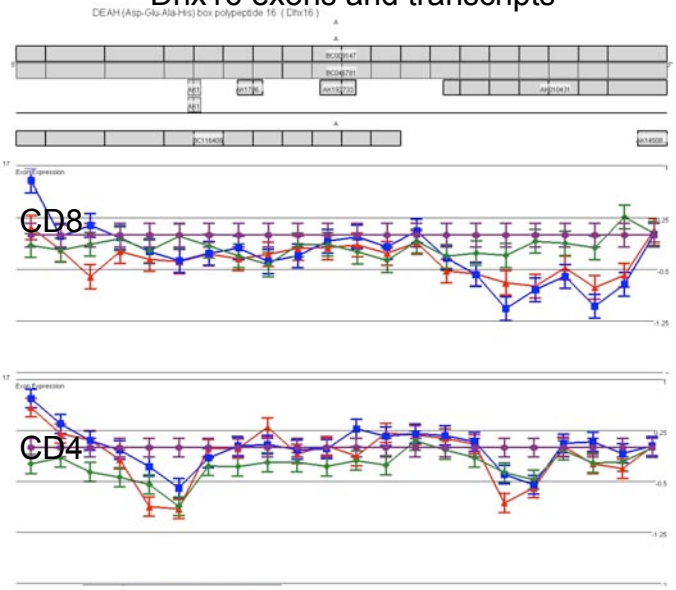
Chuk exons and transcripts



Col6a3 = LOC674521

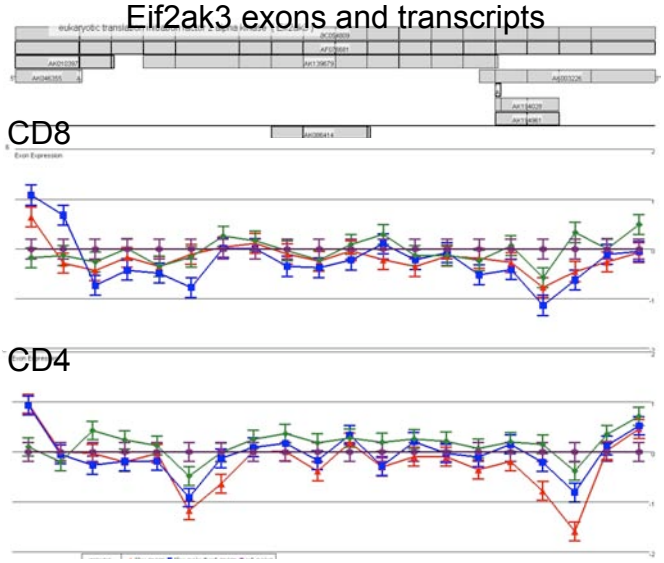


Dhx16 exons and transcripts

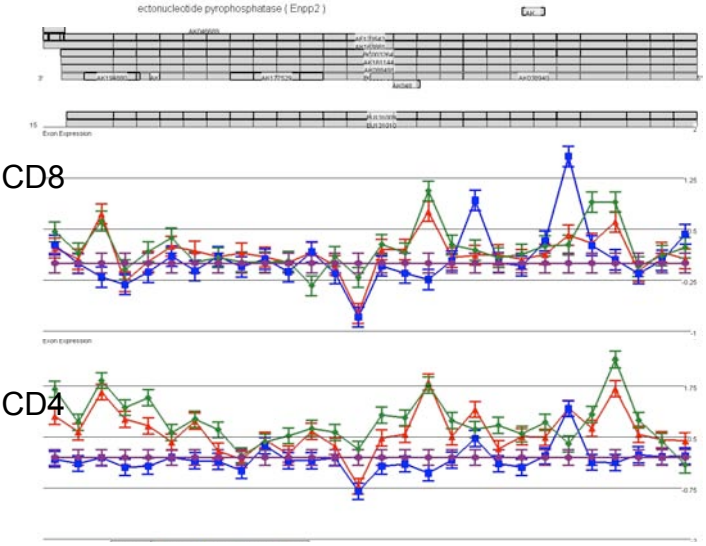


Group D Genes

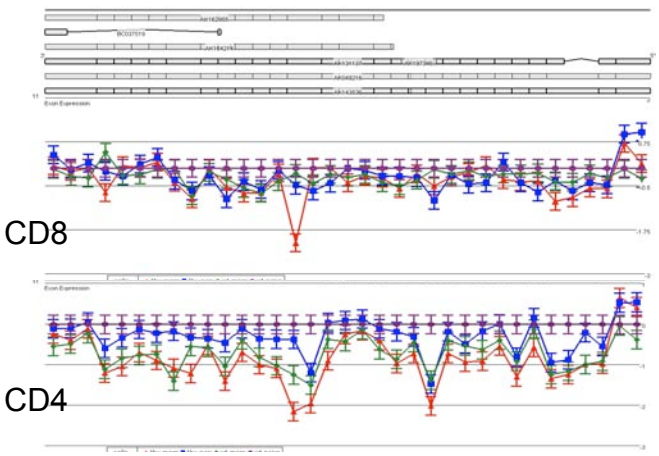
Eif2ak3 exons and transcripts



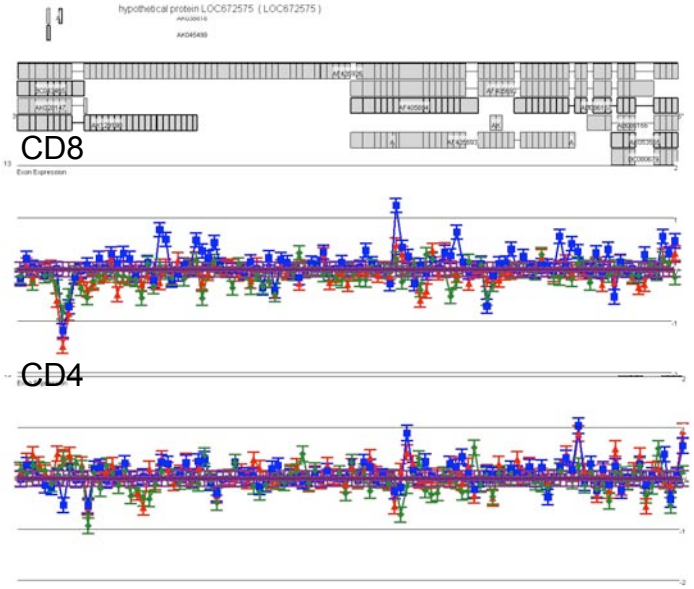
Enpp2 exons and transcripts



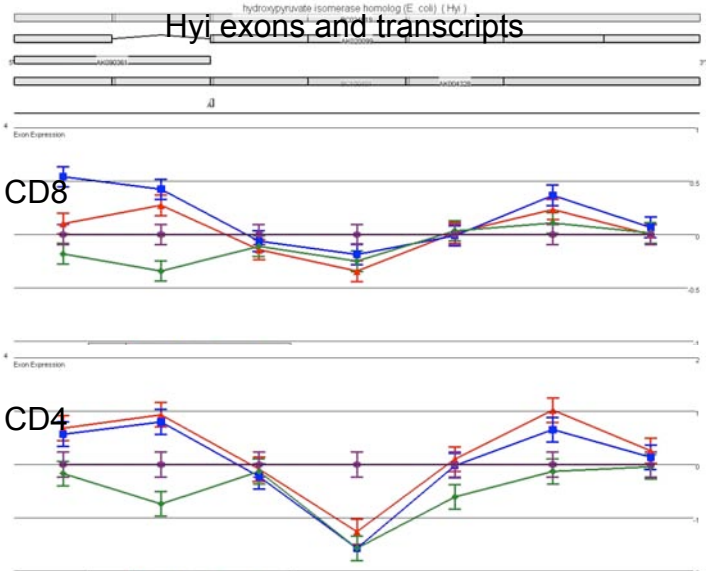
Gemin5 exons and transcripts



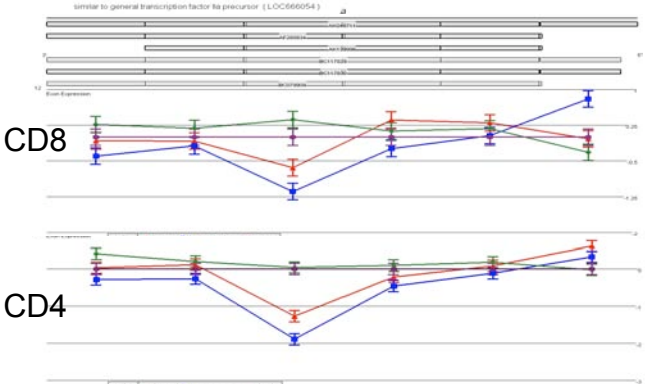
LOC672575 =Gpr98



Hyi exons and transcripts

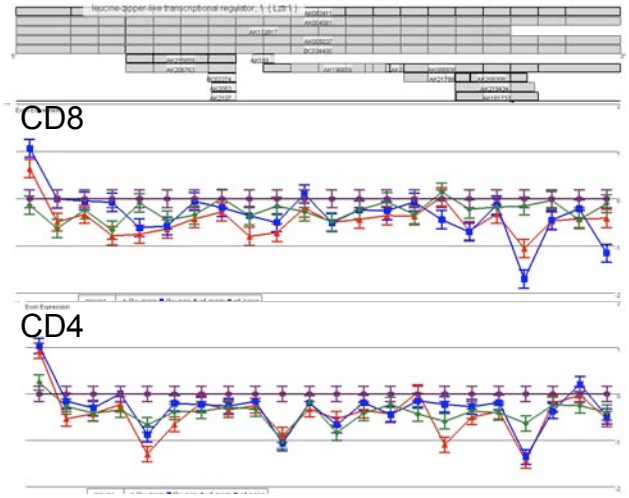


LOC666054 = Gtf2a1

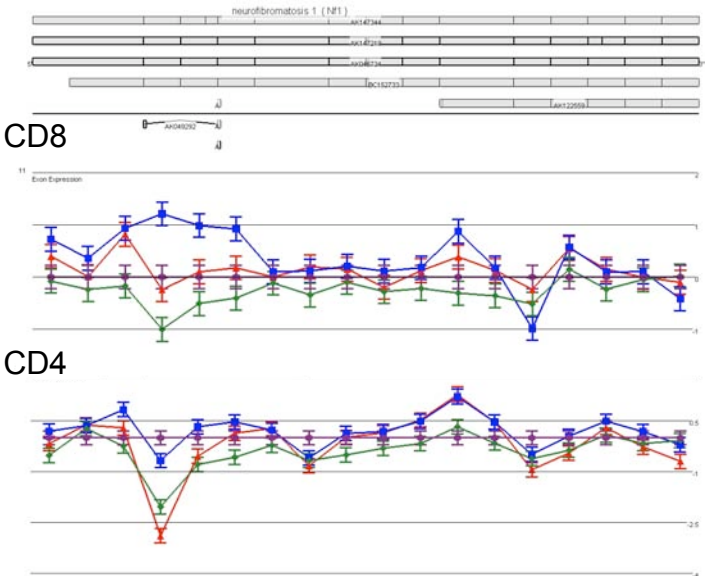


Group D Genes

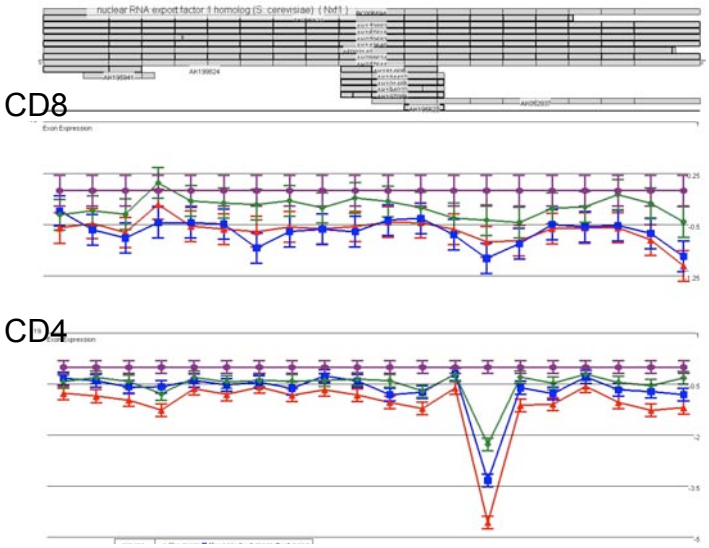
Lztr1 exons and transcripts



Nf1 exons and transcripts

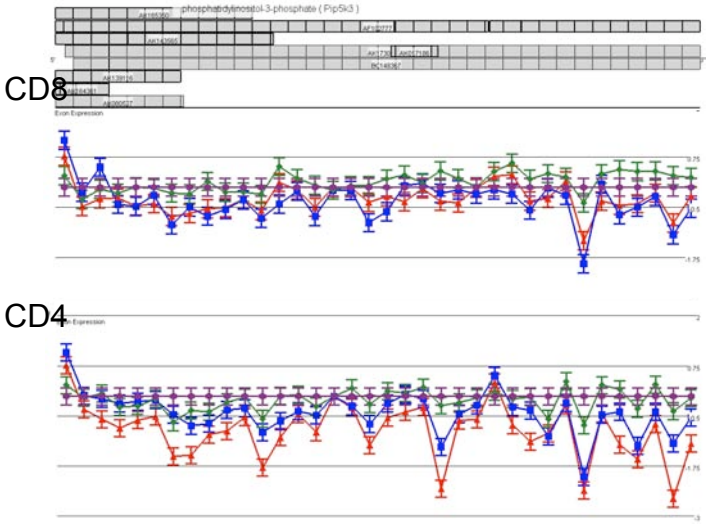


Nxf1 exons and transcripts

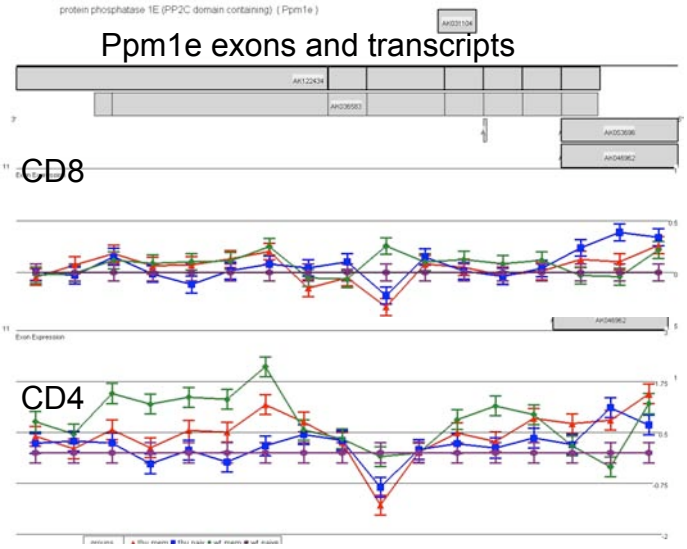


Group D Genes

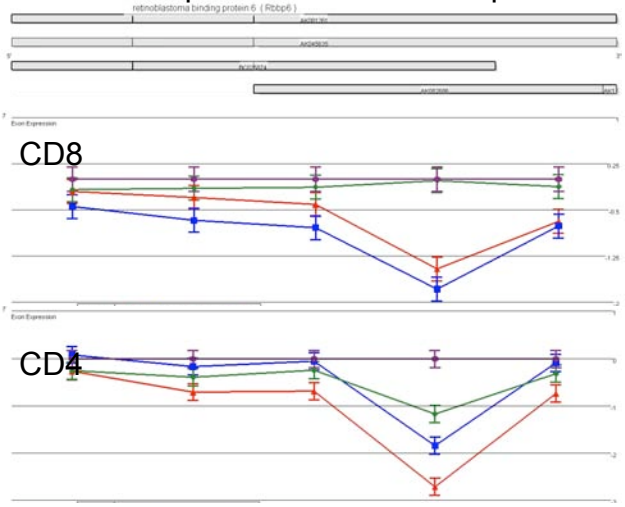
Pip5k3 exons and transcripts



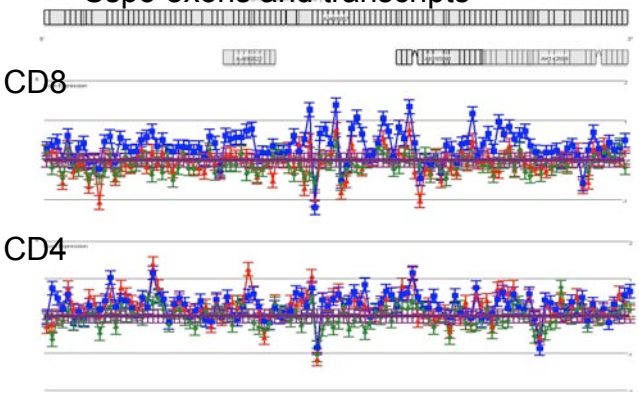
Ppm1e exons and transcripts



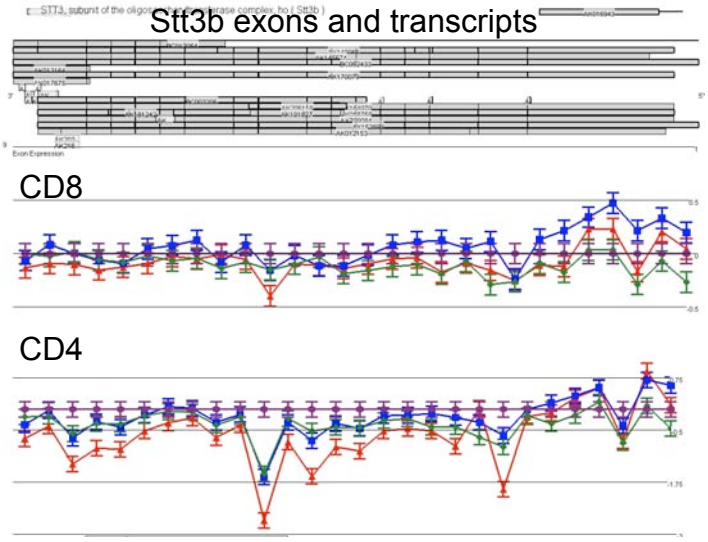
Rbbp6 exons and transcripts



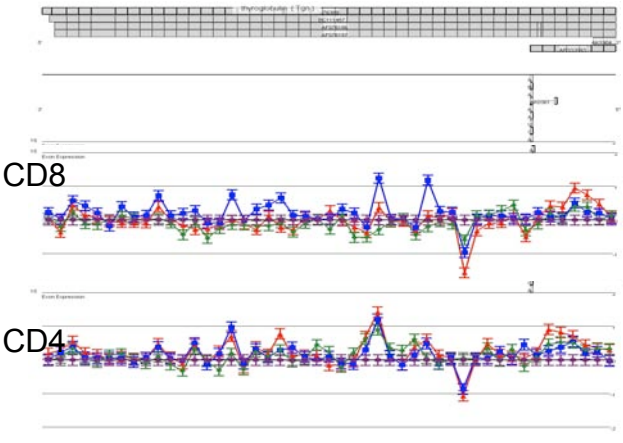
Sspo exons and transcripts



Stt3b exons and transcripts

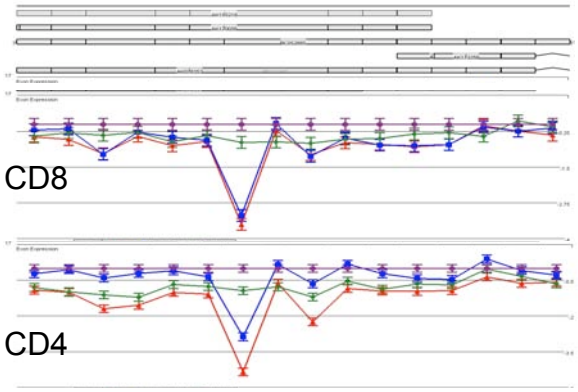


Tgn exons and transcripts

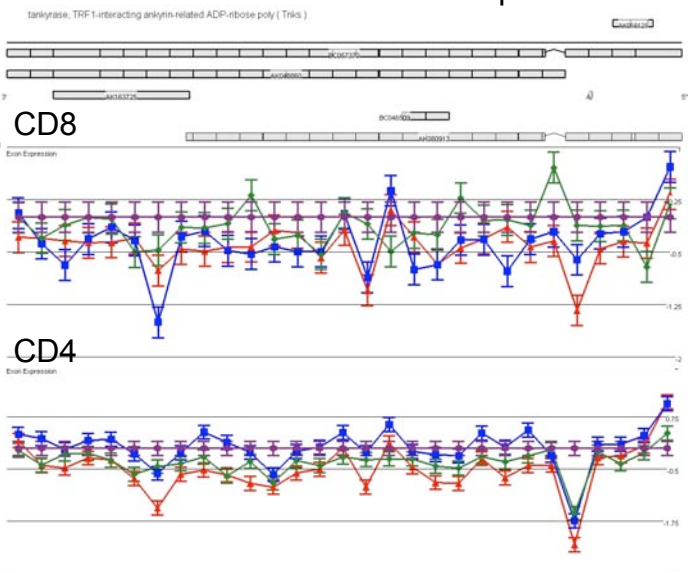


Group D Genes

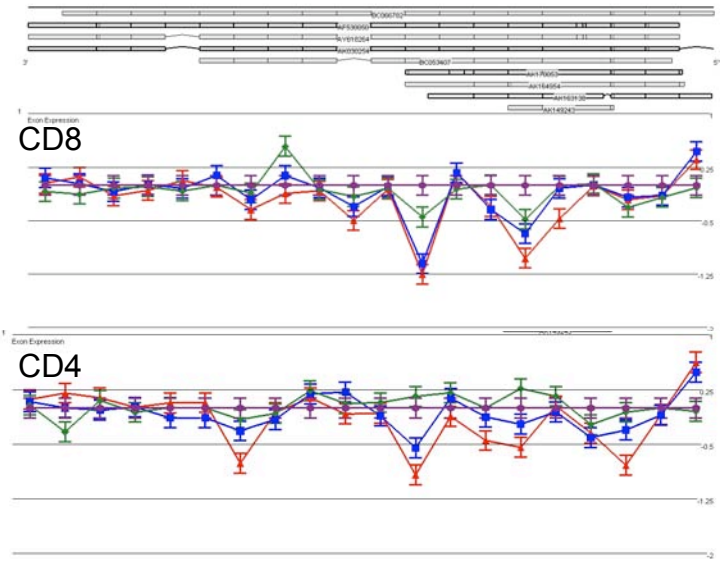
Thada exons and transcripts



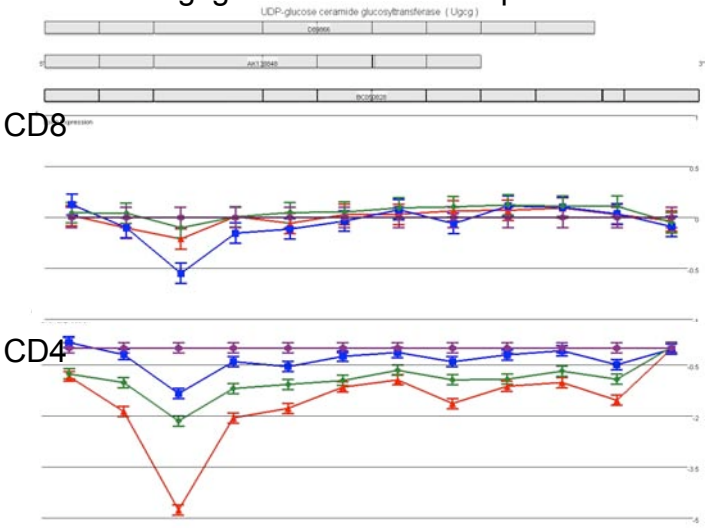
Tnks exons and transcripts



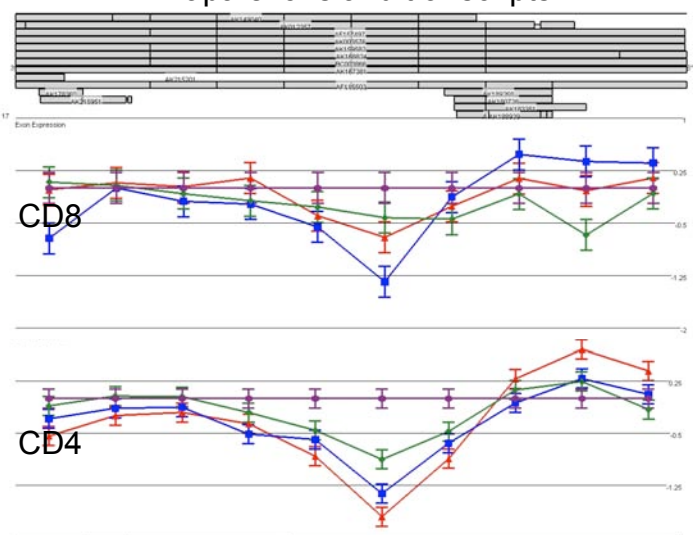
Tsga10 exons and transcripts



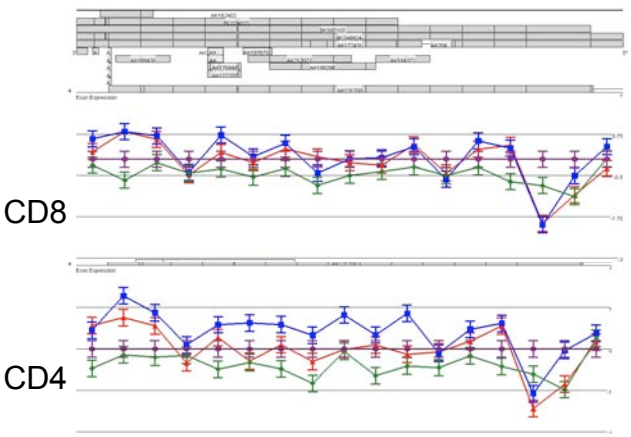
Ugcg exons and transcripts



Vapa exons and transcripts

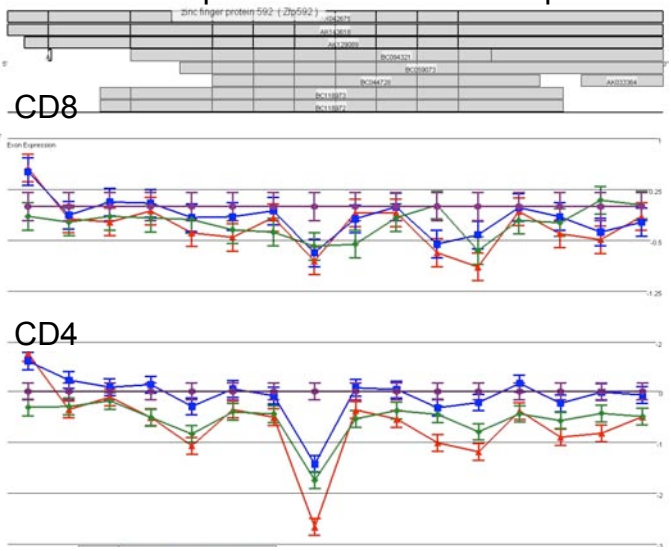


Wdtc1 exons and transcripts

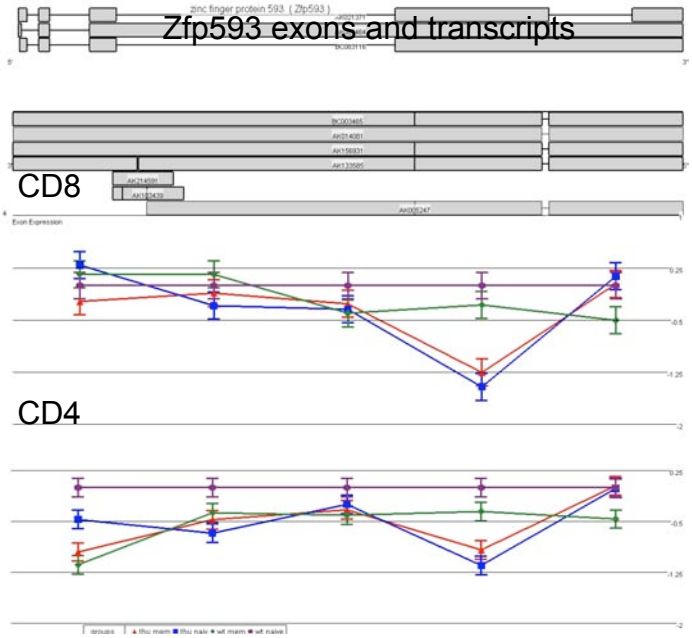


Group D Genes

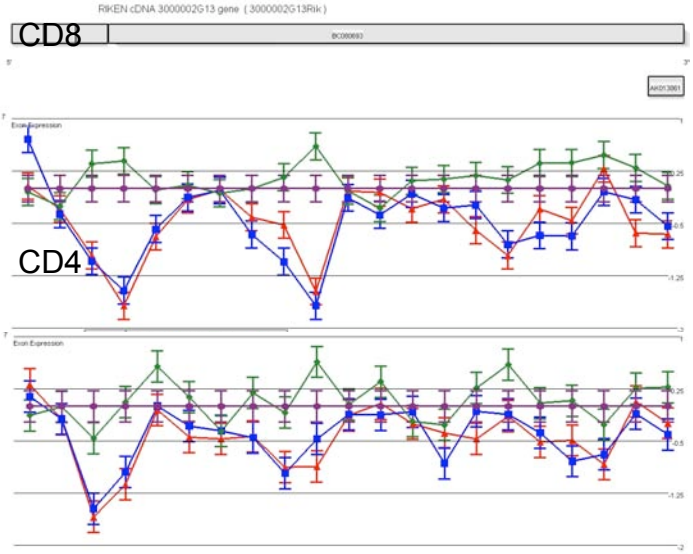
Zfp592 exons and transcripts



Zfp593 exons and transcripts

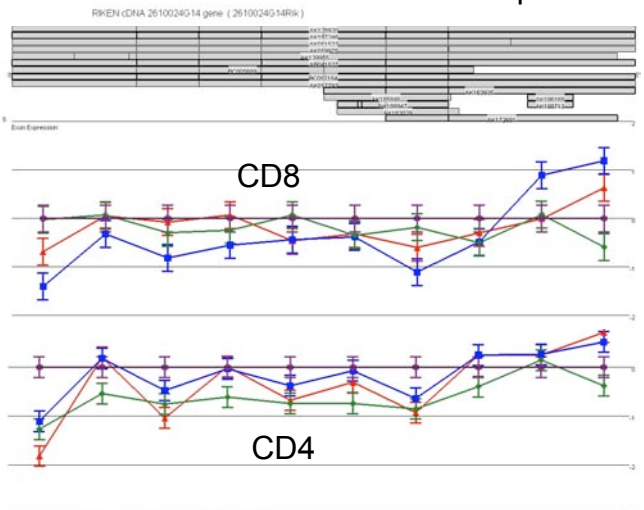


3000002G13Rik = Zfp619

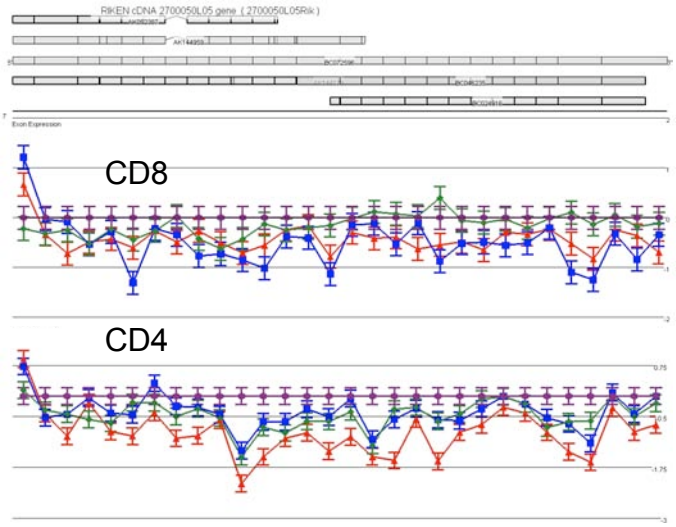


Group E Gene Views

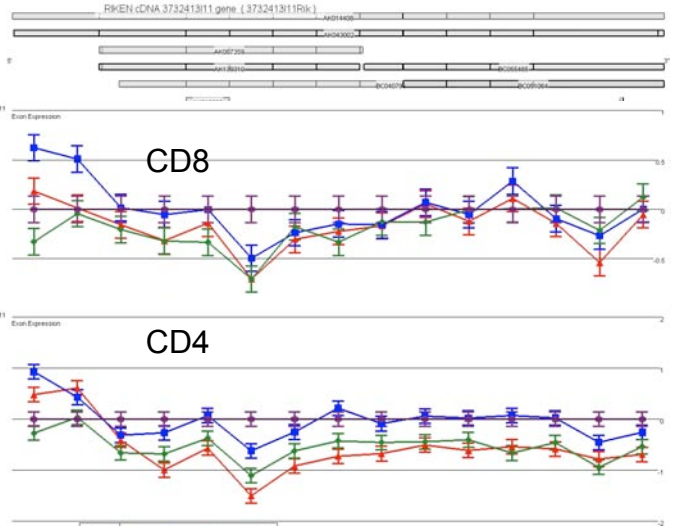
2610024G14Rik exons and transcripts



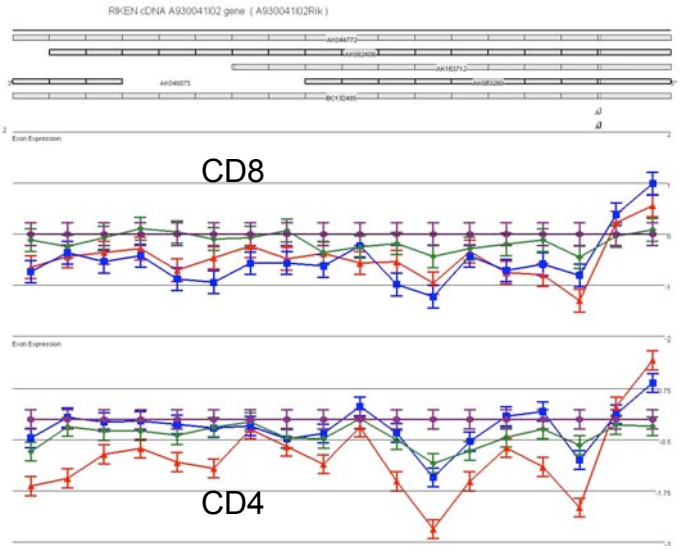
2700050L05Rik exons and transcripts



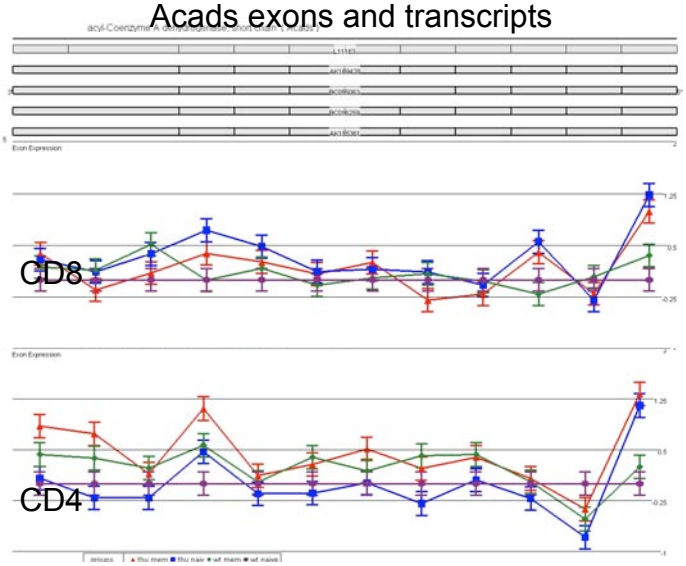
3732413I11Rik exons and transcripts



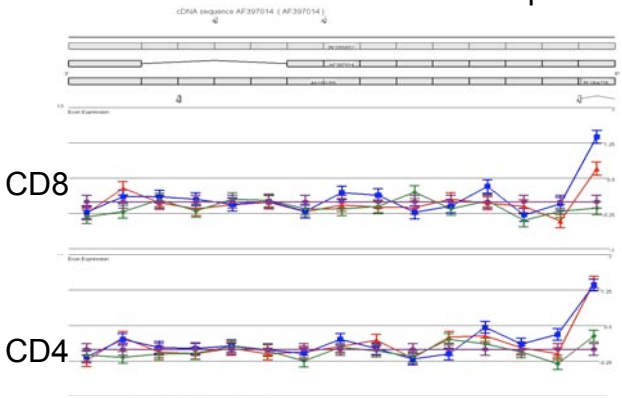
A930041I02Rik exons and transcripts



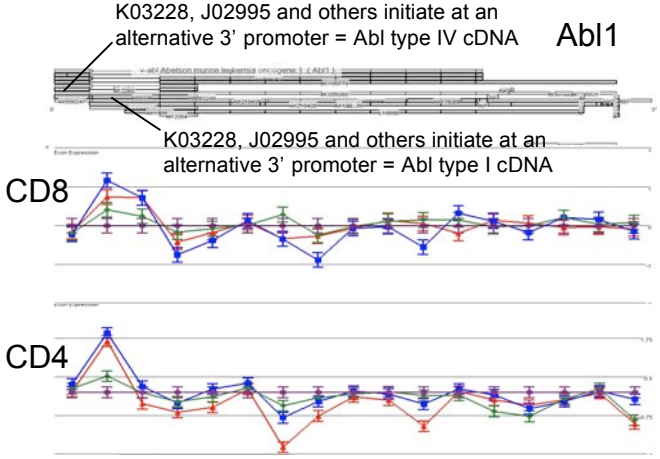
Acads exons and transcripts



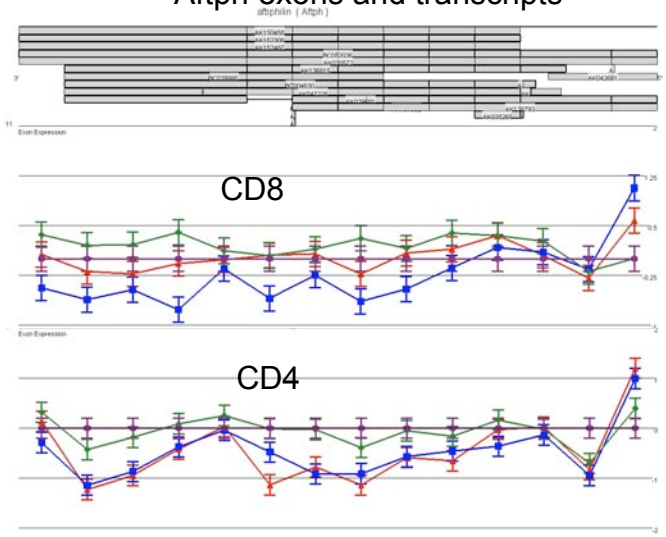
Af397014 exons and transcripts



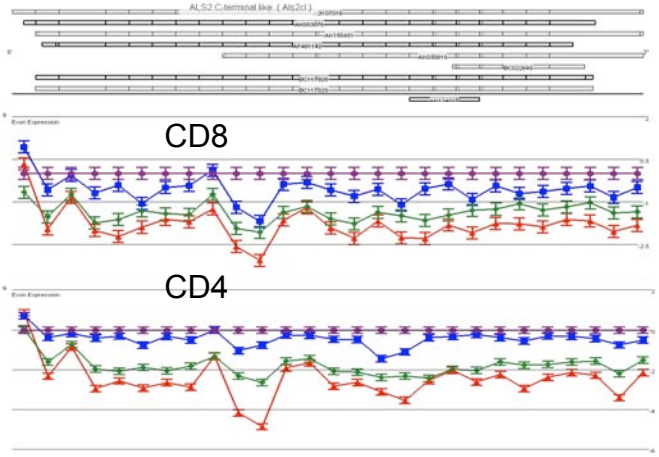
Abl1



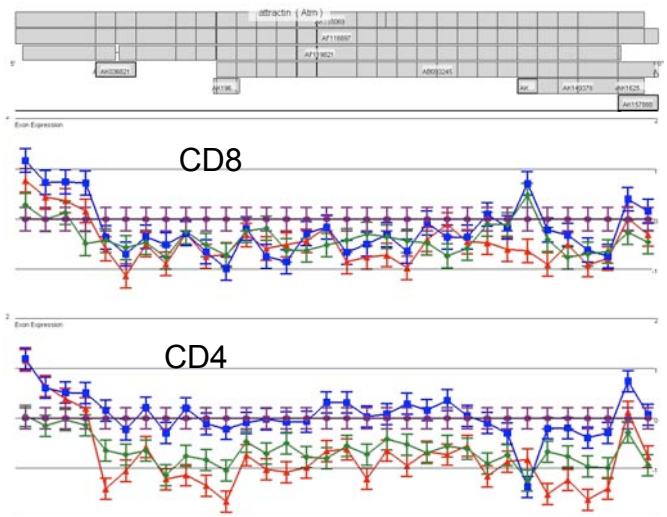
Aftph exons and transcripts



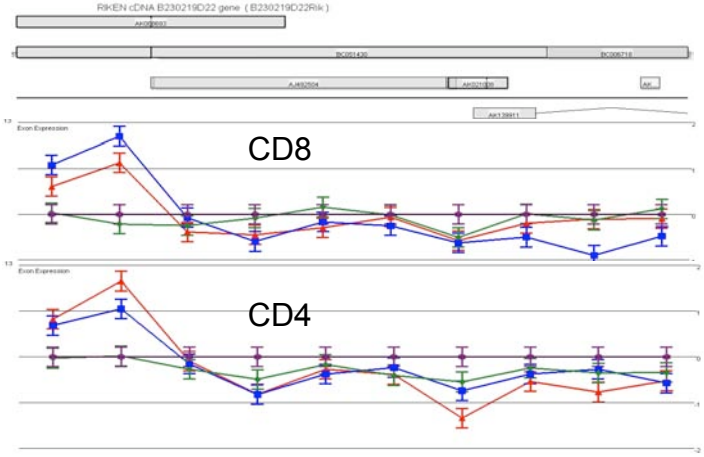
Als2cl exons and transcripts



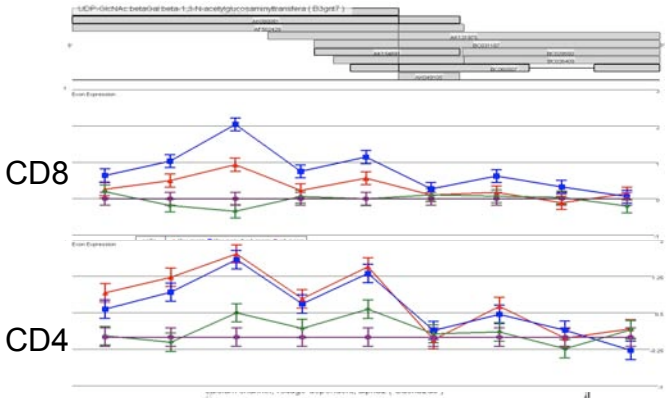
Atrn exons and transcripts



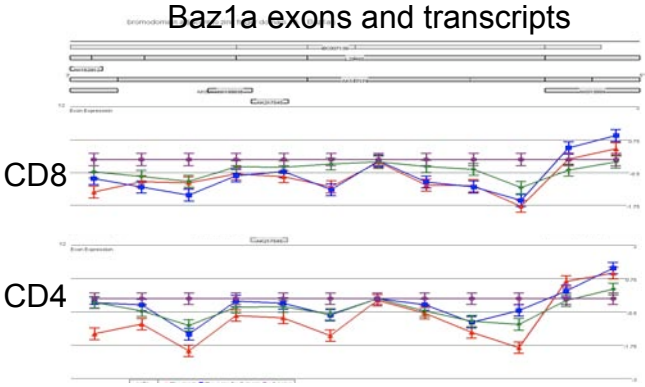
B230219D22Rik exons and transcripts



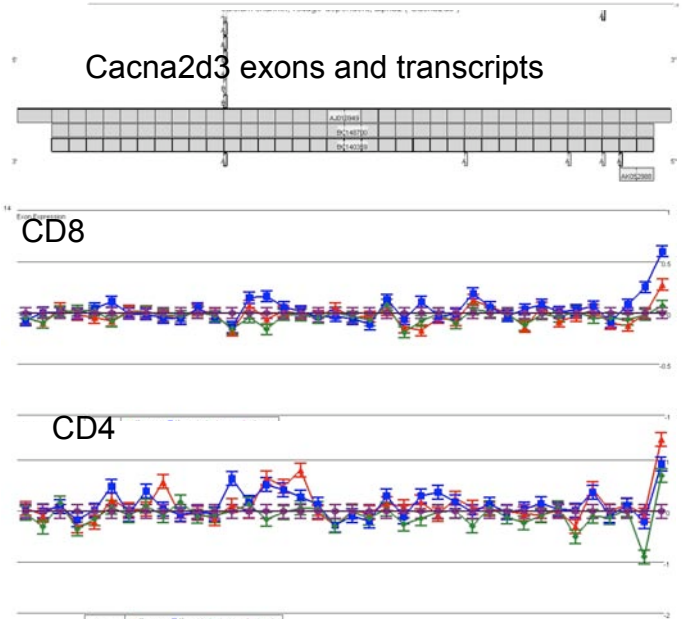
B3gnt7 exons and transcripts



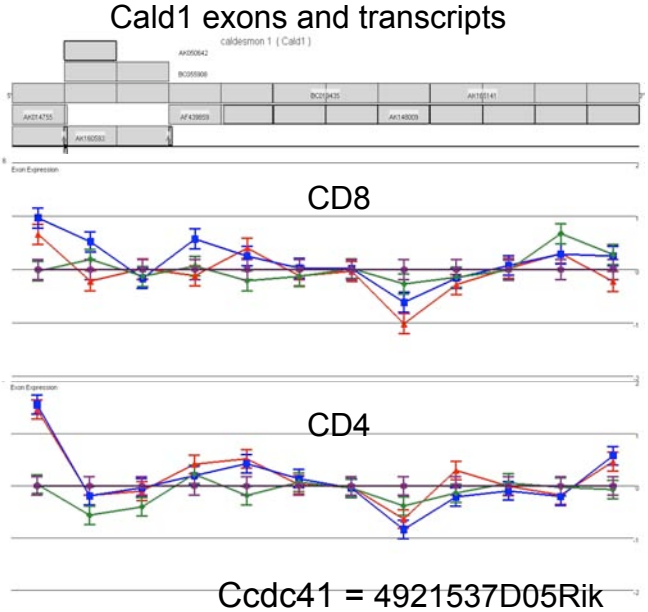
Baz1a exons and transcripts



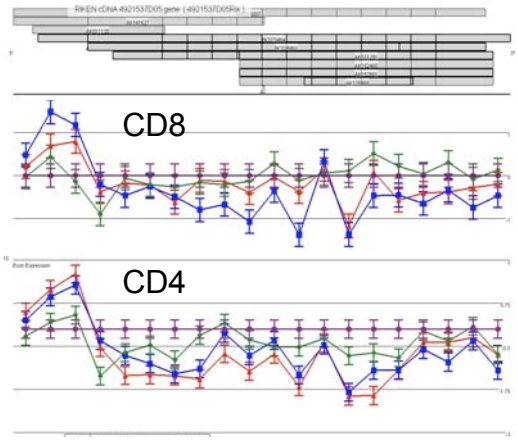
Cacna2d3 exons and transcripts



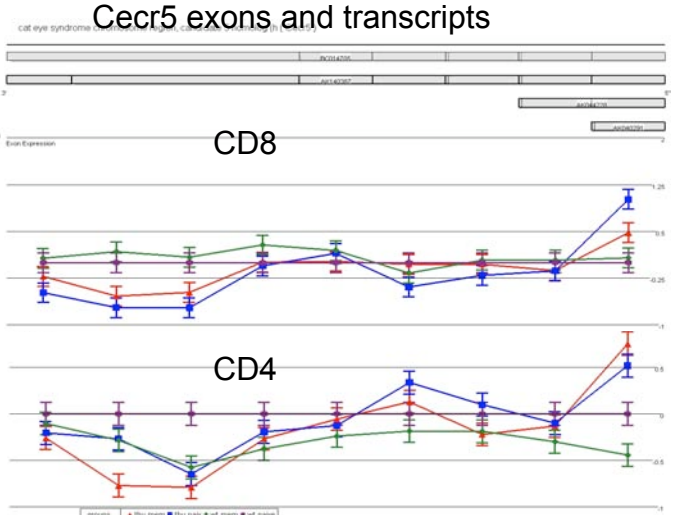
Cald1 exons and transcripts



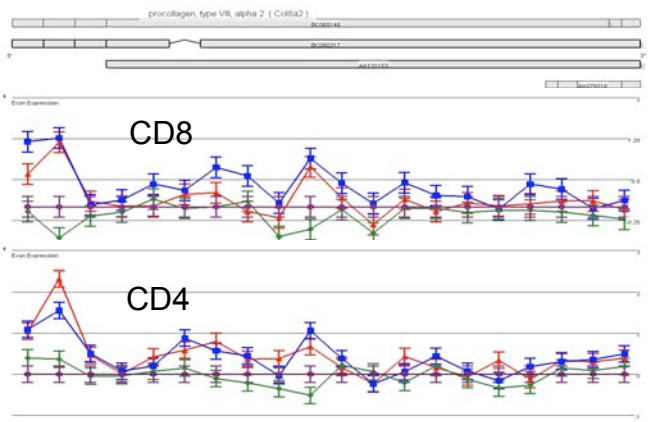
Ccdc41 = 4921537D05Rik



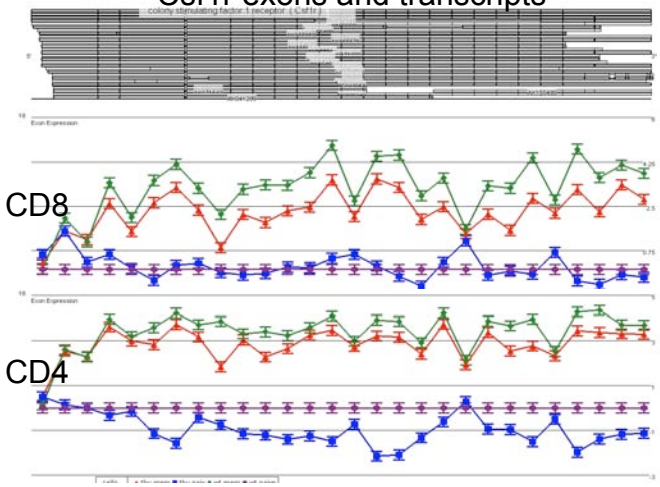
Cecr5 exons and transcripts



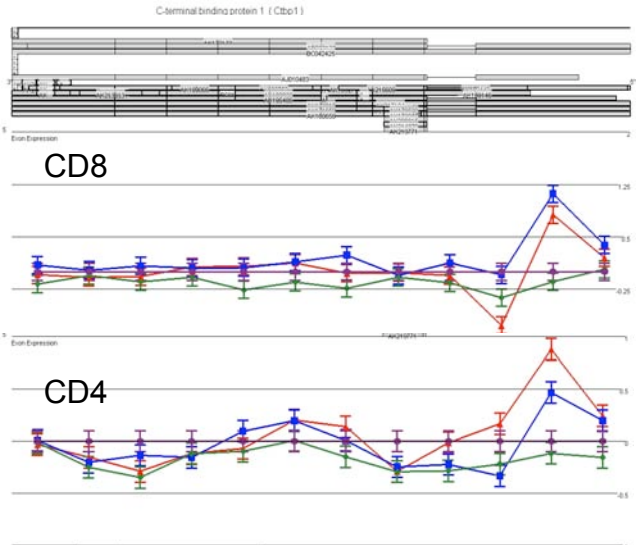
Col8a2 exons and transcripts



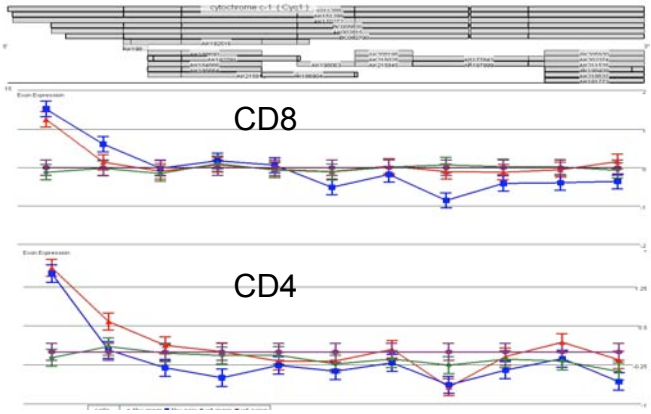
Csf1r exons and transcripts



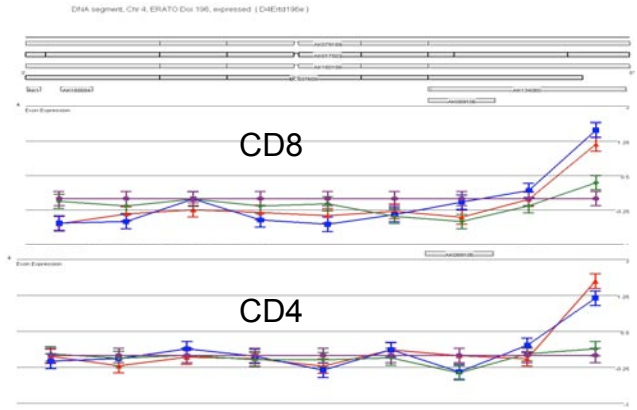
Ctbp1 exons and transcripts



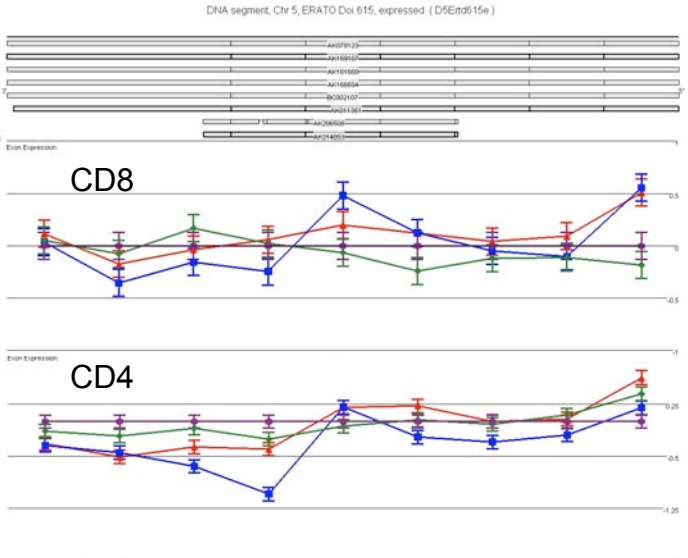
Cyc1 exons and transcripts



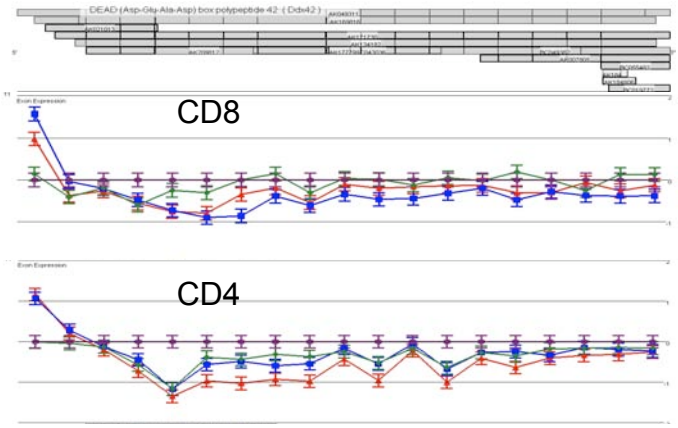
D4Ertd196e exons and transcripts



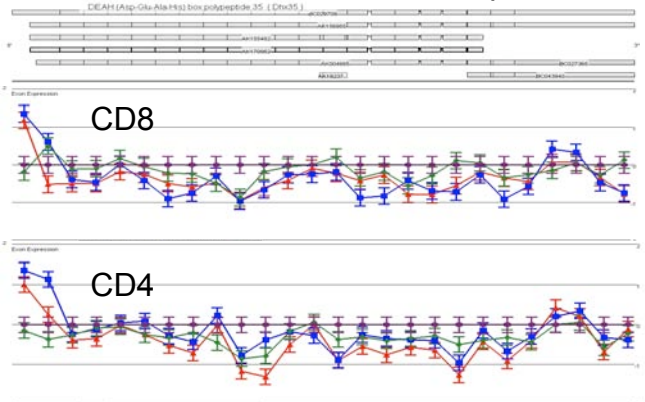
D5Ertd615e exons and transcripts



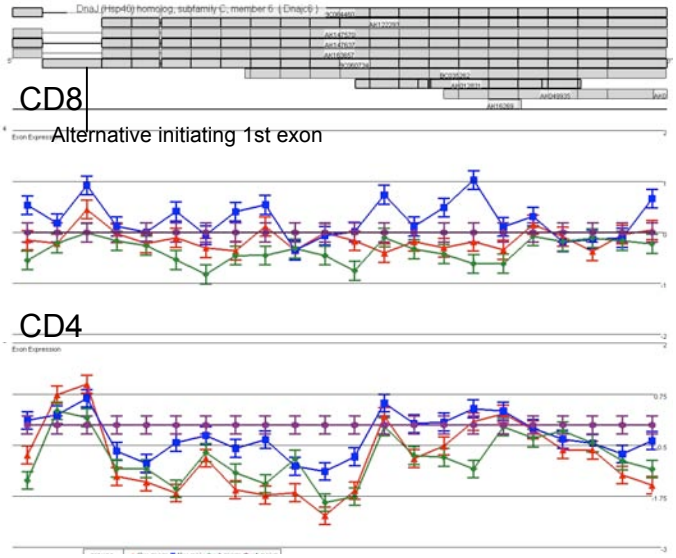
