

Fig 1. Clustering of human cDNA sequences

(A) Estimation of the number of human genes from full-length cDNAs and ESTs. Outline of our gene prediction method from the human full-length cDNAs and ESTs mapped to human genome is schematically shown. For each one of the predicted genes, classification reliability was evaluated manually. (B) Cover rate of FLJ EST sequences and (C) Cover rate of FLJ full-length sequenced cDNAs. Results of reliability analysis according to category based on the cover rates of 1.45 million of ESTs (B) and 55 thousand full-length cDNAs (C).

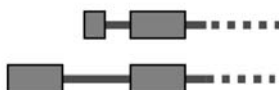
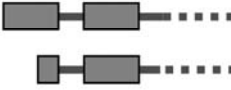
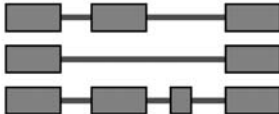
<b>Type A</b> <b>3,403</b> <b>(29%)</b>	<b>A1</b> <b>1,962</b>	<b>First exon variations</b> 
	<b>A2</b> <b>1,441</b>	<b>other types</b> 
<b>Type B</b> <b>8,277</b> <b>(70%)</b>	<b>B</b> <b>8,277</b>	<b>Alternative splicing</b> 

Fig 2. Classifications of the 11,769 full-length cDNAs based on splicing patterns.

11,769 of human full-length cDNAs were classified according to their TSS utilization. Type A: these cDNAs were derived from transcripts which were generated utilizing a TSS different than the previously analyzed TSS of the gene. Type A1: cDNAs contained a sequence variation known as FEV (First exon variation). Type A2: this class of cDNAs did not have the FEV feature. Type B: these cDNAs were derived from transcripts which were generated utilizing the same TSS as the previously analyzed TSS, but were found to be alternatively spliced. We could not classify 89 cDNAs because they coded for newly identified proteins.

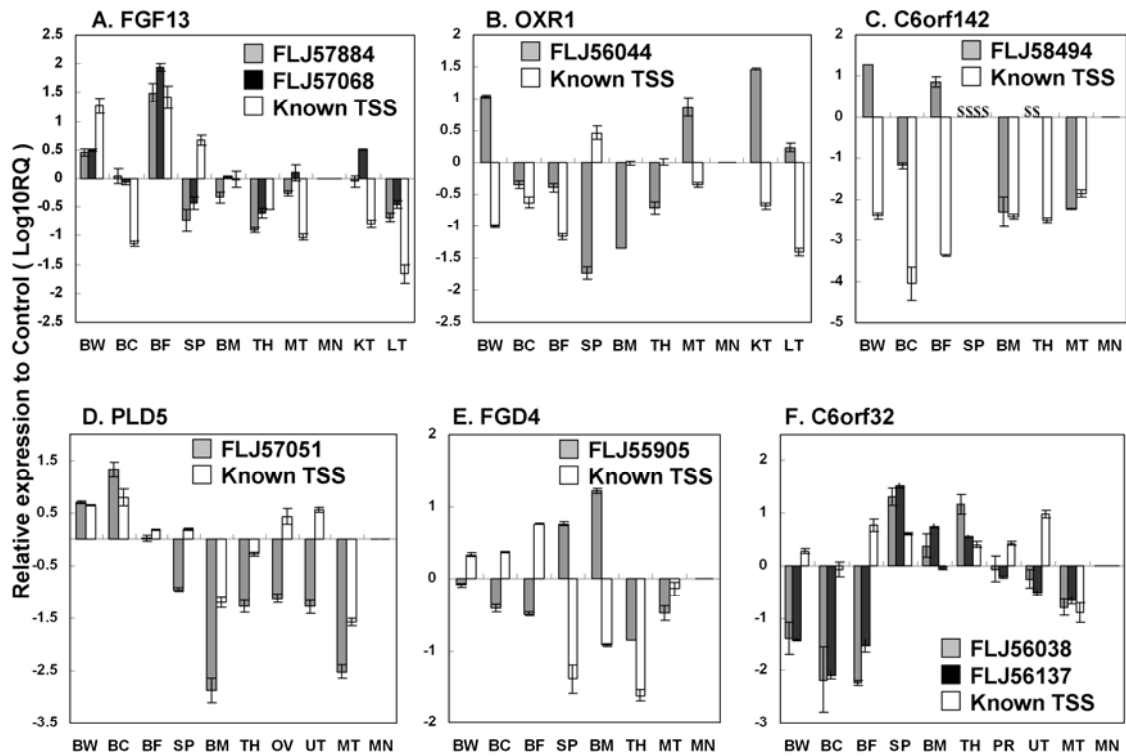


Fig 3. Quantitative evaluation of selected genes by real-time PCR.

Expression levels of the first exon regions of the selected genes were analyzed by real-time PCR. The data were normalized with respect to that of the human GAPDH as described in Materials and methods. The expression levels of genes were represented in log10 base. Expression levels of cDNAs labeled "\$\$" represent very low expression level or undetected. (A) FGF13, (B) OXR1, (C) C6orf142, (D) PLD5, (E) FGD4, (F) C6orf32

\* BW: brain, whole; BC: brain, cerebellum; BF: fetal brain; SP: spleen; BM: bone marrow; TH: thymus; OV: ovary; PR: prostate; UT: uterus; MT: mixture of tumor human tissues; MN: control, mixture of normal human tissues; KT: kidney tumor; LT: lung tumor