Ddx42 exons and transcripts



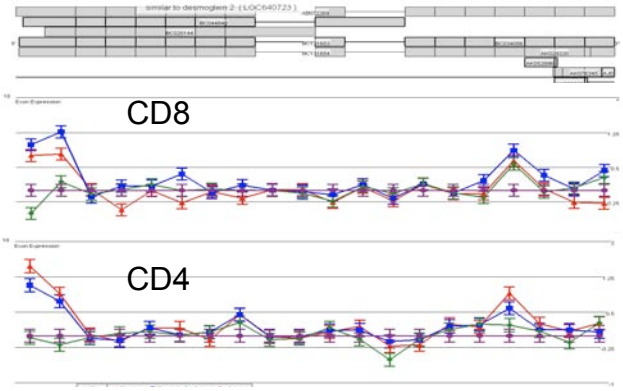
Dhx35 exons and transcripts



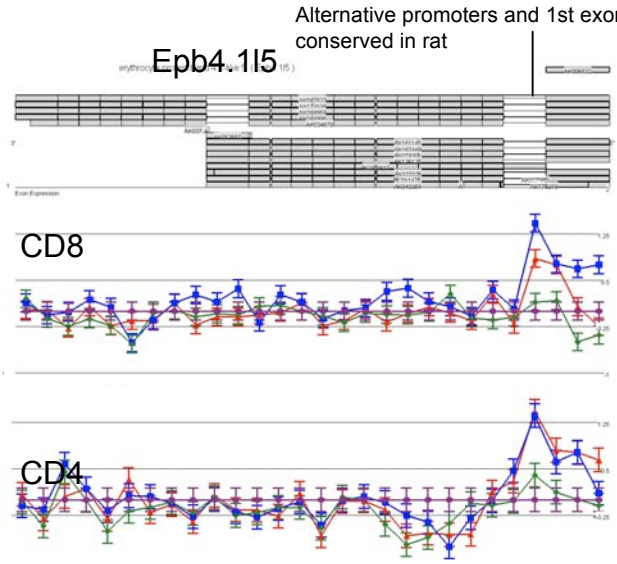
Dnajc6 exons and transcripts



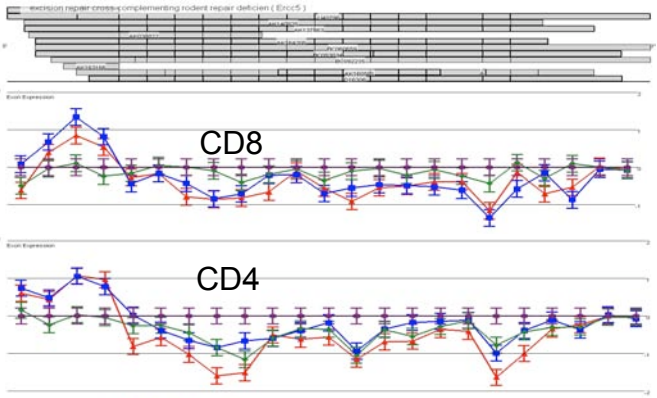
Dsg2 = LOC640723



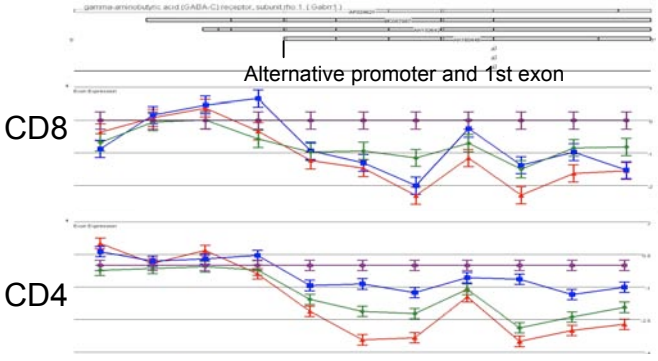
Epb4.1l5



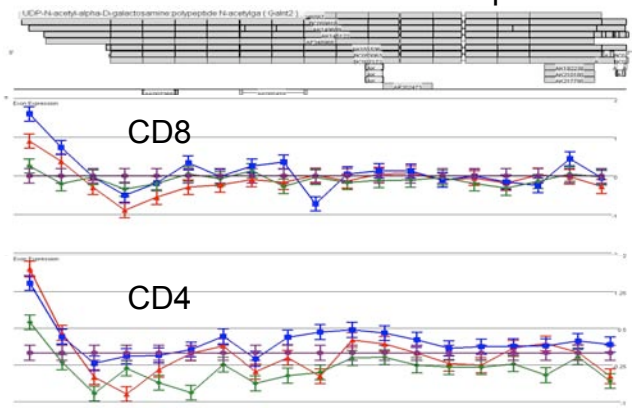
Ercc5 exons and transcripts



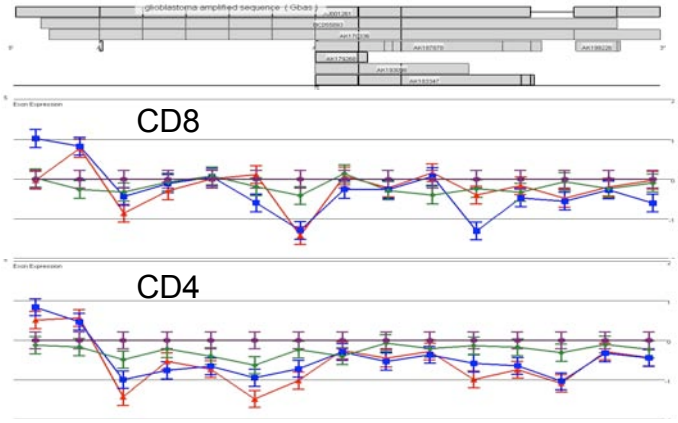
Gabrr1



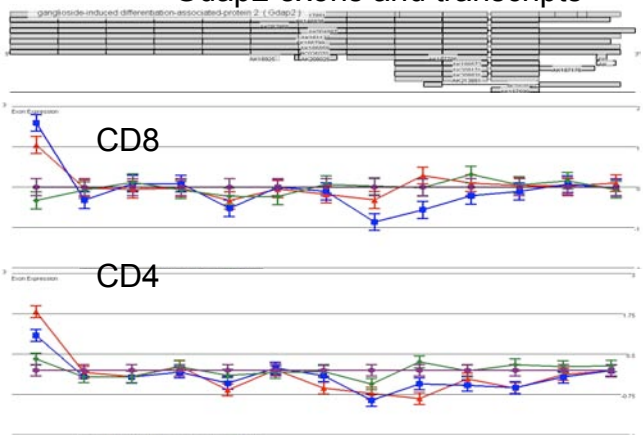
Galnt2 exons and transcripts



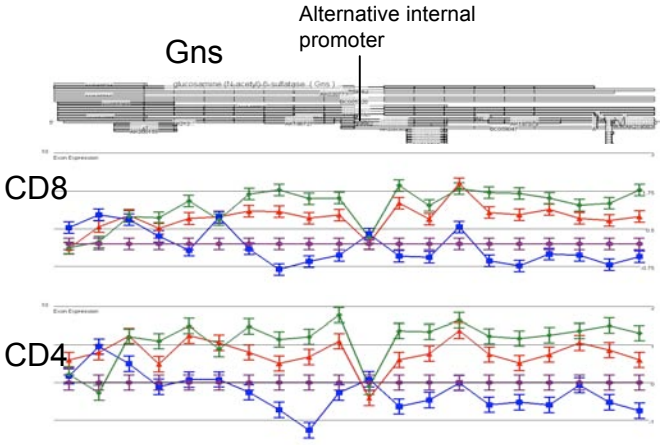
Gbas exons and transcripts



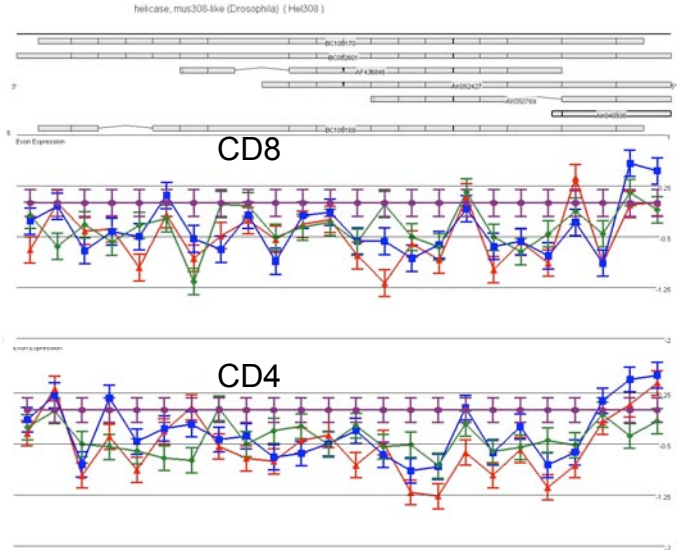
Gdap2 exons and transcripts



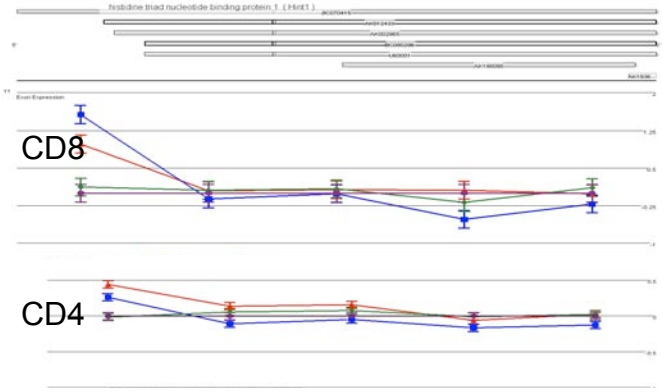
Gns



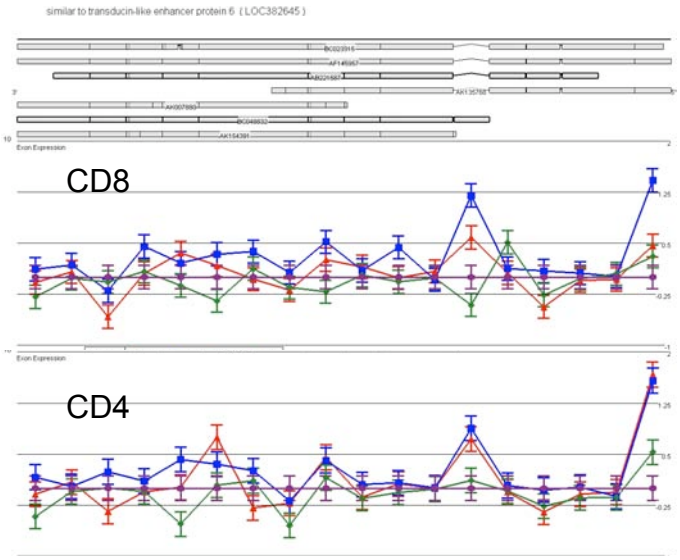
Hel308 exons and transcripts



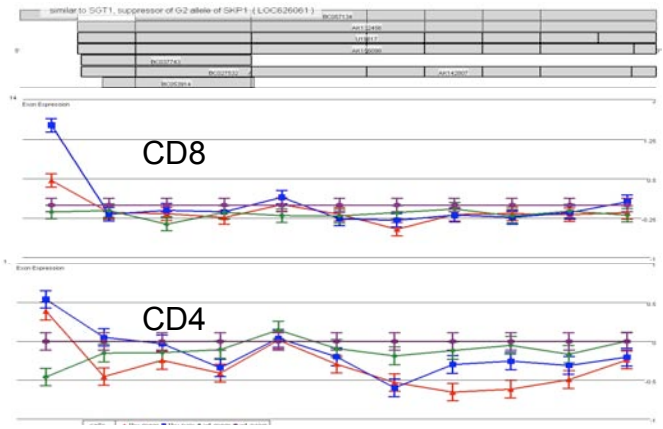
Hint1 exons and transcripts



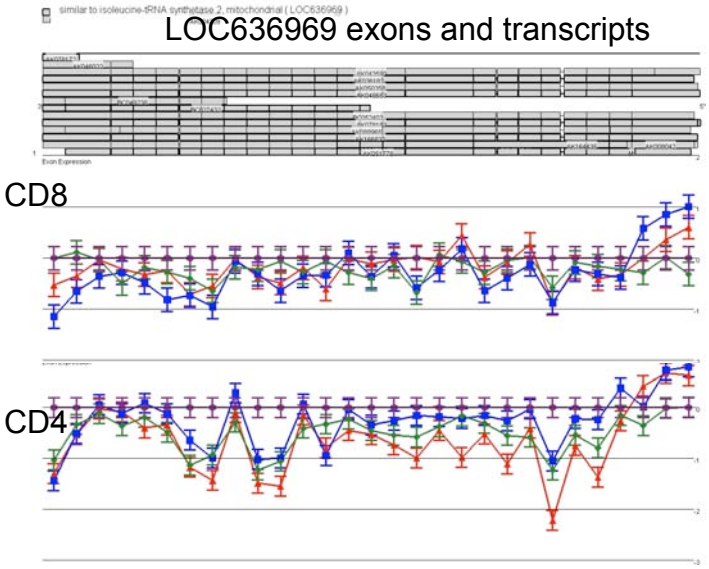
LOC382645 = Tle6



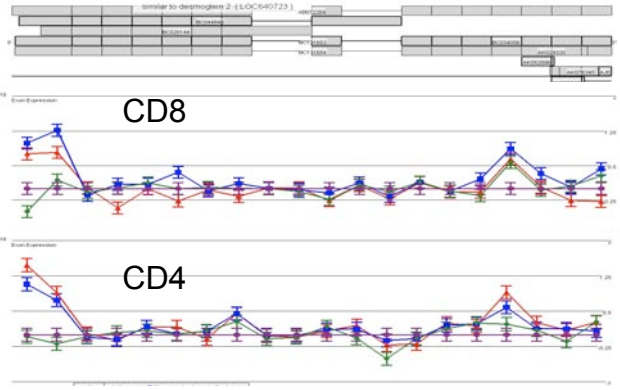
LOC626061 exons and transcripts



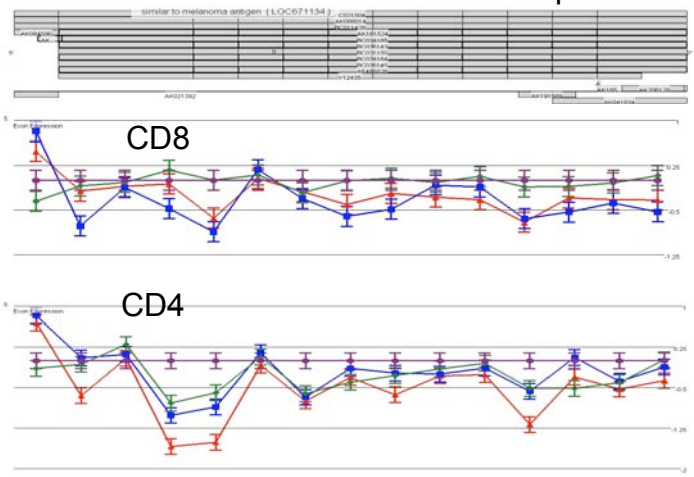
LOC636969 exons and transcripts



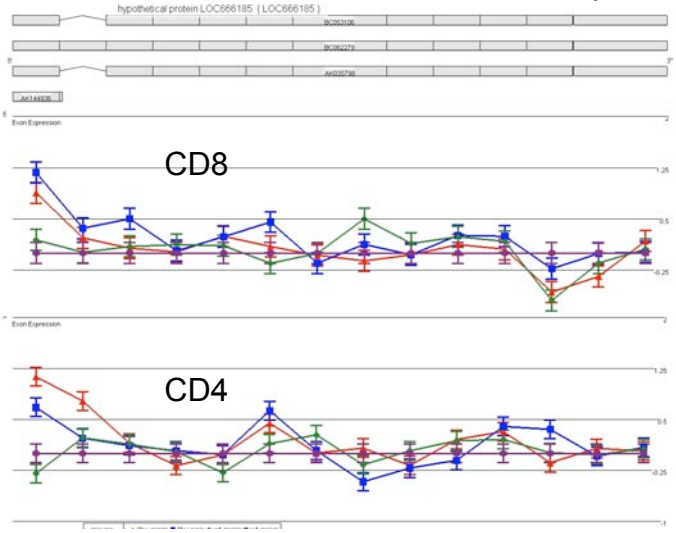
Dsg2 = LOC640723



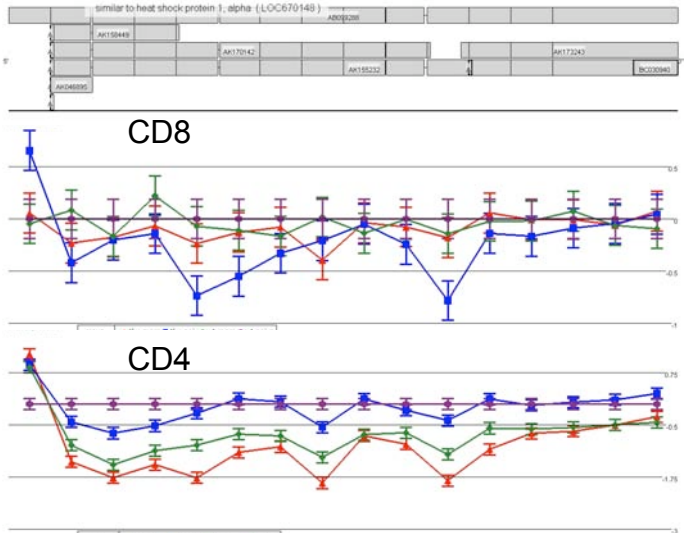
LOC671134 exons and transcripts



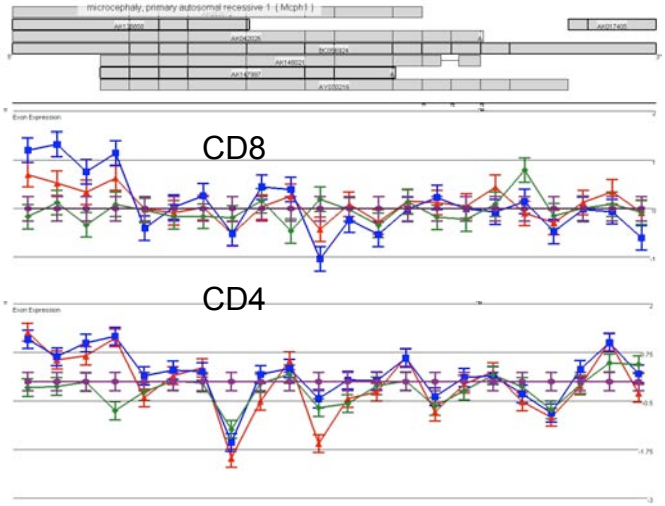
LOC666185 exons and transcripts



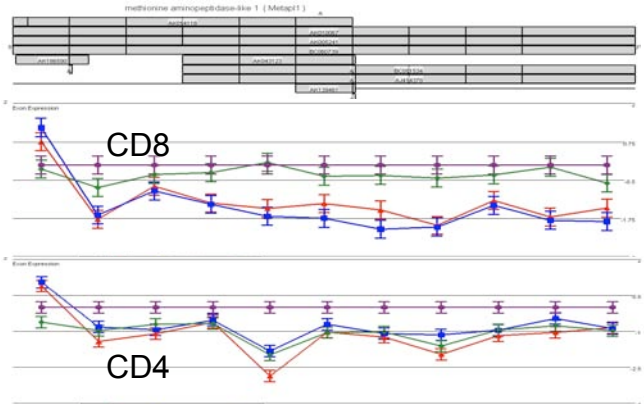
LOC670148 exons and transcripts



Mcph1 exons and transcripts

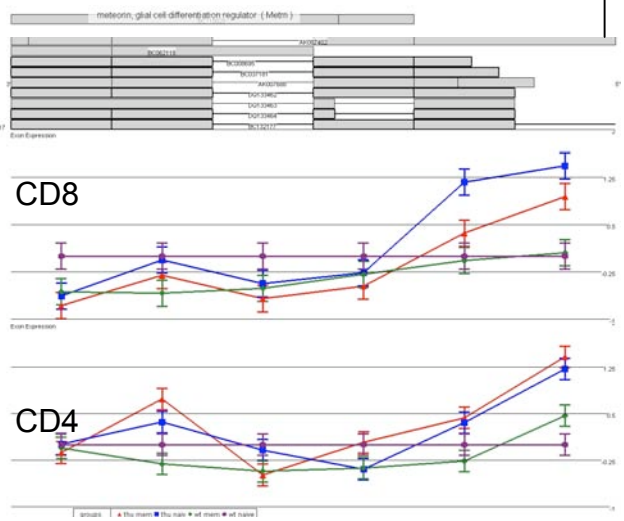


Metapl1 exons and transcripts

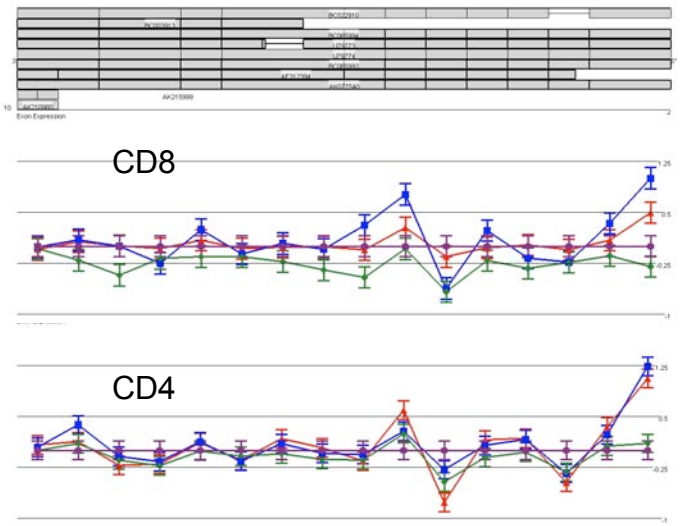


Metrn

Alternative 3' promoter in 1st intron of BC132177 & other ESTs like it



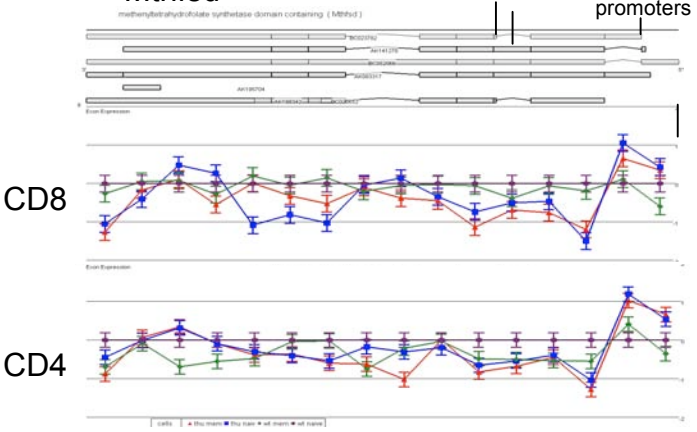
Nnp1 exons and transcripts



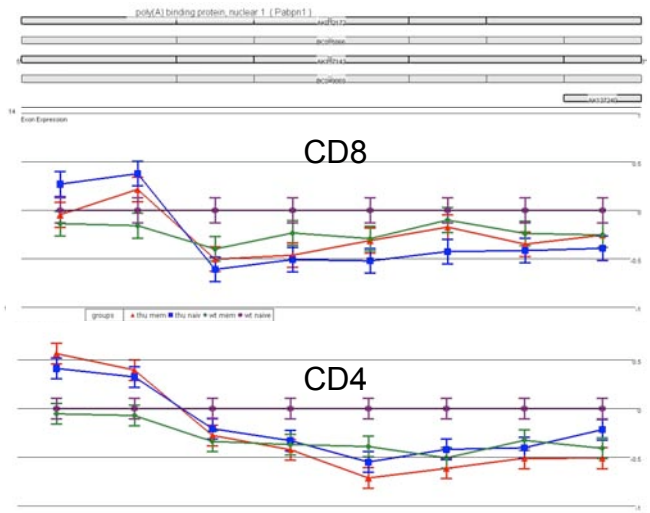
Mthfsd

Alternative exon 3

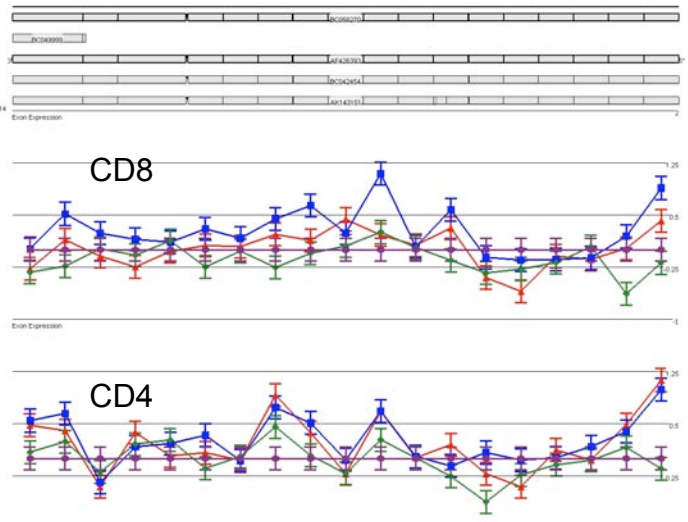
Alternative promoters



Pabpn1 exons and transcripts

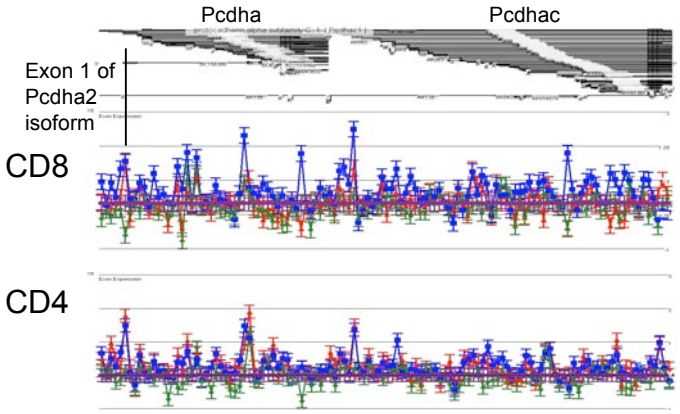


Pcdh21 exons and transcripts

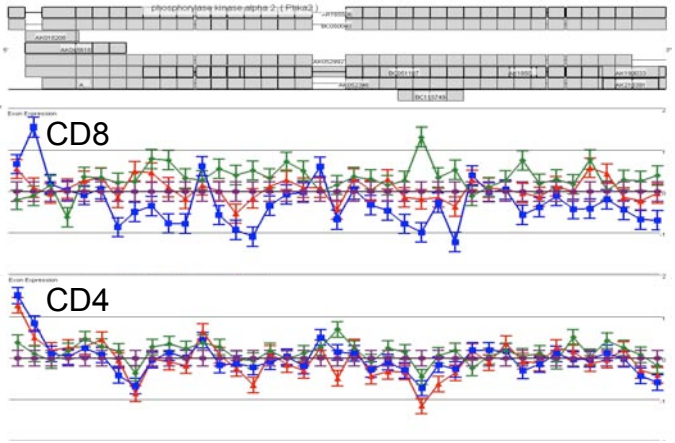


Pcdhac1 locus

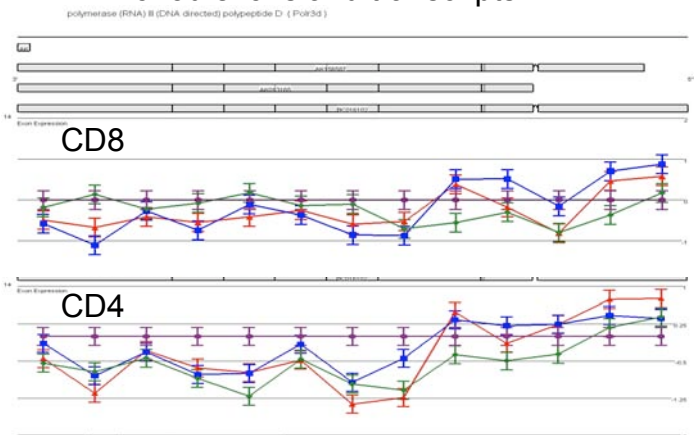
Staggered transcript starts represent alternative promoters and first exons



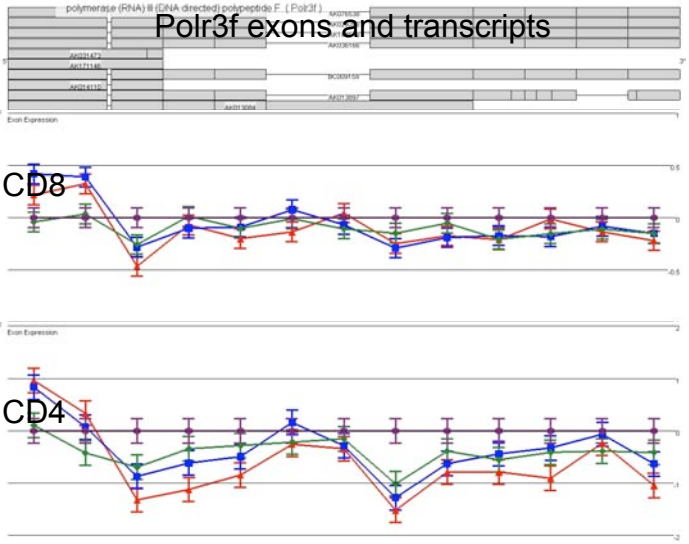
Phka2 exons and transcripts



Polr3d exons and transcripts

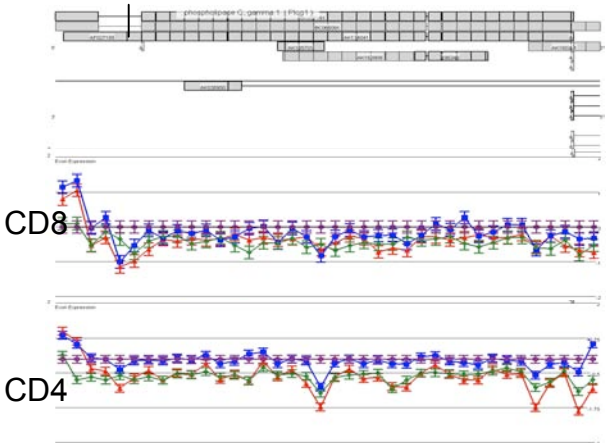


Polr3f exons and transcripts

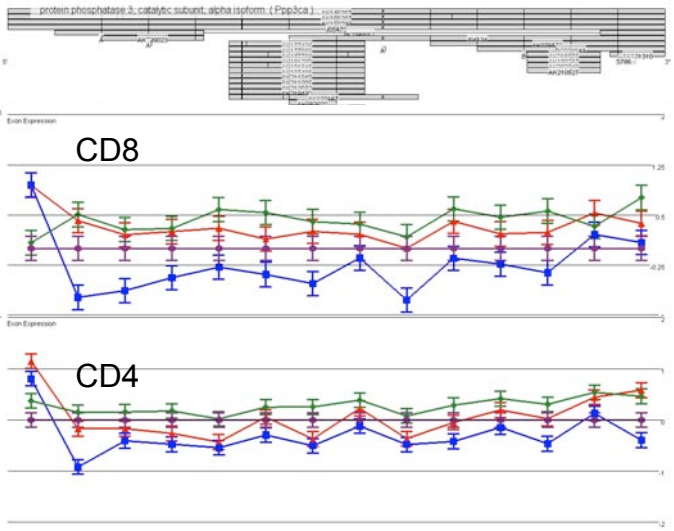


Plcg1

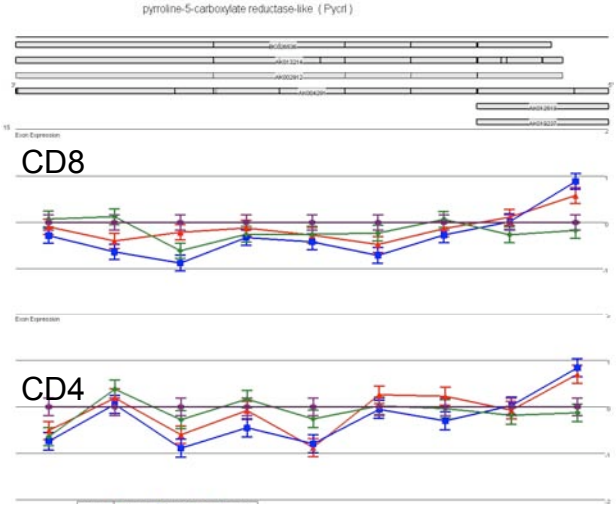
Alternative promoter & 1st exon, initiates at internal Met & lack PH domain



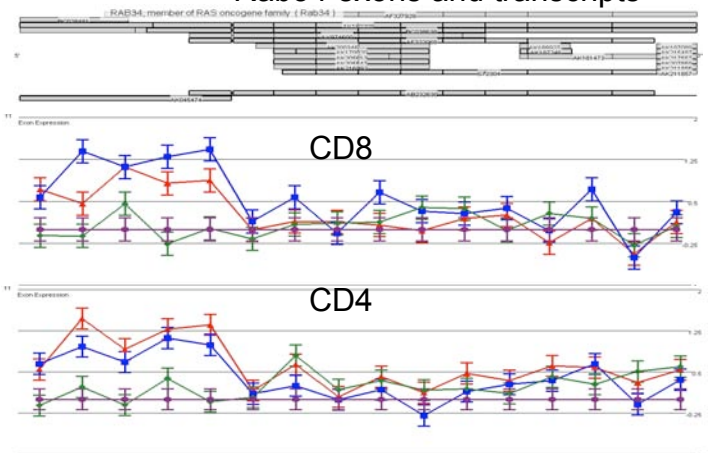
Ppp3ca exons and transcripts



Pycrl exons and transcripts



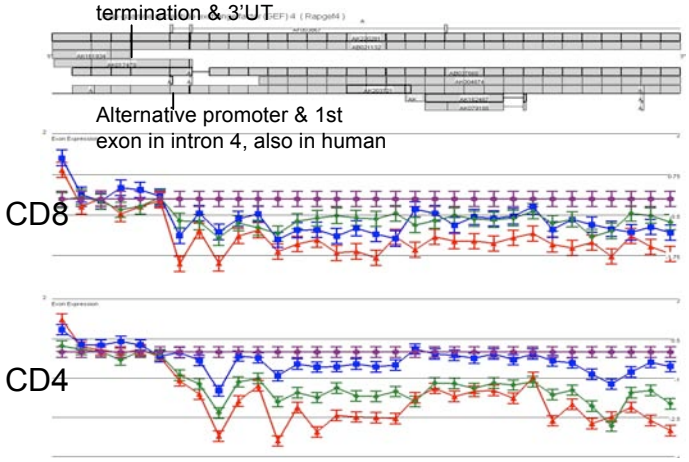
Rab34 exons and transcripts



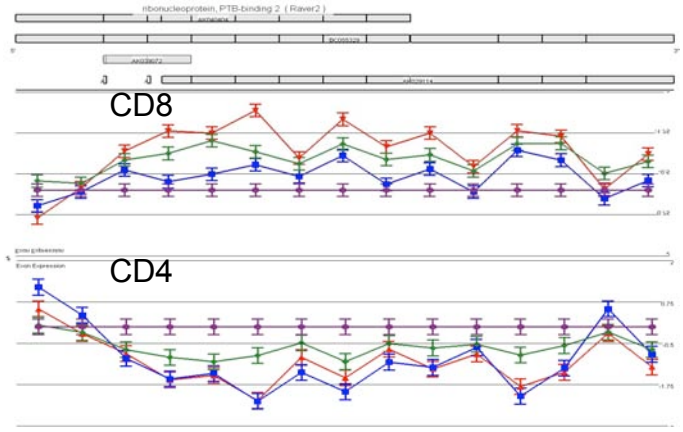
Rapgef4

Silencing of splice donor - termination & 3'UT

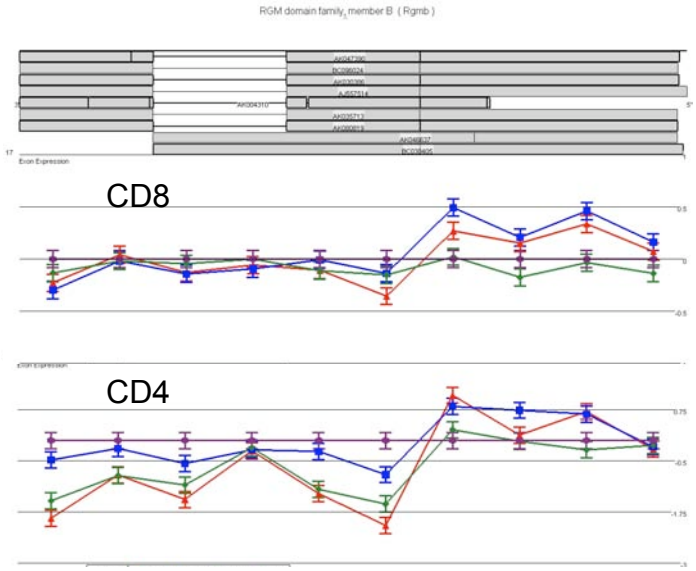
Alternative promoter & 1st exon in intron 4, also in human



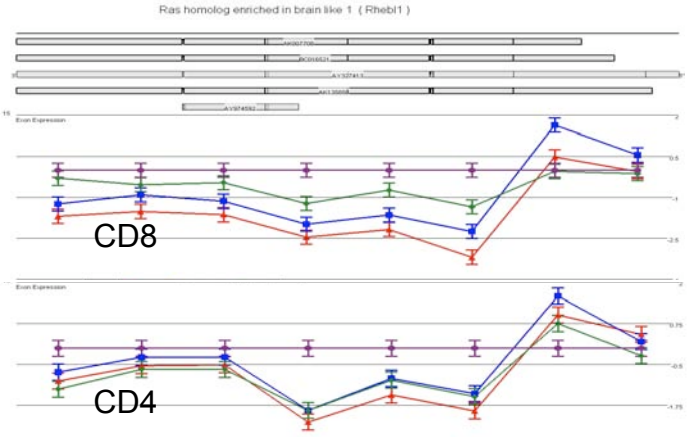
Raver2 exons and transcripts



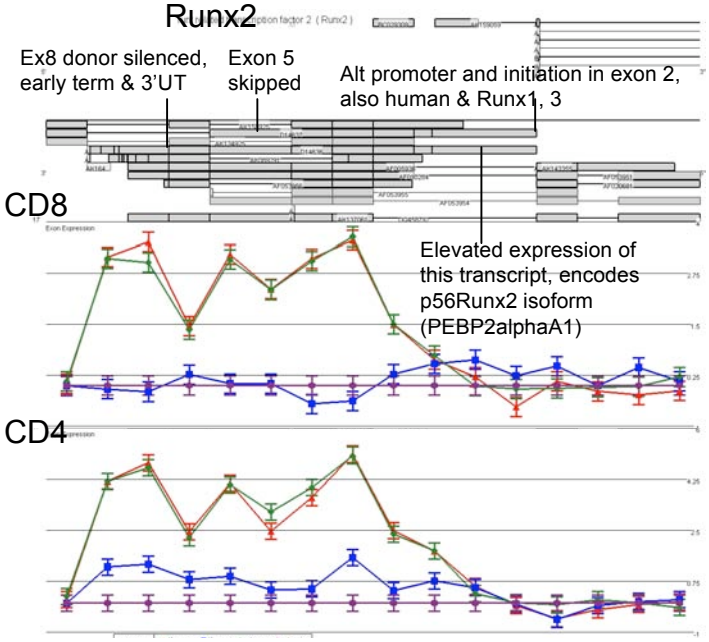
Rgmb exons and transcripts



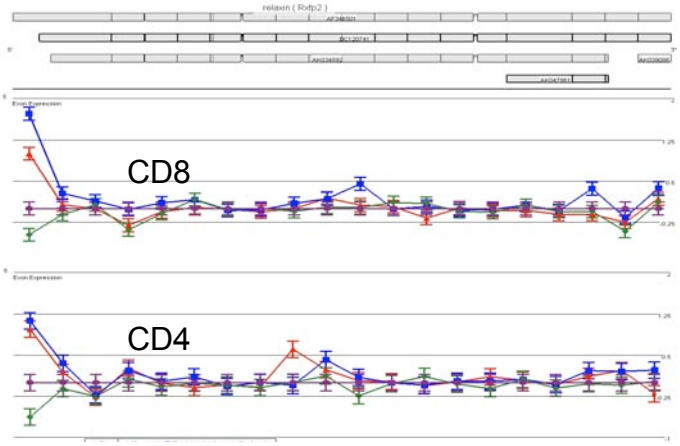
Rhebl1 exons and transcripts



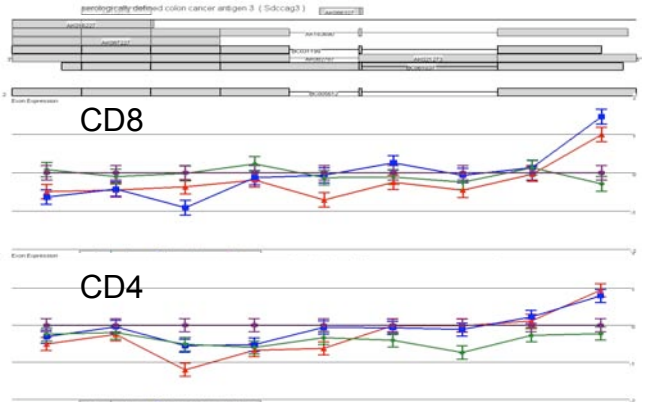
Runx2



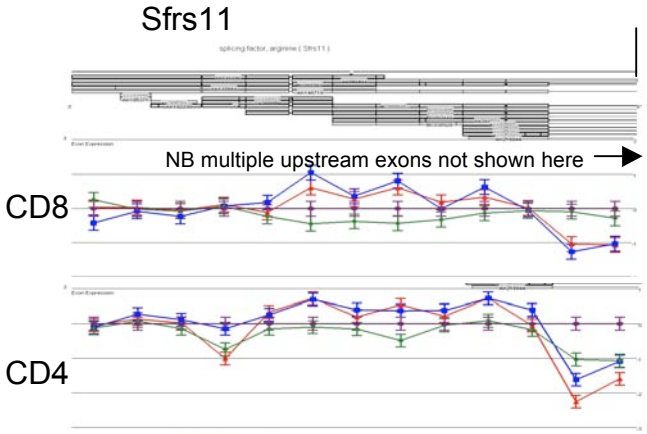
Rxfp2 exons and transcripts



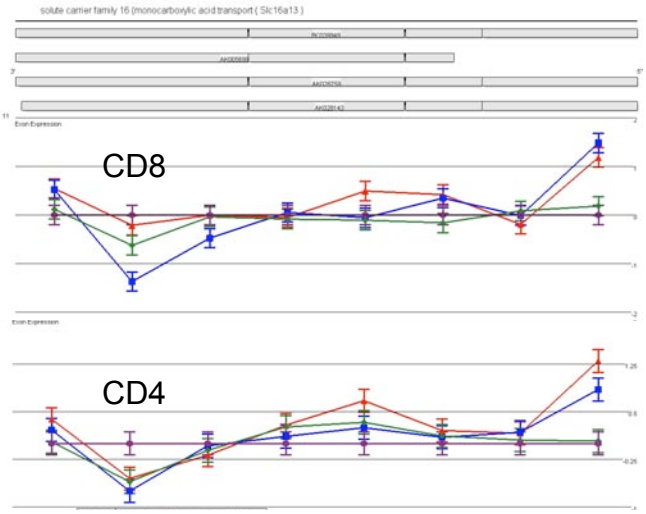
Sdccag3 exons and transcripts



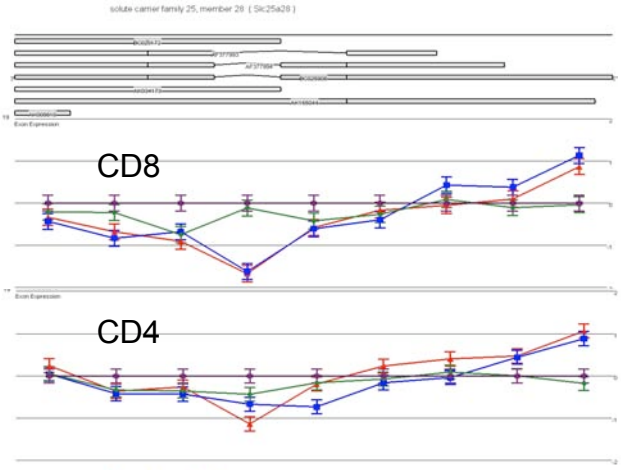
Sfrs11



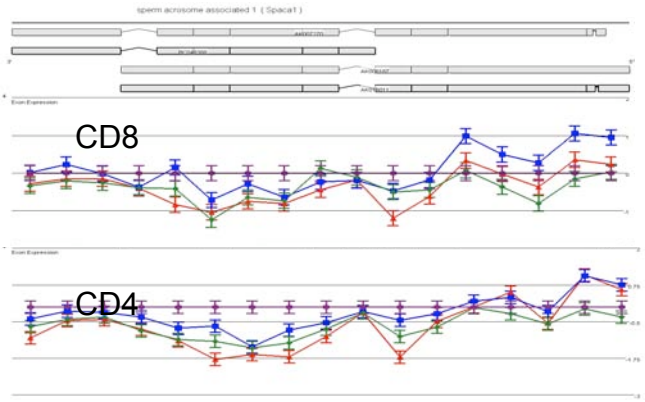
Slc16a13 exons and transcripts



Slc25a28 exons and transcripts

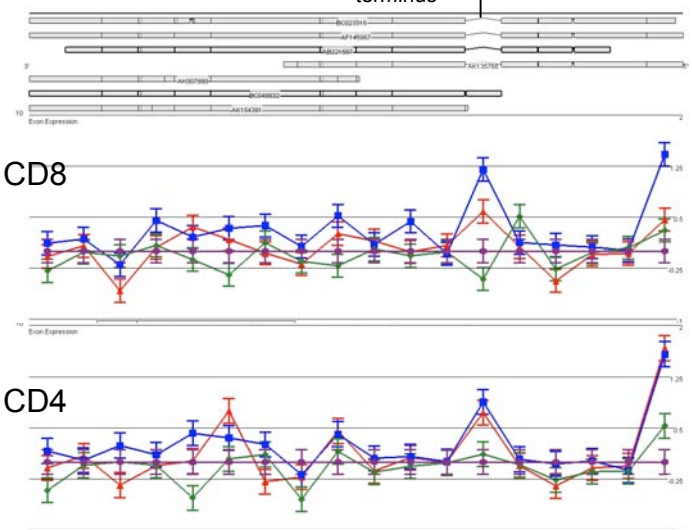


Spaca1 exons and transcripts

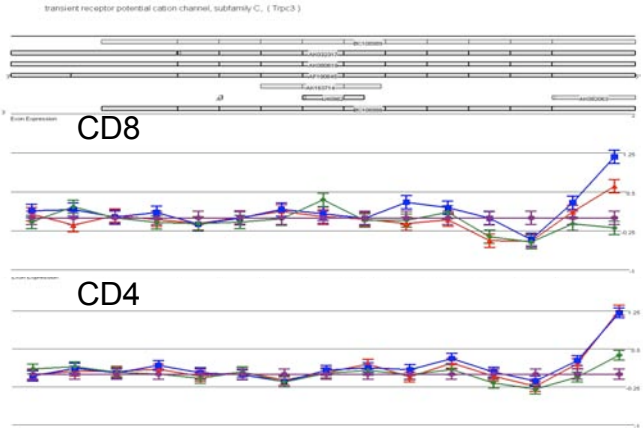


LOC382645 = Tle6

Alternative promoter/1st exon, multiple transcripts, shortened N-terminus

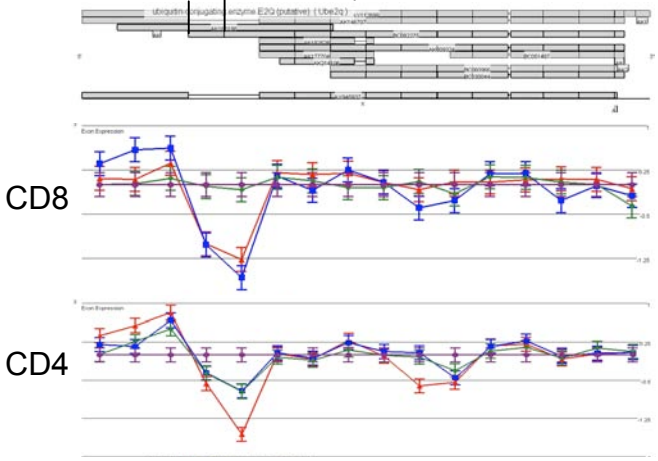


Trpc3 exons and transcripts

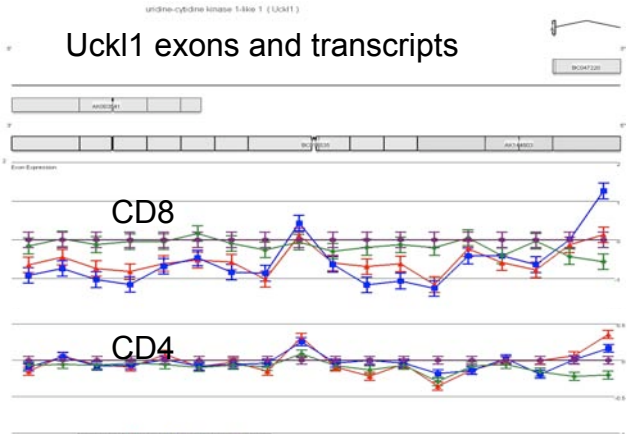


Ube2q

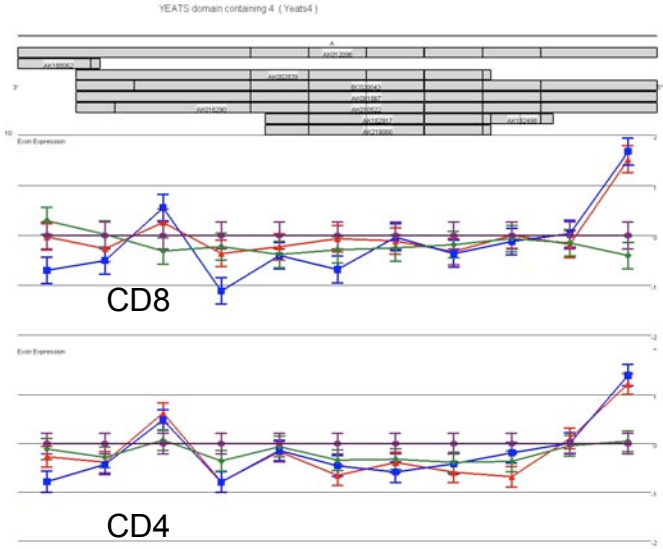
Alternative promoter and intronic exon 2 acceptor



Uckl1 exons and transcripts

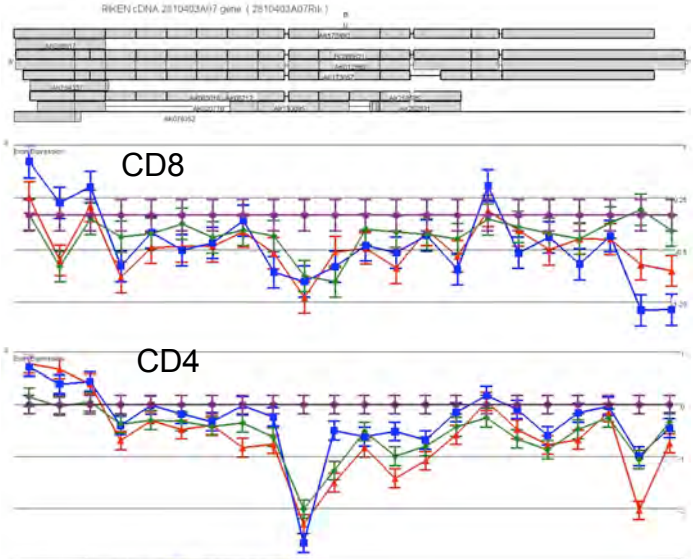


Yeats4 exons and transcripts

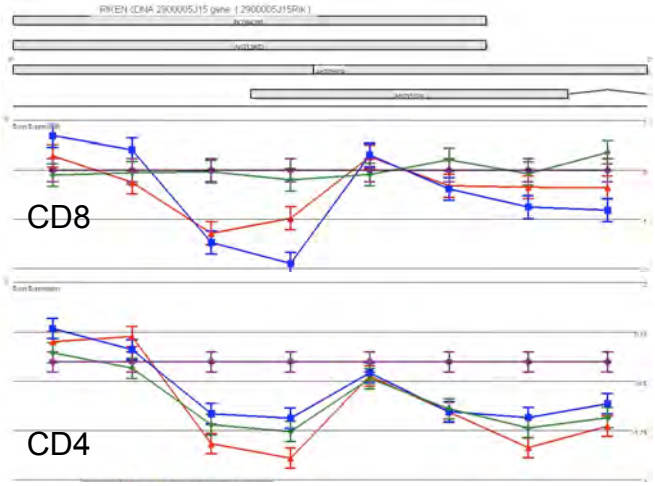


Group F Gene Views

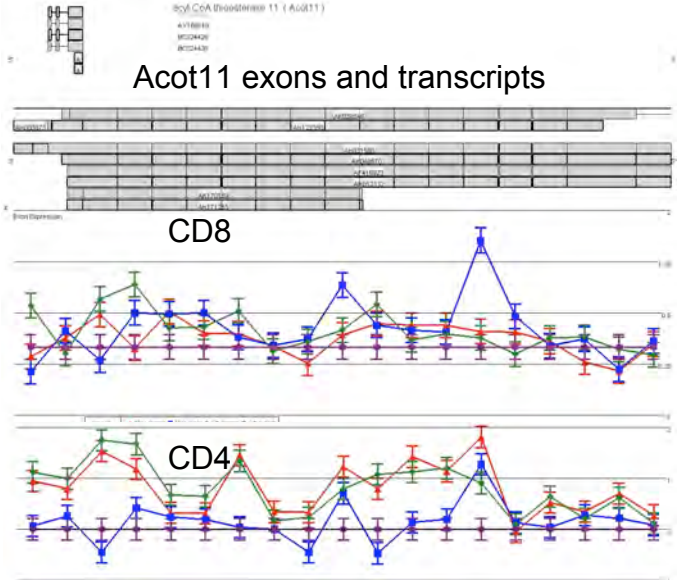
2810403A07Rik exons and transcripts



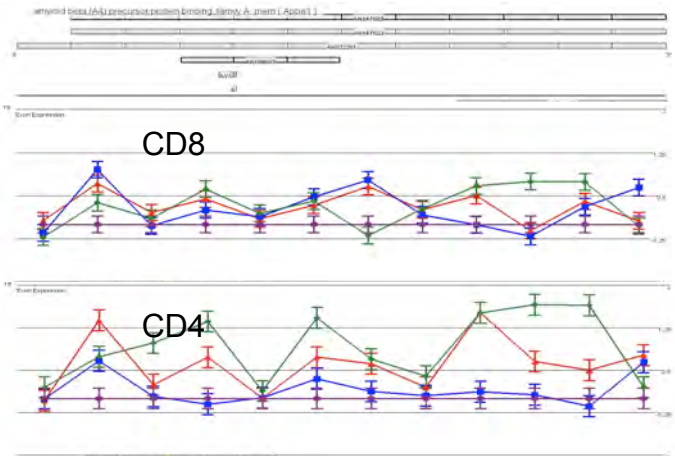
2900005J15 exons and transcripts



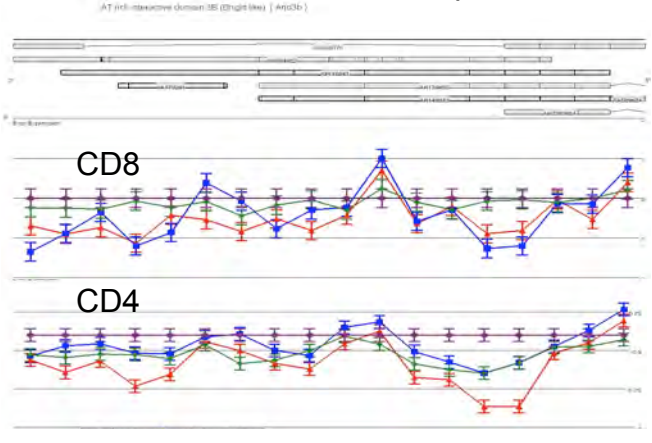
Acot11 exons and transcripts



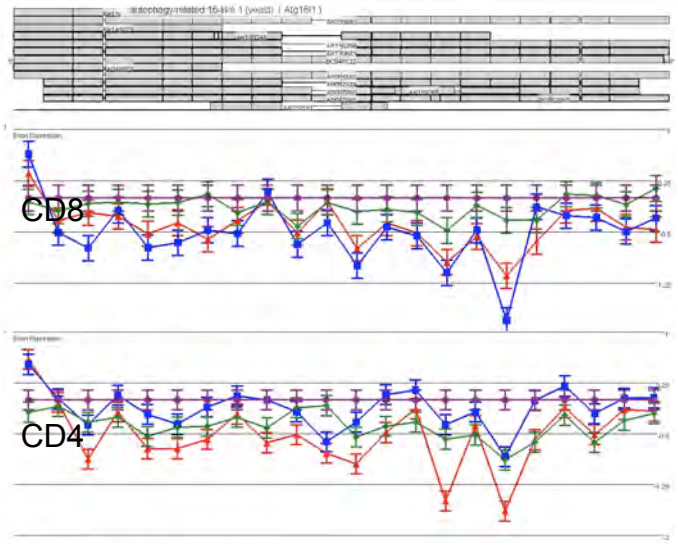
Apba1 exons and transcripts



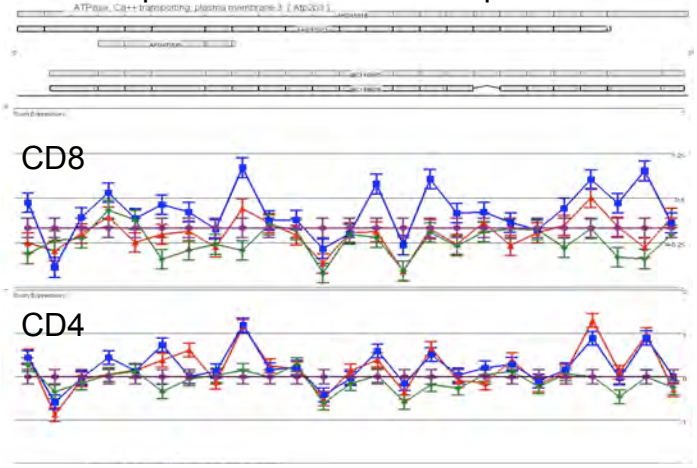
Arid3b exons and transcripts



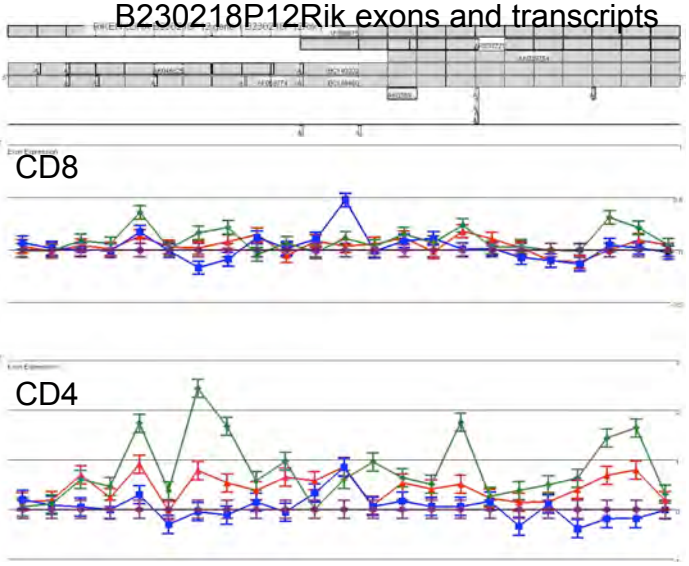
Atg16l1 exons and transcripts



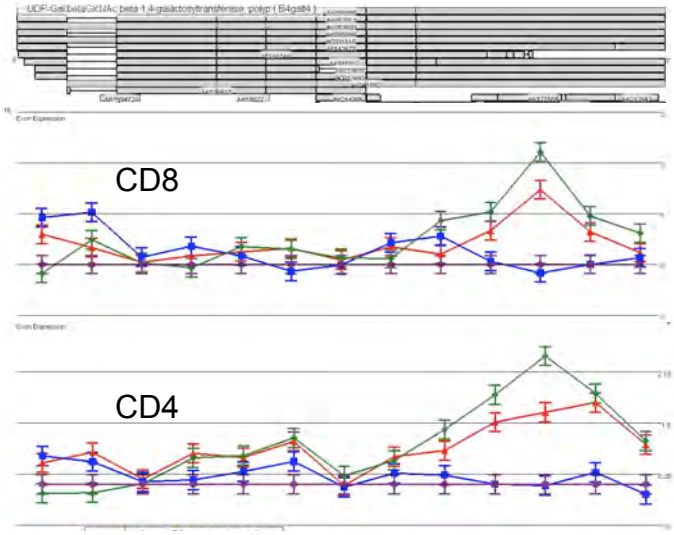
Atp2b3 exons and transcripts



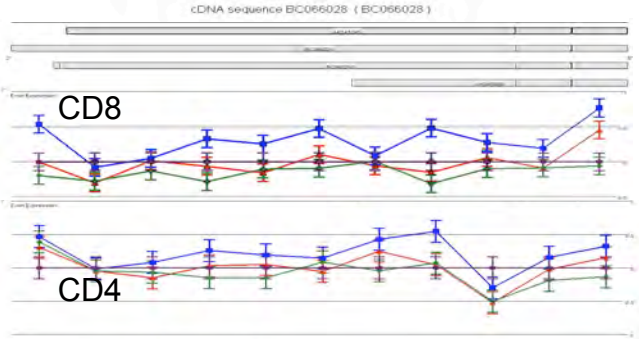
B230218P12Rik exons and transcripts



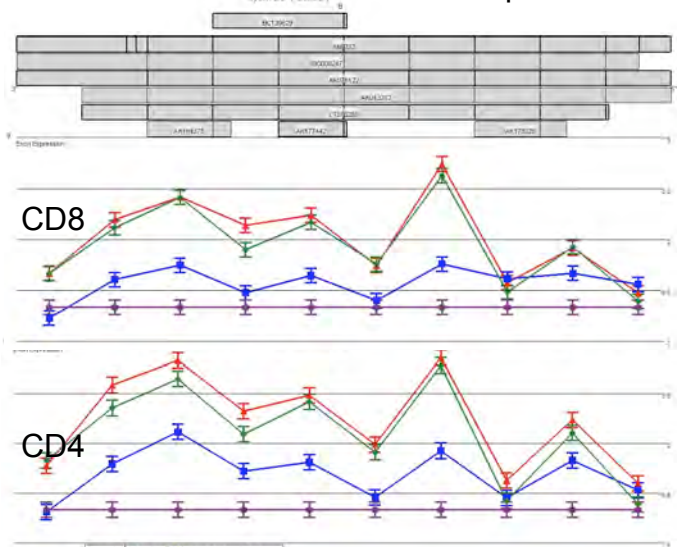
B4galt4 exons and transcripts



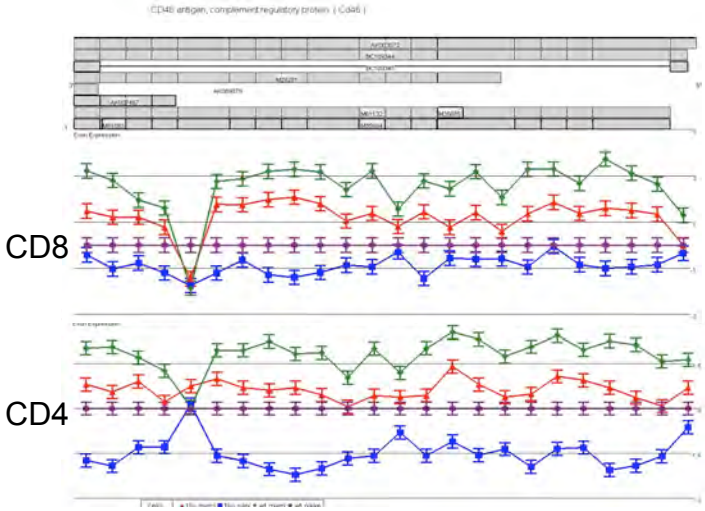
BC066028 exons and transcripts



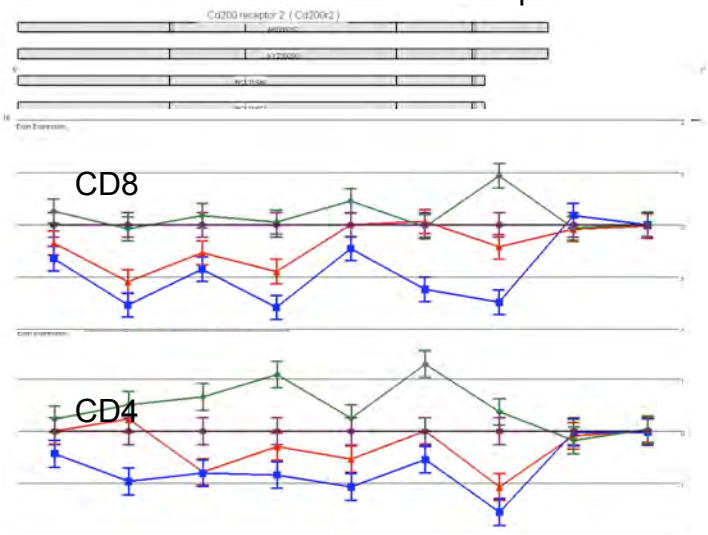
Ccnb2 exons and transcripts



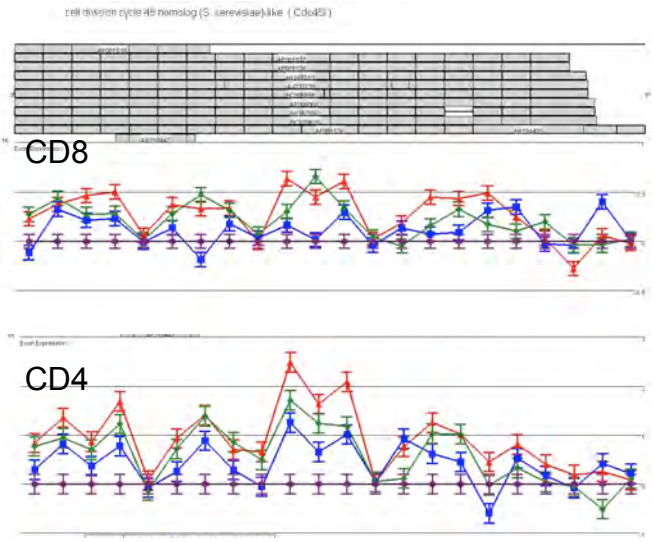
Cd46 exons and transcripts



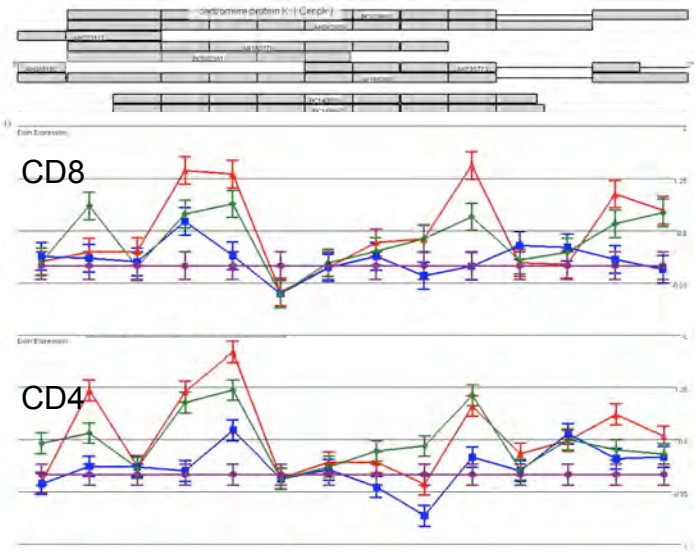
Cd200r2 exons and transcripts



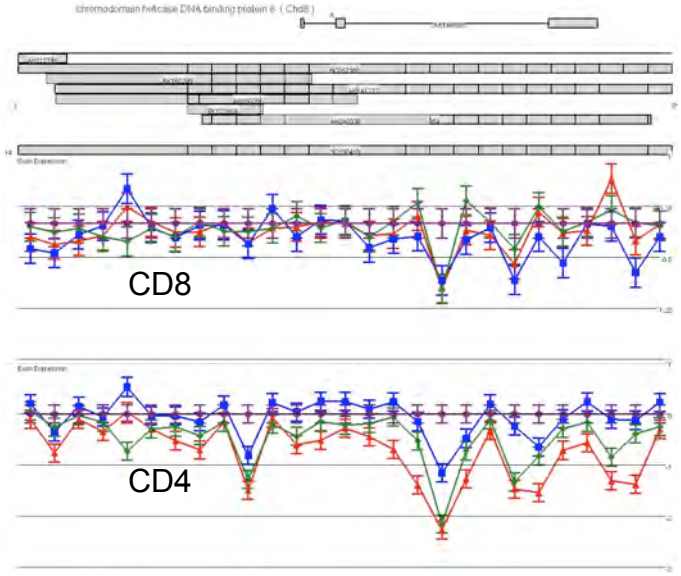
Cdc45l exons and transcripts



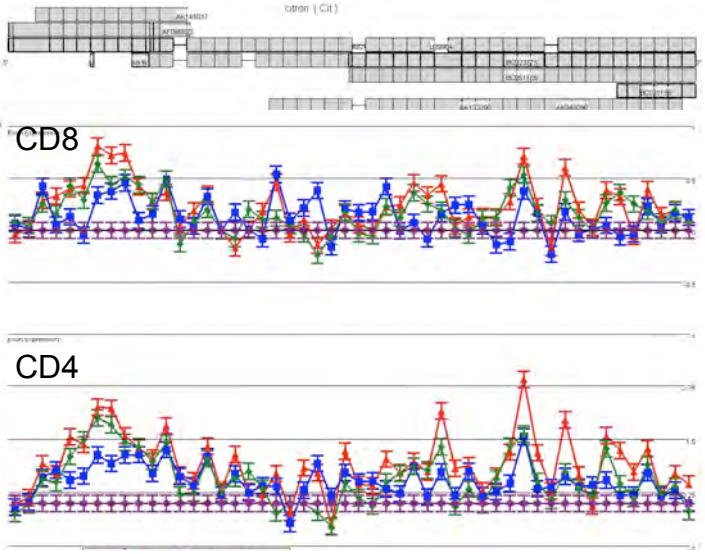
Cenpk exons and transcripts



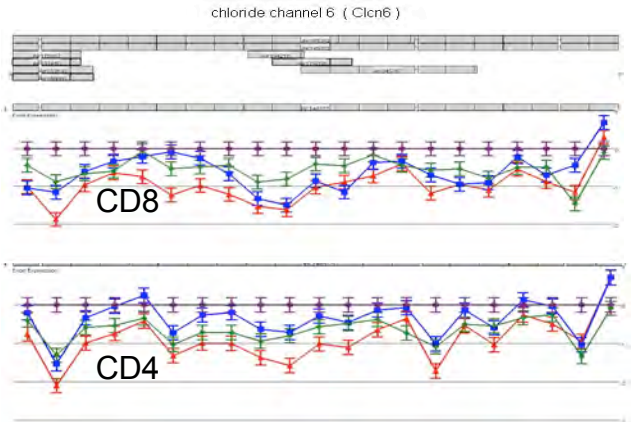
Chd8 exons and transcripts



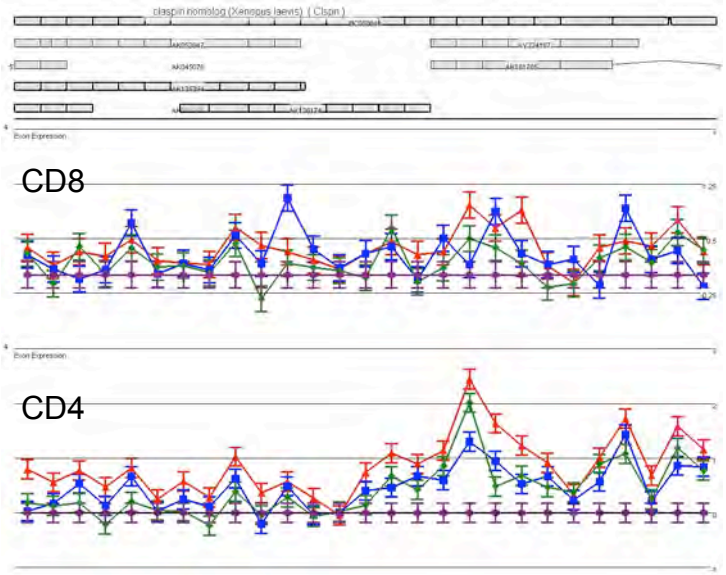
Cit exons and transcripts



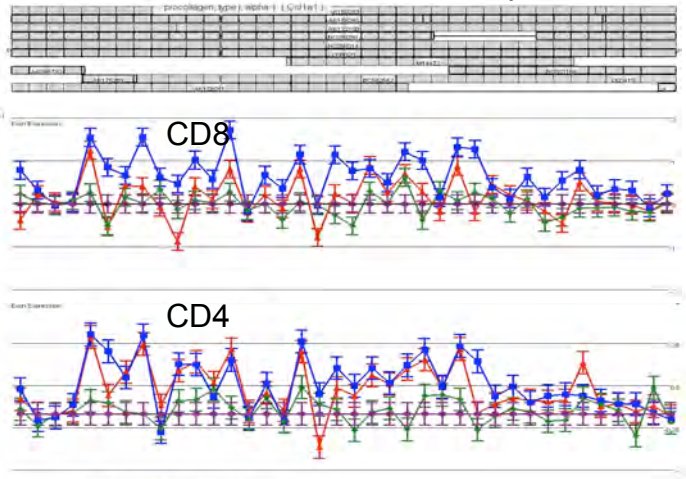
Clcn6 exons and transcripts



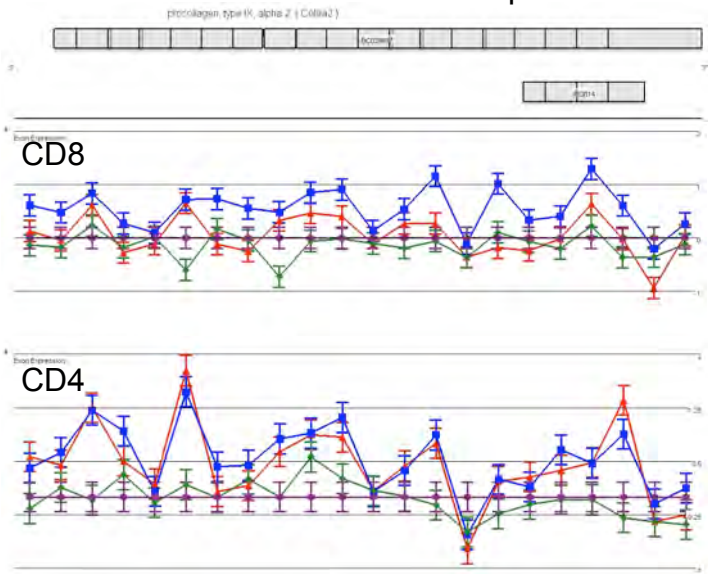
Clspn exons and transcripts



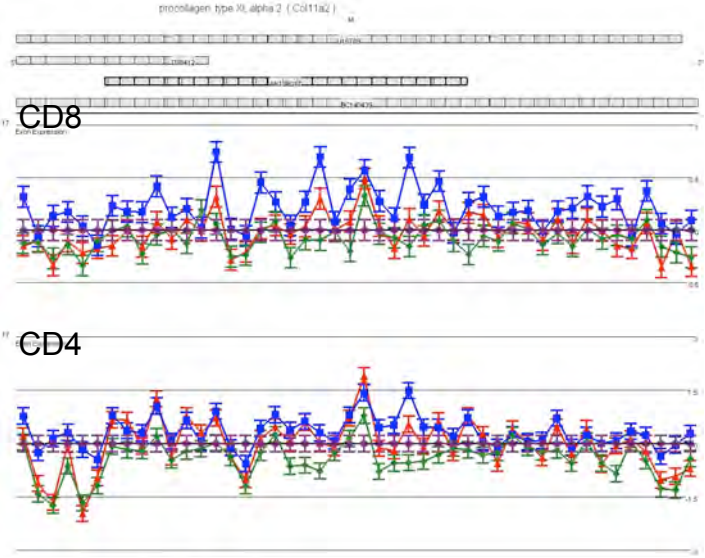
Col1a1 exons and transcripts



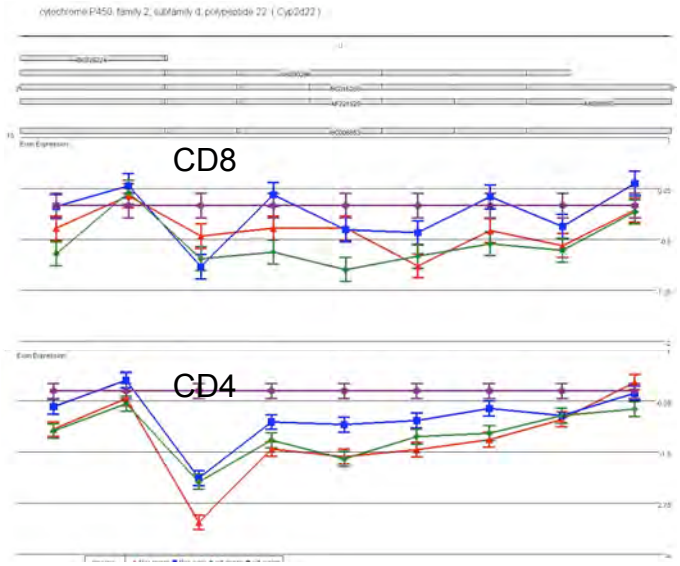
Col9a2 exons and transcripts



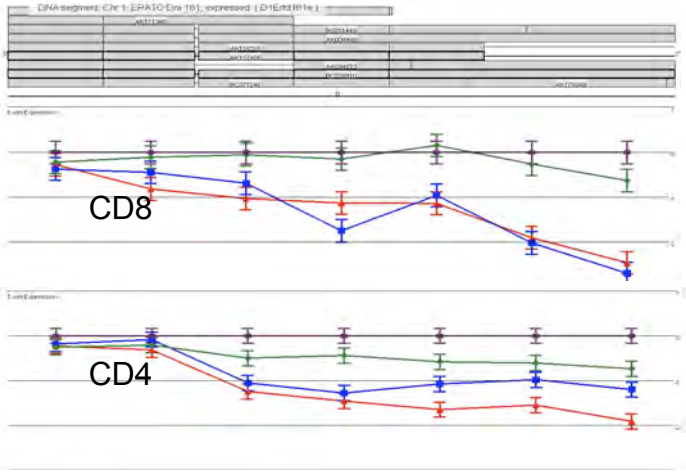
Col11a2 exons and transcripts



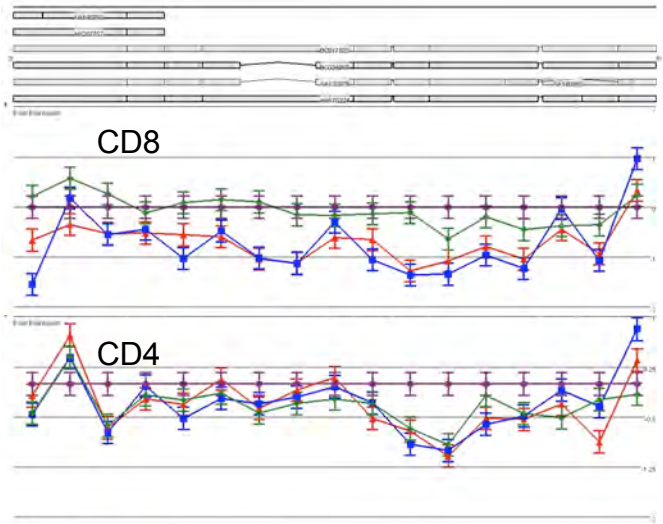
Cyp2d22 exons and transcripts



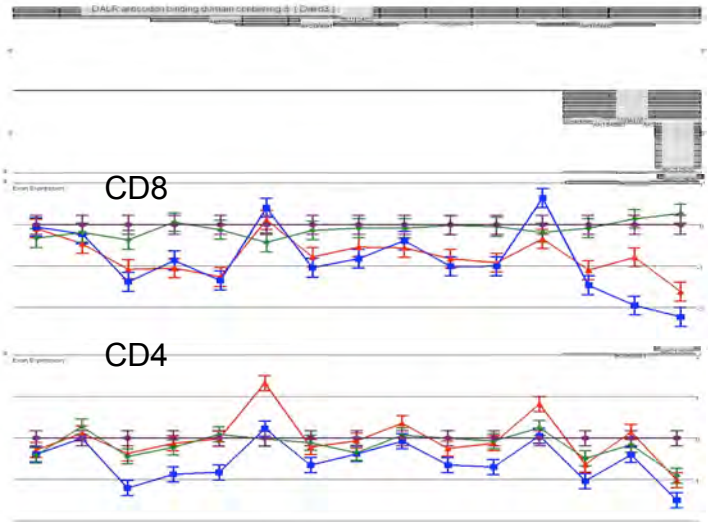
D1Ertd161e exons and transcripts



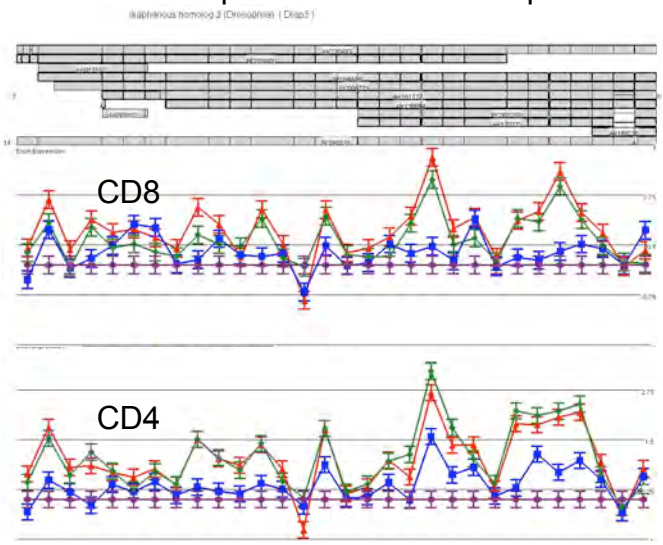
D4Wsu114e exons and transcripts



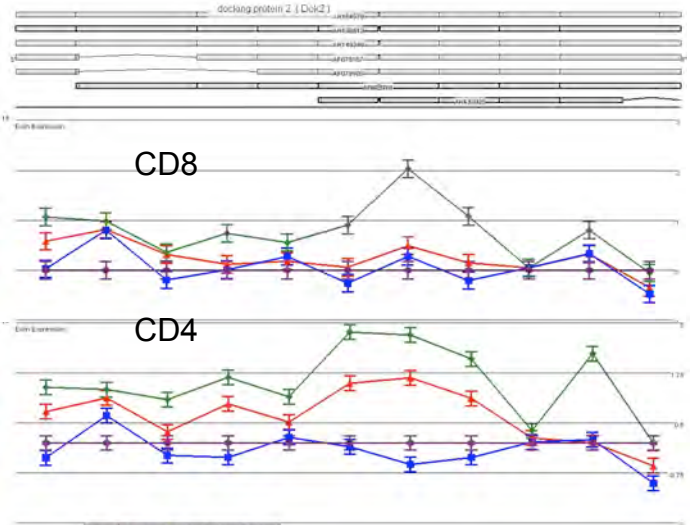
Dalrd3 exons and transcripts



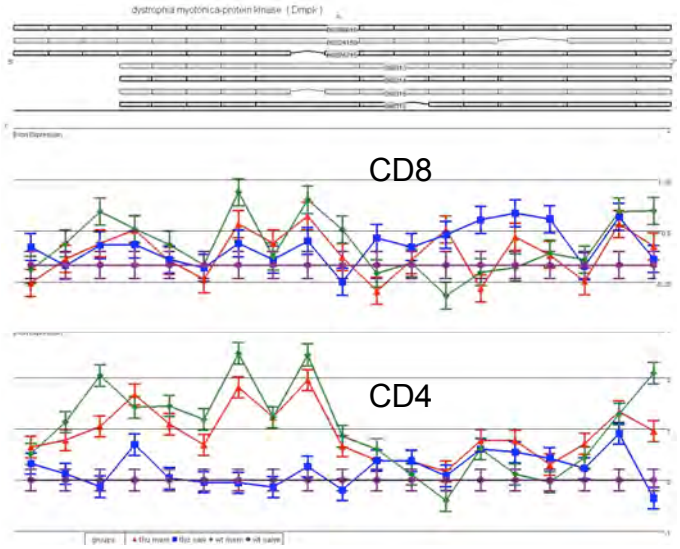
Diap3 exons and transcripts



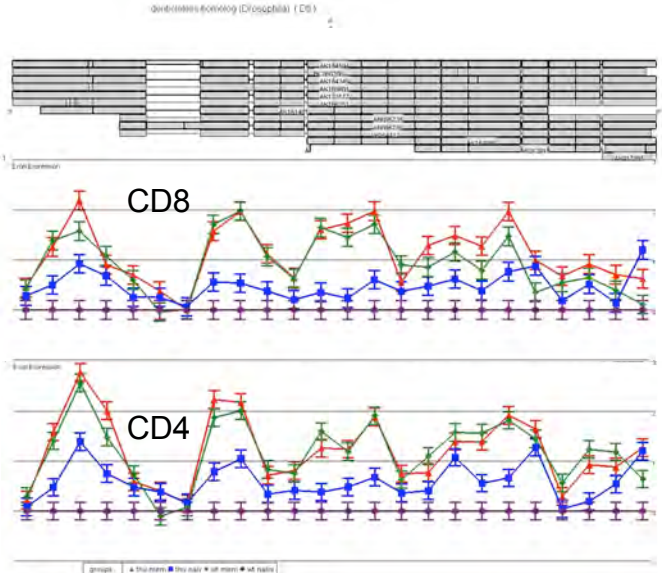
Dok2 exons and transcripts



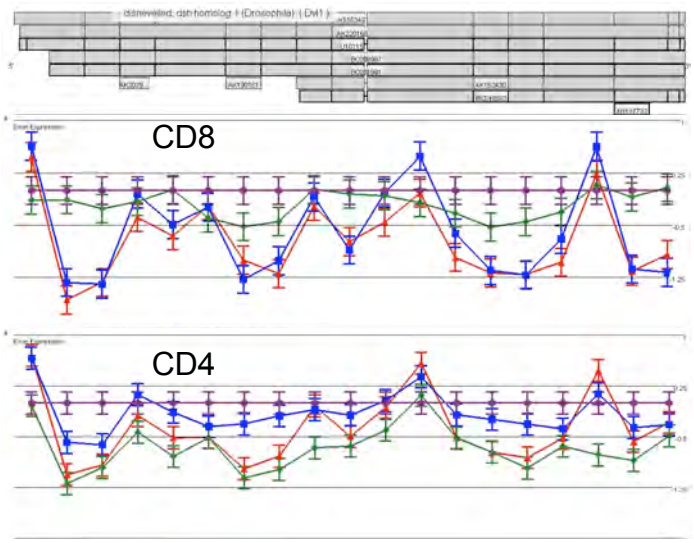
Dmpk exons and transcripts



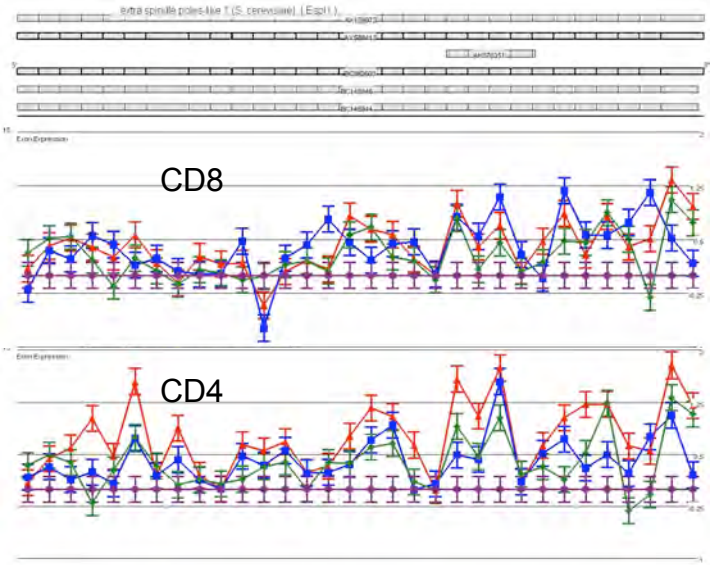
Dtl exons and transcripts



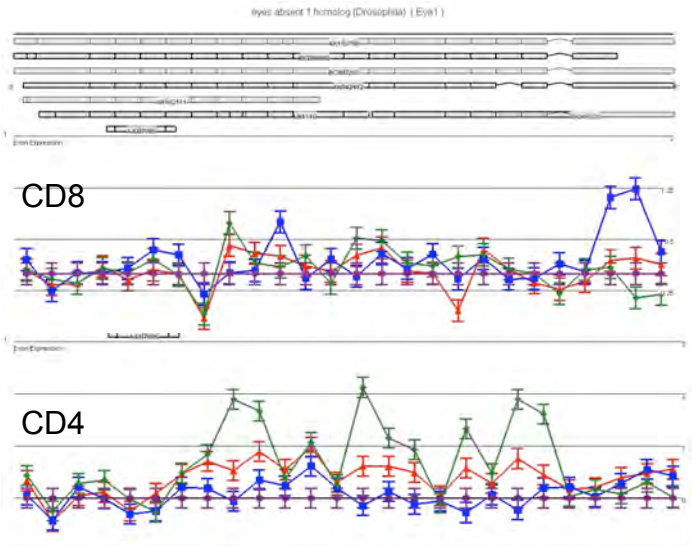
Dvl1 exons and transcripts



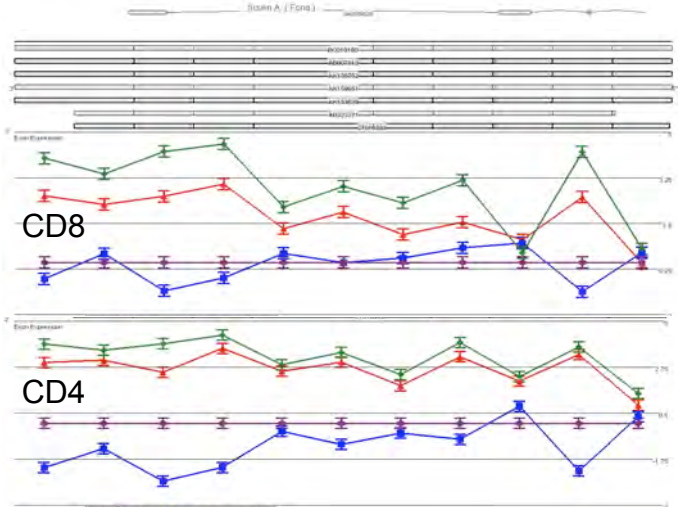
Espl1 exons and transcripts



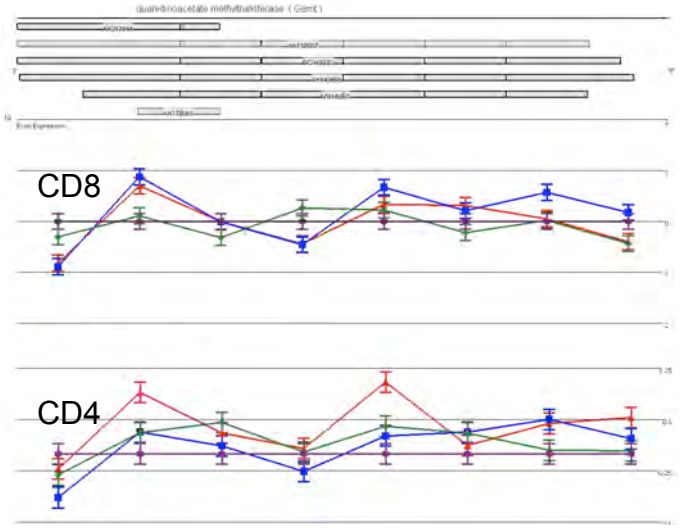
Eya1 exons and transcripts



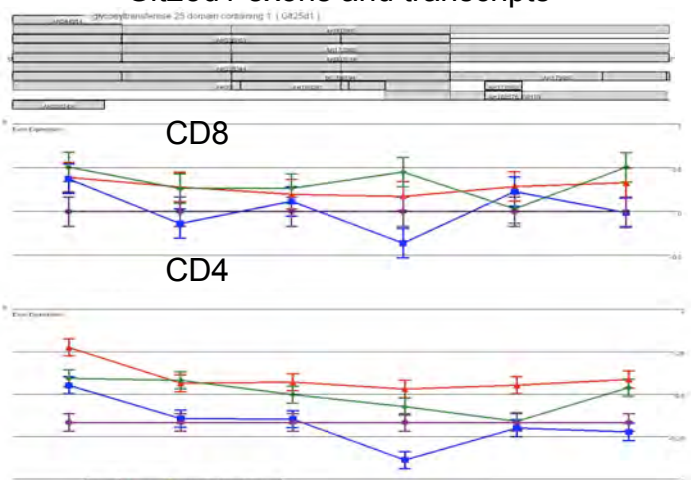
Fcna exons and transcripts



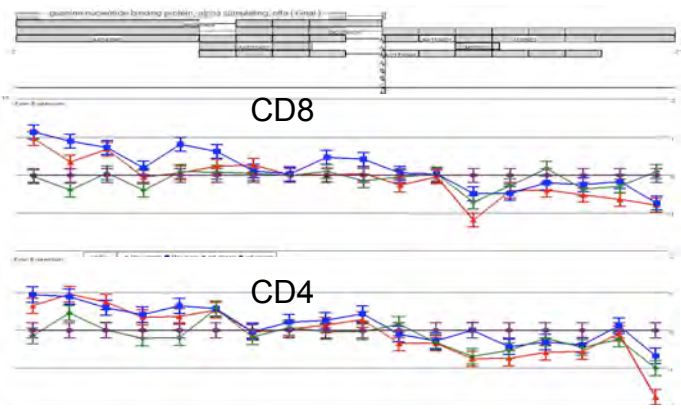
Gamt exons and transcripts



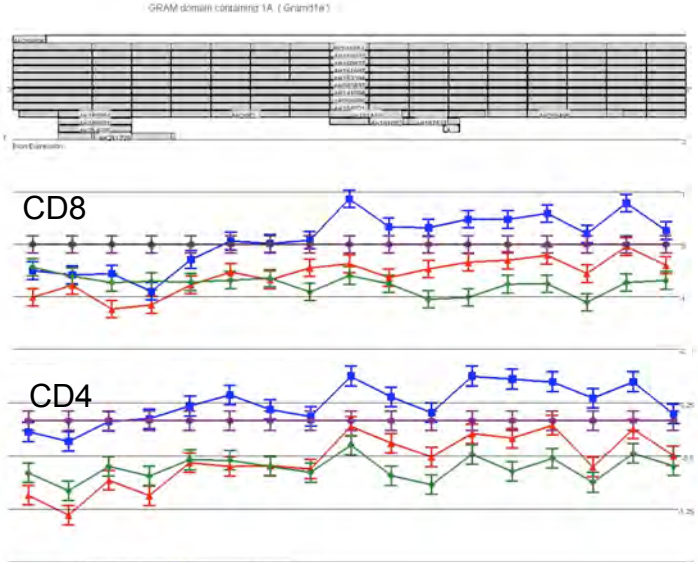
Glt25d1 exons and transcripts



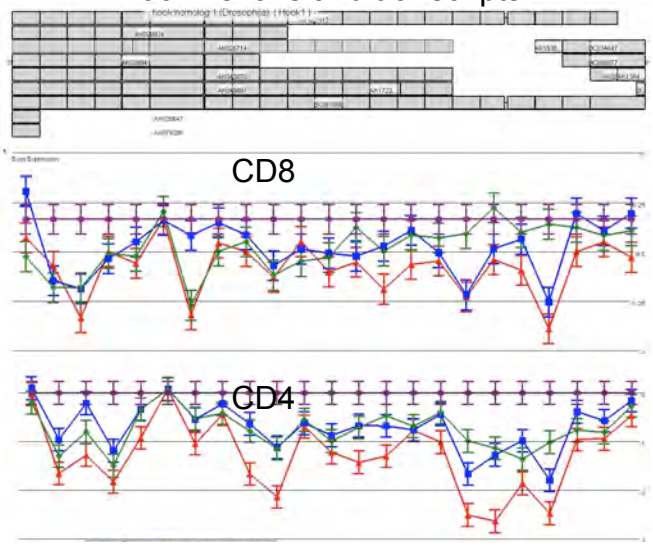
Gnal exons and transcripts



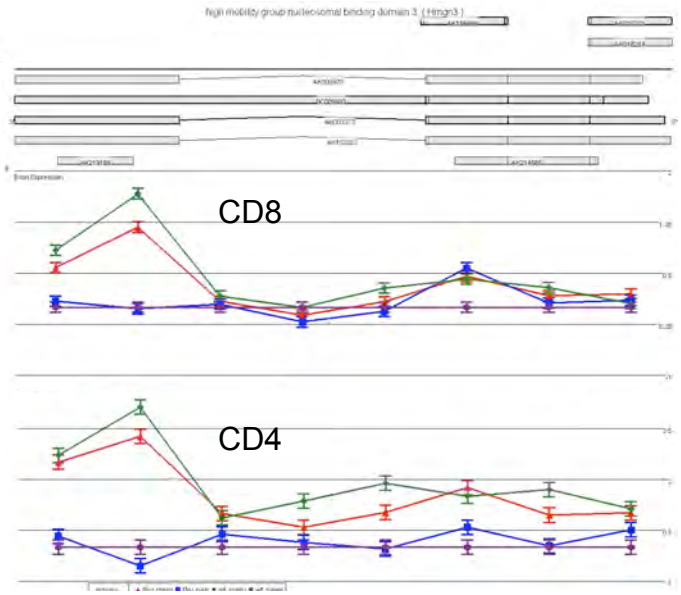
Gramd1a exons and transcripts



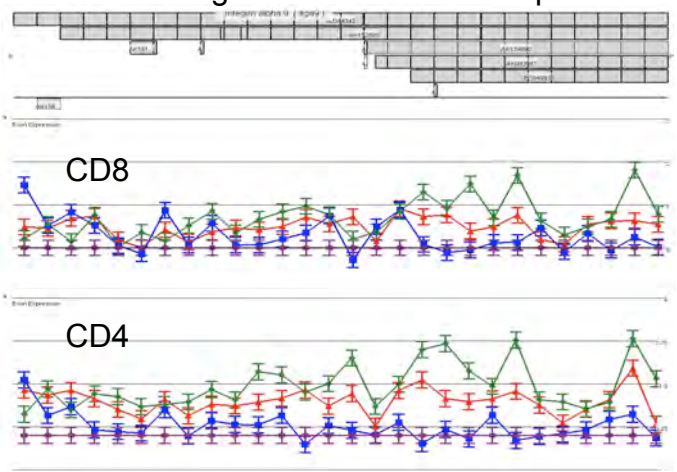
Hook1 exons and transcripts



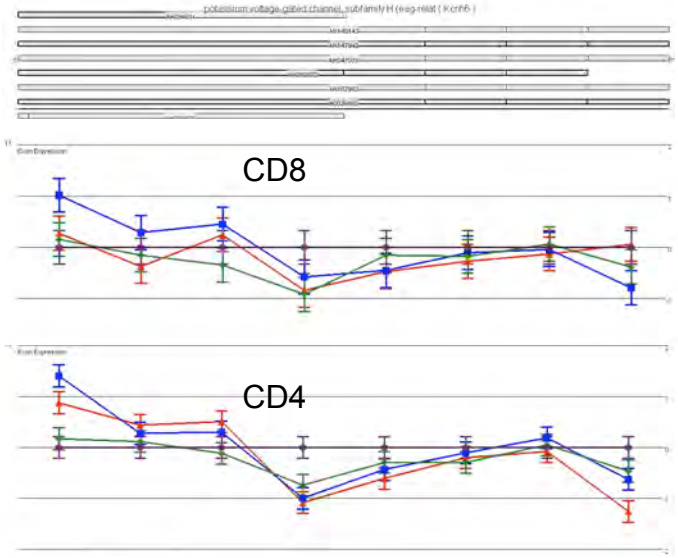
Hmgn3 exons and transcripts



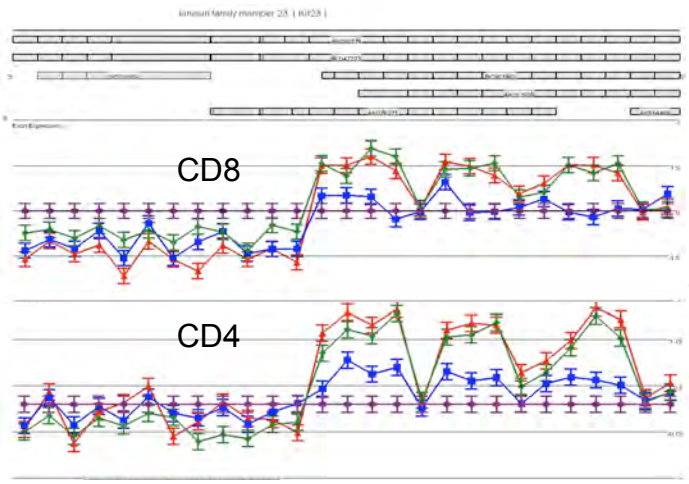
Itga9 exons and transcripts



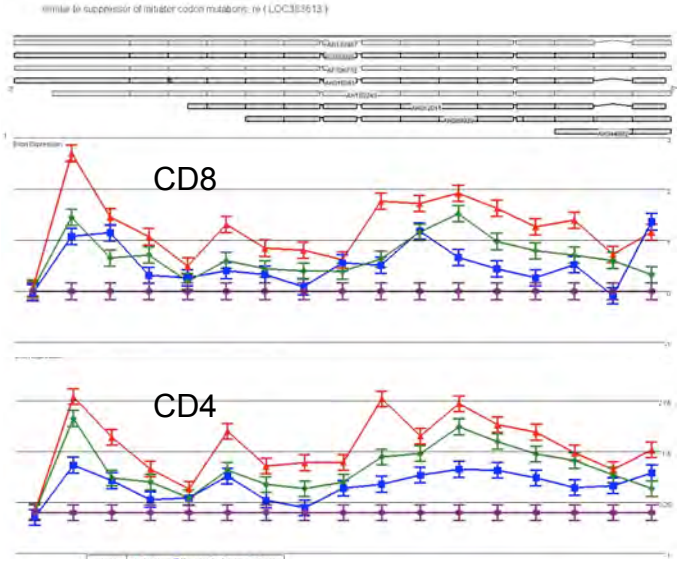
Kcnh6 exons and transcripts



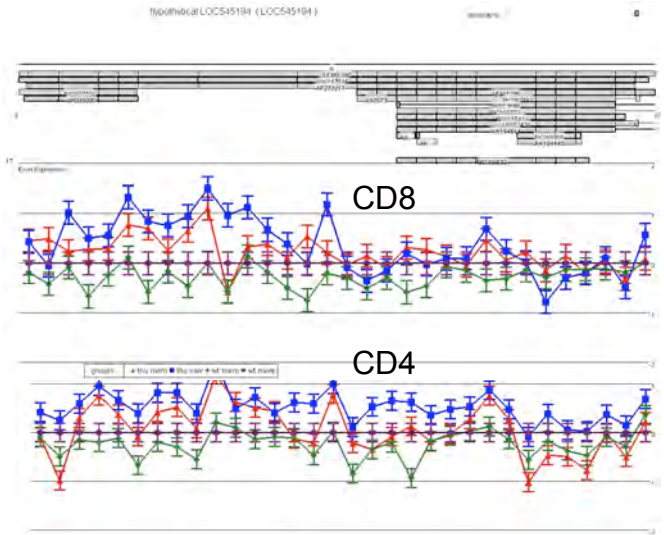
Kif23 exons and transcripts



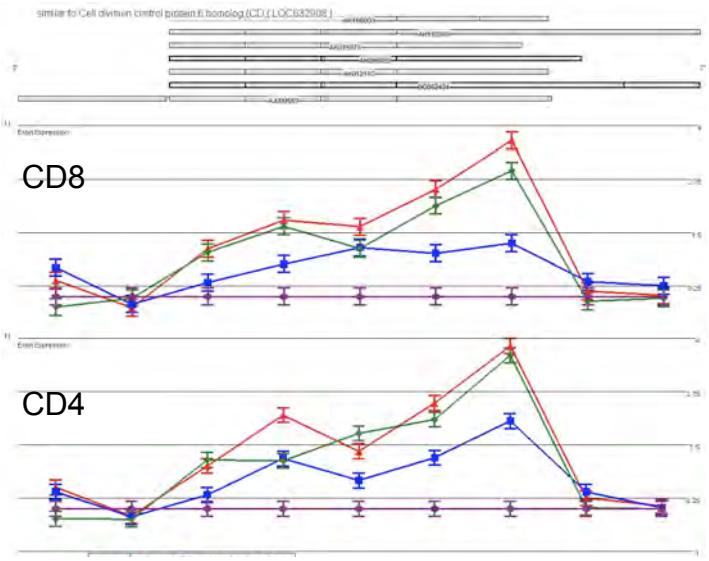
LOC383613 exons and transcripts



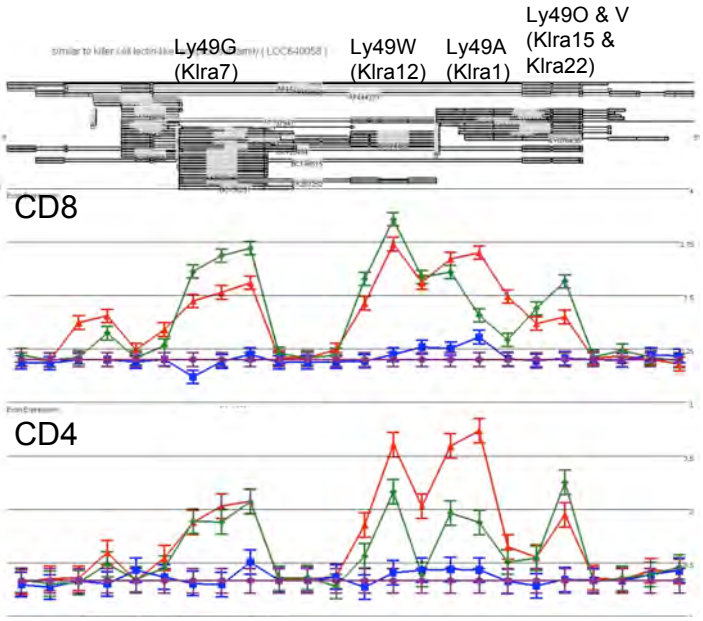
LOC545194 exons and transcripts



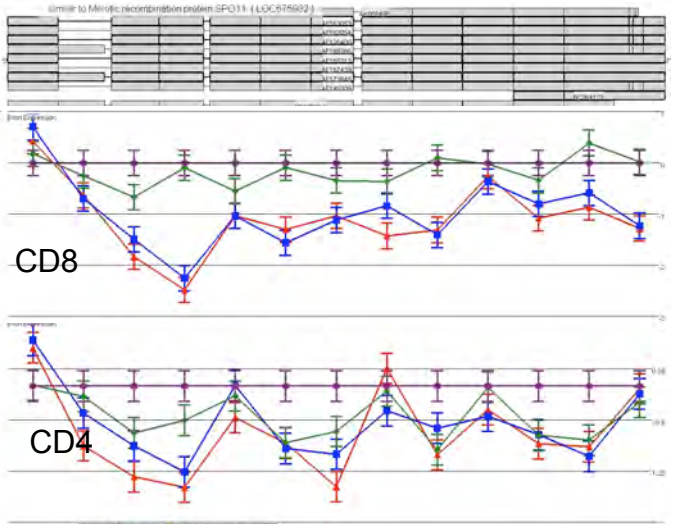
LOC632908 exons and transcripts



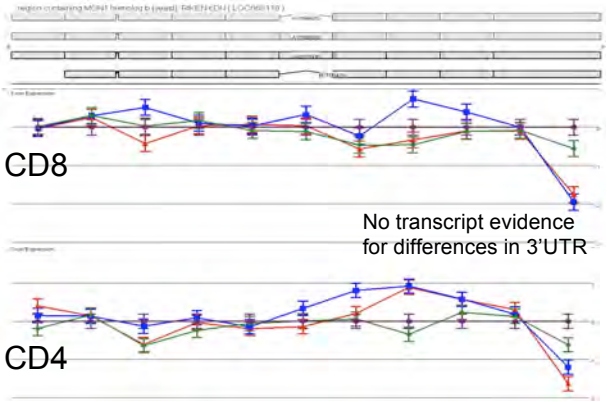
LOC640058 = Ly49 locus = Klra cluster



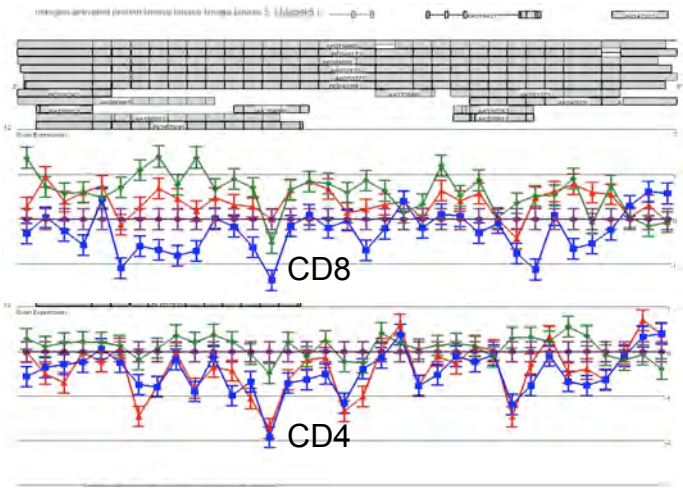
LOC675982 exons and transcripts



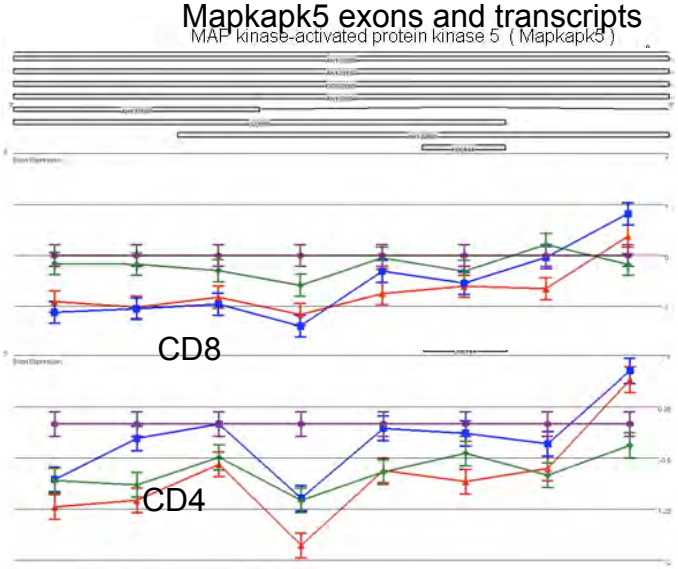
LOC668110 = 4930481F22Rik



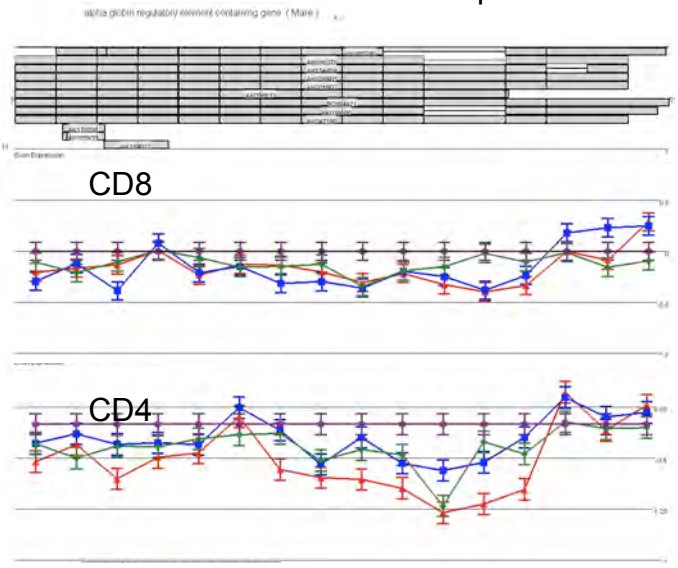
Map4k5 exons and transcripts



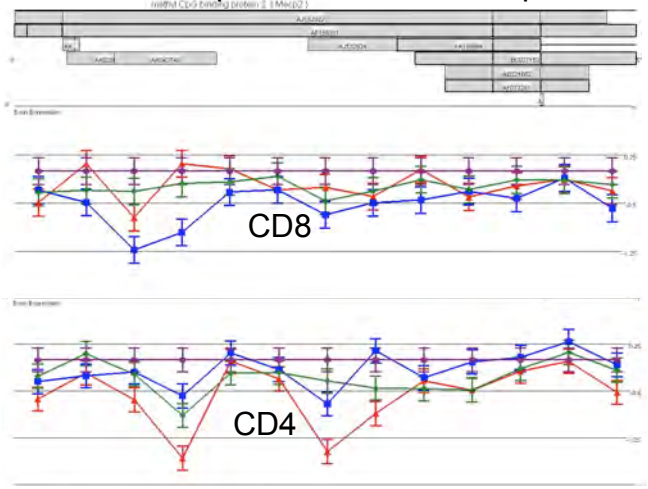
Mapkapk5 exons and transcripts



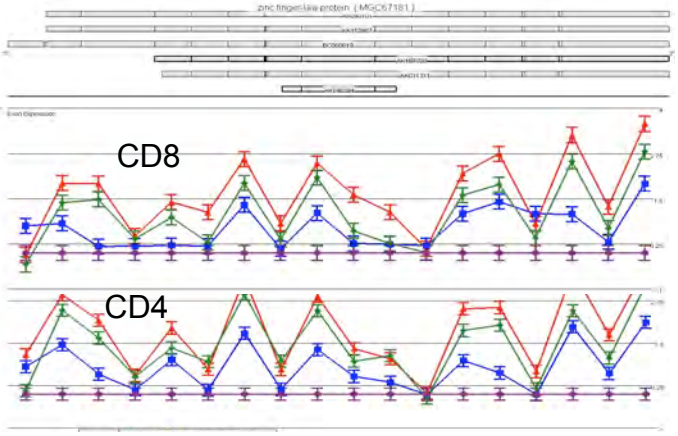
Mare exons and transcripts



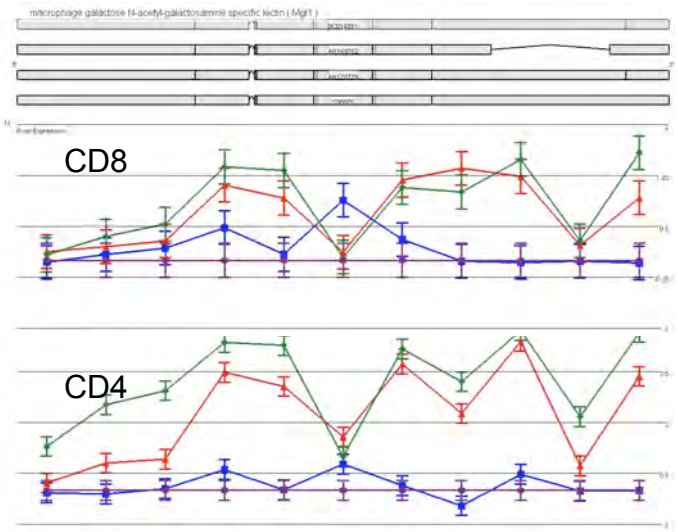
Mecp2 exons and transcripts



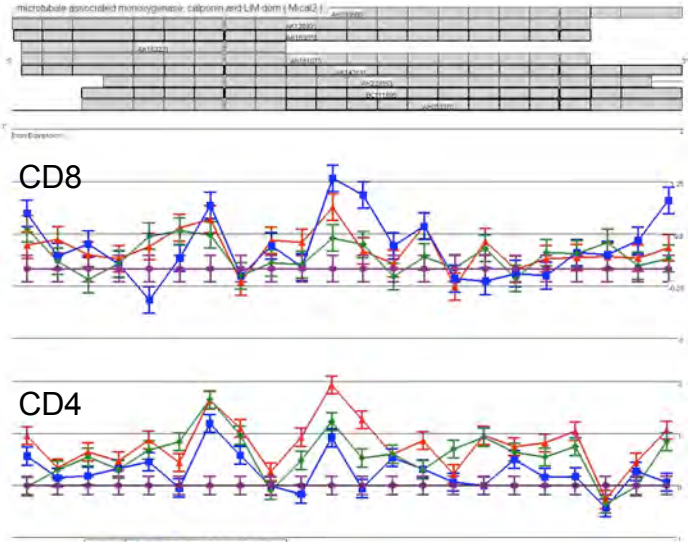
MGC67181 exons and transcripts



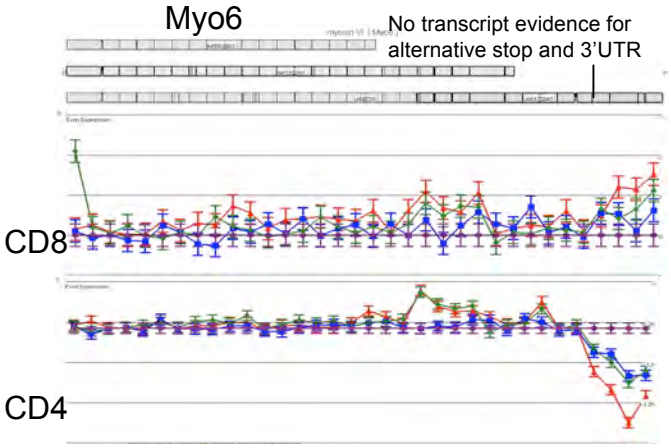
Mgl1 exons and transcripts



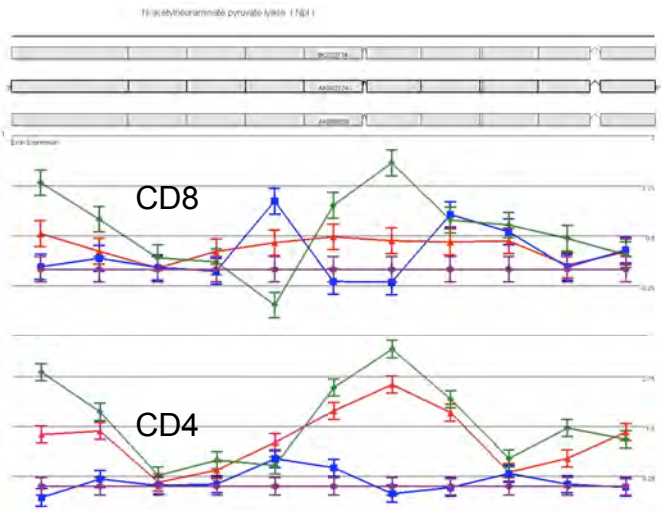
Mical2 exons and transcripts



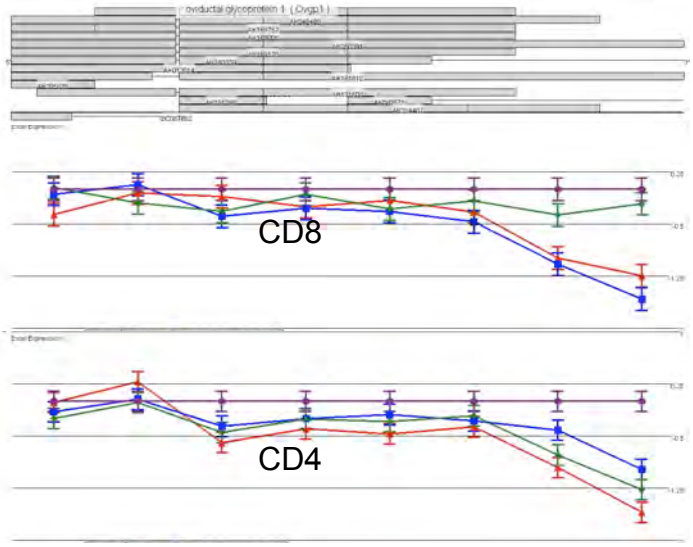
Myo6



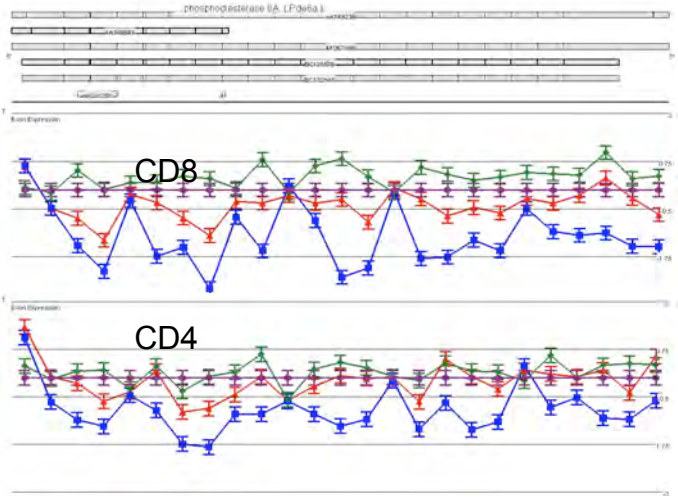
Npl exons and transcripts



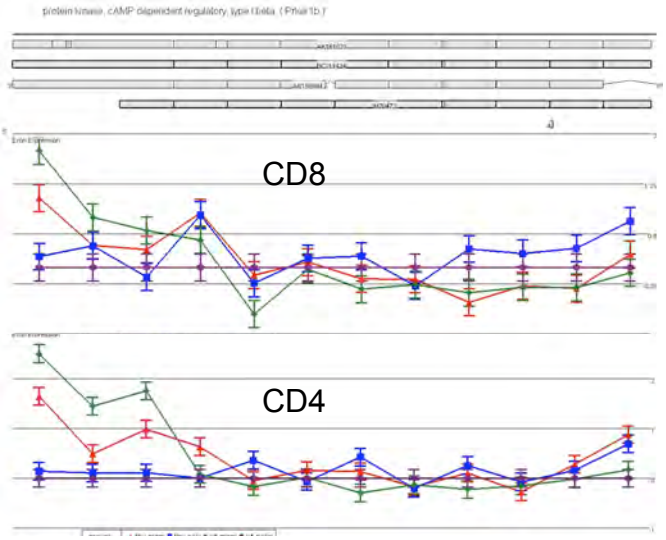
Ovgp1 exons and transcripts



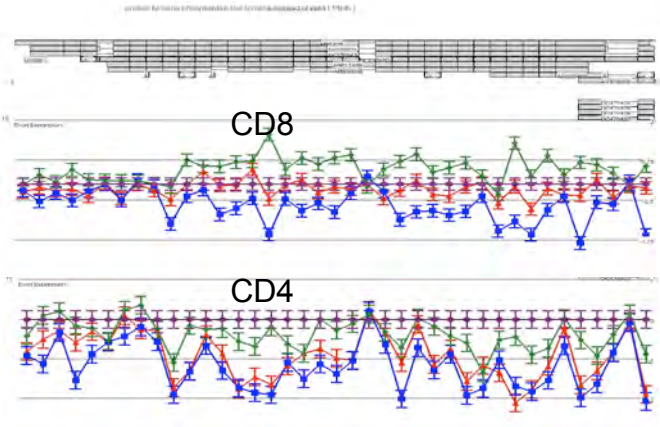
Pde8a exons and transcripts



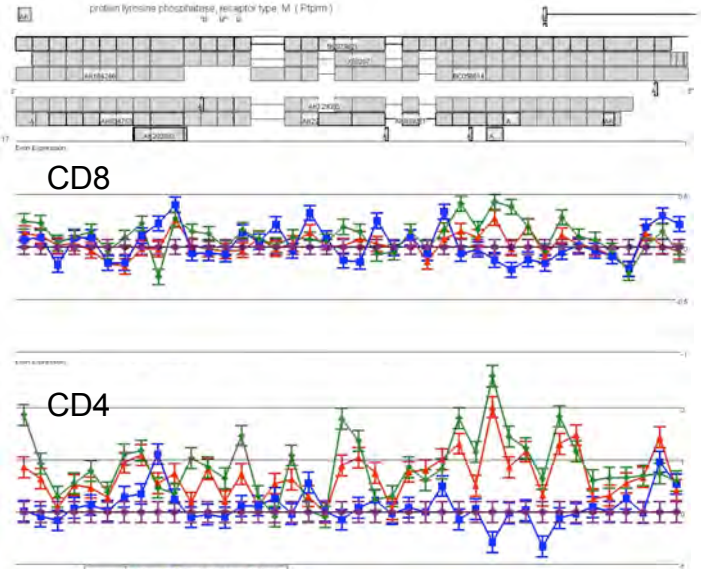
Prkar1b exons and transcripts



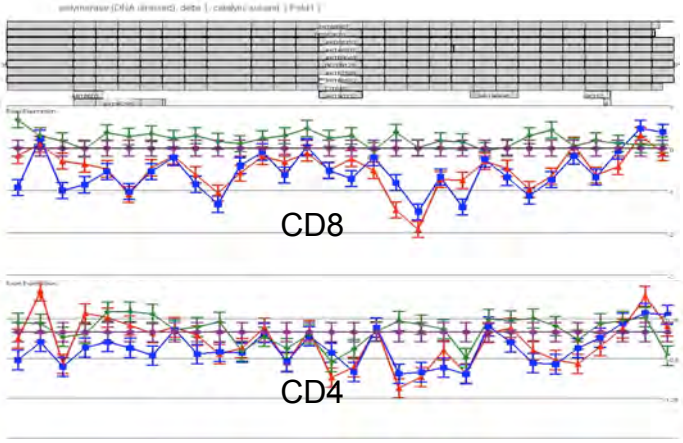
Ptplb exons and transcripts



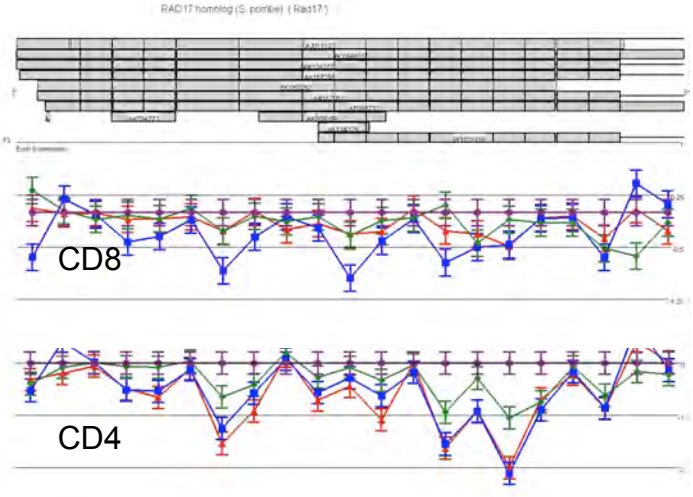
Ptprm exons and transcripts



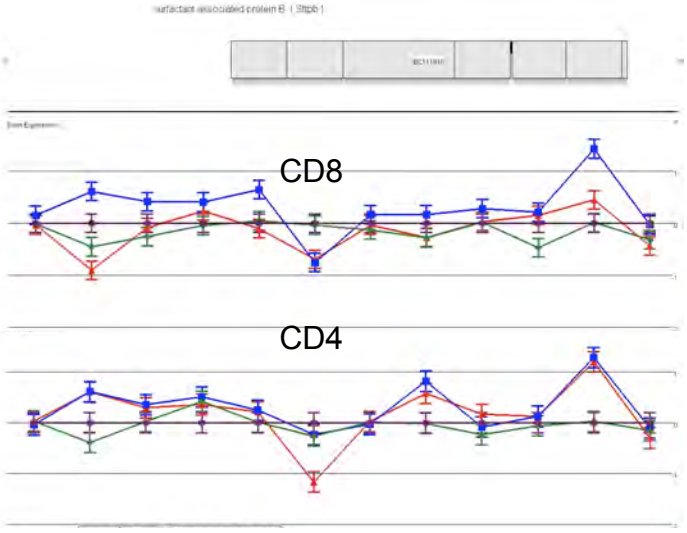
Pold1 exons and transcripts



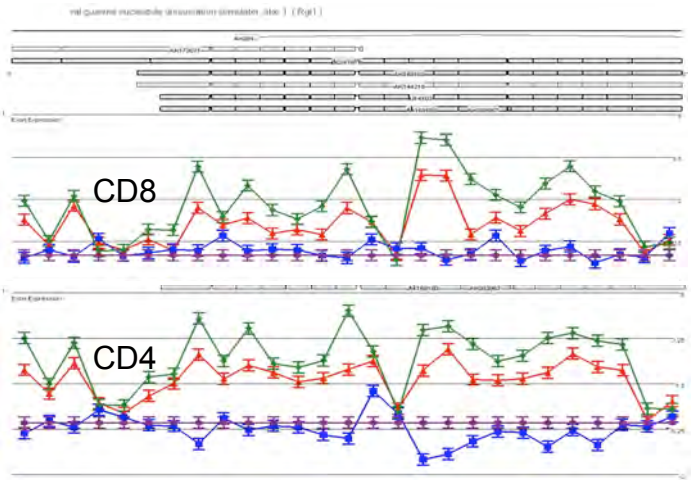
Rad17 exons and transcripts



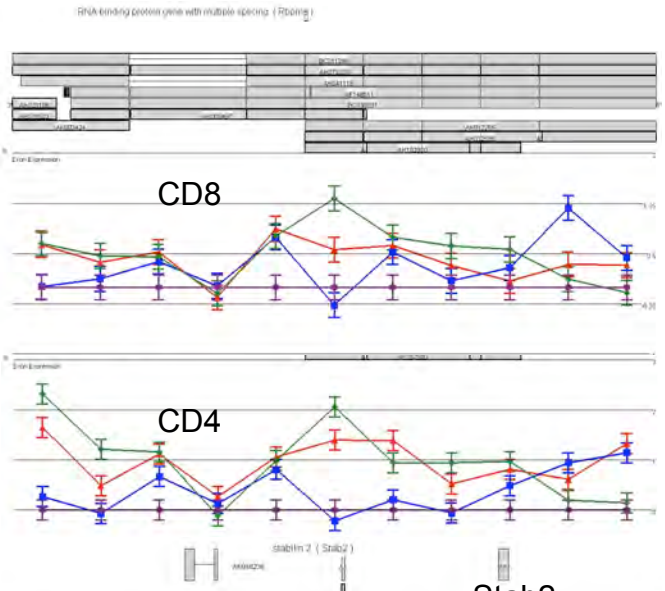
Sftpb exons and transcripts



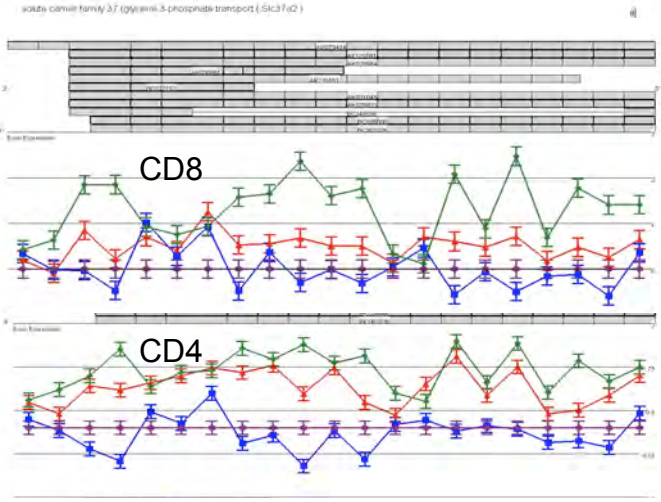
Rgl1 exons and transcripts



Rbpms exons and transcripts

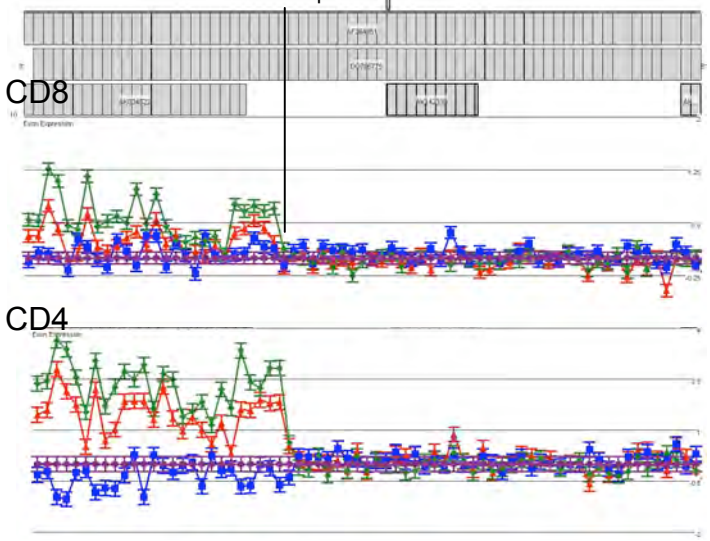


Slc37a2 exons and transcripts

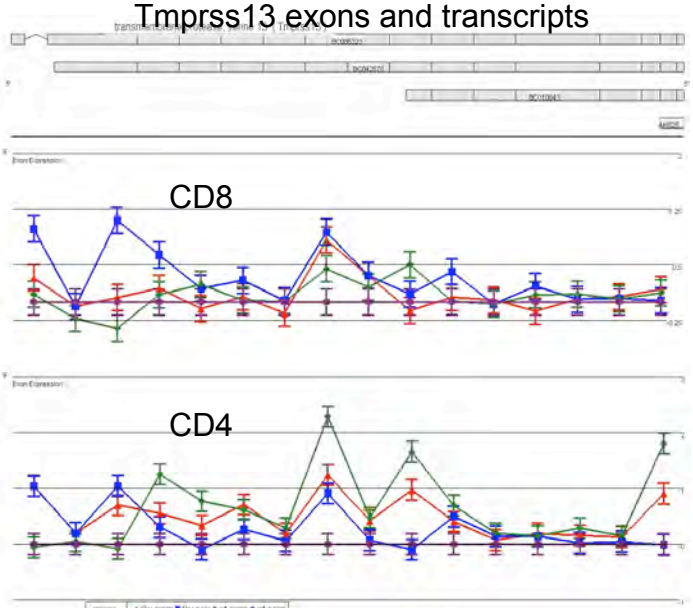


No transcript evidence for alternative start or stop, however few ESTs and transcripts in dbase

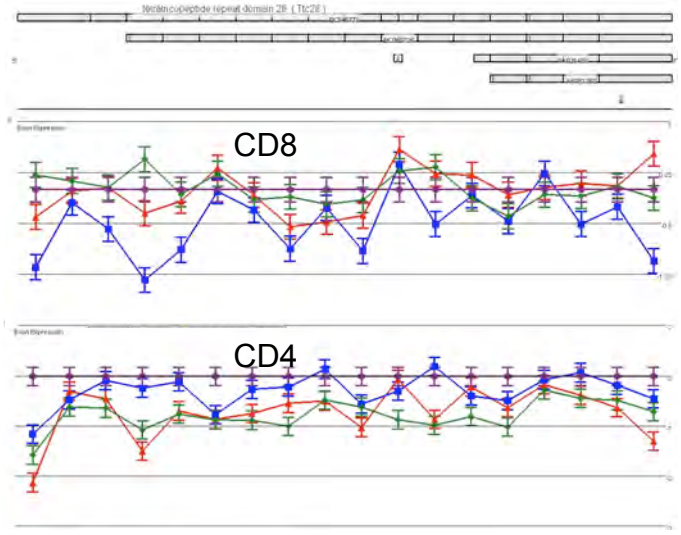
Stab2



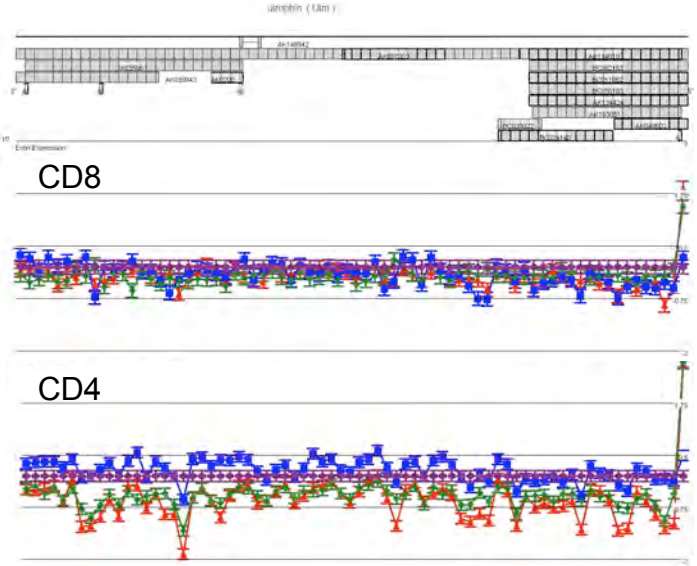
Tmprss13 exons and transcripts



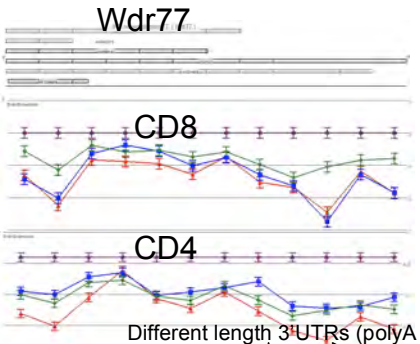
Ttc28 exons and transcripts



Utrn exons and transcripts



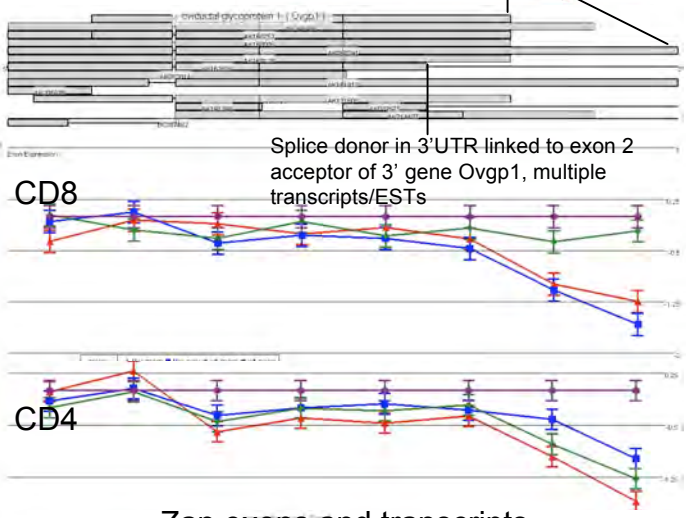
Wdr77



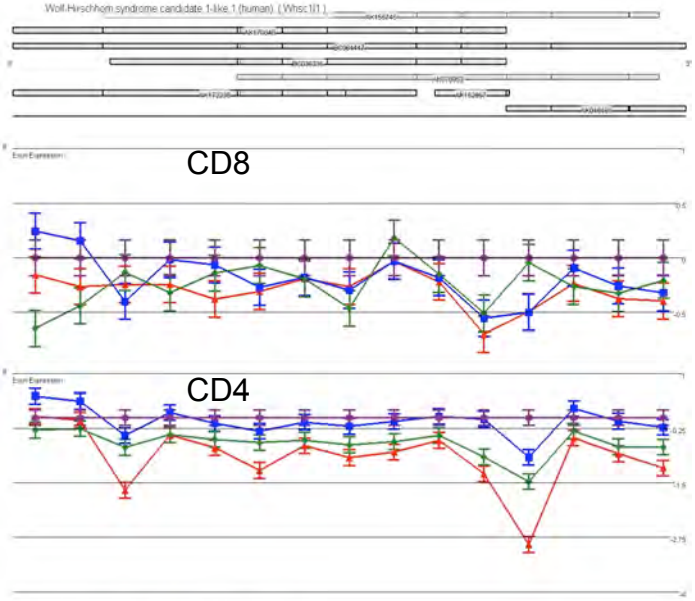
Wdr77

Different length 3'UTRs (polyA choice?

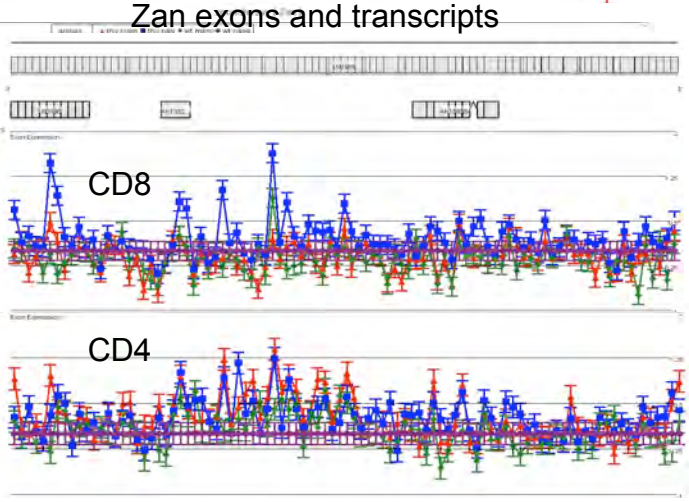
Splice donor in 3'UTR linked to exon 2 acceptor of 3' gene Ovgp1, multiple transcripts/ESTs



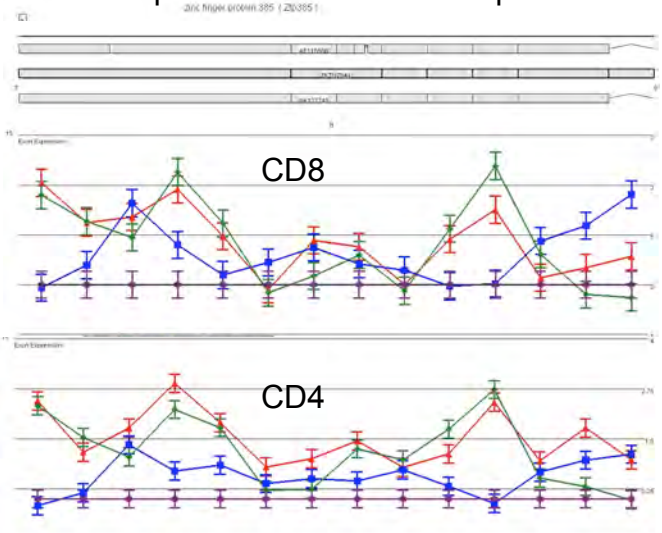
Whsc1l1 exons and transcripts



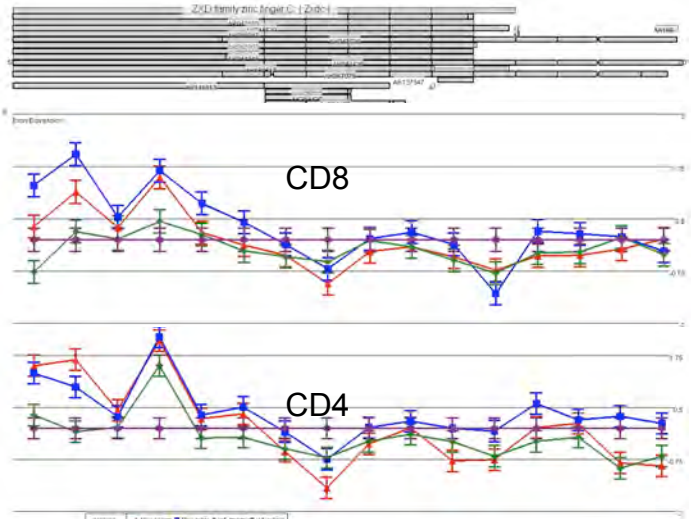
Zan exons and transcripts



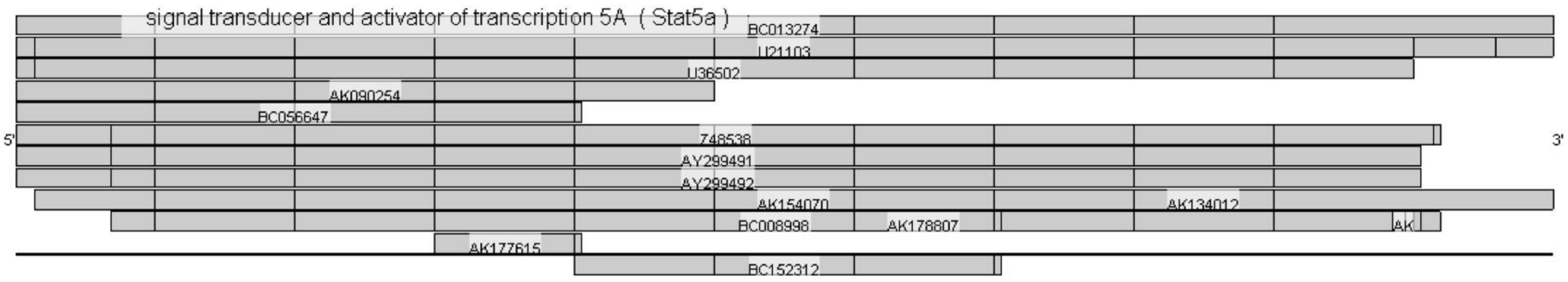
Zfp385 exons and transcripts



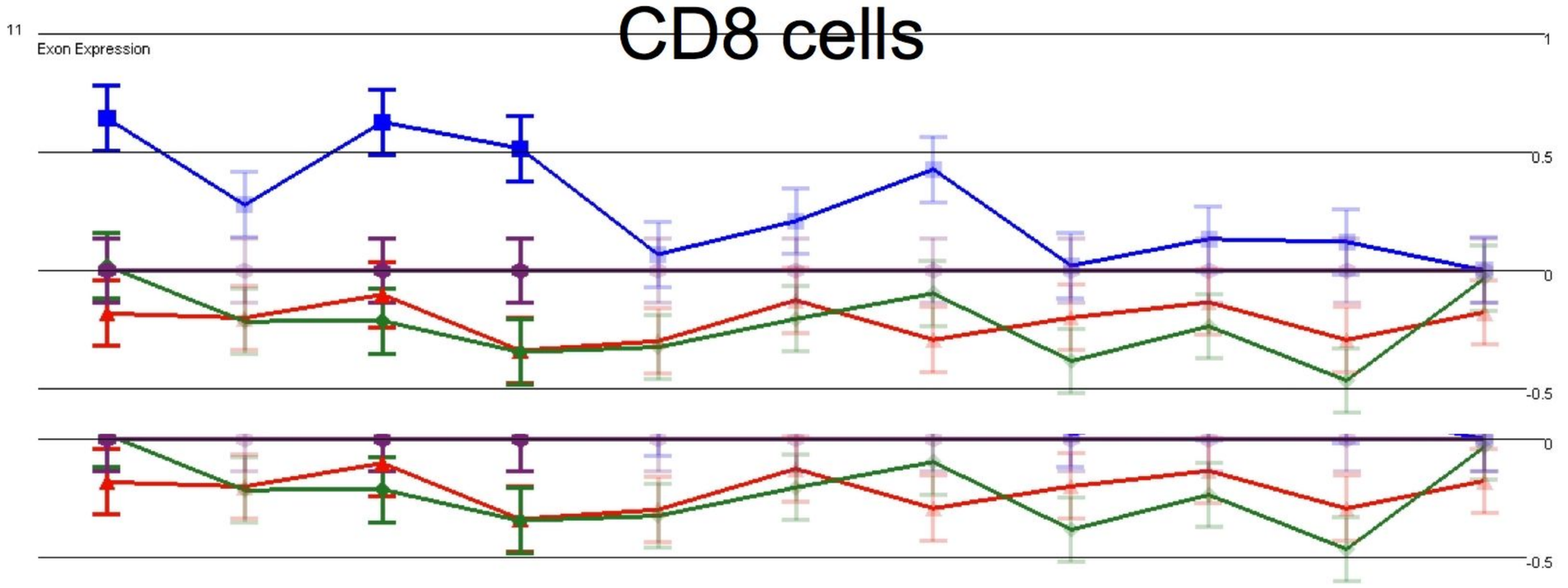
Zxdc exons and transcripts



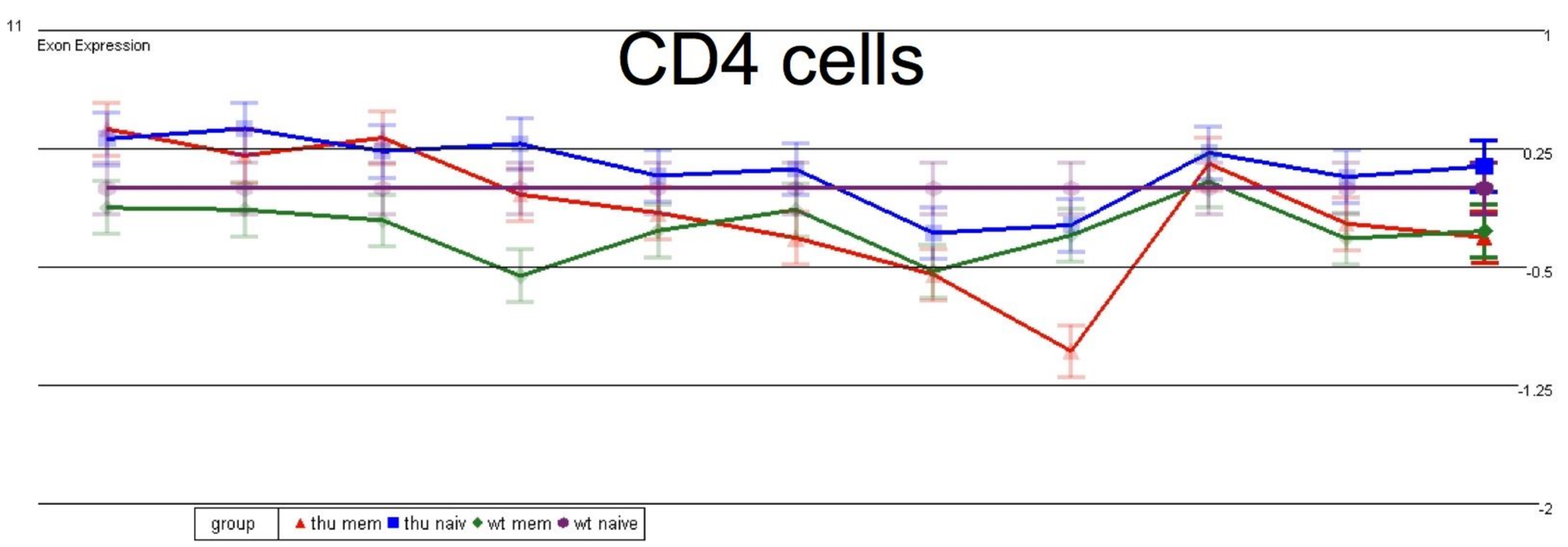
Stat5a



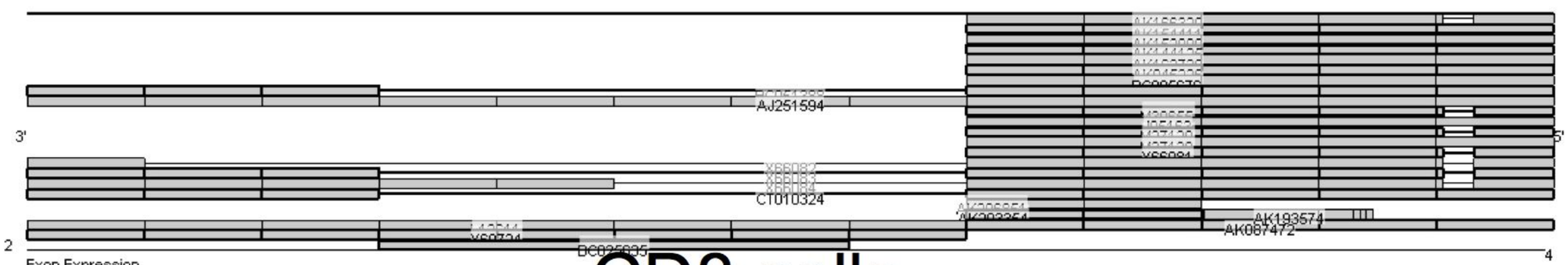
CD8 cells



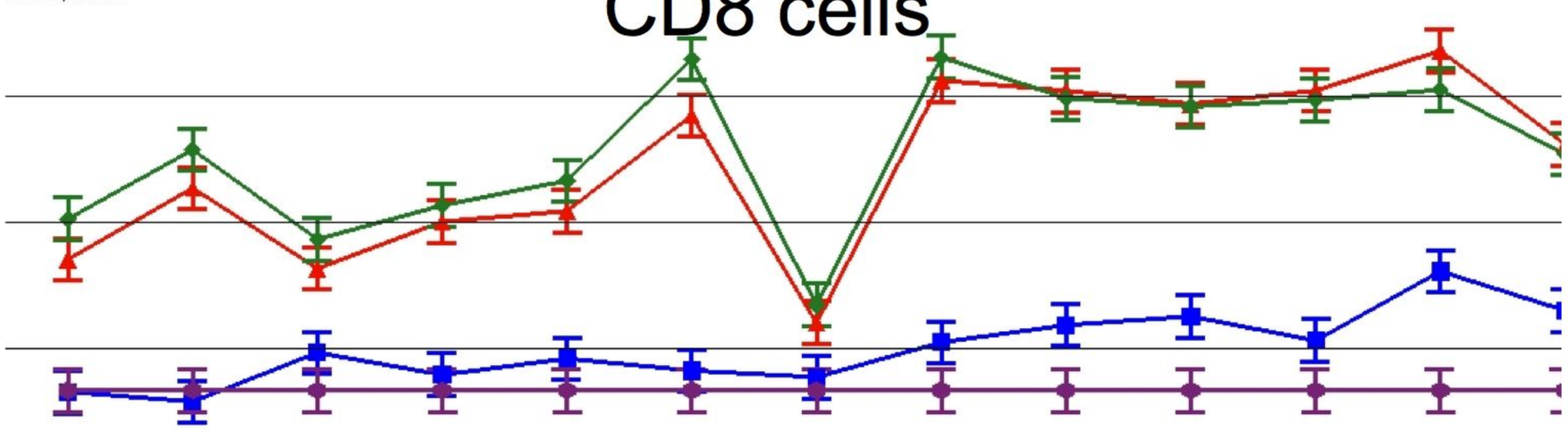
CD4 cells



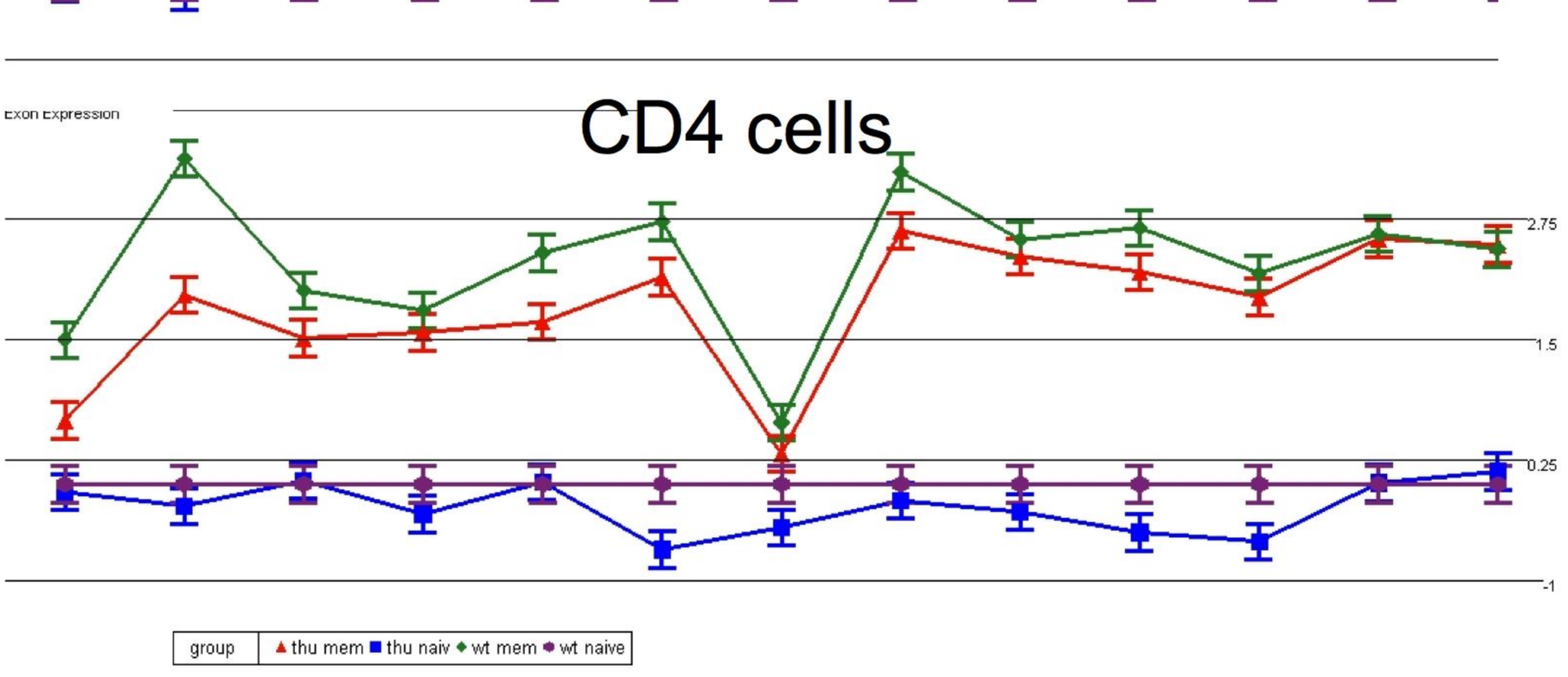
CD44



CD8 cells



CD4 cells



Cell type comparison	ANOVA P value (probeset x genotype)	FDR by step-up method (Benjamini & Hochberg, 1995)
naïve CD4 thu vs wt	0.001	0.012
naïve CD4 thu vs wt	0.010	0.061
naïve CD4 thu vs wt	0.050	0.189
naïve CD8 thu vs wt	0.001	0.029
naïve CD8 thu vs wt	0.010	0.117
naïve CD8 thu vs wt	0.050	0.306
memory CD4 thu vs wt	0.001	0.048
memory CD4 thu vs wt	0.010	0.162
memory CD4 thu vs wt	0.050	0.369
memory CD8 thu vs wt	0.001	0.059
memory CD8 thu vs wt	0.010	0.208
memory CD8 thu vs wt	0.050	0.463
wt naïve v memory CD4	0.001	0.036
wt naïve v memory CD4	0.010	0.224
wt naïve v memory CD4	0.050	0.693
wt naïve v memory CD8	0.001	0.026
wt naïve v memory CD8	0.010	0.182
wt naïve v memory CD8	0.050	0.607

Wu Supp Table 1: False Discovery Rates

Gene Symbol	Protein Name	Function	Neural Splicing Published	Conserved in human	No cell types p<0.05	No cell types p<0.001	p value wt v thu memory CD4	P value wt vs thu naïve CD4	p value wt vs thu memory CD8	p value wt vs thu naïve CD8	p value wt vs thu thymus CD8SP	p value wt vs thu thymus CD4SP	P-value WT naïve v WT memory CD4	P-value WT naïve v WT memory CD8
***Ptprc* (CD45)**	protein tyrosine phosphatase receptor C (CD45)	TCR signalling	no	yes	5	5	6.30E-06	1.87E-21	0.000626	0.416078	0.00075274	6.23E-13	0.261298	0.996406
Leng8	leukocyte receptor cluster (LRC) member 8	signalling?	no	?	4	3	0.006587	0.0002871	3.60E-07	0.00033574	0.954812	0.97432	0.0495943	0.94133
Mobp	myelin-associated oligodendrocytic basic protein	myeilin sheath	yes	yes	4	3	2.52E-05	0.0006332	0.0002979	0.00579835	0.951833	0.555057	0.145585	0.500468
Smn1	survival motor neuron 1	splicing factor	yes	yes	4	3	0.000365	1.56E-05	5.49E-05	0.00533134	0.890278	0.700891	0.545616	0.967199
Kif21a	kinesin family member 21A	intracellular transport	yes	yes	4	3	3.01E-09	3.94E-07	0.0149672	1.28E-05	0.991992	0.823922	0.975008	0.487182
Grip1	glutamate receptor interacting protein 1	intracellular adaptor & targetting, adhesion, neural dendrites	yes	yes	4	3	3.91E-10	0.0374348	3.67E-05	3.91E-08	0.95244	0.960712	0.987967	0.277761
Itsn1	intersectin 1 (SH3 domain protein 1A)	endocytosis & signalling	yes	yes	4	2	0.0225853	5.55E-05	0.0089652	2.70E-05	0.982168	0.962889	9.58E-25	3.31E-13
Ptprt	protein tyrosine phosphatase receptor rho	Homophilic cell-cell signalling	yes	yes	4	2	0.0017147	1.43E-06	7.65E-05	0.0400525	0.981391	0.99595	5.33E-14	0
Unk	unkempt homolog, C3H zinc finger protein 5	unknown	no	?	4	2	0.0039706	0.0001061	0.0038239	4.92E-05	0.998236	0.992372	0.0009566	0.988875
Pigw	phosphatidylinositol glycan anchor biosynthesis W	GPI synthesis, membrane protein maturation	no	?	4	2	0.0089198	0.0001904	2.38E-05	0.00333339	0.926295	0.949688	0.015284	0.755938
Gabbr1	gamma-aminobutyric acid (GABA-B) receptor	GABA signalling	yes	yes	4	2	0.0033755	0.003652	4.72E-05	5.23E-07	0.954084	0.941142	0.022304	0.0503995
Rps6kb2	S6Kb, ribosomal protein S6 kinase beta	intracellular signalling	?	?	4	2	0.0050055	5.66E-05	0.0145281	3.69E-06	0.999531	0.973144	0.0809972	0.992271
Cadps2	Ca2+-dependent activator protein for secretion 2	vesicle transport and targetting, calcium response	yes	yes	4	2	0.000229	0.0021705	0.0086284	9.41E-10	0.402832	0.987836	0.111572	0.244762
Otof	otoferlin	calcium signalling, hearing	yes	yes	4	2	0.0001678	0.0038077	0.0182671	1.45E-05	0.999529	0.999868	0.344466	0.920943
Cobl	cordon-bleu	actin cytoskeleton, neurite growth & branching	yes	yes	4	2	3.24E-05	2.10E-06	0.0270493	0.00220533	0.566075	0.986604	0.399575	0.886559
Foxd1	forkhead box D1, BF-2, brain factor 2	transcription, neural development	no	no	4	2	0.0018806	0.0008678	0.0602315	0.0149609	0.0005307	0.591378	0.448742	0.563712
Ttc23	tetratricopeptide repeat domain 23	unknown	no	?	4	2	0.0004533	0.0086577	0.0002727	0.00461283	0.924113	0.131236	0.64848	0.857539
***Sorbs2*)**	ArgBP2, Arg-binding protein 2, sorbin and SH3 domain cor	Abl/Arg signalling, cytoskeleton, neurite growth	yes	yes	4	2	0.0007947	0.0103388	0.0250927	0.00060531	0.999802	0.998977	0.665895	0.97655
Mlx	MAX-like protein X , BigMAX	Myc/MAX transcription	yes	yes	4	2	0.0001335	0.0005111	0.0016486	0.00196302	0.899736	0.370709	0.782533	0.938605
Ube1l	ubiquitin-activating enzyme E1-like	ISGylation, interferon response	no	?	4	2	0.0099546	4.32E-05	0.0141134	2.07E-06	0.997997	0.998421	0.967224	0.961086
Osbpl1a	oxysterol binding protein-like 1A	endocytosis & lipid metabolism	yes	yes	4	1	0.0142533	0.0001024	0.0425186	0.00422772	0.999987	0.991067	5.68E-09	0.0006192
Pctk1	PCTAIRE-motif protein kinase 1	neurite growth, vesicle transport	yes	?	4	1	0.0174091	0.0106006	0.0140768	0.00027502	0.999094	0.890876	0.0286842	0.644743
Scube3	signal peptide, CUB domain, EGF-like 3	unknown	no	?	4	1	0.0065946	0.0011077	0.0268514	0.00027474	0.858606	0.952558	0.0553291	0.0033521
Slc38a11	solute carrier family 38, member 11	putative transporter	no	?	4	1	0.0004687	0.0228921	0.0278889	0.0484205	0.844192	0.916058	0.0630483	0.0430996
Cpsf4	cleavage and polyadenylation specific factor 4 (30kD)	poly A site recognition, transcriptional pausing & cleavage, splic	no	yes	4	1	0.0022052	0.019108	8.67E-05	0.0446332	0.998448	0.911022	0.0635906	0.789204
Nphp4	nephronophthisis 4	primary cilium, transport & cytoskeleton	?	yes	4	1	0.001545	4.49E-06	0.0114074	0.00477619	0.681373	0.997214	0.108076	0.477469
Dync1i1	dynein cytoplasmic 1 intermediate chain 1	intracellular transport	yes	yes	4	1	0.0413449	7.71E-08	0.0032208	0.0331637	0.275046	0.999718	0.169231	0.882967
Ptprk	protein tyrosine phosphatase receptor type K	Homophilic cell-cell signalling	yes	yes	3	3	0.0004069	0.0530596	1.54E-05	1.11E-12	0.99994	0.990722	0.991378	2.07E-05
1520401A03Rik	unknown	unknown	no	?	3	2	0.0506364	0.0032833	0.0005668	0.00023328	0.970987	0.930584	8.16E-07	0.0739707
Kcnip2	KCHIP2, Kv channel-interacting protein 2	K channel & calcium signalling	yes	yes	3	2	0.0002644	0.0021719	0.350955	1.66E-08	0.853563	0.995556	0.0006754	0.84625
Nfatc1	nuclear factor of activated T-cells cytoplasmic 1	transcription, TCR signalling	?	yes	3	2	0.0071347	0.0001307	0.24282	0.000668	0.998839	0.999167	0.0007609	0.45094
Atp13a2	ATPase type 13A2	P5 family of ATPase - possibly calcium pump	yes	?	3	1	0.0761357	0.0094429	5.53E-05	0.00403596	0.986225	0.99991	7.90E-07	0.0002027
Rhbdf1	rhomboid family 1	TGFalpha/EGFR signalling, possible 7TM protease	no	?	3	1	0.0116824	4.59E-05	0.48761	0.00327903	0.981617	0.996603	2.05E-06	0.127328
Mus81	MUS81 endonuclease homolog	DNA damage repair & tolerance	no	?	3	1	0.0013879	0.0003322	0.489488	0.0146492	0.987895	0.772735	0.0005865	0.52523

Wu Supp Table 2:
Group A Genes

Gene Symbol	Protein Name	Function	Neural Splicing Published	Conserved in human	No cell types p<0.05	No cell types p<0.001	p value wt v thu memory CD4	P value wt vs thu naïve CD4	p value wt vs thu memory CD8	p value wt vs thu naïve CD8	p value wt vs thu thymus CD8SP	p value wt vs thu thymus CD4SP	P-value WT naïve v WT memory CD4	P-value WT naïve v WT memory CD8
Nlrp1a	Nalp1, NACHT leucine rich repeat and PYD containing 1	pathogen response, inflammation, apoptosis	no	yes	6	5	0.0333499	0.000103	4.22E-18	1.16E-05	3.42E-08	1.35E-21	0.994205	0.528491
Slc15a2	solute carrier family 15 (H+/peptide transporter)	di- and tri-peptide transport across membranes	no	?	4	4	2.70E-05	9.26E-06	8.30E-14	4.96E-10	0.491784	0.447847	4.83E-14	0.980473
Spnb2	spectrin beta 2	membrane cytoskeleton & cell shape	yes	yes	4	4	0.0006838	8.74E-15	4.35E-13	9.37E-06	0.931711	0.958482	1.36E-07	0.887075
Tle1	Grg1, Groucho-related gene 1, transducin-like enhancer of sp	transcription, TCF/LEF & RUNX corepressor, Wnt signalli	no	yes	4	4	0.000236	3.52E-13	5.09E-05	8.73E-14	0.270171	0.337821	0.01134	0.023389
Thnsl1	threonine synthase-like 1	metabolism? Phospho-lyase?	no	?	4	3	0.0198129	9.02E-06	1.43E-05	0.0001755	0.67467	0.715841	0.171314	0.697691
Zfp692	zinc finger protein 692	transcription, AMPK signalling, gluconeogenesis	no	?	4	3	0.0001167	0.000927	0.0195465	3.62E-07	0.998356	0.916524	0.358552	0.821561
Man2a2	mannosidase 2, alpha 2	N-linked glycan processing	no	?	4	3	0.0135135	0.000412	1.85E-06	0.0001982	0.926074	0.884787	0.670316	0.850888
Cd247	CD247, CD3 zeta, theta, eta & iota	TCR assembly, signalling, trafficking	no	no	4	2	0.0187033	0.008873	0.0001376	4.73E-07	0.985803	0.941899	3.38E-06	0.664266
Fermt3	Kindlin3	integrin adhesion, cytoskeleton, platelet activation	yes	yes	4	2	0.0016346	0.000557	1.04E-05	0.007918	0.629075	0.927073	0.000991	0.999769
Prpf4b	PRP4, pre-mRNA processing factor 4 homolog B	splicing, phosphorylation of non-SR splicing factors, trans	no	?	4	2	0.0017652	1.32E-07	0.0002783	0.0017653	0.99762	0.97813	0.0013	0.945416
Tpr	translocated promoter region	nuclear pore complex & nuclear architecture, nucleocytop	no	?	4	2	0.033201	2.05E-11	9.40E-05	0.0450269	0.999995	0.750315	0.002053	0.993357
Bcl2l11	Bim, BCL2-like 11	triggering of apoptosis	no	similar	4	2	1.52E-05	9.87E-05	0.0163195	0.0059476	0.995467	0.911252	0.004227	0.966022
Ccnl2	cyclin L2	splicing, neuronal plasticity & development	yes	yes	4	2	0.0090295	7.69E-09	1.80E-07	0.0151181	0.785663	0.340948	0.017629	0.975545
Ankra2	ankyrin repeat family A (RFXANK-like) 2	BK Ca channel and LDL receptor modulation, histone ace	no	?	4	2	0.0013483	0.006665	0.0005089	2.67E-06	0.992153	0.564046	0.037152	0.923175
Gtf3c2	general transcription factor IIIC polypeptide 2 beta	RNA Pol3 transcription	no	?	4	2	0.0065141	7.41E-05	0.001792	0.0003842	0.785439	0.993458	0.051885	0.657976
Pax6	paired box gene 6	transcription, eye & CNS development	yes	yes	4	2	0.0380728	0.225075	0.0143756	8.83E-05	0.000772	0.670191	0.054004	0.428765
Ufc1	USP21 ubiquitin-fold modifier conjugating enzyme 1	de-ubiquitination & de-NEDDylation	no	?	4	2	0.0048312	4.79E-10	0.0001928	0.0266233	0.994788	0.99796	0.188442	0.0598025
Rpain	RIP, RPA interacting protein	DNA replication, repair & recombination	no	yes	4	2	0.0330886	6.54E-05	0.0215877	0.161365	0.481821	0.000586	0.243155	0.941839
Map3k7	TAK1, Tgf-beta activated kinase 1	TCR, TLR, NFkB signaling	no	yes	4	2	0.000923	0.048793	0.0058786	0.0001172	0.720198	0.245233	0.243486	0.30049
Lrdd	PIDD, p53-induced death domain, leucine-rich and death dom	p53 induced apoptosis and NFkB induction	no	?	4	1	0.0018979	5.52E-05	0.0410679	0.0138404	0.996805	0.546922	0.003715	0.022405
Thrap3	Trap150, thyroid hormone receptor associated protein	RNA Pol2 transcription, possibly mRNA processing	no	?	4	1	0.0068828	0.000282	0.0013589	0.0273623	0.962685	0.996527	0.003987	0.958073
Tbxa2r	thromboxane A2 receptor	arachidonic acid signalling	no	?	4	1	0.0349245	0.001453	0.0006109	0.0088525	0.83153	0.373352	0.012264	0.749185
Rragc	Ras-related GTP binding C	nucleocytoplasmic transport	no	?	4	1	0.0110702	1.67E-05	0.0071533	0.0464435	0.997568	0.930173	0.028377	0.840267
Cnot10	Lin41, Trim71 tripartite motif-containing 71	cell fate determination, developmental timing	no	?	4	1	0.0098115	0.000123	0.001012	0.0016787	0.999963	0.830551	0.077588	0.242782
Gigyf2	Grb10 interacting GYF protein 2, trinucleotide repeat containir	IGF-Grb10 signalling?	no	?	4	1	0.0018973	5.42E-11	0.035495	0.0074257	0.999993	0.971116	0.087662	0.950211
Zfp598	zinc finger protein 598	unknown, possible ubiquitin E3 ligase (RING finger)	no	yes	4	1	0.0173252	0.041011	0.021954	0.000306	0.980152	0.747752	0.087871	0.0167374
Abcc5	ATP-binding cassette sub-family C (CFTR/MRP) member 5	ABC-type putative transporter, unknown function	yes	yes	4	1	0.0331469	0.00381	0.0173842	1.35E-05	0.999986	0.980666	0.111378	0.40764
Dido1	death inducer-obliterator 1	myeloproliferation control, centrosome function	no	yes	4	1	0.0076678	4.27E-06	0.0157789	0.0331595	0.997733	0.985032	0.127523	0.997798
Fbxw2	F box protein W2	SCF E3 ubiquitin ligase complex substrate recognition	no	?	4	1	0.0394073	0.000199	0.0070373	0.0035636	0.854372	0.788792	0.135513	0.834747
Mtap4	microtubule-associated protein 4	microtubules, cytoskeleton	yes	?	4	1	0.0052277	4.13E-05	0.0039474	0.0293176	0.999665	0.868981	0.199743	0.944898
Htf9c	HpaII tiny fragments locus 9c	possibly RNA splicing (RRM and RNA methylase domains	no	?	4	1	0.019489	0.000311	0.0472203	0.0038403	0.991432	0.977075	0.2261	0.315114
Tpd52l1	D53, tumor protein D52-like 1	vesicle traffic & exocytosis? ASK1-JNK signaled apoptosi	yes	yes	4	1	0.0285687	0.001222	0.0032334	1.37E-05	0.317225	0.706502	0.248273	0.990948
Kif23	kinesin family member 23	motor protein for cytokinesis and chromosome spindle	no	yes	3	3	0.72759	1.02E-08	0.0002938	1.33E-06	0.872051	0.919268	7.15E-38	8.06E-34
4932412H11Rik	unknown	unknown, leucine rich repeat protein	no	?	3	3	0.667096	5.64E-13	9.91E-06	3.65E-05	0.997287	0.999961	0.003934	0.803059
Myo7a	myosin VIIa	cytoskeleton, motor protein, hearing & retina	yes	yes	3	2	0.0009409	0.492642	0.0022582	0.0002608	0.999262	0.989901	1.32E-08	0.63375
Mid1	midline 1	unknown	no	?	3	1	0.0091484	7.63E-05	0.325107	0.0034906	0.94791	0.516562	1.16E-05	0.566906
Madd	MAP-kinase activating death domain, DENN, 1G20	apoptosis, sensitivity to TRAIL & TNF, NFkB signalling	yes	yes	3	1	0.045361	0.000738	0.0326003	0.153639	0.999907	0.999903	1.90E-05	0.994131
Efr3a	similar to cyclin-dependent kinase 6	unknown	no	?	3	1	0.0096951	1.96E-05	0.277575	0.0266789	0.943638	0.811493	2.61E-05	0.979653
Fam48a	P38 interacting protein	NIK-p38 signalling, mesoderm development, gastrulation	no	?	3	1	0.007641	5.40E-06	0.003058	0.344594	0.82577	0.990673	0.0002	0.989506
Nisch	nischarin, imidazoline receptor antisera-selected protein	integrin-cytoskeleton links, endocytosis, imadazoline rece	yes	yes	3	1	0.0971994	8.97E-06	0.0039309	0.015207	0.998755	0.99982	0.000367	0.900738
Wdr42a	WD repeat domain 42A	unknown	no	yes	3	1	0.153213	4.47E-08	0.014842	0.0441146	0.976652	0.548838	0.000581	0.760396
Tcf7l2	TCF4, transcription factor 7-like 2 T-cell specific HMG-box	transcription, WNT & Pax6 signalling, brain development,	yes	yes	3	0	0.263496	0.009493	0.0098694	0.0011678	0.97697	0.898989	6.46E-12	1.15E-16
Aprt	adenine phosphoribosyl transferase	nucleotide synthesis & metabolism	no	?	2	2	0.14511	8.62E-05	0.179658	7.20E-11	0.895476	0.993468	3.21E-07	0.0015497
Ogdh	oxoglutarate dehydrogenase (lipoamide)	Krebs cycle metabolism	no	?	2	2	0.0001072	1.53E-05	0.908345	0.362993	0.999707	0.999837	1.30E-06	0.999369
Abhd11	abhydrolase domain containing 11	unknown	no	?	2	2	0.160038	8.54E-08	0.162781	2.48E-05	0.999824	0.729185	1.98E-05	0.492322

Wu Supp Table 3:
Group B Genes

Gene Symbol	Protein Name	# Probe Sets	Transcript ID	No cell types p<0.05	No cell types p<0.001	p value wt v thu memory CD4	P value wt vs thu naïve CD4	p value wt vs thu memory CD8	p value wt vs thu naïve CD8	p value wt vs thu thymus CD8SP	p value wt vs thu thymus CD4SP	P-value WT naïve v WT memory CD4	P-value WT naïve v WT memory CD8
Acp2	acid phosphatase 2, lysosomal	13	6879034	6	5	3.55E-09	1.15E-07	7.44E-10	9.66E-08	0.037614	1.39E-07	0.00026	1.41E-08
2310044H10Rik	unknown	11	6966869	6	3	0.00531	8.42E-07	1.18E-10	3.30E-12	0.001317	0.006016	0.668461	0.877916
Coch	coagulation factor C homolog	13	6795033	4	4	4.46E-05	6.73E-05	0.000295	3.69E-05	0.38176	0.882229	8.98E-07	0.0048621
Satb1	special AT-rich sequence binding protein 1	11	6855981	4	4	0.00046	5.10E-05	0.000279	2.18E-05	0.997319	0.776257	0.001173	0.0514125
Adam33	a disintegrin and metallopeptidase domain 33	26	6890980	4	4	0.00052	1.84E-08	0.00087	8.33E-05	0.999976	0.485605	0.187924	0.659598
Col16a1	procollagen, type XVI, alpha 1	50	6917365	4	4	0.00042	1.77E-05	5.14E-05	5.02E-06	0.998109	0.997894	0.380785	0.971232
Pla2g4e	phospholipase A2, group IVE	21	6890290	4	4	5.15E-07	0.000255	0.000283	0.000122	0.930915	0.95658	0.857535	0.112063
Evl	Ena-vasodilator stimulated phosphoprotein	16	6797917	4	3	0.02451	7.58E-06	8.76E-06	1.09E-07	0.99999	0.84087	0.01164	0.997472
Ttc26	tetratricopeptide repeat domain 26	21	6945550	4	3	8.26E-05	0.000318	5.47E-06	0.026613	0.9057	0.995678	0.026539	0.882796
Bcl7a	B-cell CLL/lymphoma 7A	16	6934226	4	3	0.00914	0.00061	1.71E-06	7.62E-07	0.990514	0.995458	0.069755	0.932057
Pdcd11	programmed cell death protein 11	44	6870130	4	3	1.08E-05	0.000704	4.36E-07	0.018202	0.880127	0.84953	0.110201	0.997336
Tchp	trichoplein, keratin filament binding	15	6933613	4	3	0.0001	0.000278	0.022294	1.31E-07	0.87233	0.999661	0.174392	0.847046
Fgb	fibrinogen, B beta polypeptide	12	6906386	4	3	3.14E-05	0.004115	2.66E-06	1.23E-07	0.999072	0.293546	0.812086	0.0944539
Plce1	phospholipase C, epsilon 1	43	6869570	4	3	0.00097	1.12E-05	9.20E-05	0.011306	0.999779	0.994721	0.954537	0.0003911
Cacna1b	calcium channel, voltage-dependent, N type, alpha 1B sub	51	6885345	4	2	0.03412	5.60E-05	0.001715	3.84E-09	0.999988	0.999994	0.035463	0.0031955
Rhbdd3	rhomboid domain containing 3	14	6778432	4	2	0.03503	2.83E-05	0.003285	6.39E-06	0.902392	0.887371	0.060173	0.22381
Txnrd3	thioredoxin reductase 3	23	6947894	4	2	0.02002	4.80E-09	0.037229	5.50E-06	0.801513	0.989678	0.119097	0.740731
Otud7a	OTU domain containing 7A	23	6961049	4	2	0.00037	0.025129	0.000322	0.004862	0.906636	0.717403	0.131647	0.955086
Hydin	hydrocephalus inducing	101	6979096	4	2	0.01877	0.000177	0.036053	8.21E-07	0.999999	0.997499	0.161698	0.305037
Heatr6	HEAT repeat containing 6, cDNA 2700008B19Rik	25	6783045	4	2	0.00479	1.59E-07	0.02108	0.000172	0.99996	0.687526	0.177058	0.424668
Fras1	Fraser syndrome 1 homolog (human)	80	6932601	4	2	0.00399	2.71E-05	0.016222	0.000458	1	0.997925	0.181219	0.171349
Abca4	ATP-binding cassette, sub-family A (ABC1), member 4	59	6900982	4	2	0.00841	0.003197	8.66E-05	0.000799	0.999808	0.999723	0.198347	0.663792
Lama1	laminin alpha 1	73	6851848	4	2	0.00035	0.002056	0.026727	1.30E-06	0.999877	0.984322	0.258986	0.229054
BC025575	unknown, cDNA BC025575	15	6791569	4	2	0.00094	0.001147	0.022877	3.42E-05	0.813427	0.278902	0.336114	0.221047
Ptprg	protein tyrosine phosphatase receptor type G	36	6816951	4	2	0.01845	0.00014	0.00014	0.003224	0.998949	1	0.417795	0.0011789
Abca8a	ATP-binding cassette, sub-family A (ABC1) member 8a	39	6792114	4	2	0.001	0.029078	0.010504	1.35E-07	0.999942	0.610321	0.451863	0.990562
Uvrag	UV radiation resistance associated gene	19	6969690	4	2	0.00817	0.000438	0.000229	0.029848	0.933951	0.805425	0.453869	0.246073
Cnksr2	connector enhancer of kinase suppressor of Ras 2	25	7020425	4	2	0.00042	0.004937	0.001809	4.08E-09	0.052771	0.483997	0.616792	0.991069
Itgb4	integrin beta 4	41	6785183	4	2	0.00055	0.006062	0.004291	8.72E-05	0.961729	0.990076	0.739942	0.942278
Klhl11	kelch-like 11	6	6791427	4	2	8.64E-05	0.018402	2.26E-05	0.029778	0.502179	0.368681	0.925406	0.954212
Dmgdh	dimethylglycine dehydrogenase precursor	19	6809022	4	2	1.30E-06	0.022918	0.017235	0.000254	0.841052	0.925352	0.978268	0.258026
Ttpa	tocopherol (alpha) transfer protein	13	6912040	4	2	0.00054	0.000114	0.025288	0.001478	0.336247	0.745782	0.985509	0.992006
Fgf8	fibroblast growth factor 8	8	6873363	4	2	0.00118	0.000878	0.010628	0.000817	0.474431	0.600826	0.995704	0.554256
Vill	villin-like	23	6992866	4	1	0.00458	0.00014	0.027079	0.028022	0.997946	0.984554	6.47E-06	4.31E-05
Hcn1	hyperpolarization-activated, cyclic nucleotide-gated K+ 1	14	6810563	4	1	0.02225	8.69E-07	0.027549	0.001808	0.460408	0.967329	0.005775	0.704039
Evc2	Ellis van Creveld syndrome 2 homolog, sim to reproductive hom	21	6930023	4	1	0.00607	0.000397	0.044179	0.020977	0.99573	0.989122	0.041486	0.568218
Tgfa	transforming growth factor alpha	8	6947596	4	1	0.00462	0.019451	0.000729	0.004116	0.952421	0.993197	0.070742	0.309718
Col4a1	procollagen, type IV, alpha 1	50	6980364	4	1	7.14E-06	0.028994	0.007242	0.007377	0.940386	0.999949	0.073734	0.92711
Gldn	gliomedin	12	6989195	4	1	0.00077	0.025964	0.011502	0.021624	0.849549	0.507773	0.080347	0.369239
Galnt5	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylga	14	6877356	4	1	0.00492	0.021661	0.000833	0.003911	0.764995	0.106831	0.094669	0.0032335
Brunol6	bruno-like 6, RNA binding protein	13	6989554	4	1	0.03677	0.018444	0.030173	0.000206	0.990822	0.950369	0.197385	0.645284
Mall	mal, T-cell differentiation protein-like	7	6890712	4	1	0.01888	0.017812	0.017758	1.45E-06	0.883629	0.670205	0.200693	0.190117
Itga6	integrin alpha 6	25	6878031	3	3	1.23E-09	4.49E-11	0.05897	6.98E-05	0.999015	0.815633	4.93E-20	0.82559
Adamtsl4	ADAMTS-like 4	26	6907202	3	3	0.00098	1.86E-06	0.313182	1.58E-05	0.724192	0.79873	0.010361	0.0916315
Smarca4	SWI/SNF related matrix associated actin dependent regulator of	34	6987396	3	3	1.19E-06	0.000281	0.07361	0.000142	0.999017	0.999994	0.027492	0.761547
Notch4	Notch4, LOC674800	42	6850025	3	3	0.00042	1.28E-06	0.217374	0.000328	0.998328	0.999082	0.033234	0.0046917
Gucy2e	guanylate cyclase 2e	20	6789269	3	3	0.87107	4.54E-05	0.000746	0.000615	0.941905	0.18046	0.852533	0.0006216
Sept8	septin 8	14	6780983	3	2	0.08131	0.001559	0.000122	1.04E-07	0.731703	0.765659	9.75E-12	0.251945
Ptprv	protein tyrosine phosphatase, receptor type, V	42	6762395	3	2	0.00318	0.00032	0.125057	5.72E-05	0.574698	0.983431	3.71E-09	0.805486
Aff4	AF4/FMR2 family member 4	24	6780963	3	2	3.78E-05	1.05E-06	0.067882	0.004936	0.988461	0.747193	0.000588	0.604988
Asx12	additional sex combs like 2, polycomb group protein	17	6792935	3	2	0.00027	2.73E-06	0.111457	0.009564	0.980718	0.987215	0.000706	0.903749
Ryr1	ryanodine receptor 1, skeletal muscle	104	6966183	3	1	0.02471	1.16E-05	0.710008	0.004171	0.999665	0.999976	6.59E-44	1.00E-33
Rhbdl1	rhomboid, veinlet-like 1	20	6854447	3	1	0.37629	0.000471	0.043908	0.033522	0.999341	0.999992	3.21E-05	0.862986
Grhl2	grainyhead-like 2	19	6829871	3	1	0.02357	0.340328	1.72E-05	0.018335	0.683245	0.97185	0.000214	0.0007302
Arhgef10	Rho guanine nucleotide exchange factor (GEF) 10	31	6974254	3	0	0.0304	0.3885	0.657	0.0469	0.0487	0.371327	6.78E-14	0.022115
Aspm	asp (abnormal spindle)-like microcephaly associated	36	6753592	2	2	0.1429	#####	0.1104	0.0002	1	0.77895	9.66E-14	4.87E-13
Fads3	fatty acid desaturase 3	18	6868018	2	2	0.367	0.0002	0.2809	#####	0.8155	0.730998	8.51E-06	0.510816
Slc6a13	solute carrier family 6 (neurotransmitter transporter, GABA)	17	6949591	2	2	0.66326	0.000319	0.405947	5.49E-05	0.767896	0.98972	2.17E-05	0.0090303

Wu Supp Table 4: Group C Genes

Gene Symbol	Gene Description	# Probe Sets	Transcript ID	No cell types p<0.05	No cell types p<0.001	p value wt v thu memory CD4	P value wt vs thu naïve CD4	p value wt vs thu memory CD8	p value wt vs thu naïve CD8	p value wt vs thu thymus CD8SP	p value wt vs thu thymus CD4SP	P-value WT naïve v WT memory CD4	P-value WT naïve v WT memory CD8
Thada	thyroid adenoma associated	16	6857799	6	5	2.74E-12	1.80E-20	1.99E-13	2.88E-07	0.031018	9.71E-06	0.000108	0.396876
Gtf2a1	general transcription factor IIa precursor	6	6802744	6	3	0.000139	2.01E-07	0.000735	0.001687	0.024167	0.002873	0.424674	0.0856947
Wdtc1	WD and tetratricopeptide repeats 1	17	6925885	5	3	0.00251	6.59E-06	9.97E-06	0.000148	0.964049	0.0134	0.228673	0.372625
Hyi	hydroxypyruvate isomerase homolog	7	6916705	5	1	0.018832	0.000135	0.002988	0.006195	0.241555	0.024535	0.050161	0.332321
Col6a3	alpha 3 type VI collagen isoform 1 precursor	33	6760642	4	4	0.000637	7.45E-07	5.95E-07	4.70E-06	0.999974	0.99853	0.08005	0.16909
Cdc42bpb	Cdc42 binding protein kinase beta	42	6803821	4	3	0.002345	1.94E-05	1.22E-10	2.96E-06	0.632693	0.778043	3.80E-12	4.53E-32
Gemin5	gem (nuclear organelle) associated protein 5	35	6788563	4	3	0.328907	0.007471	3.30E-05	0.257472	0.000262	4.73E-05	0.007367	0.997164
Tgn	thyroglobulin	47	6831162	4	3	0.010252	0.000225	0.000154	3.61E-07	0.999974	0.819483	0.019818	0.0039451
Apaf1	apoptotic peptidase activating factor 1	31	6775977	4	3	0.000183	0.009543	0.000668	0.000436	0.960658	0.451155	0.633855	0.254134
Zfp593	zinc finger protein 593	5	6926011	4	2	0.040332	0.000355	9.62E-05	0.012223	0.176683	0.823282	0.006894	0.140325
Sspo	SCO-spondin	111	6946002	4	2	0.002922	8.49E-06	0.002457	3.03E-09	0.960717	0.999944	0.027791	0.621996
Lztr1	leucine-zipper-like transcriptional regulator 1	22	6839746	4	2	0.019525	1.04E-08	0.003098	0.000339	0.800564	0.698674	0.051847	0.73952
Eif2ak3	eukaryotic translation initiation factor 2 alpha kinase	20	6946894	4	2	0.003643	0.000737	0.147458	0.000737	0.027537	0.4341	0.120515	0.549851
2700097O09Rik	similar to translocase of outer mitochondrial membrane	9	6800912	4	2	0.004485	0.000363	0.000604	0.00344	0.95469	0.923812	0.238637	0.921178
Pip5k3	phosphatidylinositol-3-phosphate phosphatidylinositol 5-kinase 3	36	6750086	4	2	6.59E-06	2.56E-10	0.038541	0.005043	0.977783	0.419111	0.293821	0.816606
Zfp619	zinc finger protein 619, cDNA 3000002G13Rik	21	6959982	4	2	0.03107	0.033634	4.28E-11	4.78E-05	0.904491	0.078013	0.392625	0.616247
Tsga10	testis specific 10	20	6757981	4	2	9.42E-07	0.031585	0.001538	0.000322	0.999993	0.774362	0.805501	0.389058
C130038G02Rik	cDNA C130038G02Rik	10	6935630	4	2	0.000433	0.130217	0.006991	4.98E-08	0.989332	0.010323	0.982525	0.443985
Rbbp6	retinoblastoma binding protein 6	5	6964053	4	1	0.01015	1.09E-05	0.014633	0.014496	0.803824	0.923236	0.008043	0.998185
Dhx16	DEAH (Asp-Glu-Ala-His) box polypeptide 16	22	6850200	4	1	0.021933	0.00088	0.026731	0.003585	0.98944	0.979834	0.032356	0.919382
Gpr98	G protein coupled receptor 98, LOC672575	109	6814656	4	1	0.004228	0.010336	0.011774	1.12E-08	0.938951	0.591597	0.034144	0.106507
6330407J23Rik	cDNA 6330407J23Rik	14	6766862	4	1	0.037255	0.035789	0.003704	1.44E-06	0.718736	0.727404	0.054818	0.86556
Chuk	conserved helix-loop-helix ubiquitous kinase	22	6873249	4	1	0.003638	0.000269	0.015138	0.008379	0.985745	0.857927	0.183847	0.962477
Ppm1e	protein phosphatase 1E (PP2C domain containing)	17	6790547	4	0	0.004088	0.003577	0.0131	0.05816	0.016306	0.991338	0.000155	0.506241
Ugcg	UDP-glucose ceramide glucosyltransferase	12	6913876	3	2	1.14E-07	1.07E-06	0.984149	0.441943	0.035818	0.442723	3.56E-08	0.856078
Nf1	neurofibromatosis 1	18	6782777	3	2	0.000173	0.008598	0.81788	0.00015	0.999496	0.999366	6.66E-08	0.697022
Stt3b	STT3, subunit of the oligosaccharyltransferase complex	28	6999196	3	2	4.97E-06	5.40E-08	0.035675	0.395544	0.99966	0.953227	8.78E-05	0.394808
Vapa	vesicle-associated membrane protein associated protein A	10	6856674	3	2	0.000704	4.56E-07	0.019151	0.073088	0.886404	0.958838	0.000836	0.223607
Adcy7	Adenylate cyclase 7, cDNA 5730445M16Rik	32	6977975	3	1	0.030736	9.76E-12	0.007086	0.111921	0.999871	0.97381	2.24E-06	0.858573
Tnks	tankyrase, TRF1-interacting ankyrin-related ADP-ribose poly	29	6981664	3	1	0.00688	6.79E-11	0.002795	0.100459	0.523076	0.909521	0.000121	0.453867
Nxf1	nuclear RNA export factor 1 homolog	20	6867957	3	1	0.008888	6.20E-15	0.999984	0.997694	0.999976	0.029662	0.000478	0.999953
Enpp2	ectonucleotide pyrophosphatase/phosphodiesterase 2	28	6835759	2	2	0.084918	0.000192	0.596624	3.09E-10	0.10694	0.982837	1.02E-08	8.95E-05
Zfp592	zinc finger protein 592	16	6962043	2	2	0.000136	0.000293	0.050179	0.696043	0.999901	0.994832	1.44E-06	0.711474

Wu Supp Table 5:
Group D Genes

Gene Symbol	Protein name	Confirmed alt promoter by transcript dbase?	# Probe Sets	Transcript ID	No cell types p<0.05	No cell types p<0.001	p value wt v thu memory CD4	P value wt vs thu naïve CD4	p value wt vs thu memory CD8	p value wt vs thu naïve CD8	p value wt vs thu thymus CD8SP	p value wt vs thu thymus CD4SP	P-value WT naïve v WT memory CD4	P-value WT naïve v WT memory CD8
Rxfp2	relaxin/insulin-like family peptide receptor 2		20	6935732	5	3	0.0001876	0.0175927	6.30E-10	3.04E-06	0.976794	0.012318	0.503963	0.172189
Cyc1	cytochrome c-1		11	6831651	5	3	2.40E-06	7.67E-06	2.98E-05	0.00101392	0.0134932	0.415121	0.970021	0.999567
Polr3f	polymerase (RNA) III (DNA directed) polypeptide F		13	6881834	5	0	0.0308398	0.0011177	0.014061	0.0280623	0.834537	0.0149122	0.41785	0.710731
Csf1r	colony stimulating factor 1 receptor		28	6861358	4	4	1.22E-05	1.73E-08	4.10E-07	4.19E-05	0.998956	0.998607	4.72E-14	2.56E-40
Ddx42	DEAD (Asp-Glu-Ala-Asp) box polypeptide 42		19	6784605	4	4	0.0002757	2.63E-07	1.66E-05	4.91E-05	0.819968	0.810998	0.00022393	0.234056
Mthfsd	methenyltetrahydrofolate synthetase domain containing	yes	16	6985859	4	4	1.14E-06	9.03E-05	3.11E-05	7.30E-06	0.996842	0.993185	0.0156708	0.883064
B3gnt7	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7		9	6751271	4	4	0.0003739	0.0007317	0.0005443	0.00073231	0.921624	0.966651	0.317943	0.802308
Slc25a28	Mitoferrin, solute carrier family 25, member 28	yes	9	6873230	4	4	0.0004378	7.20E-05	5.08E-05	5.44E-08	0.583722	0.790922	0.476029	0.357451
Rapgef4	Rap guanine nucleotide exchange factor (GEF) 4	yes	32	6878038	4	3	1.05E-06	2.91E-08	6.86E-07	0.00588248	0.999935	0.851507	5.15E-30	0.0045142
Gabrr1	gamma-aminobutyric acid (GABA-C) receptor, subunit rho 1	yes	11	6912520	4	3	5.97E-05	3.83E-05	0.0101164	0.00012187	0.770133	0.888659	2.83E-13	0.0908446
Plcg1	phospholipase C, gamma 1	yes	38	6882880	4	3	3.40E-08	3.76E-06	0.0066863	0.00022206	0.994722	0.992793	1.88E-06	0.562914
Polr3d	polymerase (RNA) III (DNA directed) polypeptide D		13	6825696	4	3	0.0127043	0.0002309	0.0004395	0.00064851	0.974835	0.990505	2.18E-05	0.425928
Sfrs11	splicing factor, arginine/serine-rich 11	yes	13	6910948	4	3	1.01E-05	2.60E-08	1.03E-08	0.00131578	0.956719	0.862144	0.00040848	0.873607
Pcdhac1	protocadherin alpha subfamily C, 1		120	6860138	4	3	0.0037683	2.53E-05	0.0003161	2.25E-08	1	0.999958	0.00118589	0.423601
Raver2	ribonucleoprotein, PTB-binding 2		15	6915818	4	3	2.85E-05	0.0002585	0.0356753	2.70E-10	0.656187	0.806031	0.00187054	0.220527
Baz1a	bromodomain adjacent to zinc finger domain 1A		12	6800906	4	3	0.0020681	9.79E-06	0.0007142	0.00069192	0.963184	0.667374	0.00613519	0.343038
	XM_994297 // LOC671134 // similar to melanoma antigen // XM_98171		15	6929366	4	3	7.00E-06	8.55E-07	0.0005184	0.00740371	0.991377	0.491434	0.0111994	0.961717
Ccdc41	4921537D05Rik		20	6769934	4	3	1.19E-06	1.83E-07	0.0034376	0.0007082	0.997725	0.856656	0.0133556	0.66045
Ercc5	excision repair cross-complementing rodent repair deficien		23	6749108	4	3	0.0045689	5.86E-08	9.36E-05	0.00090151	0.957616	0.0621897	0.070508	0.993315
Trpc3	transient receptor potential cation channel, subfamily C,		15	6904303	4	3	0.0006637	1.01E-06	0.0219834	0.00044582	0.716438	0.584469	0.134451	0.548268
	XM_912301 // LOC626061 // similar to SGT1, suppressor of G2 allele of SKP1		11	6820573	4	3	0.0002042	5.22E-05	0.0146618	1.97E-06	0.99977	0.984377	0.136415	0.98898
Uckl1	uridine-cytidine kinase 1-like 1		18	6894303	4	3	4.35E-06	0.0056647	5.36E-05	3.08E-05	0.998187	0.977702	0.14181	0.650657
Dsg2	LOC640723 desmoglein 2		20	6859285	4	3	2.67E-08	0.0022421	9.96E-07	1.62E-06	0.981682	0.119586	0.155582	0.13622
Gdap2	ganglioside-induced differentiation-associated-protein 2		13	6899921	4	3	1.04E-06	0.0009352	0.0060361	0.00066973	0.704882	0.242501	0.348842	0.901437
	NM_138654 // AF397014 // cDNA sequence AF397014		15	6811260	4	3	0.0005562	6.36E-06	0.0151055	2.26E-05	0.471636	0.475204	0.407451	0.824375
Phka2	phosphorylase kinase alpha 2		39	7014855	4	3	0.0017287	7.30E-05	1.55E-06	0.00082221	0.993944	0.960764	0.545825	0.0159969
Sdccag3	serologically defined colon cancer antigen 3		9	6885485	4	3	0.0001327	0.0121751	0.0001002	6.05E-05	0.828575	0.957572	0.598131	0.859609
Col8a2	procollagen, type VIII, alpha 2		20	6917153	4	3	0.0001387	0.0006588	0.0051641	0.00093243	0.78761	0.98962	0.696427	0.900618
	NM_181278 // B230219D22Rik		10	6807263	4	3	4.51E-06	0.0023015	0.0003045	1.02E-05	0.950613	0.120176	0.798172	0.793686
	XM_975887 // D4Ertd196e // DNA segment, Chr 4, ERATO Doi 196, expressed		9	6925880	4	3	1.82E-05	0.0009389	0.0054725	7.52E-07	0.868466	0.248111	0.8174	0.224956
Hint1	histidine triad nucleotide binding protein 1		5	6781083	4	3	0.0009562	0.0001943	0.0401752	0.00010327	0.79189	0.719453	0.918734	0.802668
Als2cl	ALS2 C-terminal like		27	6992475	4	2	8.95E-11	1.78E-08	0.0064204	0.00774164	0.999891	0.994688	9.24E-21	0.0031278
Rgmb	RGM domain family, member B		10	6853960	4	2	0.0313629	4.50E-07	0.0043692	3.90E-05	0.936921	0.83838	8.02E-08	0.840093
Gns	glucosamine (N-acetyl)-6-sulfatase	yes	20	6771224	4	2	0.005151	1.39E-05	0.0050028	3.87E-06	0.787379	0.486948	1.92E-07	9.51E-09
Galnt2	UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylga		19	6979772	4	2	0.0060867	0.0003345	0.0004923	0.00437458	0.999448	0.994643	2.25E-07	0.907301
Mcph1	microcephaly, primary autosomal recessive 1		22	6974480	4	2	1.86E-07	6.81E-07	0.0013538	0.00983987	0.996643	0.998361	0.00468492	0.875964
Metapl1	methionine aminopeptidase-like 1		11	6878016	4	2	0.003966	7.78E-05	0.0073619	1.24E-05	0.996054	0.607996	0.0267787	0.964285
Metrn	meteorin, glial cell differentiation regulator	yes	6	6854438	4	2	0.0105823	0.0008754	0.0461535	6.93E-05	0.844835	0.895902	0.0314022	0.364898
Pabpn1	poly(A) binding protein, nuclear 1		8	6819200	4	2	8.12E-05	0.0007419	0.0274772	0.0183231	0.994086	0.931362	0.0745585	0.864162
	NM_019836 // 2610024G14Rik		10	6939353	4	2	0.0377564	0.0002675	0.0002514	0.0106942	0.99054	0.52889	0.100119	0.917167
Rab34	RAB34, member of RAS oncogene family		16	6782680	4	2	0.0005791	0.0010054	0.0024256	0.00010187	0.99935	0.937618	0.213101	0.753626
Epb4.1l5	erythrocyte protein band 4.1-like 5	yes	28	6761675	4	2	0.0006306	0.0102827	0.04819	1.61E-06	0.15906	0.808541	0.278742	0.48428
Pycrl	pyrroline-5-carboxylate reductase-like		9	6836782	4	2	0.044896	0.0004196	0.0267182	6.51E-05	0.997682	0.813127	0.29277	0.455301
Cecr5	cat eye syndrome chromosome region, candidate 5 homolog		9	6956965	4	2	1.35E-08	0.00786	0.0181941	0.00013287	0.388461	0.930669	0.38108	0.926588
Dhx35	DEAH (Asp-Glu-Ala-His) box polypeptide 35		26	6882775	4	2	0.0017253	1.45E-07	0.0031002	9.42E-05	0.94974	0.971585	0.531181	0.715982
Ctbp1	C-terminal binding protein 1		12	6937228	4	2	4.76E-05	0.0216973	0.0061325	0.00018874	0.980415	0.959718	0.610921	0.924623
Gbas	glioblastoma amplified sequence		15	6934634	4	2	0.0003989	0.0029814	6.26E-05	0.00172896	0.922763	0.993106	0.800142	0.892479
Ppp3ca	protein phosphatase 3, catalytic subunit, alpha isoform		14	6901657	4	2	0.0008701	0.00047	0.0060752	0.0208899	0.990273	0.987578	0.807642	0.672586
Yeats4	YEATS domain containing 4		11	6777305	4	2	0.0304279	4.00E-05	2.21E-05	0.00777506	0.069241	0.464103	0.920013	0.883919
	XM_912548 // LOC636969 // similar to isoleucine-tRNA synthetase 2, mitochondrial		29	6764854	4	1	0.0067455	2.11E-10	0.0103932	0.021326	0.984232	0.551871	0.00182041	0.964557
Ube2q	ubiquitin-conjugating enzyme E2Q	yes	16	6899268	4	1	0.0048442	0.0011766	0.0075415	7.78E-05	0.96423	0.74167	0.00361281	0.988456
Aftph	aftiphilin		14	6786504	4	1	0.0226456	2.21E-06	0.0040058	0.0432138	0.990803	0.298187	0.00699185	0.748002
	NM_178115 // 2700050L05Rik		30	6964730	4	1	0.0329994	6.92E-05	0.0037104	0.0396391	0.961581	0.941164	0.00980534	0.896795
	XM_990825 // D5Ertd615e		9	6937867	4	1	0.0071772	5.01E-06	0.0089517	0.0395924	0.86879	0.659747	0.0217957	0.818083
Acads	acyl-Coenzyme A dehydrogenase, short chain		12	6941249	4	1	0.0358481	0.0002104	0.0086474	0.00447091	0.766612	0.940242	0.0404615	0.644627
Pcdh21	protocadherin 21		19	6824004	4	1	0.0250167	0.0009928	0.0435933	0.0166968	0.987906	0.994756	0.0555238	0.546052
Nnp1	novel nuclear protein 1		16	6775257	4	1	0.0289576	2.81E-05	0.0338675	0.0317165	0.973807	0.980083	0.0799901	0.17599
	NM_178778 // A930041I02Rik		18	6886203	4	1	4.35E-07	0.0013063	0.0242268	0.0240633	0.999765	0.823036	0.114017	0.870373
Tle6	LOC382645 // transducin-like enhancer protein 6	yes	18	6775478	4	1	0.0099313	0.0018994	0.0059514	0.0009422	0.957418	0.837404	0.124348	0.467535
Atrn	attractin		32	6881231	4	1	0.0003662	0.0025046	0.0115452	0.00146076	0.999852	0.809604	0.130495	0.599825
Cald1	caldesmon 1		12	6945335	4	1	0.0012246	3.49E-06	0.0220124	0.00617109	0.912485	0.830389	0.138803	0.589023
	XM_982175 // LOC666185		14	6929407	4	1	1.24E-05	0.0024435	0.0151472	0.0189562	0.956636	0.908618	0.148673	0.0515091
Slc16a13	solute carrier family 16 (monocarboxylic acid transport		8	6789395	4	1	0.0264394	0.0110842	0.0128182	3.13E-05	0.868764	0.934737	0.16826	0.502255
Hel308	helicase, mus308-like		24	6940366	4	1	0.0058894	1.43E-05	0.0049821	0.0115674	0.99843	0.931787	0.189157	0.0957669
Rhebl1	Ras homolog enriched in brain like 1		8	6838410	3	3	0.22817	9.03E-07	2.01E-06	9.49E-07	0.479696	0.972368	7.97E-06	7.82E-05
Abl1	v-abl Abelson murine leukemia oncogene 1	yes	17	6876154	3	3	2.72E-06	1.91E-09	0.235988	0.00097564	0.827551	0.937206	0.0011918	0.906496
Spaca1	sperm acrosome associated 1		17	6920713	3	3	0.0004115	5.73E-05	0.14551	7.81E-05	0.962961	0.999665	0.00403352	0.0753392
Cacna2d3	calcium channel, voltage-dependent, alpha2/delta subuni		38	6823635	3	2	0.0077438	0.0003213	0.889429	0.0001041	0.992349	0.976029	8.27E-05	0.881583
	XM_977418 // 3732413I11Rik		15	6780438	3	2	0.0003456	0.0005272	0.037211	0.0886833	0.999198	0.995796	0.00020356	0.0939623
Dnajc6	DnaJ (Hsp40) homolog, subfamily C, member 6	yes	22	6915844	3	1	0.0043807	7.31E-05	0.680249	0.00315822	0.980083	0.99341	1.70E-09	0.145697
Runx2	runt related transcription factor 2	yes	16	6855567	1	1	0.995697	6.27E-07	0.98955	0.237407	0.55498	0.77547	5.51E-23	1.64E-25
	XR_002827 // LOC670148 // similar to heat shock protein 1, alpha		16	6782563	1	1	0.13382	2.57E-16	0.313256	0.543857	0.995017	0.42808	9.89E-13	0.996349

Wu Supp Table 6:
Group E Genes

Gene Description	# Probe Sets	Transcript ID	No cell types p<0.05	No cell types p<0.001	p value wt v thu memory CD4	P value wt vs thu naïve CD4	p value wt vs thu memory CD8	p value wt vs thu naïve CD8	p value wt vs thu thymus CD8SP	p value wt vs thu thymus CD4SP	P-value WT naïve v WT memory CD4	P-value WT naïve v WT memory CD8
NM_020258 // Slc37a2 // solute carrier family 37 (glycerol-3-phosph	21	6994678	4	3	5.80E-05	2.35E-06	8.41E-10	0.0014634	0.78845	0.139502	1.40E-07	7.79E-10
NM_027432 // Wdr77 // WD repeat domain 77 // 3 F2.2 // 70465 /// N	12	6900237	4	3	0.001887	0.0007489	0.0008583	7.65E-08	0.521317	0.488372	0.0049533	0.149482
NM_011131 // Pold1 // polymerase (DNA directed), delta 1, catalytic	30	6966875	4	3	2.38E-05	0.0459173	1.31E-08	5.29E-06	0.606914	0.875409	0.0881439	0.99658
AK171818 // Dalrd3 // DALR anticodon binding domain containing 3	15	6992350	4	3	0.0069003	8.49E-05	1.32E-09	2.31E-05	0.997368	0.999753	0.135712	0.773668
XM_983775 // Pde8a // phosphodiesterase 8A // 7 D3 // 18584 /// NM	25	6962054	4	3	0.002718	7.63E-12	4.15E-06	3.35E-16	0.99185	0.845566	0.149993	0.137729
XM_001004675 // LOC668110 // region containing MON1 homolog I	11	6979269	4	3	2.89E-05	0.0003374	0.009316	0.000138	0.56866	0.998366	0.204611	0.664106
NM_198619 // MGC67181 // zinc finger-like protein // 4 E2 // 242747	28	7011014	4	3	0.0002585	1.07E-06	0.002071	1.88E-10	0.998975	0.931176	0.234584	0.447653
NM_010788 // Mecp2 // methyl CpG binding protein 2 // X A7.3|X 29	24	7017609	4	3	0.000869	4.40E-09	0.0116912	3.29E-05	1	0.975643	0.517057	0.82235
NM_024275 // Map4k5 // mitogen-activated protein kinase kinase kii	35	6801493	4	3	2.95E-06	0.0001163	0.000314	0.0049671	0.745386	0.926058	0.784085	7.70E-06
NM_146211 // Glt25d1 // glycosyltransferase 25 domain containing '	16	6977093	4	3	0.0056956	1.15E-06	6.48E-05	6.86E-09	0.867205	0.736075	0.986169	0.949855
NM_007708 // Cit // citron // 5 F // 12704 /// AK047757 // C030025P1	50	6933697	4	2	0.0138462	5.43E-08	0.0254447	1.12E-07	0.999615	0.998915	4.43E-28	4.34E-13
NM_010071 // Dok2 // docking protein 2 // 14 D2-D3|14 39.0 cM // 1	11	6820113	4	2	0.0004076	0.0006989	0.0162168	0.0094636	0.92603	0.5397	1.13E-11	0.0001568
NM_133721 // Itga9 // integrin alpha 9 // 9 F3|9 67.0 cM // 104099 ///	28	6992855	4	2	0.001191	0.0045358	0.0003231	0.0001679	0.514082	0.898803	1.53E-08	9.50E-08
NM_177034 // Apba1 // amyloid beta (A4) precursor protein binding,	12	6868728	4	2	0.0018924	0.0314338	0.0004241	0.000463	0.87496	0.64571	0.0001982	0.0036286
NM_010091 // Dvl1 // dishevelled, dsh homolog 1 (Drosophila) // 4 E	19	6919191	4	2	0.0206097	0.0404153	3.29E-07	4.17E-07	0.991568	0.810865	0.0012298	0.791104
NM_011741 // Zan // zonadhesin // 5 G2|5 78.0 cM // 22635 /// AK15	93	6942550	4	2	0.0294309	7.15E-05	0.0042885	1.83E-06	0.999955	0.985493	0.0065626	0.0603288
NM_011929 // Clcn6 // chloride channel 6 // 2 // 26372 /// AK220319 /	21	6926864	4	2	0.0176975	1.43E-07	0.0013601	4.37E-05	0.734862	0.999989	0.0097403	0.017672
NM_177137 // Gnal // guanine nucleotide binding protein, alpha stim	18	6861688	4	2	0.0037998	0.0025926	0.000359	3.05E-07	0.243299	0.897172	0.0099961	0.525469
NM_019689 // Arid3b // AT rich interactive domain 3B (Bright like) //	18	6995938	4	2	0.0019189	5.02E-07	0.0373534	2.38E-05	0.994611	0.943201	0.0261903	0.971649
NM_001025365 // D4Wsu114e // DNA segment, Chr 4, Wayne State	17	6926851	4	2	0.0687375	0.0003094	0.0203408	0.0002032	0.960609	0.015469	0.0299223	0.0414677
XM_992707 // LOC545194 // hypothetical LOC545194 // 17 B1 // 54	32	6854868	4	2	0.0005956	0.0369169	1.78E-05	0.0016723	0.999389	0.995876	0.121257	0.673717
XM_986470 // LOC675982 // similar to Meiotic recombination protei	13	6883640	4	2	0.0345598	0.0028129	0.0001856	0.00015	0.93955	0.953032	0.151038	0.490712
NM_010255 // Gamt // guanidinoacetate methyltransferase // 10 C1|	8	6775355	4	2	0.0488341	0.0065428	6.12E-06	0.000327	0.909526	0.777417	0.18449	0.233957
NM_010765 // Mapkapk5 // MAP kinase-activated protein kinase 5 /	8	6941681	4	2	0.0127348	2.90E-05	0.0221248	0.0004145	0.697439	0.787898	0.20777	0.67675
NM_027898 // Gramd1a // GRAM domain containing 1A // 7 B1|7 13	17	6966358	4	2	0.0183432	0.0092926	4.71E-09	0.0002896	0.973267	0.997871	0.431197	0.709427
AK004948 // D1Ertd161e // DNA segment, Chr 1, ERATO Doi 161, e	7	6750600	4	2	0.0081315	0.0007745	0.0001079	0.0036826	0.24429	0.468454	0.505626	0.111451
NM_001001180 // BC066028 // cDNA sequence BC066028 // 7 F5 /	10	6978892	4	2	0.0055394	0.0003291	3.96E-05	0.002611	0.67496	0.49015	0.644293	0.402828
NM_007741 // Col9a2 // procollagen, type IX, alpha 2 // 4 D2.2|4 53.	22	6916849	4	2	0.0001418	0.0005472	0.0013238	0.0251239	0.946822	0.984439	0.750898	0.598067
NM_007742 // Col1a1 // procollagen, type I, alpha 1 // 11 D|11 56.0	38	6783685	4	2	0.0058713	2.26E-07	0.0013053	2.15E-06	0.996113	0.916291	0.756922	0.655957
NM_147779 // Sftpb // surfactant associated protein B // 6 C1|6 31.0	12	6946968	4	2	1.11E-05	0.0485727	0.0372773	0.0002637	0.778462	0.966752	0.818777	0.771412
XM_896593 // B230218P12Rik // RIKEN cDNA B230218P12 gene /	23	6962502	4	1	7.46E-05	0.0363629	0.041808	0.0287643	0.929404	0.323202	1.34E-12	0.0119229
NM_206535 // Cd200r2 // Cd200 receptor 2 // 16 B4 // 271375 /// NM	9	6841140	4	1	0.0469569	0.0011224	0.0873431	0.003935	0.115403	0.0002121	0.0133992	0.290122
NM_024477 // Ttc28 // tetratricopeptide repeat domain 28 // 5 F // 20	18	6933474	4	1	0.0348897	0.0193654	0.0089276	0.0004221	0.484135	0.926197	0.0930573	0.65102
NM_029846 // Atg16l1 // autophagy-related 16-like 1 (yeast) // 1 D //	22	6751345	4	1	8.90E-05	0.0026704	0.0072461	0.0088352	0.88051	0.84218	0.153497	0.910805
NM_007630 // Ccnb2 // cyclin B2 // 9 D // 12442 /// XM_974848 // LO	13	6996649	4	1	4.86E-05	0.0320455	0.0043949	0.0401636	0.907814	0.983863	0.215049	0.622312
NM_001037712 // Kcnh6 // potassium voltage-gated channel, subfa	8	6784591	4	1	0.028874	1.36E-05	0.933712	0.0131087	0.364476	0.0053863	0.225372	0.920292
NM_175074 // Hmgn3 // high mobility group nucleosomal binding dc	8	6997349	4	0	0.156763	0.013771	0.0091699	0.0011569	0.927912	0.0434	1.99E-07	4.85E-15
NM_181569 // Mare // alpha globin regulatory element containing ge	16	6787102	4	0	0.0016278	0.0266672	0.0280811	0.0026882	0.991018	0.933747	0.0009609	0.693719
NM_138673 // Stab2 // stabilin 2 // 10 C1 // 192188 /// AK164675 // S	69	6775762	3	3	1.60E-08	7.09E-05	3.99E-08	0.395984	0.999999	0.987122	0	0
NM_024245 // Kif23 // kinesin family member 23 // 9 B // 71819 /// BI	27	6996186	3	3	0.72759	1.02E-08	0.0002938	1.33E-06	0.872051	0.919268	7.15E-38	8.06E-34
XM_001001170 // Rgl1 // ral guanine nucleotide dissociation stimula	27	6763090	3	3	9.22E-07	2.26E-06	3.43E-07	0.137445	0.999729	0.747762	1.52E-37	1.99E-41
NM_001039546 // Myo6 // myosin VI // 9 E1|9 44.0 cM // 17920	34	6990774	3	3	2.35E-07	1.00E-05	0.819889	0.575904	0.994545	2.18E-09	1.49E-22	0.374682
NM_010778 // Cd46 // CD46 antigen, complement regulatory protein	24	6765460	3	3	9.05E-05	0.0002934	0.0004712	0.943393	0.99631	0.990708	9.32E-09	4.22E-10
NM_007995 // Fcna // ficolin A // 2 A3 // 14133 /// AK133539 // Fcna	11	6885432	3	3	0.0506094	6.26E-07	8.13E-05	5.76E-05	0.950293	0.934335	2.63E-05	2.67E-15
NM_023587 // Ptplb // protein tyrosine phosphatase-like (proline inst	39	6836602	3	3	0.0546053	5.25E-19	9.72E-05	1.23E-08	0.998248	0.898594	0.0255978	0.005623
NM_177236 // Atp2b3 // ATPase, Ca++ transporting, plasma membr	25	7011937	3	3	9.80E-07	5.83E-06	0.210673	2.38E-05	0.998699	0.951771	0.11136	0.0396572
NM_198619 // MGC67181 // zinc finger-like protein // 4 E2 // 242747	18	6882333	3	2	0.0412797	2.20E-15	0.749923	5.82E-06	0.965982	0.969615	3.53E-21	1.58E-16
XR_003333 // LOC640058 // similar to killer cell lectin-like receptor,	24	6957435	3	2	0.113112	0.776013	5.94E-07	0.0004763	0.999974	0.0057036	7.06E-15	0
NM_008984 // Ptprm // protein tyrosine phosphatase, receptor type,	40	6856756	3	2	0.0768984	1.60E-05	0.0234164	8.80E-05	0.981948	0.981333	3.56E-11	4.89E-07
NM_009926 // Col11a2 // procollagen, type XI, alpha 2 // 17 B1|17 1	46	6850001	3	2	0.0155003	0.000916	0.259966	0.0006087	0.993039	0.997954	2.02E-10	0.487467
NM_009862 // Cdc45l // cell division cycle 45 homolog (S. cerevisiae	22	6844359	3	2	0.530433	0.0001927	0.0014021	5.51E-05	0.999195	0.746951	5.48E-08	0.0001272
NM_030260 // Zxdc // ZXD family zinc finger C // 6 D2 // 80292 /// NI	16	6947929	3	2	0.001602	0.0001354	0.136204	7.68E-07	0.999361	0.985882	0.0001104	0.44012
NM_026059 // 2900005J15Rik // RIKEN cDNA 2900005J15 gene //	8	6929349	3	2	0.112743	0.0001039	0.0274056	5.99E-07	0.999219	0.994536	0.000167	0.929038
NM_030014 // Hook1 // hook homolog 1 (Drosophila) // 4 C5|4 45.5	23	6915567	3	2	0.0114408	0.0007128	0.0004055	0.180503	0.512633	0.486719	0.000419	0.0053384

Wu Supp Table 7:
Group F Genes

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Gene Description**	**# Probe Sets**	**Transcript ID**	**Number of cell types p<0.05**	**Number of cell types p<0.001**	**p value wt v thu memory CD4**	**P value wt vs thu naïve CD4**	**p value wt vs thu memory CD8**	**p value wt vs thu naïve CD8**	**p value wt vs thu thymus CD8SP**	**p value wt vs thu thymus CD4SP**	**P-value WT naïve v WT memory CD4**	**P-value WT naïve v WT memory CD8**
2	NM_138673 // Stab2 // stabilin 2 // 10 C1 // 192188 /// AK164675 // Sta	69	6775762	3	3	1.60E-08	7.09E-05	3.99E-08	0.395984	0.999999	0.987122	0	0
3	NM_011682 // Utrn // utrophin // 10 A1|10 3.0 cM // 22288 /// AK04143	72	6772193	2	2	0.159699	1.27E-16	8.71E-05	0.522169	0.999996	0.999116	0	2.17E-05
4	NM_009109 // Ryr1 // ryanodine receptor 1, skeletal muscle // 7 A2-B3	104	6966183	3	1	0.0247109	1.16E-05	0.710008	0.0041708	0.999665	0.999976	6.59E-44	1.00E-33
5	XM_898936 // LOC434328 // hypothetical LOC434328 // 8 B3.3 // 434	72	6940431	3	2	0.0103945	8.64E-06	0.0622501	2.34E-07	0.973513	0.811818	2.11E-39	0
6	NM_010553 // Il18rap // interleukin 18 receptor accessory protein // 1	15	6748893	1	0	0.820427	0.0030746	0.944638	0.548918	1	0.844381	6.82E-39	3.61E-19
7	NM_024245 // Kif23 // kinesin family member 23 // 9 B // 71819 /// BC0	27	6996186	3	3	0.72759	1.02E-08	0.0002938	1.33E-06	0.872051	0.919268	7.15E-38	8.06E-34
8	NM_009778 // C3 // complement component 3 // 17 E1-E3|17 34.3 cM	39	6856290	3	2	4.35E-06	2.88E-05	0.0011648	0.0987891	0.998447	0.96805	1.02E-37	5.24E-14
9	XM_001001170 // Rgl1 // ral guanine nucleotide dissociation stimulato	27	6763090	3	3	9.22E-07	2.26E-06	3.43E-07	0.137445	0.999729	0.747762	1.52E-37	1.99E-41
10	NM_008512 // Lrp1 // low density lipoprotein receptor-related protein 1	92	6777957	3	0	0.0375073	0.196676	0.0480725	0.0321942	0.694163	0.948985	7.09E-37	1.75E-38
11	XM_910432 // LOC209589 // similar to 40S ribosomal protein S7 (S8)	17	6954415	1	0	0.363248	0.025593	0.985096	0.0768466	0.998866	0.997807	9.63E-37	1.21E-18
12	NM_053214 // Myo1f // myosin IF // 17 B-C|17 17.5 cM // 17916 /// NM	31	6849951	1	0	0.453443	0.0074038	0.151558	0.939408	0.99991	0.967273	1.97E-36	1.26E-06
13	XR_004438 // LOC670718 // similar to PRAME family member 10 // //	26	6953507	2	1	0.117389	0.0004429	0.273319	0.0355063	0.371439	0.99643	6.81E-36	7.51E-41
14	NM_010766 // Marco // macrophage receptor with collagenous structu	19	6761701	3	2	1.85E-05	7.59E-05	0.0027874	0.0550443	0.950442	0.941175	1.04E-34	1.09E-14
15	NM_023122 // Gpm6b // glycoprotein m6b // X F5 // 14758 /// AF2548	19	7015061	0	0	0.378249	0.467384	0.271461	0.149154	0.941756	0.998698	1.00E-33	2.25E-05
16	XM_001001359 // LOC668417 // similar to Ig heavy chain V region M	64	6803946	1	0	0.0466881	0.770816	0.128173	0.490651	1	0.991496	2.30E-33	4.84E-21
17	NM_033078 // Klrk1 // killer cell lectin-like receptor subfamily K, memb	13	6957406	0	0	0.0552567	0.588688	0.150833	0.142591	0.995939	0.92472	7.84E-31	1.81E-18
18	NM_013813 // Epb4.1l3 // erythrocyte protein band 4.1-like 3 // 17 E1.	21	6851897	3	2	8.85E-06	0.0010443	0.0001243	0.634407	0.89185	0.973813	3.79E-30	2.98E-22
19	NM_009782 // Cacna1e // calcium channel, voltage-dependent, R type	47	6763196	1	0	0.937593	0.661418	0.0187135	0.461113	0.911714	0.595	1.98E-29	1.26E-44
20	XM_980588 // LOC670310 // hypothetical protein LOC670310 // // 67	42	6847748	1	0	0.999466	0.984045	0.0072723	0.156007	0.999897	0.999893	2.11E-29	0.00066412
21	NM_019507 // Tbx21 // T-box 21 // 11 D // 57765	10	6791171	1	0	0.364406	0.0081552	0.541437	0.352382	0.999638	0.996226	8.56E-29	5.30E-19
22	NM_029600 // Abcc3 // ATP-binding cassette, sub-family C (CFTR/MR	36	6790944	4	1	0.0326833	3.03E-05	0.026733	0.0109198	0.970653	0.998085	2.33E-28	1.29E-22
23	NM_007708 // Cit // citron // 5 F // 12704 /// AK047757 // C030025P15	50	6933697	4	2	0.0138462	5.43E-08	0.0254447	1.12E-07	0.999615	0.998915	4.43E-28	4.34E-13
24	NM_019984 // Tgm1 // transglutaminase 1, K polypeptide // 14 C1 // 2	18	6824838	3	1	6.89E-05	0.0450446	0.0469582	0.0565813	0.891307	0.972709	4.61E-28	5.77E-09
25	XM_910215 // LOC635217 // similar to leucine-rich repeat kinase 1 //	47	6967969	1	0	0.334209	0.373059	0.0763829	0.0381961	0.924452	0.97671	3.01E-27	0
26	NM_008396 // Itga2 // integrin alpha 2 // 13 D2.2|13 64.0 cM // 16398	26	6816247	1	0	0.231084	0.0130303	0.386708	0.138414	0.864085	0.561884	4.40E-27	9.52E-18
27	NM_008399 // Itgae // integrin, alpha E, epithelial-associated // 11 B4|	34	6782286	3	1	0.244301	0.0045476	0.0024081	0.113782	2.48E-05	0.909793	9.36E-27	2.83E-05
28	NM_172285 // Plcg2 // phospholipase C, gamma 2 // 8 E1 // 234779	34	6979439	0	0	0.120552	0.651429	0.268655	0.262649	0.835133	0.999997	3.72E-26	9.72E-22
29	NM_019670 // Diap3 // diaphanous homolog 3 (Drosophila) // 14 D3 //	30	6826546	1	1	0.462031	0.0001635	0.436757	0.0628106	0.99973	0.99677	1.08E-25	5.36E-09
30	NM_138301 // Trpm2 // transient receptor potential cation channel, su	38	6775236	2	0	0.959002	0.0040308	0.30893	0.0103137	0.640473	0.996668	2.50E-25	0
31	XR_005042 // LOC676711 // similar to MORF-related gene 15 isoform	19	6869503	1	1	0.174996	4.70E-09	0.594723	0.28185	0.999804	0.999936	3.80E-25	3.34E-18
32	NM_010587 // Itsn1 // intersectin 1 (SH3 domain protein 1A) // 16 C3.3	43	6843198	4	2	0.0225853	5.55E-05	0.0089652	2.70E-05	0.982168	0.962889	9.58E-25	3.31E-13
33	NM_146064 // Soat2 // sterol O-acyltransferase 2 // 15 F3|15 61.7 cM	19	6833394	3	0	0.0251807	0.0262788	0.175356	0.0200463	0.995049	0.992311	1.91E-24	5.61E-07
34	NM_008905 // Ppfibp2 // protein tyrosine phosphatase, receptor-type,	28	6963271	0	0	0.641385	0.192807	0.0780662	0.85174	0.942276	0.289105	4.31E-24	0
35	NM_018797 // Plxnc1 // plexin C1 // 10 C3 // 54712 /// XM_622776 // P	29	6776156	2	0	0.0367115	0.148694	0.0432887	0.671045	0.997026	0.887155	5.51E-24	8.26E-36
36	NM_013839 // Nr1h3 // nuclear receptor subfamily 1, group H, membe	12	6888752	2	0	0.0162327	0.0057317	0.53222	0.636925	0.891437	0.985738	5.73E-24	1.24E-06
37	NM_199240 // Sema6d // sema domain, transmembrane domain (TM)	24	6880776	2	0	0.121058	0.0085771	0.0018627	0.216257	0.992406	0.204383	8.11E-24	1.21E-23
38	NM_028341 // 2810439F02Rik // RIKEN cDNA 2810439F02 gene // 11	13	6858910	1	0	0.904461	0.0370393	0.117949	0.154325	0.898593	0.869242	1.34E-23	4.45E-05
39	XM_986599 // LOC666183 // hypothetical protein LOC666183 // 15 B3	35	6811724	1	0	0.0373188	0.717877	0.0774006	0.0545691	0.975854	0.796488	2.06E-23	1.38E-28
40	NM_009820 // Runx2 // runt related transcription factor 2 // 17 B3|17 2	16	6855567	1	1	0.995697	6.27E-07	0.98955	0.237407	0.55498	0.77547	5.51E-23	1.64E-25
41	NM_173437 // Nav1 // neuron navigator 1 // 1 E4 // 215690 /// BC0288	33	6762429	1	0	0.991536	0.89767	0.445484	0.0103129	0.893627	0.997375	1.29E-22	5.18E-43
42	XM_001002503 // LOC668592 // hypothetical protein LOC668592 // 1	34	6856231	0	0	0.400684	0.0633128	0.373424	0.297145	0.998684	0.997022	1.74E-22	4.08E-15
43	NM_013689 // Tec // cytoplasmic tyrosine kinase, Dscr28C related (Dr	23	6939077	1	0	0.856515	0.0381141	0.195075	0.778324	0.998133	0.675696	2.60E-22	2.02E-05
44	NM_146001 // Hip1 // huntingtin interacting protein 1 // 5 F-G2|5 75.0	33	6942441	0	0	0.477525	0.590469	0.388991	0.0977287	0.892995	0.998572	3.40E-22	5.36E-22
45	AK080042 // Timd4 // T-cell immunoglobulin and mucin domain contair	12	6780572	2	0	0.0022584	0.0073877	0.958543	0.797921	0.991205	0.783873	7.76E-22	3.12E-08
46	NM_008479 // Lag3 // lymphocyte-activation gene 3 // 6 F2 // 16768	16	6957144	1	0	0.375641	0.314116	0.804002	0.0038405	0.524867	0.964221	1.05E-21	7.21E-21
47	NM_198619 // MGC67181 // zinc finger-like protein // 4 E2 // 242747 //	18	6882333	3	2	0.0412797	2.20E-15	0.749923	5.82E-06	0.965982	0.969615	3.53E-21	1.58E-16
48	NM_133348 // Acot7 // acyl-CoA thioesterase 7 // 4 E2 // 70025 /// ABC	14	6919012	1	0	0.930541	0.375569	0.0327396	0.448251	0.828811	0.897344	5.83E-21	2.11E-11
49	NM_145150 // Prc1 // protein regulator of cytokinesis 1 // 7 D2|7 38.0	13	6961987	2	2	0.749792	2.57E-08	0.649827	5.01E-05	0.99982	0.99758	1.12E-20	9.74E-14
50	NM_010620 // Kif15 // kinesin family member 15 // 9 F4 // 209737 /// A	31	6993098	3	1	0.189231	1.15E-06	0.0493535	0.0055144	0.999997	0.999965	1.16E-20	8.32E-17
51	NM_133762 // Luzp5 // leucine zipper protein 5 // 12 F2 // 76044 /// NN	31	6798403	1	0	0.529504	0.0054249	0.287531	0.227871	0.999999	0.977289	2.35E-20	1.35E-18
52	XM_001003501 // Arhgap21 // Rho GTPase activating protein 21 // 2 A	25	6885195	2	0	0.0079024	0.0239037	0.612141	0.309582	0.970353	0.912205	2.46E-20	6.94E-17
53	NM_010580 // Itgb5 // integrin beta 5 // 16 B3 // 16419	17	6840637	3	1	0.0108366	8.61E-10	0.0112335	0.299047	0.998826	0.982242	4.53E-20	2.21E-17
54	XM_981660 // LOC666111 // similar to 60S ribosomal protein L29 // 18	9	6816161	2	1	0.378008	0.0088307	0.991772	2.33E-08	0.835094	0.740917	6.25E-20	1.17E-24
55	NM_181072 // Myo1e // myosin IE // 9 D|9 41.0 cM // 71602 /// AK029	30	6990263	0	0	0.221131	0.921942	0.638589	0.832616	0.989289	0.966401	1.30E-19	3.81E-13
56	NM_173762 // Cenpe // centromere protein E // 3 H2 // 229841 /// AK1	43	6901592	1	0	0.866774	0.0299485	0.534503	0.26722	0.995279	0.99849	2.29E-19	4.77E-10
57	NM_013470 // Anxa3 // annexin A3 // 5 E3|5 54.0 cM // 11745 /// XM_9	14	6932603	3	2	0.0004609	0.0010954	7.97E-06	0.0587045	0.734463	0.913759	2.53E-19	1.14E-23
58	NM_007403 // Adam8 // a disintegrin and metallopeptidase domain 8	22	6972106	1	0	0.521156	0.0202059	0.0903424	0.108468	0.9191	0.943855	2.90E-19	6.64E-41
59	XM_897921 // Slc43a1 // solute carrier family 43, member 1 // 2 E1 // 1	19	6878712	1	0	0.288088	0.259083	0.456387	0.0270844	0.999972	0.922638	6.00E-19	2.33E-14
60	NM_013862 // Rabgap1l // RAB GTPase activating protein 1-like // 1 H	26	6763495	1	0	0.918466	0.0053322	0.516199	0.506	0.967067	0.983042	8.70E-19	8.95E-06
61	XM_981230 // Atp2b4 // ATPase, Ca++ transporting, plasma membran	18	6762321	1	1	0.681759	1.42E-07	0.899212	0.405945	0.729198	0.641887	1.44E-18	4.00E-10
62	NM_013464 // Ahr // aryl-hydrocarbon receptor // 12 A3|12 18.0 cM //	14	6800060	0	0	0.0650828	0.271184	0.760514	0.152377	0.907919	0.920033	2.01E-18	4.68E-14
63	NM_011498 // Bhlhb2 // basic helix-loop-helix domain containing, clas	8	6948913	1	0	0.273225	0.861	0.0119766	0.405555	0.796547	0.850541	2.41E-18	1.20E-09
64	NM_173006 // Pon3 // paraoxonase 3 // 6 A1|6 0.5 cM // 269823 /// BC	15	6951401	0	0	0.094078	0.105558	0.505858	0.508239	0.984908	0.845589	2.50E-18	3.43E-06
65	NM_147100 // Olfr614 // olfactory receptor 614 // // 259104 /// NM_17	10	6976765	0	0	0.125494	0.0900823	0.123112	0.842066	0.995021	0.840461	4.06E-18	5.48E-20
66	NM_007754 // Cpd // carboxypeptidase D // 11 B5|11 46.0 cM // 12874	22	6789830	0	0	0.906165	0.255634	0.118139	0.576861	0.401646	0.36325	7.10E-18	3.14E-06
67	NM_011370 // Cyfip1 // cytoplasmic FMR1 interacting protein 1 // 7 B5	34	6960836	0	0	0.258943	0.800731	0.88364	0.130637	0.954654	0.983413	1.18E-17	1.78E-25
68	NM_008374 // Il9r // interleukin 9 receptor // 11 A4 // 16199 /// XM_977	12	6787099	0	0	0.304041	0.249983	0.352459	0.312815	0.302568	0.941771	1.21E-17	1.50E-06

Wu Supp Table 8: Highest ranked wt memory vs naive cells Page 1

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Gene Description**	**# Probe Sets**	**Transcript ID**	**Number of cell types p<0.05**	**Number of cell types p<0.001**	**p value wt v thu memory CD4**	**P value wt vs thu naïve CD4**	**p value wt vs thu memory CD8**	**p value wt vs thu naïve CD8**	**p value wt vs thu thymus CD8SP**	**p value wt vs thu thymus CD4SP**	**P-value WT naïve v WT memory CD4**	**P-value WT naïve v WT memory CD8**
69	NM_147218 // Abca6 // ATP-binding cassette, sub-family A (ABC1), m	46	6792122	2	2	0.403196	0.32678	0.0009088	0.0001424	1	0.998482	1.27E-17	3.77E-08
70	NM_182809 // Ntrk3 // neurotrophic tyrosine kinase, receptor, type 3 //	23	6968643	2	1	0.124627	4.46E-05	0.984968	0.0012984	0.99463	0.995026	1.57E-17	7.09E-17
71	NM_029766 // Dtl // denticleless homolog (Drosophila) // 1 H6 // 76843	24	6765235	1	1	0.527012	3.79E-06	0.644849	0.110798	0.999991	0.971977	1.95E-17	9.46E-11
72	NM_007615 // Ctnnd1 // catenin (cadherin associated protein), delta 1	25	6888307	1	0	0.118223	0.634409	0.0184886	0.119119	0.95786	0.997951	2.44E-17	2.66E-16
73	NM_177353 // Slc9a7 // solute carrier family 9 (sodium/hydrogen exch	20	7015941	2	0	0.0911122	0.136488	0.0184892	0.0034303	0.999998	0.978249	2.76E-17	6.14E-14
74	NM_173051 // Serpinb1c // serine (or cysteine) peptidase inhibitor, cla	16	6812178	1	0	0.242115	0.0255545	0.183821	0.99568	0.996859	0.741448	4.13E-17	9.82E-15
75	NM_011597 // Tjp2 // tight junction protein 2 // // 21873 /// AK076689 /	24	6872250	0	0	0.367341	0.937586	0.98468	0.596253	0.548885	0.995699	5.00E-17	1.94E-16
76	XM_622573 // Anxa2 // annexin A2 // 9 C|9 37.0 cM // 12306 /// NM_0	15	6990216	1	0	0.180536	0.0276161	0.990005	0.952589	0.767704	0.850836	5.50E-17	2.24E-15
77	NM_028072 // Sulf2 // sulfatase 2 // 2 H3 // 72043 /// AK129316 // Sulf	24	6893057	0	0	0.0543707	0.291321	0.104564	0.106127	0.977561	0.973333	5.97E-17	1.00E-22
78	NM_010790 // Melk // maternal embryonic leucine zipper kinase // 4 B	24	6913080	3	2	0.0866639	0.0003267	0.0110967	0.0006279	0.996733	0.822457	7.92E-17	9.00E-20
79	NM_011125 // Pltp // phospholipid transfer protein // 2 H3|2 93.0 cM //	14	6892964	0	0	0.059598	0.24571	0.473189	0.693006	0.954592	0.627444	9.01E-17	7.51E-05
80	XM_913343 // LOC637609 // similar to receptor-like tyrosine kinase //	18	6991991	0	0	0.090156	0.218542	0.486874	0.0969846	0.831839	0.909947	1.02E-16	0.00033746
81	XR_003401 // LOC671381 // similar to prothymosin alpha // // 671381	22	6805180	2	0	0.0013441	0.235281	0.10337	0.0287251	0.993955	0.984475	1.03E-16	0.00016139
82	NM_009772 // Bub1 // budding uninhibited by benzimidazoles 1 homo	24	6890715	1	0	0.647613	0.0038388	0.451474	0.0756314	0.97943	0.804034	1.46E-16	3.33E-13
83	XM_917417 // LOC640461 // similar to eukaryotic translation initiation	25	6880451	1	0	0.505088	0.0695661	0.0468822	0.291741	0.999944	0.999823	1.46E-16	1.07E-13
84	XM_001003800 // Plxnb2 // plexin B2 // 15 43.0 cM // 140570 /// NM_0	43	6837787	2	0	0.144673	0.203358	0.0395882	0.0265042	0.999992	0.418486	1.76E-16	4.39E-20
85	NM_025282 // Mef2c // myocyte enhancer factor 2C // 13 C3|13 45.0 c	17	6808609	1	0	0.0043331	0.156645	0.116024	0.112253	0.997218	0.753733	1.80E-16	7.78E-16
86	NM_177192 // D030011O10Rik // RIKEN cDNA D030011O10 gene //	17	6958439	1	0	0.0333296	0.0604442	0.519182	0.0779539	0.934848	0.997988	2.70E-16	3.02E-08
87	XM_907196 // LOC632908 // similar to Cell division control protein 6 h	9	6784054	3	1	0.0329392	6.16E-10	0.474982	0.0031607	0.999586	0.81319	3.05E-16	2.18E-11
88	NM_008534 // Ly9 // lymphocyte antigen 9 // 1 H3|1 93.3 cM // 17085	11	6764093	0	0	0.548256	0.465809	0.909913	0.943256	0.956245	0.993951	3.93E-16	1.90E-20
89	NM_145148 // Frmd4b // FERM domain containing 4B // 6 D3 // 23228	23	6955778	1	0	0.189809	0.108221	0.0177604	0.0752124	0.938519	0.726671	4.04E-16	4.53E-14
90	NM_146023 // Evi2b // ecotropic viral integration site 2b // 11 B5 // 216	16	6790046	0	0	0.340891	0.0542889	0.270646	0.398427	0.999954	0.99759	4.41E-16	2.18E-07
91	XM_908720 // Gm1276 // gene model 1276, (NCBI) // 19 A // 383435 //	10	6871545	1	0	0.216074	0.0013965	0.329563	0.0803065	0.523219	0.994779	6.76E-16	8.32E-08
92	NM_008969 // Ptgs1 // prostaglandin-endoperoxide synthase 1 // 2 B2	13	6876430	4	1	0.0133948	0.0008156	0.0026003	0.0229953	0.817433	0.986077	8.71E-16	5.58E-11
93	NM_183222 // Fcrl5 // Fc receptor-like 5 // 3 F1 // 329693 /// AK08975	15	6899000	2	0	0.0112823	0.253855	0.0150613	0.312311	0.703259	0.885901	8.99E-16	1.02E-12
94	NM_013471 // Anxa4 // annexin A4 // 6 D1|6 38.0 cM // 11746 /// NM_0	15	6955228	0	0	0.34321	0.139689	0.0984367	0.16494	0.143616	0.897197	9.44E-16	5.47E-27
95	NM_007421 // Adssl1 // adenylosuccinate synthetase like 1 // 12 F1 //	15	6798271	1	0	0.123609	0.178199	0.0727601	0.0034485	0.687657	0.958027	1.09E-15	1.46E-07
96	XM_980183 // 1700020I22Rik // RIKEN cDNA 1700020I22 gene // 9 D	8	6996678	1	0	0.593483	0.0895526	0.0091392	0.36151	0.893179	0.255888	1.17E-15	4.87E-09
97	NM_023044 // Slc15a3 // solute carrier family 15, member 3 // 19 B //	13	6868058	2	0	0.0388865	0.0733956	0.434195	0.0231613	0.679651	0.211348	1.19E-15	1.69E-08
98	NM_011623 // Top2a // topoisomerase (DNA) II alpha // 11 D|11 57.0 c	34	6791298	2	1	0.0978802	5.35E-15	0.124839	0.0027654	0.999969	0.998402	1.35E-15	2.56E-12
99	NM_027918 // 1300017J02Rik // RIKEN cDNA 1300017J02 gene // 9	14	6998398	0	0	0.0577516	0.0965191	0.434474	0.442219	0.818644	0.64803	1.56E-15	0.00036803
100	AK046303 // Cd33 // CD33 antigen // 7 B3|7 23.0 cM // 12489 /// AK08	11	6966808	1	1	0.0005728	0.132044	0.206949	0.282492	0.977608	0.768371	2.17E-15	1.48E-05
101	NM_009465 // Axl // AXL receptor tyrosine kinase // 7 A3-B1|7 6.0 cM	23	6965982	2	0	0.0021497	0.0057682	0.14408	0.290831	0.940576	0.999985	3.86E-15	2.99E-26
102	NM_008923 // Prkar1b // protein kinase, cAMP dependent regulatory,	12	6942655	1	1	1.19E-05	0.482075	0.17497	0.190223	0.961381	0.632848	4.00E-15	1.33E-06
103	NM_181277 // Col14a1 // procollagen, type XIV, alpha 1 // 15 D // 128	45	6830506	3	0	0.281587	0.0061287	0.0458286	0.00898	0.979459	0.999716	6.46E-15	2.35E-10
104	XR_003333 // LOC640058 // similar to killer cell lectin-like receptor, su	24	6957435	3	2	0.113112	0.776013	5.94E-07	0.0004763	0.999974	0.0057036	7.06E-15	0
105	NM_153804 // Plekhg3 // pleckstrin homology domain containing, fam	15	6796193	1	0	0.958483	0.0058731	0.580021	0.233552	0.214956	0.412951	7.98E-15	1.96E-06
106	XM_986085 // LOC671299 // similar to constitutive photomorphogenic	10	6957465	2	0	0.0370076	0.227158	0.0276641	0.813414	0.929048	0.939161	8.27E-15	8.04E-06
107	NM_010578 // Itgb1 // integrin beta 1 (fibronectin receptor beta) // 8 E2	18	6980032	1	1	0.0003699	0.230942	0.310869	0.997632	0.993433	0.773793	9.21E-15	1.23E-06
108	NM_177025 // Cobll1 // Cobl-like 1 // 2 C3 // 319876 /// AK052939 // C	18	6887286	2	0	0.139956	0.0020128	0.373171	0.0373085	0.946894	0.791024	1.01E-14	3.17E-21
109	AJ010751 // Klrg1 // killer cell lectin-like receptor subfamily G, membe	7	6957025	1	0	0.269208	0.870773	0.587001	0.0063354	0.502413	0.959044	1.03E-14	1.81E-13
110	NM_009516 // Wee1 // wee 1 homolog (S. pombe) // 7 E3 // 22390	15	6963418	1	0	0.0782962	0.0084571	0.217214	0.406959	0.988383	0.998497	1.15E-14	9.96E-06
111	XM_001001776 // LOC668491 // hypothetical protein LOC668491 // 14	9	6825872	0	0	0.566095	0.0500051	0.144421	0.860899	0.688662	0.899501	1.34E-14	1.37E-06
112	NM_144553 // Dlg7 // discs, large homolog 7 (Drosophila) // 14 C1|14	20	6824346	1	0	0.413782	0.0021759	0.92505	0.2725	0.999982	0.994736	1.37E-14	2.80E-12
113	NM_145437 // 4732429D16Rik // RIKEN cDNA 4732429D16 gene // 1	8	6792371	1	0	0.0724003	0.982255	0.828407	0.0082844	0.161219	0.963465	3.30E-14	0.00010591
114	NM_008401 // Itgam // integrin alpha M // 7 F4 // 16409 /// AK089521 /	31	6964380	0	0	0.853981	0.126937	0.504408	0.0751195	0.786636	0.980826	3.67E-14	1.18E-22
115	NM_007779 // Csf1r // colony stimulating factor 1 receptor // 18 D|18 3	28	6861358	4	4	1.22E-05	1.73E-08	4.10E-07	4.19E-05	0.998956	0.998607	4.72E-14	2.56E-40
116	NM_181318 // Rasgef1b // RasGEF domain family, member 1B // 5 E3	16	6940303	2	1	0.0002294	0.0146426	0.521085	0.512663	0.998576	0.981235	4.77E-14	1.28E-09
117	XM_986599 // LOC666183 // hypothetical protein LOC666183 // 15 B3	31	6875181	4	2	0.0017147	1.43E-06	7.65E-05	0.0400525	0.981391	0.99595	5.33E-14	0
118	NM_007680 // Ephb6 // Eph receptor B6 // 6 B2.1 // 13848	22	6945730	1	0	0.765052	0.0013613	0.257973	0.277039	0.991084	0.97331	6.39E-14	0.00016966
119	NM_175172 // 4930506M07Rik // RIKEN cDNA 4930506M07 gene //	21	6874085	1	0	0.955692	0.241643	0.389342	0.0053483	0.997212	0.716725	7.23E-14	7.42E-38
120	NM_010187 // Fcgr2b // Fc receptor, IgG, low affinity IIb // 1 H3|1 92.3	8	6764038	1	0	0.187009	0.0771745	0.421957	0.0016062	0.932596	0.926573	7.23E-14	7.60E-14
121	XM_916960 // Mlkl // mixed lineage kinase domain-like // 8 D3 // 7456	11	6985361	0	0	0.327461	0.10005	0.164698	0.198486	0.924925	0.541239	7.45E-14	0.00060935
122	NM_013484 // C2 // complement component 2 (within H-2S) // 17 B1|19	19	6855061	3	1	0.0494736	0.0006291	0.203242	0.0230004	0.575715	0.883897	8.41E-14	2.87E-10
123	NM_009791 // Aspm // asp (abnormal spindle)-like, microcephaly assc	36	6753592	2	2	0.142883	3.44E-07	0.1104	0.0001528	0.999996	0.77895	9.66E-14	4.87E-13
124	NM_019418 // Tnfsf14 // tumor necrosis factor (ligand) superfamily, me	6	6856288	1	0	0.803464	0.423693	0.001538	0.338165	0.90837	0.846106	1.01E-13	4.94E-06
125	NM_173014 // Aytl1 // acyltransferase like 1 // 8 C5 // 270084 /// XM_9	16	6978232	2	0	0.0086361	0.0024696	0.0647031	0.437003	0.999144	0.885737	1.12E-13	3.91E-23
126	NM_009015 // Rad54l // RAD54 like (S. cerevisiae) // 4 D1 // 19366 ///	24	6924763	1	1	0.414162	0.0001641	0.151141	0.136684	0.999995	0.513611	1.23E-13	9.55E-08
127	NM_011539 // Tbxas1 // thromboxane A synthase 1, platelet // 6 F1-pt	16	6945584	1	0	0.930001	0.139711	0.0114101	0.849791	0.132767	0.881294	1.39E-13	8.95E-21
128	NM_009396 // Tnfaip2 // tumor necrosis factor, alpha-induced protein	17	6798213	3	0	0.028506	0.0013735	0.0259098	0.168581	0.746521	0.685551	1.55E-13	1.71E-10
129	NM_007630 // Ccnb2 // cyclin B2 // 9 D // 12442	10	6996646	2	1	0.84453	1.13E-07	0.624927	0.036678	0.999494	0.806591	1.67E-13	1.85E-13
130	NM_018763 // Chst2 // carbohydrate sulfotransferase 2 // // 54371 ///	7	6997990	1	1	0.47496	0.32498	0.745838	8.70E-05	0.914831	0.345344	1.82E-13	0.00039394
131	NM_175155 // Sash1 // SAM and SH3 domain containing 1 // 10 A1 //	23	6772009	3	1	0.002396	0.0577467	0.012163	0.000883	0.999758	0.981881	2.02E-13	1.27E-15
132	NM_019933 // Ptpn4 // protein tyrosine phosphatase, non-receptor typ	24	6761680	1	1	0.388567	1.05E-06	0.982637	0.438267	0.999997	0.131334	2.14E-13	4.29E-19
133	NM_172616 // C330027C09Rik // RIKEN cDNA C330027C09 gene //	26	6841372	0	0	0.0511945	0.173973	0.552806	0.79583	0.999471	0.974545	2.23E-13	8.51E-09
134	NM_008869 // Pla2g4a // phospholipase A2, group IVA (cytosolic, calc	18	6762944	3	0	0.0127936	0.0318231	0.0010899	0.185499	0.995553	0.78953	2.60E-13	1.56E-15
135	XM_988224 // Fmnl2 // formin-like 2 // 2 C1.1 // 71409 /// XM_901889	15	6877139	2	1	0.111907	0.0002709	0.379921	0.004008	0.998573	0.850563	3.05E-13	1.09E-30

Wu Supp Table 8: Highest ranked wt memory vs naive cells Page 2

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Gene Description**	**# Probe Sets**	**Transcript ID**	**Number of cell types p<0.05**	**Number of cell types p<0.001**	**p value wt v thu memory CD4**	**P value wt vs thu naïve CD4**	**p value wt vs thu memory CD8**	**p value wt vs thu naïve CD8**	**p value wt vs thu thymus CD8SP**	**p value wt vs thu thymus CD4SP**	**P-value WT naïve v WT memory CD4**	**P-value WT naïve v WT memory CD8**
136	XM_001003509 // LOC629432 // hypothetical LOC629432 // 16 C1.3 /	19	6846571	0	0	0.414933	0.385145	0.329274	0.866225	0.994714	0.772756	3.23E-13	1.57E-05
137	NM_008599 // Cxcl9 // chemokine (C-X-C motif) ligand 9 // 5 E2|5 53.(	7	6939987	0	0	0.663821	0.143864	0.117175	0.711835	0.267152	0.959925	3.44E-13	4.50E-15
138	NM_001037136 // Centg2 // centaurin, gamma 2 // 1 D // 347722 /// NI	20	6751442	1	0	0.335516	0.0118422	0.451294	0.212088	0.921662	0.911533	4.17E-13	0.00033703
139	AK004232 // Pla2g2d // phospholipase A2, group IID // 4 D3|4 67.0 cM	7	6918015	2	0	0.0447952	0.0493694	0.436076	0.523101	0.877136	0.567472	4.26E-13	0.00019494
140	NM_008408 // Stt3a // STT3, subunit of the oligosaccharyltransferase	14	6994666	1	0	0.316988	0.0084814	0.605725	0.176072	0.9419	0.933917	4.46E-13	2.88E-08
141	NM_178045 // Rassf4 // Ras association (RalGDS/AF-6) domain famil	18	6956765	0	0	0.213668	0.0781927	0.094175	0.103378	0.78664	0.984447	5.13E-13	1.04E-21
142	NM_001035509 // Zcchc18 // zinc finger, CCHC domain containing 18	9	7014048	1	1	0.0811219	0.683685	0.0592339	2.90E-07	0.69307	0.276015	7.47E-13	2.48E-10
143	NM_139232 // Fgd4 // FYVE, RhoGEF and PH domain containing 4 //	21	6844210	2	0	0.398982	0.370092	0.0012616	0.0529363	0.0011222	0.946226	7.71E-13	5.23E-08
144	NM_008278 // Hpgd // hydroxyprostaglandin dehydrogenase 15 (NAD	8	6976237	2	1	0.0052925	9.20E-05	0.197758	0.311704	0.890865	0.956703	8.02E-13	4.21E-08
145	NM_018775 // Tbc1d8 // TBC1 domain family, member 8 // 1 B // 5461	26	6758027	1	0	0.680063	0.213385	0.622699	0.0097005	0.717136	0.998773	8.52E-13	9.16E-39
146	NM_001034098 // Tnfsf12-tnfsf13 // tumor necrosis factor (ligand) sup	8	6789329	3	1	0.0027076	0.0007948	0.0477096	0.142093	0.816306	0.468497	8.85E-13	2.12E-06
147	NM_010703 // Lef1 // lymphoid enhancer binding factor 1 // 3 G3|3 61.	20	6901380	2	0	0.0997992	0.0466652	0.0033829	0.90426	0.99987	0.999489	9.07E-13	5.40E-14
148	NM_010304 // Gna15 // guanine nucleotide binding protein, alpha 15 /	8	6775472	0	0	0.736786	0.981813	0.67737	0.0571631	0.676891	0.992567	1.08E-12	2.49E-09
149	XM_990117 // Rxra // retinoid X receptor alpha // 2 A3|2 17.0 cM // 201	14	6875901	1	0	0.200563	0.644058	0.885963	0.0322425	0.572681	0.866395	1.44E-12	0.00011768
150	NM_013710 // Fgd2 // FYVE, RhoGEF and PH domain containing 2 //	19	6849622	2	1	0.199926	0.0698197	0.0050888	0.0004778	0.990788	0.99883	1.61E-12	1.85E-21
151	NM_031376 // Pik3ap1 // phosphoinositide-3-kinase adaptor protein 1	17	6873133	3	2	0.0202289	0.0003455	0.220762	0.0002266	0.998616	0.83738	2.00E-12	2.94E-18
152	NM_133221 // Slc24a6 // solute carrier family 24 (sodium/potassium/c	19	6933973	3	0	0.546702	0.0159327	0.0129952	0.0055624	0.999902	0.880476	2.35E-12	1.04E-15
153	NM_015736 // Galnt3 // UDP-N-acetyl-alpha-D-galactosamine:polypep	14	6887337	2	1	0.620042	0.0167103	9.12E-09	0.104238	0.924417	0.852396	3.08E-12	1.44E-16
154	NM_028749 // Npl // N-acetylneuraminate pyruvate lyase // 1 G2 // 740	11	6763146	3	1	0.0304618	0.0655201	0.0007402	0.0244784	0.829143	0.825002	3.70E-12	1.11E-06
155	NM_183016 // Cdc42bpb // Cdc42 binding protein kinase beta // 12 F1	42	6803821	4	3	0.0023447	1.94E-05	1.22E-10	2.96E-06	0.632693	0.778043	3.80E-12	4.53E-32
156	NM_010233 // Fn1 // fibronectin 1 // 1 C1-C5|1 36.1 cM // 14268 /// AK	51	6759621	2	1	0.943506	0.0018586	0.75564	0.0002776	0.991441	0.741068	3.82E-12	4.74E-05
157	NM_126166 // Tlr3 // toll-like receptor 3 // 8 B2 // 142980	17	6982102	0	0	0.907714	0.732985	0.745422	0.331723	0.852978	0.542176	4.41E-12	4.85E-25
158	NM_080448 // Srgap3 // SLIT-ROBO Rho GTPase activating protein 3	24	6956509	2	0	0.0110131	0.162468	0.823326	0.034761	0.175087	0.953871	4.64E-12	8.18E-11
159	NM_008528 // Blnk // B-cell linker // 19 C3|19 31.0 cM // 17060 /// AK0	18	6873111	1	1	0.526092	0.0002373	0.187623	0.0924298	0.882716	0.937928	5.47E-12	2.71E-10
160	NM_175181 // 2600010E01Rik // RIKEN cDNA 2600010E01 gene // 2	12	6889197	2	0	0.0650541	0.047913	0.513575	0.0103357	0.767964	0.885687	6.23E-12	1.16E-14
161	NM_017407 // Spag5 // sperm associated antigen 5 // 11 B5|11 44.95	26	6782691	1	0	0.0684361	0.027508	0.551579	0.076701	0.999634	0.421019	6.42E-12	7.65E-09
162	NM_009333 // Tcf7l2 // transcription factor 7-like 2, T-cell specific, HM	18	6870580	3	0	0.263496	0.0094932	0.0098694	0.0011678	0.97697	0.898989	6.46E-12	1.15E-16
163	NM_023132 // Renbp // renin binding protein // X A7.3|X 29.53 cM // 1	13	7017604	0	0	0.0757889	0.555189	0.426313	0.500863	0.975904	0.99183	7.15E-12	2.03E-10
164	NM_021475 // Adamdec1 // ADAM-like, decysin 1 // 14 D1 // 58860 ///	15	6825600	2	0	0.0695356	0.0031835	0.0052477	0.202115	0.962597	0.0502927	1.05E-11	8.41E-16
165	NM_010071 // Dok2 // docking protein 2 // 14 D2-D3|14 39.0 cM // 134	11	6820113	4	2	0.0004076	0.0006989	0.0162168	0.0094636	0.92603	0.5397	1.13E-11	0.00015676
166	NM_007646 // Cd38 // CD38 antigen // 5 B3|5 28.0 cM // 12494 /// AK0	11	6930383	3	1	0.0008098	0.0833889	0.0333641	0.0136903	0.96322	0.935661	1.33E-11	2.26E-05
167	XM_001000290 // Gm150 // gene model 150, (NCBI) // 5 F // 208908 /	30	6934262	3	2	0.0026479	0.937567	4.19E-08	7.01E-07	0.997365	0.999387	1.39E-11	1.17E-05
168	NM_009185 // Stil // Scl/Tal1 interrupting locus // 4 D1 // 20460	19	6916483	1	0	0.752335	0.0039687	0.613065	0.2868	0.998661	0.99701	1.43E-11	8.79E-05
169	NM_010567 // Inppl1 // inositol polyphosphate phosphatase-like 1 //	32	6969886	1	0	0.0374578	0.240882	0.252823	0.070865	0.953983	0.972745	1.50E-11	8.64E-08
170	NM_019793 // Tspan3 // tetraspanin 3 // 9 C // 56434 /// AF242591 // T	10	6995825	0	0	0.495498	0.379261	0.825345	0.996584	0.775304	0.819476	1.53E-11	0.00071374
171	XM_903338 // LOC630222 // similar to Interleukin-2 precursor (IL-2) (T	6	6904309	0	0	0.641808	0.384488	0.356861	0.961858	0.958368	0.217572	1.58E-11	0.00010119
172	NM_013612 // Slc11a1 // solute carrier family 11 (proton-coupled dival	14	6750546	3	2	0.164034	8.35E-07	0.0002412	0.0348006	0.955049	0.771423	1.72E-11	9.15E-31
173	NM_013737 // Pla2g7 // phospholipase A2, group VII (platelet-activatin	14	6850534	3	1	0.0037431	0.0003181	0.0261846	0.592068	0.590442	0.847008	1.76E-11	6.82E-14
174	XM_911132 // LOC635918 // similar to glucosaminyl (N-acetyl) transfe	8	6871874	0	0	0.757326	0.462006	0.89618	0.412394	0.926874	0.971112	1.76E-11	2.65E-09
175	NM_133185 // Rogdi // rogdi homolog (Drosophila) // 16 A1|16 3.3 cM	14	6843680	1	0	0.141349	0.0692976	0.0706415	0.0079296	0.996521	0.969164	1.81E-11	9.46E-12
176	NM_007575 // C2ta // class II transactivator // 16 B1 // 12265 /// XM_9	20	6839340	1	0	0.0708402	0.0154076	0.0614994	0.153188	0.999067	0.995384	1.82E-11	7.64E-21
177	NM_013866 // Zfp385 // zinc finger protein 385 // 15 F3 // 29813	14	6838808	3	1	0.0008734	0.0242476	0.14829	0.0071522	0.975523	0.280583	1.97E-11	2.38E-12
178	NM_009004 // Kif20a // kinesin family member 20A // 18 B1|18 17.0 cM	24	6859953	2	0	0.789057	0.001682	0.669678	0.0230449	0.999975	0.997824	2.00E-11	6.05E-12
179	NM_010612 // Kdr // kinase insert domain protein receptor // 5 C3.3|5	33	6939241	1	0	0.180317	0.241824	0.997819	0.0027968	0.999993	0.994977	2.63E-11	2.16E-06
180	NM_010186 // Fcgr1 // Fc receptor, IgG, high affinity I // 3 F2.1|3 45.2	14	6907262	4	0	0.0418203	0.0163604	0.026965	0.0030235	0.678862	0.947431	3.22E-11	1.04E-05
181	NM_008984 // Ptprm // protein tyrosine phosphatase, receptor type, M	40	6856756	3	2	0.0768984	1.60E-05	0.0234164	8.80E-05	0.981948	0.981333	3.56E-11	4.89E-07
182	XM_487967 // Clec2f // C-type lectin domain family 2, member f // 6 F3	9	6950115	0	0	0.394114	0.128265	0.97722	0.878932	0.178306	0.922892	3.63E-11	4.01E-08
183	NM_015753 // Zfhx1b // zinc finger homeobox 1b // 2 C1 // 24136 /// N	13	6886356	2	0	0.0202895	0.0115824	0.185955	0.718224	0.903494	0.534624	3.90E-11	2.75E-20
184	NM_027290 // Mcm10 // minichromosome maintenance deficient 10 (S	22	6884441	2	0	0.647035	0.0029846	0.242878	0.0121774	0.999485	0.734001	3.98E-11	1.90E-08
185	XM_905874 // LOC631361 // similar to T-cell receptor gamma chain C	7	6805106	0	0	0.808847	0.728264	0.702134	0.988821	0.755465	0.584423	4.23E-11	0.00015385
186	NM_013542 // Gzmb // granzyme B // 14 D3|14 20.5 cM // 14939	7	6824880	2	0	0.225355	0.0052904	0.499649	0.0110681	0.755621	0.376832	6.33E-11	4.82E-07
187	NM_011146 // Pparg // peroxisome proliferator activated receptor gam	12	6949202	1	0	0.245756	0.0048417	0.052568	0.0524844	0.994634	0.934755	6.81E-11	3.69E-09
188	XM_001002757 // Nfam1 // Nfat activating molecule with ITAM motif 1	18	6837415	3	0	0.0018394	0.275764	0.0303097	0.0025194	0.79099	0.994378	6.95E-11	0.00031461
189	NM_016707 // Bcl11a // B-cell CLL/lymphoma 11A (zinc finger protein)	14	6779432	2	0	0.0089637	0.301342	0.223827	0.0151715	0.926861	0.934441	7.91E-11	2.89E-13
190	NM_145491 // Rhoq // ras homolog gene family, member Q // 17 E4|1	11	6852845	2	0	0.0277939	0.350497	0.455223	0.019049	0.744142	0.853942	1.00E-10	5.41E-14
191	NM_024469 // Bhlhb3 // basic helix-loop-helix domain containing, clas	10	6958256	0	0	0.18545	0.648797	0.243894	0.598588	0.99439	0.914296	1.00E-10	0.00030374
192	XR_003523 // LOC545866 // similar to RNA binding motif protein 17 //	19	6884709	1	1	0.124663	0.0003726	0.152897	0.456932	0.884377	0.990117	1.06E-10	1.55E-10
193	NM_007547 // Sirpa // signal-regulatory protein alpha // 2 F3|2 73.1 cM	12	6881139	0	0	0.10752	0.904056	0.195988	0.645144	0.906004	0.984513	1.09E-10	9.00E-20
194	NM_130859 // Card10 // caspase recruitment domain family, member	25	6837046	2	1	0.07879	0.0139246	0.766158	1.33E-06	0.964258	0.377521	1.14E-10	3.00E-06
195	NM_010705 // Lgals3 // lectin, galactose binding, soluble 3 // 14 C1 //	8	6818696	1	0	0.850487	0.0713303	0.0110107	0.540085	0.948321	0.627317	1.29E-10	1.87E-14
196	NM_019804 // B4galt4 // UDP-Gal:betaGlcNAc beta 1,4-galactosyltran	13	6840877	3	0	0.0099424	0.444539	0.0489157	0.0013994	0.471799	0.994776	1.29E-10	1.91E-07
197	XM_001001358 // Pigs // phosphatidylinositol glycan anchor biosynthe	8	6782697	3	0	0.052431	0.0257114	0.004937	0.0470622	0.535716	0.920188	1.41E-10	1.25E-05
198	NM_145141 // Fcrla // Fc receptor-like A // 1 H3 // 98752 /// BC064708	10	6764036	0	0	0.958899	0.651628	0.754256	0.693662	0.915492	0.753394	1.83E-10	2.20E-10
199	NM_008446 // Kif4 // kinesin family member 4 // X C3|X 39.5 cM // 165	34	7012842	2	0	0.515631	0.0087752	0.31516	0.0012469	0.999226	0.969769	2.20E-10	1.98E-13
200	NM_011170 // Prnp // prion protein // 2 F1|2 75.0 cM // 19122 /// NM_0	8	6881291	0	0	0.863841	0.701089	0.977226	0.205062	0.63009	0.860667	2.20E-10	5.05E-06
201	XM_001005892 // LOC673581 // similar to Baculoviral IAP repeat-con	9	6815523	0	0	0.499337	0.778286	0.891543	0.416177	0.862789	0.36551	2.22E-10	1.30E-08
202	BC026950 // Kynu // kynureninase (L-kynurenine hydrolase) // 2 C1 //	17	6876735	0	0	0.777306	0.855641	0.861433	0.1754	0.553674	0.861406	2.30E-10	6.02E-16

Wu Supp Table 8: Highest ranked wt memory vs naive cells Page 3

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Gene Description**	**# Probe Sets**	**Transcript ID**	**Number of cell types p<0.05**	**Number of cell types p<0.001**	**p value wt v thu memory CD4**	**P value wt vs thu naïve CD4**	**p value wt vs thu memory CD8**	**p value wt vs thu naïve CD8**	**p value wt vs thu thymus CD8SP**	**p value wt vs thu thymus CD4SP**	**P-value WT naïve v WT memory CD4**	**P-value WT naïve v WT memory CD8**
203	NM_008088 // Gas7 // growth arrest specific 7 // 11 B3 // 14457 /// XM	16	6781885	0	0	0.958778	0.200806	0.990599	0.557923	0.994454	0.95636	2.32E-10	3.06E-11
204	NM_172900 // Siglecg // sialic acid binding Ig-like lectin G // 7 B3 // 24	16	6960140	3	0	0.0732904	0.0436334	0.0185481	0.0055906	0.990751	0.971415	2.39E-10	1.60E-10
205	NM_007599 // Capg // capping protein (actin filament), gelsolin-like // 9	9	6946993	0	0	0.587148	0.785043	0.997735	0.355891	0.789898	0.999539	2.98E-10	7.33E-07
206	XM_986015 // Tmem163 // transmembrane protein 163 // 1 E3 // 7216	11	6761964	1	0	0.904324	0.0516241	0.255932	0.0101186	0.994419	0.512417	3.08E-10	1.20E-14
207	NM_027810 // Bbs7 // Bardet-Biedl syndrome 7 // 3 B // 71492 /// NM	12	6904300	2	0	0.858072	0.0050781	0.633768	0.0482439	0.992429	0.995213	3.08E-10	2.12E-15
208	NM_177686 // Clec12a // C-type lectin domain family 12, member a //	10	6950137	1	1	0.159209	0.0002407	0.156531	0.767062	0.658356	0.767777	3.18E-10	3.53E-17
209	NM_026082 // Dock7 // dedicator of cytokinesis 7 // 4 C6 // 67299 /// B	36	6923694	1	0	0.0946116	0.0589488	0.0750527	0.0072558	0.983448	0.324563	3.63E-10	4.38E-21
210	NM_009917 // Ccr5 // chemokine (C-C motif) receptor 5 // 9 72.0 cM //	4	6993154	0	0	0.647902	0.0941468	0.958936	0.507434	0.832815	0.947186	3.72E-10	6.10E-08
211	NM_146208 // Neil3 // nei like 3 (E. coli) // 8 B1.3 // 234258 /// NM_02	14	6982550	1	0	0.993455	0.0235519	0.821426	0.567473	0.996063	0.9749	3.76E-10	4.97E-07
212	XM_892197 // LOC627525 // similar to protein tyrosine phosphatase,	17	6888696	3	2	0.0001098	0.0063688	0.13226	8.17E-05	0.654229	0.825822	4.57E-10	5.15E-22
213	NM_019467 // Aif1 // allograft inflammatory factor 1 // 17 B1|17 19.05	8	6855084	1	0	0.742226	0.0024844	0.814166	0.696998	0.525169	0.992875	4.74E-10	7.28E-18
214	NM_007602 // Capn5 // calpain 5 // 7 E1 // 12337 /// BC014767 // Cap	18	6969640	1	1	0.142357	0.533906	0.746811	0.0001424	0.971951	0.405737	4.92E-10	4.20E-06
215	NM_016704 // C6 // complement component 6 // 15 A1|15 3.0 cM // 12	16	6828403	2	2	0.178143	2.36E-05	0.0001083	0.690885	0.895964	0.971878	5.13E-10	3.72E-17
216	NM_144538 // Rab3il1 // RAB3A interacting protein (rabin3)-like 1 // 19	13	6868017	0	0	0.562286	0.0950776	0.969676	0.0857001	0.872991	0.924648	5.14E-10	1.84E-06
217	NM_009848 // Entpd1 // ectonucleoside triphosphate diphosphohydro	11	6869635	1	0	0.496951	0.720694	0.0186916	0.458105	0.765599	0.998259	6.57E-10	2.36E-11
218	XM_357154 // LOC383613 // similar to suppressor of initiator codon m	17	6763972	3	1	0.207308	0.001724	0.027536	6.39E-06	0.991616	0.597859	6.64E-10	4.34E-05
219	NM_144831 // Dhx8 // DEAH (Asp-Glu-Ala-His) box polypeptide 8 // 1	4	6784290	1	0	0.955309	0.625979	0.657386	0.024972	0.894388	0.646627	6.91E-10	1.16E-10
220	NM_194344 // Sh3tc1 // SH3 domain and tetratricopeptide repeats 1 //	19	6937442	2	1	0.0846438	0.441989	0.0033732	3.60E-05	0.984102	0.862429	7.26E-10	1.59E-05
221	NM_009689 // Birc5 // baculoviral IAP repeat-containing 5 // 11 E2 // 1	9	6785307	1	1	0.467963	0.0001602	0.841429	0.33671	0.998537	0.511848	7.43E-10	3.38E-12
222	NM_027758 // Tbc1d9 // TBC1 domain family, member 9 // 8 C3 // 713	22	6977648	1	0	0.209024	0.800047	0.561102	0.0490662	0.950275	0.299129	7.54E-10	8.37E-29
223	NM_029478 // Tmem49 // transmembrane protein 49 // 11 C // 75909 //	11	6790508	1	0	0.0516815	0.00295	0.805226	0.448259	0.636303	0.58301	7.82E-10	0.00014484
224	NM_010653 // Klrc2 // killer cell lectin-like receptor subfamily C, memb	5	6957421	1	1	0.227882	0.201242	0.224311	0.295033	0.950901	0.0005664	7.85E-10	4.80E-08
225	NM_010229 // Flt3 // FMS-like tyrosine kinase 3 // 5 G3|5 82.0 cM // 14	26	6943142	1	0	0.831299	0.66123	0.0029101	0.244798	0.760181	0.99988	8.04E-10	0
226	NM_177713 // 9830130M13Rik // RIKEN cDNA 9830130M13 gene //	17	6816025	1	0	0.568257	0.0866083	0.986967	0.0038232	0.20664	0.773898	8.38E-10	2.01E-07
227	NM_009275 // Srprb // signal recognition particle receptor, B subunit //	15	6998397	2	0	0.274208	0.0025307	0.112118	0.0118407	0.621677	0.969236	8.41E-10	5.61E-25
228	NM_133198 // Pygl // liver glycogen phosphorylase // 12 C2|12 30.0 cM	20	6801506	2	1	0.223332	5.04E-07	0.308693	0.0021732	0.87803	0.947206	8.46E-10	1.62E-07
229	NM_010208 // Fgr // Gardner-Rasheed feline sarcoma viral (Fgr) onco	14	6917549	2	0	0.2097	0.309645	0.0337974	0.0407962	0.983549	0.904903	8.50E-10	3.90E-09
230	NM_011369 // Shcbp1 // Shc SH2-domain binding protein 1 // 8 A1.2 //	13	6980158	1	0	0.956831	0.0120019	0.0738787	0.265212	0.251167	0.744183	9.17E-10	5.32E-06
231	NM_013515 // Stom // stomatin // 2 B-C1|2 24.0 cM // 13830	12	6886039	0	0	0.790255	0.290585	0.373562	0.13355	0.816598	0.253974	1.04E-09	1.08E-14
232	XM_989161 // LOC671796 // hypothetical protein LOC671796 // // 67	16	6900994	1	0	0.159164	0.106641	0.403876	0.0090545	0.913549	0.923188	1.05E-09	7.83E-07
233	NM_008903 // Ppap2a // phosphatidic acid phosphatase 2a // 13 D2.2	11	6810299	2	0	0.305635	0.0213926	0.306101	0.0018218	0.773683	0.661119	1.09E-09	9.43E-18
234	XM_001003962 // LOC673056 // similar to Baculoviral IAP repeat-con	11	6815535	0	0	0.884985	0.161712	0.863374	0.078333	0.445118	0.137119	1.14E-09	5.68E-10
235	NM_029394 // Snx24 // sorting nexing 24 // 18 D1 // 69226 /// AK0176	13	6860929	2	0	0.0172004	0.296106	0.0058551	0.745323	0.985718	0.964354	1.15E-09	5.98E-07
236	XM_974248 // Klri2 // killer cell lectin-like receptor family I member 2 //	7	6957428	0	0	0.745971	0.723569	0.247287	0.327663	0.401382	0.899398	1.33E-09	3.58E-08
237	NM_146061 // Arhgap8 // Rho GTPase activating protein 8 // 15 E // 1	11	6832321	0	0	0.0739209	0.353677	0.550484	0.0893989	0.993132	0.56637	1.45E-09	2.25E-11
238	NM_001039150 // Cd44 // CD44 antigen // 2 E2|2 56.0 cM // 12505 ///	13	6889258	0	0	0.0683225	0.278255	0.795578	0.0531878	0.958071	0.397417	1.48E-09	3.16E-10
239	NM_011352 // Sema7a // sema domain, immunoglobulin domain (Ig),	15	6989438	2	2	0.106419	0.0007779	0.11323	0.0001122	0.553097	0.827743	1.54E-09	1.92E-10
240	NM_008587 // Mertk // c-mer proto-oncogene tyrosine kinase // 2 F1 //	21	6881087	3	1	0.457598	0.000168	0.0098912	0.0018009	0.999913	0.98589	1.94E-09	2.45E-12
241	NM_054054 // Brdt // bromodomain, testis-specific // 5 E5 // 114642 ///	19	6933251	2	1	0.001386	0.0652291	0.0004564	0.3103	0.996075	0.99993	1.98E-09	0.0009583
242	NM_011578 // Tgfbr3 // transforming growth factor, beta receptor III //	18	6940841	1	1	0.135101	5.77E-05	0.881369	0.925943	0.627743	0.871823	2.19E-09	5.29E-05
243	XM_986629 // Itgad // integrin, alpha D // 7 F3 // 381924 /// NM_02133	28	6964382	0	0	0.607762	0.995075	0.628049	0.422906	0.987345	0.989061	2.66E-09	6.26E-43
244	NM_007548 // Prdm1 // PR domain containing 1, with ZNF domain // 1	14	6773655	0	0	0.194587	0.0871177	0.871806	0.78173	0.989757	0.98612	2.73E-09	1.59E-05
245	NM_053099 // Setbp1 // SET binding protein 1 // 18 E3 // 240427 /// AI	10	6866862	0	0	0.439557	0.35766	0.643767	0.291012	0.96402	0.819705	3.10E-09	7.68E-05
246	NM_030249 // Cttnbp2nl // CTTNBP2 N-terminal like // 3 F2.2 // 80281	14	6907888	0	0	0.242528	0.269923	0.191014	0.621281	0.881746	0.926953	3.30E-09	5.56E-11
247	NM_001035531 // Adrbk2 // adrenergic receptor kinase, beta 2 // 5 F|5	18	6941146	1	0	0.916826	0.994978	0.0144194	0.637177	0.80395	0.715237	3.90E-09	5.74E-38
248	XM_988271 // LOC626703 // hypothetical LOC626703 // 6 G1 // 6267	7	6957679	3	0	0.0010889	0.0372855	0.0347552	0.630487	0.0567214	0.991875	3.90E-09	4.57E-11
249	NM_133809 // Kmo // kynurenine 3-monooxygenase (kynurenine 3-hy	19	6755378	0	0	0.408514	0.975627	0.522416	0.147415	0.984479	0.998423	4.00E-09	6.42E-14
250	NM_139138 // Emr4 // EGF-like module containing, mucin-like, hormo	17	6851186	1	0	0.14269	0.0327035	0.287358	0.794281	0.990394	0.896706	4.28E-09	9.52E-22
251	NM_080555 // Ppap2b // phosphatidic acid phosphatase type 2B // 4 C	8	6916023	1	1	0.0001973	0.0823957	0.094227	0.349406	0.253466	0.377552	4.42E-09	0.00066042
252	NM_009915 // Ccr2 // chemokine (C-C motif) receptor 2 // 9 F|9 71.9 c	4	6993153	0	0	0.174835	0.481603	0.509098	0.260985	0.669269	0.513395	4.49E-09	9.69E-05
253	NM_011216 // Ptpro // protein tyrosine phosphatase, receptor type, O	32	6950539	0	0	0.992264	0.671914	0.89669	0.666387	0.868022	0.989429	4.97E-09	3.91E-14
254	NM_019753 // Cdh17 // cadherin 17 // 4 A2 // 12557	20	6911729	1	0	0.472785	0.142383	0.595863	0.0014745	0.993177	0.674558	5.25E-09	1.22E-05
255	NM_025730 // Lrrk2 // leucine-rich repeat kinase 2 // 15 F1 // 66725 ///	48	6832713	3	1	0.0464883	0.0228264	0.0002728	0.0718531	0.999914	0.995776	5.28E-09	3.72E-33
256	NM_144549 // Trib1 // tribbles homolog 1 (Drosophila) // 15 D1 // 2117	18	6830770	1	0	0.0633948	0.531349	0.974889	0.0413616	0.858316	0.820891	5.41E-09	3.93E-14
257	NM_207530 // Osbpl1a // oxysterol binding protein-like 1A // 18 A2 // 6	37	6863467	4	1	0.0142533	0.0001024	0.0425186	0.0042277	0.999987	0.991067	5.68E-09	0.00061917
258	NM_010872 // Birc1b // baculoviral IAP repeat-containing 1b // 13 D1	15	6815522	1	0	0.137493	0.045431	0.15079	0.602588	0.831882	0.913084	5.90E-09	2.69E-06
259	NM_013484 // C2 // complement component 2 (within H-2S) // 17 B1	17	6855062	1	0	0.394845	0.969267	0.295377	0.0410734	0.713794	0.419703	6.46E-09	0.00029201
260	XM_891957 // Cdkn3 // cyclin-dependent kinase inhibitor 3 // 14 C1 //	10	6818647	1	0	0.206224	0.0128032	0.860434	0.0562001	0.964657	0.98025	6.93E-09	5.43E-12
261	XM_982884 // LOC666300 // similar to thymosin, beta 10 // 19 A // 666	15	6876543	1	0	0.355464	0.309685	0.287219	0.0477367	0.995634	0.996444	7.91E-09	1.51E-21
262	NM_008505 // Lmo2 // LIM domain only 2 // 2 E2|2 60.0 cM // 16909 //	8	6879637	0	0	0.0602338	0.395419	0.594352	0.310697	0.997291	0.839408	8.34E-09	3.23E-07
263	NM_023066 // Asph // aspartate-beta-hydroxylase // 4 A1 // 65973 /// N	34	6919596	2	1	0.375284	1.62E-05	0.580082	0.001215	0.873401	0.952357	8.64E-09	1.45E-07
264	NM_001024617 // Inpp4b // inositol polyphosphate-4-phosphatase, ty	25	6977589	0	0	0.714462	0.131931	0.930215	0.984623	0.999078	0.824406	9.12E-09	1.36E-15
265	NM_010778 // Cd46 // CD46 antigen, complement regulatory protein //	24	6765460	3	3	9.05E-05	0.0002934	0.0004712	0.943393	0.99631	0.990708	9.32E-09	4.22E-10
266	XM_978684 // LOC670042 // similar to Probable ubiquitin carboxyl-ter	11	6979527	2	0	0.254457	0.0033621	0.0068403	0.84622	0.515636	0.999452	9.33E-09	1.94E-08
267	NM_010652 // Klrc1 // killer cell lectin-like receptor subfamily C, memb	4	6957423	1	0	0.206927	0.0487724	0.289551	0.75655	0.802425	0.450853	9.51E-09	9.30E-09
268	NM_015744 // Enpp2 // ectonucleotide pyrophosphatase/phosphodies	28	6835759	2	2	0.0849183	0.0001924	0.596624	3.09E-10	0.10694	0.982837	1.02E-08	8.95E-05
269	NM_029612 // Slamf9 // SLAM family member 9 // 1 H3 // 98365	7	6755240	2	0	0.556011	0.0044831	0.667501	0.0144959	0.753126	0.922888	1.05E-08	0.00042809

Wu Supp Table 8: Highest ranked wt memory vs naive cells Page 4

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Gene Description**	**# Probe Sets**	**Transcript ID**	**Number of cell types p<0.05**	**Number of cell types p<0.001**	**p value wt v thu memory CD4**	**P value wt vs thu naïve CD4**	**p value wt vs thu memory CD8**	**p value wt vs thu naïve CD8**	**p value wt vs thu thymus CD8SP**	**p value wt vs thu thymus CD4SP**	**P-value WT naïve v WT memory CD4**	**P-value WT naïve v WT memory CD8**
270	NM_172578 // C79407 // expressed sequence C79407 // 12 C1 // 217	18	6801335	1	0	0.24907	0.0104123	0.377806	0.191983	0.997181	0.999911	1.09E-08	2.06E-05
271	NM_175088 // Mdfic // MyoD family inhibitor domain containing // 6 A1	8	6944293	1	0	0.317706	0.941624	0.764597	0.0190324	0.306998	0.713366	1.24E-08	2.32E-09
272	XM_920821 // Ppfibp1 // PTPRF interacting protein, binding protein 1	30	6951118	1	0	0.755318	0.134425	0.0442067	0.0561783	0.989446	0.766949	1.25E-08	0.00021242
273	NM_008337 // Ifng // interferon gamma // 10 D2|10 67.0 cM // 15978	5	6771052	0	0	0.506075	0.107592	0.0761136	0.136946	0.57624	0.92522	1.29E-08	1.45E-06
274	NM_134250 // Havcr2 // hepatitis A virus cellular receptor 2 // 11 B1.1	11	6780551	0	0	0.383644	0.88648	0.224724	0.389806	0.204036	0.429682	1.47E-08	3.73E-24
275	NM_172488 // 9030625A04Rik // RIKEN cDNA 9030625A04 gene // 1	8	6826116	0	0	0.169843	0.059797	0.0721224	0.279031	0.211722	0.887787	1.48E-08	4.34E-17
276	NM_181547 // Nostrin // nitric oxide synthase trafficker // 2 C2 // 3294	15	6877901	3	0	0.0275379	0.048719	0.0055929	0.684342	0.134328	0.547083	1.50E-08	7.02E-10
277	NM_012057 // Irf5 // interferon regulatory factor 5 // 6 A3.3|6 7.1 cM //	11	6945011	1	1	0.48642	0.0009979	0.654803	0.373851	0.973492	0.997344	1.52E-08	2.37E-09
278	NM_133721 // Itga9 // integrin alpha 9 // 9 F3|9 67.0 cM // 104099 /// A	28	6992855	4	2	0.001191	0.0045358	0.0003231	0.0001679	0.514082	0.898803	1.53E-08	9.50E-08
279	NM_207245 // BC066107 // cDNA sequence BC066107 // 17 B1 // 240	7	6854907	0	0	0.491024	0.72886	0.878453	0.278494	0.979254	0.765697	1.81E-08	2.97E-07
280	XM_983668 // LOC675618 // similar to ribosomal protein L31 // // 675	8	6998919	1	1	0.0700539	0.854022	0.592602	2.94E-05	0.996478	0.848882	1.93E-08	6.33E-07
281	NM_021451 // Pmaip1 // phorbol-12-myristate-13-acetate-induced pro	7	6861662	1	0	0.264225	0.452303	0.329532	0.0103802	0.957008	0.816837	2.12E-08	6.16E-10
282	NM_011820 // Ggtla1 // gamma-glutamyltransferase-like activity 1 // 1	18	6768868	0	0	0.0656955	0.0588271	0.813426	0.190252	0.95677	0.938634	2.34E-08	0.00023413
283	NM_175384 // Cdca2 // cell division cycle associated 2 // 14 D1 // 108	14	6825544	2	0	0.0015082	0.0030175	0.14689	0.248982	0.999844	0.985112	2.42E-08	6.09E-06
284	XM_001000500 // Cd244 // CD244 natural killer cell receptor 2B4 // 1	10	6755189	0	0	0.691086	0.99867	0.828468	0.192175	0.895828	0.408799	2.69E-08	2.03E-18
285	NM_010188 // Fcgr3 // Fc receptor, IgG, low affinity III // 1 H3|1 92.3 c	6	6764040	1	0	0.220739	0.0405213	0.221995	0.81563	0.941702	0.284291	2.80E-08	1.25E-07
286	NM_013584 // Lifr // leukemia inhibitory factor receptor // 15 A1|15 4.6	21	6828522	1	0	0.572433	0.0209196	0.634198	0.124009	0.999767	0.967044	2.82E-08	1.67E-08
287	NM_008175 // Grn // granulin // 11 D|11 60.0 cM // 14824	16	6784345	1	0	0.899021	0.0940463	0.0441412	0.760131	0.915236	0.773287	2.85E-08	5.73E-11
288	NM_022420 // Gprc5b // G protein-coupled receptor, family C, group 5	8	6970857	0	0	0.342292	0.488012	0.516837	0.165961	0.846652	0.394897	2.88E-08	7.25E-07
289	NM_024124 // Hdac9 // histone deacetylase 9 // 12 A3 // 79221 /// AK1	12	6800020	2	0	0.0300861	0.311517	0.0386505	0.080619	0.673868	0.978736	2.88E-08	1.97E-11
290	NM_152803 // Hpse // heparanase // 5 E4 // 15442 /// AK087283 // Hp	14	6940363	2	0	0.4979	0.0096244	0.342231	0.0286106	0.997963	0.951877	2.97E-08	2.48E-08
291	NM_134471 // Kif2c // kinesin family member 2C // 4 D2.1 // 73804 ///	14	6924834	0	0	0.349274	0.123869	0.314781	0.187393	0.992049	0.989912	3.17E-08	3.16E-06
292	NM_181588 // 2310016A09Rik // RIKEN cDNA 2310016A09 gene // 1	7	6829549	2	0	0.129794	0.029581	0.164015	0.0037358	0.855651	0.92561	3.81E-08	2.79E-05
293	NM_011269 // Rhag // Rhesus blood group-associated A glycoprotein	12	6850412	3	0	0.0180752	0.003342	0.185555	0.0021735	0.890333	0.490652	4.04E-08	9.57E-06
294	NM_009500 // Vav2 // Vav2 oncogene // 2 A3|2 15.3 cM // 22325 /// BC	27	6885530	1	0	0.0478507	0.376982	0.244685	0.0594184	0.99096	0.90034	4.19E-08	1.74E-17
295	NM_172723 // Centa1 // centaurin, alpha 1 // 5 G2 // 231821 /// AK076	11	6942669	0	0	0.0696818	0.342989	0.334738	0.101129	0.748938	0.722387	4.52E-08	5.05E-08
296	NM_177388 // Slc41a2 // solute carrier family 41, member 2 // 10 C1 //	12	6775559	0	0	0.236407	0.590411	0.971762	0.212145	0.979368	0.829141	4.78E-08	1.64E-12
297	NM_010194 // Fes // feline sarcoma oncogene // 7 D2|7 39.0 cM // 14	21	6968780	1	0	0.717447	0.137338	0.278492	0.0133327	0.977293	0.99466	5.11E-08	6.60E-15
298	NM_176841 // A430106J12Rik // RIKEN cDNA A430106J12 gene // 11	14	6779714	0	0	0.802448	0.433814	0.245354	0.341368	0.162082	0.849467	5.25E-08	1.20E-15
299	XR_003830 // LOC623748 // similar to cytochrome P450, family 4, sut	11	6983299	0	0	0.103828	0.145025	0.890177	0.130801	0.99045	0.91553	5.29E-08	4.68E-07
300	NM_001033405 // Treml2 // triggering receptor expressed on myeloid	8	6850821	1	0	0.71892	0.084858	0.0235524	0.0790885	0.560277	0.908407	5.47E-08	3.76E-24
301	NM_009862 // Cdc45l // cell division cycle 45 homolog (S. cerevisiae)	22	6844359	3	2	0.530433	0.0001927	0.0014021	5.51E-05	0.999195	0.746951	5.48E-08	0.0001272
302	NM_010101 // Edg3 // endothelial differentiation, sphingolipid G-protei	5	6807007	0	0	0.407483	0.259325	0.118928	0.655358	0.837823	0.651982	5.90E-08	0.00082351
303	NM_145586 // Tmem159 // transmembrane protein 159 // 7 F2 // 2338	6	6963907	1	0	0.527578	0.156601	0.330042	0.0408204	0.477805	0.672405	6.50E-08	0.00028126
304	XM_908518 // 2310047M10Rik // RIKEN cDNA 2310047M10 gene // 1	13	6781970	0	0	0.380957	0.64553	0.819595	0.270794	0.970076	0.934547	6.80E-08	3.68E-05
305	NM_008946 // Psmb6 // proteasome (prosome, macropain) subunit, b	28	6782134	0	0	0.960875	0.525989	0.0514292	0.574611	0.987388	0.999115	6.81E-08	0.00010447
306	NM_028778 // Nuak2 // NUAK family, SNF1-like kinase, 2 // 1 E4 // 74	16	6753110	3	1	0.0005512	0.0487939	0.144737	0.0027577	0.953054	0.892614	7.14E-08	4.48E-08
307	NM_001004140 // Ckap2 // cytoskeleton associated protein 2 // 8 A2 //	9	6980944	0	0	0.177729	0.0749756	0.583661	0.254365	0.886926	0.924837	7.17E-08	3.79E-05
308	XM_904393 // LOC630776 // similar to Nidogen precursor (Entactin) //	24	6804893	1	0	0.516447	0.36915	0.0125615	0.263957	0.992612	0.984978	7.23E-08	2.56E-10
309	NM_018866 // Cxcl13 // chemokine (C-X-C motif) ligand 13 // 5 E3 // 5	5	6932571	1	0	0.0410132	0.788281	0.494483	0.189573	0.912201	0.873533	7.52E-08	3.23E-05
310	XM_894169 // RP23-320E6.1 // hypothetical protein LOC629303 // 11	5	6790291	0	0	0.762905	0.26647	0.921343	0.754326	0.795438	0.392084	7.59E-08	1.11E-05
311	NM_133990 // Il13ra1 // interleukin 13 receptor, alpha 1 // X A3.2|X 12	13	7010645	0	0	0.0990419	0.0789527	0.469333	0.374456	0.862424	0.795569	7.95E-08	2.91E-09
312	NM_007659 // Cdc2a // cell division cycle 2 homolog A (S. pombe) // 1	9	6774794	1	0	0.649087	0.0234552	0.183714	0.176094	0.960624	0.986629	8.02E-08	5.14E-12
313	NM_009027 // Rasgrf2 // RAS protein-specific guanine nucleotide-rele	23	6815027	1	0	0.167481	0.0153265	0.167298	0.219271	0.99587	0.828478	8.95E-08	1.47E-05
314	NM_053082 // Tspan4 // tetraspanin 4 // 7 F5 // 64540 /// AK188368 //	8	6965206	1	0	0.805383	0.452602	0.500862	0.0027486	0.963678	0.953689	9.45E-08	1.22E-13
315	NM_025831 // 1300014I06Rik // RIKEN cDNA 1300014I06 gene // 13	8	6812250	0	0	0.547028	0.539491	0.357963	0.695556	0.994731	0.931745	1.13E-07	3.20E-05
316	NM_011355 // Sfpi1 // SFFV proviral integration 1 // 2 E3|2 47.5 cM //	7	6879020	3	1	0.0046885	0.117687	0.0032918	9.54E-05	0.480792	0.671305	1.17E-07	3.69E-09
317	NM_130904 // Cd209d // CD209d antigen // 8 A1.1 // 170779	8	6980101	2	0	0.013698	0.0128575	0.997774	0.0500991	0.477001	0.145502	1.17E-07	4.60E-08
318	NM_008535 // Lyl1 // lymphoblastomic leukemia // 8 C3|8 38.5 cM // 1	8	6977758	1	0	0.0038707	0.102622	0.121457	0.0998894	0.931341	0.381236	1.28E-07	5.56E-07
319	NM_020258 // Slc37a2 // solute carrier family 37 (glycerol-3-phosphat	21	6994678	4	3	5.80E-05	2.35E-06	8.41E-10	0.0014634	0.78845	0.139502	1.40E-07	7.79E-10
320	NM_133995 // Upb1 // ureidopropionase, beta // 10 C1 // 103149	11	6768860	0	0	0.737849	0.183788	0.937786	0.0794272	0.974742	0.761783	1.43E-07	4.60E-05
321	NM_027533 // Tspan2 // tetraspanin 2 // 3 F3 // 70747 /// AK020982 //	10	6900082	0	0	0.25021	0.867022	0.177938	0.924958	0.759256	0.184201	1.49E-07	1.67E-09
322	NM_028039 // Esco2 // establishment of cohesion 1 homolog 2 (S. cer	12	6825436	2	0	0.194962	0.015808	0.830281	0.0322199	0.999913	0.620248	1.49E-07	5.74E-06
323	NM_172133 // Centa2 // centaurin, alpha 2 // 11 B5|11 47.24 cM // 216	10	6782808	1	0	0.28626	0.120173	0.0180521	0.364474	0.748554	0.48556	1.55E-07	1.11E-06
324	NM_146260 // Tmie // transmembrane inner ear // 9 F3|9 64.0 cM // 20	8	6998893	3	2	2.71E-05	0.0001712	0.196702	0.0013398	0.735279	0.499246	1.78E-07	8.48E-05
325	NM_011309 // S100a1 // S100 calcium binding protein A1 // 3 F1-F2|3	6	6906903	1	0	0.0054162	0.0907886	0.762533	0.266914	0.938545	0.944379	1.79E-07	4.65E-05
326	NM_008357 // Il15 // interleukin 15 // 8 C2|8 38.0 cM // 16168 /// AB02	10	6983744	0	0	0.0977679	0.683418	0.23672	0.937673	0.76289	0.410212	1.86E-07	9.22E-10
327	NM_011074 // Pftk1 // PFTAIRE protein kinase 1 // 5 A1|5 0.0 cM // 18	16	6935970	2	0	0.0698463	0.028562	0.0862146	0.0025996	0.243999	0.68198	1.90E-07	8.17E-28
328	NM_029364 // Gns // glucosamine (N-acetyl)-6-sulfatase // 10 D2 // 75	20	6771224	4	2	0.005151	1.39E-05	0.0050028	3.87E-06	0.787379	0.486948	1.92E-07	9.51E-09
329	NM_018755 // Pgcp // plasma glutamate carboxypeptidase // 15 B3.2	11	6829612	1	0	0.224955	0.785717	0.908875	0.0556343	0.0465979	0.111529	1.96E-07	3.77E-05
330	NM_175074 // Hmgn3 // high mobility group nucleosomal binding dom	8	6997349	4	0	0.156763	0.013771	0.0091699	0.0011569	0.927912	0.0434	1.99E-07	4.85E-15
331	NM_183318 // Rgag4 // retrotransposon gag domain containing 4 // X	11	7018576	2	1	0.0106364	0.32671	0.0548592	0.0001824	0.229188	0.354894	2.26E-07	5.52E-09
332	NM_022018 // Niban // niban protein // 1 G2 // 63913 /// AK035217 // N	14	6754027	0	0	0.724222	0.500827	0.154894	0.784164	0.999665	0.999731	2.53E-07	2.10E-10
333	NM_199225 // Cd300c // CD300C antigen // 11 E2 // 387565	5	6792368	0	0	0.993256	0.398863	0.934482	0.844977	0.700144	0.835586	2.65E-07	0.00012475
334	NM_001033435 // Gm885 // gene model 885, (NCBI) // 11 E1 // 38073	8	6923059	0	0	0.0857482	0.529514	0.872336	0.799024	0.923938	0.608123	2.70E-07	1.74E-11
335	NM_181320 // Dusp16 // dual specificity phosphatase 16 // 6 G1|6 62.	14	6957640	1	0	0.337094	0.0926108	0.36184	0.0084407	0.883751	0.408713	2.83E-07	1.43E-09
336	XM_993285 // Igk-V21-9 // immunoglobulin kappa chain variable 21 (V	6	6946835	0	0	0.282451	0.395344	0.120279	0.980688	0.935971	0.944697	2.89E-07	4.17E-06

Wu Supp Table 8: Highest ranked wt memory vs naive cells

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Gene Description**	**# Probe Sets**	**Transcript ID**	**Number of cell types p<0.05**	**Number of cell types p<0.001**	**p value wt v thu memory CD4**	**P value wt vs thu naïve CD4**	**p value wt vs thu memory CD8**	**p value wt vs thu naïve CD8**	**p value wt vs thu thymus CD8SP**	**p value wt vs thu thymus CD4SP**	**P-value WT naïve v WT memory CD4**	**P-value WT naïve v WT memory CD8**
337	NM_021542 // Kcnk5 // potassium channel, subfamily K, member 5 //	8	6823011	0	0	0.747708	0.701586	0.753441	0.113586	0.102791	0.554778	2.91E-07	1.13E-08
338	NM_011693 // Vcam1 // vascular cell adhesion molecule 1 // 3 G1|3 50	15	6908486	2	1	0.329967	0.0005771	0.420076	0.0272698	0.360455	0.999978	2.95E-07	1.97E-21
339	NM_020590 // Gabarapl1 // gamma-aminobutyric acid (GABA(A)) rece	8	6950168	1	0	0.780758	0.198413	0.964469	0.0074592	0.966109	0.283293	3.16E-07	4.09E-14
340	NM_175606 // Hod // homeobox only domain // 5 D // 74318 /// BC024	5	6939338	0	0	0.78164	0.426296	0.678831	0.779963	0.988796	0.979943	3.25E-07	0.00032475
341	NM_009824 // Cbfa2t3h // core-binding factor, runt domain, alpha sub	13	6985946	0	0	0.139598	0.0934973	0.223598	0.0856034	0.974624	0.945126	3.40E-07	5.63E-13
342	NM_175316 // Slco2b1 // solute carrier organic anion transporter famil	16	6969753	2	0	0.0228452	0.10229	0.0018864	0.0534956	0.982361	0.999895	3.59E-07	1.61E-10
343	NM_172732 // Clec9a // C-type lectin domain family 9, member a // 6	10	6950148	0	0	0.543155	0.0646563	0.787817	0.761373	0.881509	0.97585	3.61E-07	1.80E-19
344	AK045150 // B130040O20Rik // RIKEN cDNA B130040O20 gene // 11	13	6880508	2	0	0.192003	0.332706	0.0053449	0.0108234	0.78384	0.82629	4.40E-07	6.53E-27
345	NM_153513 // BC028528 // cDNA sequence BC028528 // 3 F2.1 // 22	7	6907222	2	0	0.0073699	0.0089566	0.36343	0.180795	0.405578	0.897989	4.46E-07	5.13E-09
346	NM_008360 // Il18 // interleukin 18 // 9 A5.3|9 29.0 cM // 16173 /// AK0	7	6989015	2	0	0.560971	0.0016621	0.0396334	0.629589	0.516541	0.542877	4.76E-07	3.31E-09
347	NM_177547 // Sgk3 // serum/glucocorticoid regulated kinase 3 // 1 A2	18	6747497	2	0	0.0035063	0.0417684	0.617944	0.914339	0.810939	0.22249	5.47E-07	1.04E-05
348	NM_170702 // Cd40 // CD40 antigen // 2 H3|2 97.0 cM // 21939 /// NM	13	6883132	1	0	0.059479	0.0850535	0.0944859	0.0378738	0.983278	0.945691	5.57E-07	3.71E-09
349	NM_194355 // Spire1 // spire homolog 1 (Drosophila) // 18 E1 // 68166	21	6866305	0	0	0.993887	0.790556	0.553446	0.270845	0.999883	0.985903	5.77E-07	5.80E-06
350	NM_053094 // Cd163 // CD163 antigen // 6 F2 // 93671 /// AK032012 /	20	6949766	0	0	0.53287	0.191496	0.449275	0.534403	0.988461	0.137913	5.99E-07	7.58E-12
351	NM_025422 // Cd302 // CD302 antigen // 2 C1.1 // 66205 /// XR_0039	7	6887079	1	0	0.12043	0.0964111	0.0087185	0.505794	0.995811	0.993226	6.59E-07	4.35E-09
352	NM_011520 // Sdc3 // syndecan 3 // 4 D2.3|4 60.8 cM // 20970 /// AK0	11	6917389	4	0	0.0039114	0.0081624	0.003335	0.0054826	0.626897	0.97915	6.63E-07	2.17E-24
353	XM_975975 // Coro2a // coronin, actin binding protein 2A // 4 B1 // 107	20	6921381	1	0	0.365404	0.686599	0.539221	0.0416668	0.701091	0.997018	6.64E-07	1.15E-09
354	XM_905223 // LOC631417 // similar to Interferon-activatable protein 2	7	6764246	0	0	0.213421	0.236983	0.0807404	0.871348	0.96341	0.955724	6.93E-07	1.93E-15
355	XR_002483 // LOC669350 // similar to toll-like receptor 9 // // 669350	9	6992159	3	0	0.205048	0.0078148	0.0338224	0.0033336	0.981429	0.90557	6.95E-07	3.29E-10
356	NM_008737 // Nrp1 // neuropilin 1 // 8 E|8 73.0 cM // 18186 /// AK135	20	6980016	0	0	0.386762	0.88422	0.168537	0.854488	0.998641	0.878136	7.28E-07	9.56E-19
357	NM_029494 // Rab30 // RAB30, member RAS oncogene family // 7 E1	8	6962577	0	0	0.901647	0.405793	0.265023	0.154862	0.6441	0.719818	7.41E-07	5.10E-14
358	NM_013748 // Clnk // cytokine-dependent hematopoietic cell linker // 5	13	6937617	0	0	0.12571	0.195329	0.859572	0.598134	0.828924	0.450333	7.59E-07	1.03E-13
359	NM_009613 // Adam11 // a disintegrin and metalloppeptidase domain 1	27	6784363	1	0	0.522975	0.077431	0.304958	0.0353597	0.177105	0.991495	7.89E-07	1.68E-06
360	NM_029097 // Atp13a2 // ATPase type 13A2 // 4 D3 // 74772 /// XM_9	31	6918129	3	1	0.0761357	0.0094429	5.53E-05	0.004036	0.986225	0.99991	7.90E-07	0.00020265
361	NM_013820 // Hk2 // hexokinase 2 // 6 C3|6 34.5 cM // 15277 /// X966	25	6954982	0	0	0.801752	0.669921	0.996125	0.549057	0.339011	0.882394	9.14E-07	3.55E-15
362	XM_994774 // LOC668178 // similar to UbiE-YGHL1 fusion protein //	14	6833232	0	0	0.106762	0.672371	0.0531578	0.635227	0.788888	0.790073	9.20E-07	2.07E-05
363	NM_172277 // Snx8 // sorting nexin 8 // 5 G2 // 231834 /// BC019142 /	14	6942716	1	0	0.0809447	0.0947545	0.0472038	0.312111	0.677991	0.96178	9.93E-07	6.85E-08
364	NM_026376 // Plxnd1 // plexin D1 // 6 E3 // 67784	44	6956748	3	1	0.0080783	0.0034019	0.131106	0.0001671	0.949433	1	1.00E-06	0.00033996
365	NM_147219 // Abca5 // ATP-binding cassette, sub-family A (ABC1), m	38	6792129	3	0	0.0433674	0.0095511	0.250937	0.0261285	0.905824	0.999847	1.03E-06	0.00056103
366	NM_153197 // Clec4a3 // C-type lectin domain family 4, member a3 //	9	6949727	1	0	0.075502	0.0715542	0.739809	0.760419	0.0449096	0.801989	1.06E-06	2.81E-09
367	BC096435 // Cd74 // CD74 antigen (invariant polypeptide of major hist	12	6861341	3	1	0.0148917	0.0070593	8.77E-05	0.267549	0.986192	0.597209	1.10E-06	9.49E-08
368	NM_018752 // Trpm1 // transient receptor potential cation channel, su	22	6961099	2	1	0.965731	1.68E-06	0.0089631	0.318907	0.987952	0.998682	1.12E-06	4.75E-12
369	NM_008361 // Il1b // interleukin 1 beta // 2 F|2 73.0 cM // 16176 /// AK	8	6890838	0	0	0.367828	0.675126	0.18548	0.831851	0.391991	0.750786	1.12E-06	2.35E-07
370	NM_011461 // Spic // Spi-C transcription factor (Spi-1/PU.1 related) //	7	6775864	3	1	0.0108593	0.0003339	0.0657298	0.120377	0.0229601	0.236892	1.19E-06	3.38E-08
371	NM_007897 // Ebf1 // early B-cell factor 1 // 11 B1.1|11 20.0 cM // 135	15	6780443	0	0	0.19026	0.670243	0.620477	0.751903	0.999989	0.846403	1.21E-06	2.51E-07
372	NM_023056 // 1810009M01Rik // RIKEN cDNA 1810009M01 gene // 6	10	6953429	0	0	0.709343	0.277668	0.735288	0.409462	0.99939	0.978905	1.22E-06	6.90E-10
373	NM_145402 // Tmem51 // transmembrane protein 51 // 4 E1 // 214359	5	6926504	1	0	0.213073	0.691791	0.0089992	0.0502315	0.911292	0.991652	1.25E-06	0.00011072
374	NM_031181 // Siglece // sialic acid binding Ig-like lectin E // 7 B2 // 83	12	6966818	2	0	0.0304912	0.0336889	0.0985523	0.221193	0.784465	0.972373	1.28E-06	9.01E-10
375	NM_024217 // Cmtm3 // CKLF-like MARVEL transmembrane domain	5	6978736	1	0	0.0021929	0.436129	0.0811443	0.438789	0.953934	0.631552	1.28E-06	4.93E-09
376	NM_173182 // Fndc3b // fibronectin type III domain containing 3B // 3	31	6903779	1	0	0.139912	0.174797	0.0147281	0.124523	0.987302	0.97161	1.28E-06	1.33E-07
377	XM_980282 // LOC675222 // similar to basic transcription factor 3 //	8	6785111	1	0	0.989951	0.868249	0.0050335	0.447642	0.903666	0.716455	1.32E-06	5.75E-16
378	XM_975780 // Anubl1 // AN1, ubiquitin-like, homolog (Xenopus laevis)	10	6949272	0	0	0.732317	0.279838	0.47607	0.160172	0.215562	0.363687	1.42E-06	0.00034046
379	NM_023063 // Lima1 // LIM domain and actin binding 1 // 15 F1|15 60	11	6838492	0	0	0.34063	0.0775645	0.0588416	0.17632	0.810969	0.325793	1.46E-06	4.49E-16
380	NM_133851 // Nusap1 // nucleolar and spindle associated protein 1 //	11	6880529	1	0	0.393833	0.106851	0.5892	0.0025972	0.999408	0.998579	1.51E-06	1.23E-05
381	NM_175493 // Gpr68 // G protein-coupled receptor 68 // 12 E // 23837	8	6803113	1	0	0.627221	0.0396865	0.140608	0.424865	0.966137	0.99784	1.52E-06	1.14E-06
382	NM_010404 // Hap1 // huntingtin-associated protein 1 // 11 D|11 60.0	19	6791418	2	1	0.132457	0.197653	0.0187989	3.67E-06	0.557183	0.942474	1.57E-06	1.37E-05
383	NM_010867 // Myom1 // myomesin 1 // 17 E1.3 // 17929 /// AK052539	36	6852031	1	0	0.237528	0.0013254	0.162064	0.181991	0.79813	0.620177	1.60E-06	0.0004888
384	NM_146190 // Tubgcp5 // tubulin, gamma complex associated protein	7	6960834	1	0	0.625811	0.123327	0.706589	0.340335	0.275758	0.0216782	1.61E-06	3.53E-05
385	XM_001000689 // Hes1 // hairy and enhancer of split 1 (Drosophila) //	9	6840400	4	1	0.0118164	0.0155418	0.0001622	0.0042484	0.517743	0.915257	1.61E-06	5.16E-08
386	NM_001039493 // 9430067K14Rik // RIKEN cDNA 9430067K14 gene	14	6759342	1	0	0.0122166	0.209641	0.0504258	0.461987	0.996038	0.819339	1.71E-06	9.96E-06
387	NM_207246 // Rasgrp3 // RAS, guanyl releasing protein 3 // 17 E2 // 2	20	6852235	3	1	0.0406035	0.0264113	0.269112	0.0003303	0.996484	0.992037	1.80E-06	4.02E-25
388	NM_001001932 // Eea1 // early endosome antigen 1 // 10 C3 // 21623	28	6770013	0	0	0.69117	0.970906	0.690144	0.767556	0.926684	0.64517	1.85E-06	3.92E-10
389	NM_007984 // Fscn1 // fascin homolog 1, actin bundling protein (Stror	10	6935370	1	0	0.0613657	0.0122099	0.113097	0.50678	0.853047	0.979287	2.06E-06	7.10E-07
390	NM_134248 // Havcr1 // hepatitis A virus cellular receptor 1 // 11 B1.1	6	6780570	0	0	0.405897	0.230989	0.953786	0.407422	0.979455	0.94287	2.18E-06	0.00034587
391	NM_007643 // Cd36 // CD36 antigen // 5 A3|5 2.0 cM // 12491 /// AK15	21	6936406	0	0	0.707981	0.443209	0.161243	0.991889	0.864363	0.921952	2.33E-06	6.62E-20
392	XM_975701 // LOC669547 // hypothetical protein LOC669547 // // 669	8	6768071	0	0	0.233665	0.111588	0.181751	0.512121	0.970634	0.325838	2.33E-06	3.63E-09
393	XM_884335 // Ankrd57 // ankyrin repeat domain 57 // 10 B4 // 268301	10	6768114	1	0	0.560937	0.139973	0.0038613	0.0835416	0.858953	0.896825	2.35E-06	4.27E-19
394	NM_007801 // Ctsh // cathepsin H // 9 E3.1|9 50.0 cM // 13036 /// BC0	9	6991264	1	0	0.323777	0.158349	0.0055928	0.657965	0.999785	0.983059	2.63E-06	1.89E-10
395	NM_139305 // Car9 // carbonic anhydrase 9 // 4 B1 // 230099	10	6913011	1	0	0.0545089	0.967332	0.0219821	0.400699	0.0755064	0.634485	2.91E-06	2.39E-05
396	XM_910436 // LOC629862 // similar to Protein C10orf70 // 11 E1 // 62	11	6977775	2	0	0.2075	0.0024082	0.597878	0.0140151	0.736274	0.988691	2.99E-06	2.00E-06
397	NM_013674 // Irf4 // interferon regulatory factor 4 // 13 A3.3 // 16364	13	6805825	0	0	0.681414	0.500066	0.94609	0.620848	0.999348	0.773493	3.00E-06	3.13E-10
398	NM_019472 // Myo10 // myosin X // 15 C|15 9.2 cM // 17909 /// NM_0	37	6829277	2	1	0.0005431	0.136528	0.810606	0.0316189	0.988715	0.998154	3.55E-06	0.00017448
399	NM_021273 // Ckb // creatine kinase, brain // 12 F1|12 55.0 cM // 127	7	6840849	1	0	0.0285198	0.985359	0.801091	0.218003	0.126528	0.712018	3.74E-06	3.92E-13
400	NM_144830 // Tmem106a // transmembrane protein 106A // 11 D // 21	11	6784280	2	0	0.183241	0.0047073	0.288261	0.003587	0.984664	0.876951	3.77E-06	2.62E-08
401	NM_026410 // Cdca5 // cell division cycle associated 5 // 19 A // 67849	10	6867830	1	0	0.986809	0.123587	0.825521	0.039903	0.995731	0.999678	3.89E-06	4.18E-08
402	NM_178395 // Zdhhc2 // zinc finger, DHHC domain containing 2 // 8 A	12	6975613	0	0	0.178447	0.365197	0.356829	0.389367	0.966902	0.95475	3.89E-06	2.87E-13
403	NM_172659 // Slc2a6 // solute carrier family 2 (facilitated glucose tran	14	6885514	1	0	0.138546	0.758935	0.0077259	0.214232	0.79502	0.4642	4.11E-06	2.76E-06

Wu Supp Table 8: Highest ranked wt memory vs naive cells
Page 6

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Gene Description**	**# Probe Sets**	**Transcript ID**	**Number of cell types p<0.05**	**Number of cell types p<0.001**	**p value wt v thu memory CD4**	**P value wt vs thu naïve CD4**	**p value wt vs thu memory CD8**	**p value wt vs thu naïve CD8**	**p value wt vs thu thymus CD8SP**	**p value wt vs thu thymus CD4SP**	**P-value WT naïve v WT memory CD4**	**P-value WT naïve v WT memory CD8**
404	NM_010512 // Igf1 // insulin-like growth factor 1 // 10 C1|10 48.0 cM //	8	6769597	2	0	0.0344241	0.280448	0.0479371	0.332443	0.981859	0.94052	4.26E-06	1.15E-12
405	XM_992247 // LOC676697 // hypothetical protein LOC676697 // // 67	6	6926498	0	0	0.0508137	0.152722	0.155211	0.152173	0.973511	0.719547	4.27E-06	8.63E-05
406	NM_009696 // Apoe // apolipoprotein E // 7 A3|7 4.0 cM // 11816	8	6973587	2	0	0.304721	0.0094126	0.0285458	0.235245	0.434701	0.953914	4.32E-06	2.30E-07
407	NM_029932 // 9830002I17Rik // RIKEN cDNA 9830002I17 gene // 11	13	6789544	1	0	0.141434	0.34058	0.570751	0.0106367	0.965319	0.519376	5.83E-06	2.25E-13
408	NM_011375 // St3gal5 // ST3 beta-galactoside alpha-2,3-sialyltransfer	8	6946954	0	0	0.160551	0.68791	0.809806	0.823165	0.963837	0.592152	6.01E-06	3.43E-10
409	NM_008533 // Cd180 // CD180 antigen // 13 D1 // 17079 /// AK089613	6	6809655	1	0	0.0033875	0.126504	0.627291	0.528612	0.945409	0.971198	6.46E-06	5.30E-05
410	NM_011700 // Vill // villin-like // 9 F4 // 22351 /// BC022664 // Vill // villi	23	6992866	4	1	0.0045807	0.00014	0.0270789	0.028022	0.997946	0.984554	6.47E-06	4.31E-05
411	NM_008486 // Anpep // alanyl (membrane) aminopeptidase // 7 D2 // 123	23	6968735	1	1	0.0678737	0.136549	0.278177	0.0001021	0.747367	0.191372	6.48E-06	2.95E-35
412	NM_010897 // Nf1 // neurofibromatosis 1 // 11 B4-5|11 46.06 cM // 180	63	6782776	1	1	0.0616025	6.93E-07	0.310785	0.536414	0.908732	0.987024	6.52E-06	1.52E-08
413	NM_016925 // Fanca // Fanconi anemia, complementation group A // 8	42	6985991	3	0	0.45964	0.0082493	0.03096	0.0080498	0.999517	0.907376	6.96E-06	1.21E-15
414	NM_010276 // Gem // GTP binding protein (gene overexpressed in sk	5	6911727	0	0	0.342346	0.271049	0.702273	0.3067	0.299747	0.541982	7.38E-06	3.03E-08
415	XM_001003619 // Rhebl1 // Ras homolog enriched in brain like 1 // 15	8	6838410	3	3	0.22817	9.03E-07	2.01E-06	9.49E-07	0.479696	0.972368	7.97E-06	7.82E-05
416	NM_026316 // Aldh3b1 // aldehyde dehydrogenase 3 family, member I	7	6870958	1	0	0.418379	0.004626	0.270632	0.0914116	0.871174	0.887656	8.04E-06	5.53E-05
417	NM_133685 // Rab31 // RAB31, member RAS oncogene family // 17 B	14	6856676	0	0	0.19156	0.57776	0.679167	0.971826	0.718424	0.596583	8.05E-06	6.51E-07
418	NM_009445 // Ttk // Ttk protein kinase // 9 E2 // 22137 /// AK041487 //	23	6990948	1	0	0.908375	0.001099	0.369036	0.229381	0.998712	0.951506	8.22E-06	1.58E-09
419	XM_984804 // Clec7a // C-type lectin domain family 7, member a //	5	6957410	0	0	0.560148	0.629912	0.19724	0.413919	0.872633	0.766982	8.71E-06	3.77E-05
420	NM_144818 // Brrn1 // barren homolog (Drosophila) // 2 F1 // 215387	18	6890658	1	0	0.0911757	0.0907038	0.198466	0.0185804	0.999316	0.960046	9.07E-06	0.00012709
421	NM_008464 // Klra6 // killer cell lectin-like receptor, subfamily A, mem	12	6957432	1	1	0.983757	0.111345	0.975341	2.34E-06	0.625866	0.891808	9.44E-06	0.00066332
422	NM_175688 // A530099J19Rik // RIKEN cDNA A530099J19 gene // 13	9	6811368	1	0	0.316704	0.321454	0.717328	0.0295532	0.600339	0.832604	9.83E-06	5.17E-18
423	NM_028808 // P2ry13 // purinergic receptor P2Y, G-protein coupled 13	4	6905422	1	0	0.496182	0.0112065	0.0868037	0.611918	0.915533	0.722991	1.04E-05	7.21E-08
424	NM_026560 // Cdca8 // cell division cycle associated 8 // 4 D2.2|4 57.4	13	6925254	0	0	0.154234	0.277098	0.687067	0.989117	0.997624	0.99963	1.05E-05	0.00014805
425	NM_001005423 // Mreg // melanoregulin // 1 C3 // 381269 /// AY62821	8	6759642	3	0	0.0022372	0.010755	0.061767	0.011157	0.638829	0.920955	1.05E-05	7.70E-14
426	NM_011056 // Pde4d // phosphodiesterase 4D, cAMP specific // 13 D2	17	6810067	2	0	0.0287844	0.0136342	0.584216	0.153646	0.999982	0.999879	1.24E-05	0.00065444
427	NM_011521 // Sdc4 // syndecan 4 // 2 H3|2 94.0 cM // 20971 /// BC00	6	6892905	0	0	0.11332	0.840022	0.999181	0.294067	0.929067	0.495719	1.30E-05	0.00017103
428	NM_028390 // Anln // anillin, actin binding protein (scraps homolog, D	27	6993890	2	0	0.610852	0.0459114	0.0095424	0.341802	0.999939	0.934524	1.33E-05	9.33E-05
429	NM_001007220 // Adam22 // a disintegrin and metallopeptidase doma	30	6936076	1	0	0.0994437	0.213505	0.392072	0.0026833	0.800866	0.200484	1.33E-05	0.00050033
430	XM_990644 // LOC674182 // hypothetical protein LOC674182 // // 67	14	6921131	0	0	0.536992	0.180997	0.609749	0.11837	0.95424	0.859585	1.39E-05	6.18E-05
431	NM_028279 // Naalad2 // N-acetylated alpha-linked acidic dipeptidase	20	6993587	1	0	0.995158	0.0258368	0.429904	0.983316	0.999999	0.996778	1.49E-05	0.00010592
432	NM_001001880 // Mpzl1 // myelin protein zero-like 1 // 1 H2.3 // 68481	10	6763777	3	1	0.0004135	0.0084429	0.001198	0.132662	0.665193	0.754302	1.61E-05	1.19E-06
433	NM_139065 // BC013481 // cDNA sequence BC013481 // 5 C3.1 // 24	5	6943476	0	0	0.0756174	0.859776	0.857426	0.359121	0.465101	0.388044	1.67E-05	2.81E-08
434	NM_011871 // Prkra // protein kinase, interferon inducible double stran	10	6888002	0	0	0.242591	0.868275	0.111362	0.944835	0.967828	0.859094	1.71E-05	0.00014516
435	NM_181315 // Car5b // carbonic anhydrase 5b, mitochondrial // X F5 //	8	7020676	2	0	0.734343	0.034768	0.776312	0.234712	0.0471703	0.772989	1.92E-05	8.59E-05
436	XM_286230 // E030030I06Rik // RIKEN cDNA E030030I06 gene // 10	8	6762804	1	0	0.328838	0.0109709	0.234546	0.356239	0.637302	0.992905	1.95E-05	0.00060684
437	NM_146173 // Tspan33 // tetraspanin 33 // 6 A3.3 // 232670 /// NM_17	10	6945026	1	0	0.324586	0.165497	0.430474	0.015018	0.313491	0.481122	2.13E-05	5.88E-07
438	NM_177715 // Kctd12 // potassium channel tetramerisation domain co	5	6827123	3	2	0.0155001	4.16E-05	0.447857	0.0003403	0.875552	0.724759	2.15E-05	0.00029037
439	NM_011179 // Psap // prosaposin // 10 B4|10 35.0 cM // 19156 /// AK2	15	6768155	1	0	0.737508	0.0213905	0.651105	0.480213	0.999637	0.963754	2.22E-05	1.01E-06
440	NM_153090 // Fcrl1 // Fc receptor-like 1 // 3 F1 // 229499 /// NM_1781	12	6898996	0	0	0.522611	0.198129	0.677928	0.43084	0.903873	0.886983	2.27E-05	2.39E-05
441	NM_025760 // Ptplad2 // protein tyrosine phosphatase-like A domain c	7	6923142	1	1	0.0835793	1.01E-06	0.109485	0.0628494	0.99963	0.160809	2.58E-05	0.00076463
442	NM_007995 // Fcna // ficolin A // 2 A3 // 14133 /// AK133539 // Fcna //	11	6885432	3	3	0.0506094	6.26E-07	8.13E-05	5.76E-05	0.950293	0.934335	2.63E-05	2.67E-15
443	AK011893 // Spbc25 // spindle pole body component 25 homolog (S. c	10	6887520	0	0	0.0891145	0.486534	0.309636	0.457039	0.974352	0.825304	2.65E-05	1.98E-05
444	NM_008776 // Pafah1b3 // platelet-activating factor acetylhydrolase, is	7	6965940	1	0	0.376374	0.14596	0.270816	0.017221	0.0976732	0.782135	2.73E-05	0.00010239
445	XM_974657 // 1100001H23Rik // RIKEN cDNA 1100001H23 gene // 6	13	6957744	0	0	0.98733	0.378132	0.406179	0.486649	0.994752	0.872026	2.75E-05	5.81E-20
446	NM_010370 // Gzma // granzyme A // 13 D|13 64.0 cM // 14938 /// XM	7	6816160	1	0	0.863995	0.0110009	0.368064	0.283593	0.494537	0.900684	2.78E-05	1.24E-19
447	NM_020590 // Gabarapl1 // gamma-aminobutyric acid (GABA(A)) rece	9	6950160	0	0	0.745353	0.904104	0.705126	0.780099	0.993973	0.301111	2.84E-05	0.0003956
448	NM_008916 // RP23-136K12.4 // putative phosphatase // 11 B5|11 44.	35	6782457	3	2	0.101993	0.0007145	0.0137963	0.0001263	0.979621	0.751482	2.84E-05	1.15E-10
449	NM_026785 // Ube2c // ubiquitin-conjugating enzyme E2C // 2 H3|2 93	3	6883114	0	0	0.968826	0.0543229	0.733104	0.0634772	0.946119	0.976785	2.90E-05	1.34E-05
450	XR_002190 // LOC621312 // similar to fibrillarin // 13 D1 // 621312 /// N	8	6850502	2	1	0.179729	0.00022	0.0300252	0.0734509	0.808118	0.821504	2.98E-05	1.26E-07
451	NM_009369 // Tgfbi // transforming growth factor, beta induced // 13 3	18	6807336	1	0	0.283966	0.0225978	0.783437	0.232005	0.98906	0.986653	3.16E-05	2.71E-16
452	NM_011337 // Ccl3 // chemokine (C-C motif) ligand 3 // 11 C|11 47.59	4	6790294	0	0	0.513893	0.450573	0.343484	0.229158	0.69423	0.489808	3.25E-05	5.04E-06
453	XM_985034 // LOC666623 // similar to putative retrovirus-related gag	14	6820323	0	0	0.870694	0.0700952	0.341744	0.315413	0.165837	0.735214	3.30E-05	1.17E-10
454	NM_009909 // Il8rb // interleukin 8 receptor, beta // 1 C3|1 40.0 cM // 1	4	6750519	0	0	0.550072	0.456949	0.311435	0.320497	0.701907	0.977629	3.34E-05	0.0001539
455	NM_008491 // Lcn2 // lipocalin 2 // 2 A3|2 27.0 cM // 16819	8	6885873	0	0	0.537304	0.698916	0.450749	0.381918	0.67888	0.806263	3.36E-05	0.00085775
456	NM_010730 // Anxa1 // annexin A1 // 19 B|19 18.0 cM // 16952 /// BC0	12	6872010	2	0	0.0142255	0.441299	0.0339759	0.39079	0.491396	0.841454	3.37E-05	2.73E-11
457	NM_133888 // Smpdl3b // sphingomyelin phosphodiesterase, acid-like	9	6925829	2	0	0.597208	0.0290207	0.356151	0.0294392	0.962842	0.756639	3.39E-05	1.28E-05
458	NM_007781 // Csf2rb2 // colony stimulating factor 2 receptor, beta 2, l	7	6837008	0	0	0.501332	0.456724	0.214537	0.197541	0.715517	0.991512	3.53E-05	9.36E-08
459	NM_007543 // Ceacam2 // CEA-related cell adhesion molecule 2 // 7 A	7	6965950	1	0	0.0968096	0.716	0.0421882	0.441212	0.975993	0.735803	3.55E-05	7.21E-06
460	XM_906349 // LOC632297 // similar to UDP glycosyltransferase 1 fam	9	6751362	0	0	0.999604	0.956059	0.53884	0.216666	0.193757	0.823087	3.67E-05	5.23E-12
461	NM_009853 // Cd68 // CD68 antigen // 11 B3|11 39.0 cM // 12514	9	6789325	0	0	0.142059	0.924898	0.227453	0.26368	0.899668	0.375271	3.76E-05	3.25E-05
462	NM_009980 // Ctbp2 // C-terminal binding protein 2 // 7 F3|7 66.0 cM //	17	6971714	2	0	0.062606	0.033996	0.662825	0.0111922	0.992417	0.976731	3.80E-05	0.00071516
463	NM_021439 // Chst11 // carbohydrate sulfotransferase 11 // 10 C1 // 5	4	6769366	0	0	0.140064	0.137372	0.752776	0.12223	0.768852	0.896189	3.99E-05	3.85E-06
464	XM_989083 // Tgm4 // transglutaminase 4 (prostate) // 9 F4 // 331046	7	6993099	2	0	0.702771	0.0070438	0.238579	0.0195047	0.856908	0.85294	4.14E-05	0.00014201
465	NM_027154 // Tmbim1 // transmembrane BAX inhibitor motif containir	11	6759732	1	1	0.895784	0.530005	0.317917	0.0001333	0.466817	0.851574	4.30E-05	0.00026347
466	NM_172882 // Wdfy3 // WD repeat and FYVE domain containing 3 // 5	5	6940432	0	0	0.301552	0.134802	0.0745276	0.826169	0.992915	0.90909	4.63E-05	3.22E-05
467	NM_009731 // Akr1b7 // aldo-keto reductase family 1, member B7 // 6	6	6945321	0	0	0.921251	0.286503	0.123632	0.224819	0.606491	0.190062	4.73E-05	0.00066598
468	NM_146042 // Ibrdc2 // IBR domain containing 2 // 13 A5 // 218215 ///	8	6806831	0	0	0.503162	0.131282	0.485594	0.182262	0.992196	0.437231	4.82E-05	2.52E-09
469	NM_028841 // Tspan17 // tetraspanin 17 // 13 B2 // 74257 /// XM_9838	12	6807192	2	1	0.192737	1.86E-05	0.0517541	0.0022294	0.405776	0.740504	4.99E-05	0.00069376
470	NM_021099 // Kit // kit oncogene // 5 C3.3|5 42.0 cM // 16590 /// AK04	22	6931759	0	0	0.963707	0.983732	0.944379	0.244237	0.954511	0.992049	5.09E-05	8.86E-20

Wu Supp Table 8: Highest ranked wt memory vs naive cells Page 7

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Gene Description**	**# Probe Sets**	**Transcript ID**	**Number of cell types p<0.05**	**Number of cell types p<0.001**	**p value wt v thu memory CD4**	**P value wt vs thu naïve CD4**	**p value wt vs thu memory CD8**	**p value wt vs thu naïve CD8**	**p value wt vs thu thymus CD8SP**	**p value wt vs thu thymus CD4SP**	**P-value WT naïve v WT memory CD4**	**P-value WT naïve v WT memory CD8**
471	XM_983882 // 9130218O11Rik // RIKEN cDNA 9130218O11 gene // 1	6	6836962	0	0	0.778791	0.288498	0.293732	0.906676	0.367512	0.32498	5.26E-05	2.79E-08
472	XM_924210 // LOC637082 // similar to Traf2 binding protein // // 6370	4	6901196	0	0	0.33347	0.66806	0.694893	0.419574	0.425915	0.294883	5.58E-05	7.73E-06
473	XM_982009 // Casc5 // cancer susceptibility candidate 5 // 2 E5 // 764	9	6880492	0	0	0.644364	0.8613	0.909061	0.446482	0.943744	0.998719	6.14E-05	0.00059912
474	NM_021396 // Pdcd1lg2 // programmed cell death 1 ligand 2 // 19 C2-	6	6869035	0	0	0.256013	0.776487	0.067314	0.412094	0.640129	0.962829	6.19E-05	0.00018345
475	XR_004796 // LOC629568 // similar to shugoshin-like 1 // X C2 // 6295	10	6856096	0	0	0.639583	0.53078	0.730083	0.579999	0.989732	0.489204	6.37E-05	7.60E-08
476	NM_009763 // Bst1 // bone marrow stromal cell antigen 1 // 5 B3|5 25	11	6930371	0	0	0.345986	0.10185	0.962689	0.0745292	0.999798	0.970332	6.61E-05	7.15E-10
477	NM_009229 // Sntb2 // syntrophin, basic 2 // 8 D3|8 52.0 cM // 20650	10	6978937	0	0	0.736212	0.098763	0.893478	0.199233	0.994008	0.870086	6.65E-05	0.00077731
478	NM_028760 // Cep55 // centrosomal protein 55 // 19 C3 // 74107 /// AK	15	6869537	1	0	0.640591	0.093781	0.325896	0.0121216	0.984253	0.365543	7.08E-05	2.18E-06
479	NM_009835 // Ccr6 // chemokine (C-C motif) receptor 6 // 17 A1 // 124	8	6848579	0	0	0.747691	0.432068	0.998862	0.551862	0.822956	0.929759	7.32E-05	6.99E-09
480	NM_010386 // H2-DMa // histocompatibility 2, class II, locus DMa // 17	6	6850012	0	0	0.666776	0.457381	0.204379	0.642585	0.943377	0.996492	7.74E-05	0.00039491
481	NM_011518 // Syk // spleen tyrosine kinase // 13 B-C2|13 37.0 cM // 2	15	6807074	3	1	0.0049219	0.0247353	0.000935	0.662947	0.380548	0.774092	7.85E-05	1.44E-10
482	NM_011414 // Slpi // secretory leukocyte peptidase inhibitor // // 2056	6	6892899	1	0	0.86622	0.0202567	0.453147	0.853233	0.988079	0.995819	7.96E-05	1.92E-10
483	XM_977785 // LOC669888 // similar to Lipoprotein lipase precursor (L	14	6976901	3	0	0.0032523	0.0557605	0.0033692	0.0075727	0.869599	0.928744	8.40E-05	1.89E-06
484	XM_981611 // Adam23 // a disintegrin and metallopeptidase domain 2	26	6749933	2	0	0.0250118	0.985523	0.0393186	0.0503323	0.99764	0.998415	8.66E-05	1.10E-34
485	NM_172685 // Slc25a24 // solute carrier family 25 (mitochondrial carri	17	6900450	1	1	0.199744	0.0006812	0.508393	0.574731	0.840579	0.815513	8.68E-05	2.70E-08
486	NM_029249 // 4930547N16Rik // RIKEN cDNA 4930547N16 gene // 1	17	6775830	0	0	0.798458	0.164927	0.268841	0.825459	0.999668	0.997939	9.12E-05	7.68E-08
487	NM_172275 // Trafd1 // TRAF type zinc finger domain containing 1 // 5	14	6941672	0	0	0.0894617	0.327974	0.751153	0.729899	0.795244	0.920885	0.0001011	7.61E-05
488	NM_197959 // 3000004C01Rik // RIKEN cDNA 3000004C01 gene // 1	20	6791644	2	0	0.415363	0.0083445	0.519413	0.0450868	0.999986	0.924537	0.00010554	1.18E-05
489	NM_177583 // Aph1b // anterior pharynx defective 1b homolog (C. ele	5	6996432	1	0	0.0695363	0.1462	0.563677	0.0023476	0.531394	0.985776	0.00010727	0.00056565
490	NM_008455 // Klkb1 // kallikrein B, plasma 1 // 8 B1-B3|8 26.0 cM // 1	31	6982094	1	1	0.81805	0.207767	0.0505389	3.84E-07	0.999999	0.993594	0.00011278	5.68E-05
491	NM_144559 // Fcgr3a // Fc fragment of IgG, low affinity IIIa, receptor /	5	6755146	0	0	0.196292	0.498136	0.31073	0.29722	0.764692	0.404529	0.00012216	6.61E-07
492	NM_145977 // Slc45a3 // solute carrier family 45, member 3 // 1 E4 // 1	10	6753089	2	0	0.623119	0.0825699	0.0211976	0.0469023	0.922266	0.995932	0.00012351	8.85E-06
493	NM_019684 // Srpk3 // serine/arginine-rich protein specific kinase 3 //	16	7011952	3	1	0.044773	0.0033752	0.493261	2.52E-05	0.821314	0.884109	0.00012893	1.91E-07
494	NM_001025600 // Igsf4a // immunoglobulin superfamily, member 4A //	14	6988855	1	0	0.70864	0.0240129	0.817342	0.381824	0.981826	0.792357	0.0001298	2.83E-24
495	NM_009656 // Aldh2 // aldehyde dehydrogenase 2, mitochondrial // 5	15	6941685	2	0	0.0256237	0.0324789	0.50026	0.973512	0.977196	0.721564	0.00013265	4.67E-10
496	NM_153507 // Cpne2 // copine II // 8 C5 // 234577 /// AK086165 // Cpn	15	6978319	2	2	0.799271	2.94E-05	0.770652	0.0005096	0.803709	0.743465	0.00013365	9.27E-09
497	NM_025961 // Gatm // glycine amidinotransferase (L-arginine:glycine	12	6890448	2	0	0.0348725	0.0042428	0.517315	0.131974	0.949506	0.99886	0.00013449	1.89E-12
498	NM_053166 // Trim7 // tripartite motif protein 7 // 11 B1.2 // 94089 /// A	12	6780696	0	0	0.953687	0.754727	0.248718	0.0930734	0.840083	0.970285	0.0001346	1.26E-09
499	NM_008042 // Fprl1 // formyl peptide receptor-like 1 // 17 A3.2 // 1429	7	6848972	1	0	0.420767	0.0136493	0.966936	0.578895	0.445098	0.977905	0.00013775	4.45E-06
500	NM_027373 // 2600003E23Rik // RIKEN cDNA 2600003E23 gene // 5	19	6929920	1	0	0.203817	0.955161	0.780364	0.0347079	0.999008	0.984544	0.00013834	0.00046046
501	NM_015790 // Icosl // icos ligand // 10 C1 // 50723 /// AK220220 // Icos	8	6769033	1	0	0.783187	0.945067	0.0173543	0.846528	0.974245	0.955638	0.00014402	0.00059344
502	NM_008139 // Gnaq // guanine nucleotide binding protein, alpha q pol	10	6868356	0	0	0.562804	0.814543	0.357102	0.465623	0.997059	0.998289	0.00015029	0.00030178
503	NM_009845 // Cd22 // CD22 antigen // 7 B1|7 9.0 cM // 12483 /// BC04	21	6966322	1	0	0.833415	0.0013716	0.11878	0.209108	0.998874	0.999299	0.00016118	9.49E-11
504	NM_199311 // Clec4a1 // C-type lectin domain family 4, member a1 //	10	6949722	0	0	0.564813	0.492312	0.820983	0.305333	0.991058	0.98454	0.00016497	1.85E-09
505	NM_011939 // Hsf4 // heat shock transcription factor 4 // 8 D1 // 26386	16	6978818	1	0	0.468584	0.190171	0.72541	0.0106707	0.969974	0.792233	0.00016701	0.00015755
506	NM_009987 // Cx3cr1 // chemokine (C-X3-C) receptor 1 // 9 F4 // 1305	3	6999456	0	0	0.223936	0.479113	0.361826	0.0887951	0.657158	0.468118	0.00017175	6.73E-05
507	NM_016920 // Atp6v0a1 // ATPase, H+ transporting, lysosomal V0 sub	24	6784236	1	1	0.214335	0.282891	0.341356	0.0001139	0.994003	0.99697	0.00017376	3.63E-15
508	XM_980616 // LOC627016 // similar to hypothetical LOC432650 // 7 A	8	6881556	0	0	0.411307	0.852166	0.231214	0.0578273	0.275888	0.325099	0.00018242	4.93E-10
509	NM_134158 // Cd300d // Cd300D antigen // 11 E2|11 78.0 cM // 14049	4	6792373	1	0	0.233296	0.504399	0.0042324	0.632944	0.312434	0.926393	0.00021304	4.18E-08
510	NM_026496 // Grhl2 // grainyhead-like 2 (Drosophila) // 15 C // 25297	19	6829871	3	1	0.0235652	0.340328	1.72E-05	0.0183354	0.683245	0.97185	0.00021413	0.00073016
511	NM_177265 // 6330512M04Rik // RIKEN cDNA 6330512M04 gene // 8	8	6972292	1	0	0.0447492	0.571138	0.726912	0.999114	0.962943	0.355686	0.00021625	5.28E-05
512	NM_009013 // Rad51ap1 // RAD51 associated protein 1 // 6 F3|6 61.3	9	6957252	2	0	0.725366	0.008053	0.273204	0.0273938	0.790816	0.211087	0.00022093	2.73E-05
513	NM_033612 // Ela1 // elastase 1, pancreatic // 15 56.8 cM // 109901 //	6	6838578	2	0	0.554698	0.0372509	0.0217409	0.390247	0.372703	0.844716	0.00022182	6.73E-06
514	XM_001003209 // LOC676870 // region containing RIKEN cDNA 2310	14	6763937	3	0	0.27144	0.0082374	0.0256255	0.0013859	0.966405	0.943905	0.00022406	1.61E-17
515	NM_011234 // Rad51 // RAD51 homolog (S. cerevisiae) // 2 F1|2 66.8	11	6880497	0	0	0.478237	0.253213	0.984237	0.67602	0.993641	0.855019	0.00022479	7.51E-05
516	XM_001004634 // LOC619653 // similar to Thymidine kinase, cytosolic	2	6792601	1	0	0.20195	0.0012006	0.262399	0.6991	0.936568	0.281843	0.00022656	4.97E-05
517	NM_011708 // Vwf // Von Willebrand factor homolog // 6 F3|6 60.8 cM	54	6949884	2	0	0.100154	0.149633	0.0284077	0.0191844	0.420754	0.999988	0.00022918	0.00081454
518	NM_013482 // Btk // Bruton agammaglobulinemia tyrosine kinase // X	19	7019518	0	0	0.538816	0.256349	0.0568227	0.784467	0.990584	0.937366	0.00023078	7.31E-13
519	NM_177372 // Dna2l // DNA2 DNA replication helicase 2-like (yeast) //	23	6768324	1	0	0.468117	0.0016895	0.261996	0.0758415	0.997271	0.999206	0.00023311	1.49E-15
520	NM_007932 // Eng // endoglin // 2 B|2 21.4 cM // 13805 /// AY679531 /	25	6876212	1	1	0.113364	0.315121	0.000101	0.0636827	0.545941	0.999743	0.0002503	5.55E-05
521	NM_011391 // Slc16a7 // solute carrier family 16 (monocarboxylic acid	10	6777784	0	0	0.154385	0.695587	0.193079	0.99418	0.756895	0.870275	0.00025542	0.00061277
522	NM_133994 // Gstt3 // glutathione S-transferase, theta 3 // 10 C1 // 10	8	6775149	0	0	0.994489	0.728542	0.90351	0.143191	0.994541	0.990994	0.00025888	1.27E-05
523	NM_007387 // Acp2 // acid phosphatase 2, lysosomal // 2 E1|2 52.0 cM	13	6879034	6	5	3.55E-09	1.15E-07	7.44E-10	9.66E-08	0.0376135	1.39E-07	0.0002602	1.41E-08
524	NM_021891 // Fignl1 // fidgetin-like 1 // 11 A2 // 60530 /// AK051324 //	6	6786044	1	1	0.576082	0.0006085	0.100886	0.0548241	0.956942	0.77509	0.00026688	2.36E-07
525	NM_026674 // Aph1c // anterior pharynx defective 1c homolog (C. ele	6	6996438	0	0	0.531774	0.0912421	0.183166	0.6286	0.496588	0.721747	0.00027678	2.15E-08
526	NM_010548 // Il10 // interleukin 10 // 1 E4|1 69.9 cM // 16153	7	6753014	0	0	0.373026	0.336134	0.179373	0.898168	0.997772	0.94277	0.00027797	3.90E-05
527	NM_177337 // Arl11 // ADP-ribosylation factor-like 11 // 14 D1 // 21914	3	6819602	1	0	0.873853	0.177071	0.695282	0.0210339	0.352515	0.585364	0.00029092	7.30E-06
528	NM_008832 // Phka1 // phosphorylase kinase alpha 1 // X D|X 39.0 cM	36	7018594	2	0	0.201595	0.0011016	0.0956987	0.0645977	0.999923	0.0066226	0.00031886	1.14E-07
529	NM_001037134 // Ccne2 // cyclin E2 // 4 A2 // 12448 /// NM_009830 //	9	6911682	1	0	0.46178	0.918267	0.877915	0.0026924	0.896153	0.855445	0.00033268	3.20E-06
530	NM_007900 // Ect2 // ect2 oncogene // 3 B // 13605 /// AK040174 // Ec	28	6903753	1	0	0.0894027	0.0021542	0.937484	0.831731	0.999486	0.997108	0.0003386	0.00059306
531	NM_016923 // Ly96 // lymphocyte antigen 96 // 1 A3 // 17087 /// AF168	6	6747839	0	0	0.190738	0.796326	0.706047	0.754813	0.979194	0.258733	0.00034245	0.00012875
532	NM_001002896 // Bfsp2 // beaded filament structural protein 2, phakin	9	6998405	1	0	0.113903	0.188335	0.058837	0.0249056	0.67367	0.789774	0.00035037	0.00014847
533	XM_976379 // Man1c1 // mannosidase, alpha, class 1C, member 1 //	16	6926023	2	0	0.0320229	0.303285	0.869803	0.0179644	0.454901	0.795013	0.00035069	2.49E-05
534	NM_177861 // Tmem67 // transmembrane protein 67 // 4 A1 // 329795	30	6919741	0	0	0.435969	0.0857316	0.0768875	0.0753588	0.0953345	0.916908	0.00035586	0.00025
535	NM_025968 // Ltb4dh // leukotriene B4 12-hydroxydehydrogenase // 4	12	6922026	1	0	0.0698665	0.058437	0.0177461	0.701342	0.991827	0.332334	0.00036303	2.82E-05
536	NM_018825 // Sh2b2 // SH2B adaptor protein 2 // 5 G1 // 23921 /// BC	17	6942494	1	0	0.898882	0.0288185	0.570134	0.0784987	0.974922	0.994411	0.0003634	0.00012244
537	NM_007404 // Adam9 // a disintegrin and metallopeptidase domain 9	24	6981108	0	0	0.684711	0.31392	0.213125	0.129701	0.980892	0.958135	0.00036529	0.00034418

Wu Supp Table 8: Highest ranked wt memory vs naive cells
Page 8

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Gene Description**	**# Probe Sets**	**Transcript ID**	**Number of cell types p<0.05**	**Number of cell types p<0.001**	**p value wt v thu memory CD4**	**P value wt vs thu naïve CD4**	**p value wt vs thu memory CD8**	**p value wt vs thu naïve CD8**	**p value wt vs thu thymus CD8SP**	**p value wt vs thu thymus CD4SP**	**P-value WT naïve v WT memory CD4**	**P-value WT naïve v WT memory CD8**
538	NM_153064 // Ndufs2 // NADH dehydrogenase (ubiquinone) Fe-S pro	4	6764049	0	0	0.757949	0.354186	0.932581	0.204064	0.646811	0.601837	0.00037471	0.00011527
539	NM_178754 // Arhgap6 // Rho GTPase activating protein 6 // X F5 // 1	18	7015229	1	0	0.341618	0.621834	0.323609	0.0494075	0.96997	0.990426	0.00038378	6.33E-06
540	NM_178874 // Tmcc2 // transmembrane and coiled-coil domains 2 // 1	6	6762217	0	0	0.127507	0.823332	0.700058	0.924161	0.854782	0.265062	0.00040808	3.43E-05
541	NM_023158 // Cxcl16 // chemokine (C-X-C motif) ligand 16 // 11 B4 //	6	6789420	0	0	0.950947	0.150035	0.728927	0.515511	0.727966	0.738273	0.0004214	4.56E-08
542	NM_010593 // Jup // junction plakoglobin // 11 D|11 60.0 cM // 16480 /	12	6791422	1	0	0.216358	0.0801969	0.878712	0.0404457	0.927219	0.918859	0.00042317	8.25E-07
543	NM_198654 // 4833432M17Rik // RIKEN cDNA 4833432M17 gene // 1	6	6756346	2	0	0.909929	0.0336781	0.0010501	0.287657	0.93297	0.88654	0.000483	8.47E-15
544	NM_146187 // Ffar2 // free fatty acid receptor 2 // 7 B1 // 233079 /// BC	5	6966314	0	0	0.970152	0.981181	0.398396	0.713079	0.335621	0.898866	0.00048543	1.33E-05
545	NM_024264 // Cyp27a1 // cytochrome P450, family 27, subfamily a, p	11	6750566	1	1	0.854773	0.5702	0.0009656	0.442757	0.911931	0.965107	0.00050182	1.10E-17
546	NM_019449 // Unc93b1 // unc-93 homolog B1 (C. elegans) // 19 A // 5	14	6867615	1	0	0.13364	0.0073591	0.0967954	0.0574702	0.99998	0.870608	0.00050446	3.00E-09
547	NM_008591 // Met // met proto-oncogene // 6 4.0 cM // 17295 /// AK14	22	6944380	0	0	0.690191	0.757646	0.607941	0.474336	0.832179	0.729255	0.00054149	7.07E-17
548	NM_010648 // Klra3 // killer cell lectin-like receptor, subfamily A, mem	3	6957444	0	0	0.769797	0.588959	0.702929	0.461799	0.964178	0.632023	0.00054947	2.24E-08
549	NM_001024911 // Sept10 // septin 10 // 10 B4 // 103080 /// NM_00102	11	6774226	1	0	0.358006	0.0167831	0.990493	0.118363	0.30068	0.261968	0.0005516	3.18E-13
550	NM_009502 // Vcl // vinculin // 14 A3|14 2.5 cM // 22330	24	6817393	2	0	0.0256703	0.305338	0.128232	0.0016929	0.134994	0.929374	0.00063852	9.13E-09
551	NM_145924 // Cenpi // centromere protein I // X E3 // 102920 /// XM_9	22	7013857	1	0	0.395561	0.031666	0.841178	0.457511	0.999711	0.999115	0.00067337	5.69E-05
552	NM_008506 // Mycl1 // v-myc myelocytomatosis viral oncogene homo	12	6916937	2	0	0.956846	0.805635	0.0187937	0.008379	0.752859	0.973643	0.00068286	3.02E-24
553	XM_981541 // 1300007C21Rik // RIKEN cDNA 1300007C21 gene // 1	9	6934119	0	0	0.736219	0.865954	0.28325	0.739864	0.875033	0.971187	0.00069475	6.41E-05
554	NM_025972 // Asahl // N-acylsphingosine amidohydrolase (acid ceran	11	6939985	0	0	0.682735	0.0642156	0.521141	0.139769	0.904914	0.938742	0.00069509	3.06E-25
555	NM_024223 // Crip2 // cysteine rich protein 2 // 12 F1|12 58.7 cM // 68	6	6798313	1	0	0.860248	0.0868601	0.0031559	0.185133	0.970575	0.716477	0.00072313	6.17E-10
556	NM_019789 // Kcnip3 // Kv channel interacting protein 3, calsenilin // 2	9	6890699	0	0	0.186933	0.424203	0.230074	0.547423	0.998798	0.844222	0.0007262	0.00025613
557	NM_007760 // Crat // carnitine acetyltransferase // 2 B|2 18.0 cM // 12	16	6885773	0	0	0.153762	0.606504	0.559425	0.0935806	0.344584	0.833409	0.00074197	1.27E-05
558	NM_011864 // Papss2 // 3'-phosphoadenosine 5'-phosphosulfate synt	14	6869216	1	0	0.0903236	0.0905387	0.284273	0.0255193	0.949306	0.76146	0.00076516	1.27E-08
559	NR_000040 // Tyms-ps // thymidylate synthase, pseudogene // 10 C1	7	6936981	0	0	0.607764	0.414765	0.0910457	0.0715244	0.947491	0.996669	0.00077478	6.34E-06
560	NM_008013 // Fgl2 // fibrinogen-like protein 2 // 5 A3|5 7.0 cM // 14190	6	6929119	1	0	0.587548	0.745796	0.464714	0.476527	0.292423	0.0477766	0.00081438	4.24E-08
561	NM_026393 // Nmral1 // NmrA-like family domain containing 1 // 16 A1	5	6843657	0	0	0.644763	0.510342	0.270696	0.254476	0.711811	0.99816	0.0008835	0.00082524
562	NM_173402 // Rgs12 // regulator of G-protein signaling 12 // 5 B2|5 20	23	6929861	0	0	0.225728	0.662213	0.771332	0.257695	0.953072	0.976471	0.00092946	5.48E-06
563	NM_009373 // Tgm2 // transglutaminase 2, C polypeptide // 2 H1|2 89	13	6892579	0	0	0.31819	0.601128	0.485142	0.0739961	0.444746	0.94598	0.00096774	3.46E-08
564	NM_001001495 // 9030611K07Rik // RIKEN cDNA 9030611K07 gene	9	6946749	0	0	0.547265	0.354569	0.138902	0.0524454	0.520444	0.967385	0.0009678	8.59E-05
565	NM_028810 // Rnd3 // Rho family GTPase 3 // 2 C1.1 // 74194 /// AK0	9	6886678	0	0	0.56303	0.675675	0.89519	0.230754	0.69797	0.29959	0.00099338	0.000869
566	NM_019453 // Mefv // Mediterranean fever // 16 A1|16 1.1 cM // 54483	15	6843601	2	2	0.743833	0.699137	0.0001704	0.0003659	0.999884	0.680035	0.00099938	1.29E-09

Wu Supp Table 8: Highest ranked wt memory vs naive cells
Page 9