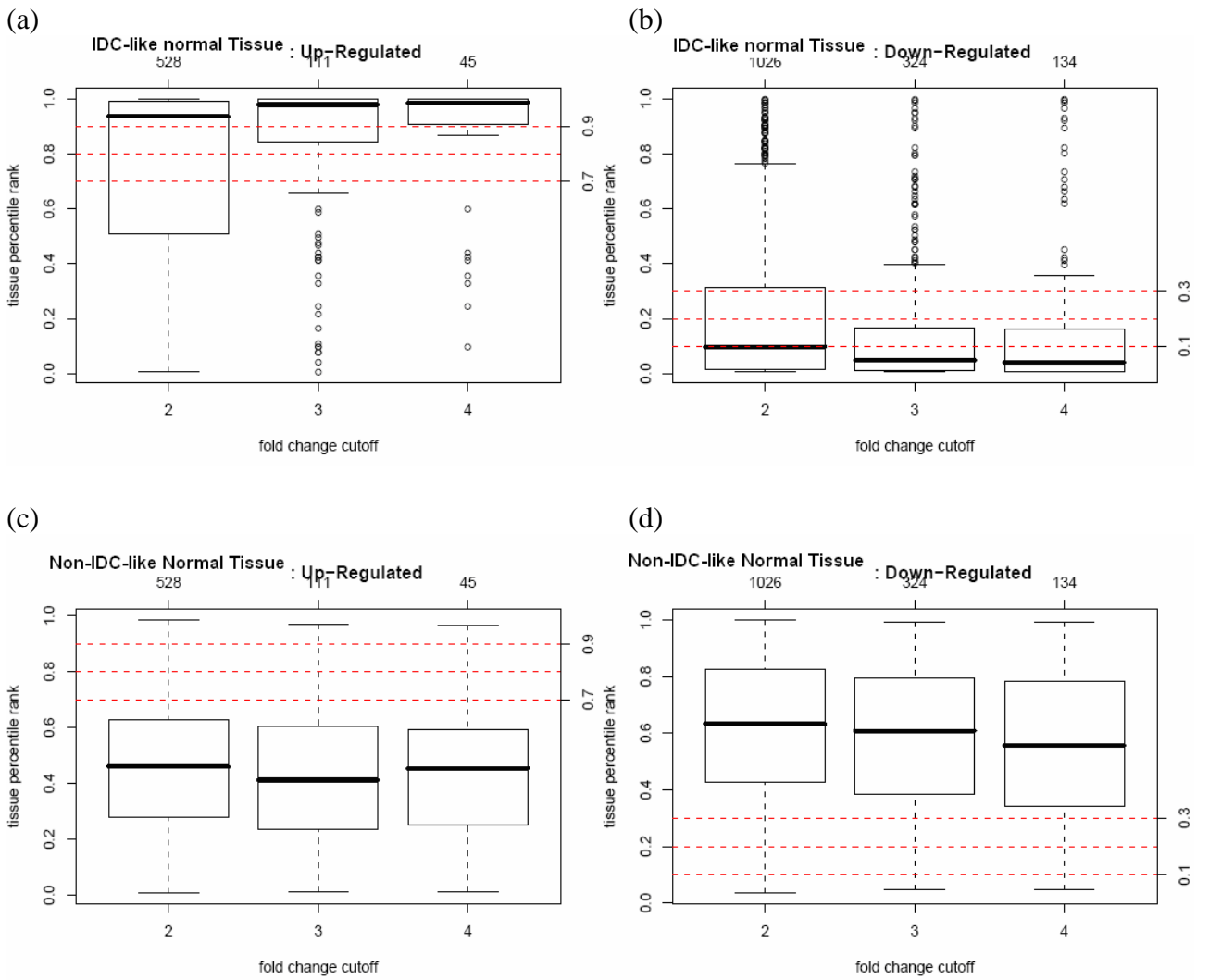


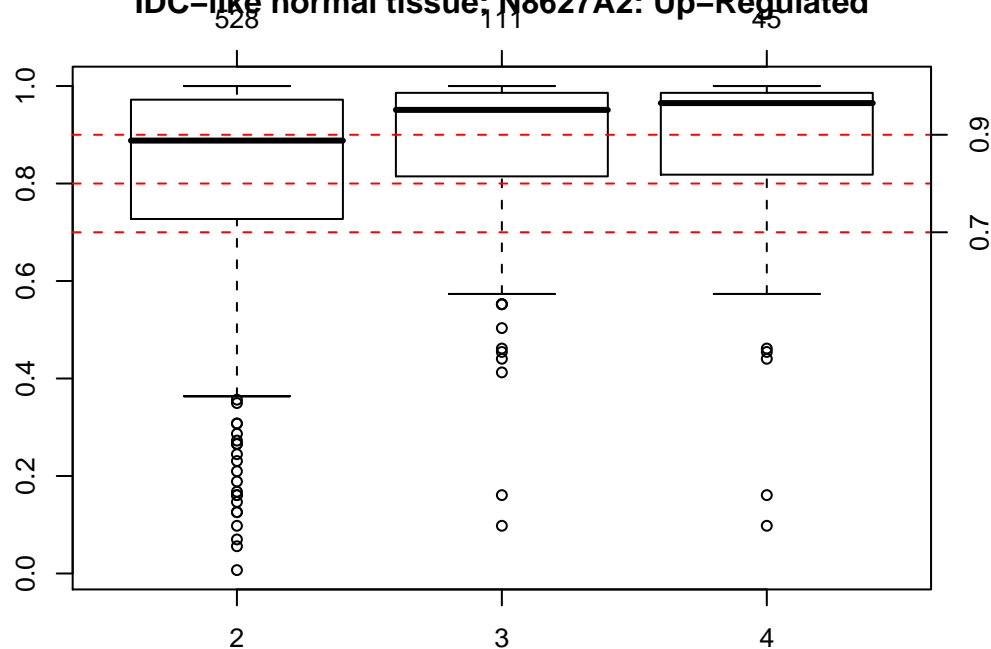
Supplementary Figure 4: Percentile rank distribution of all IDC-like normal breast tissues.

**Figure A: Representative example of the percentile rank distribution of one IDC-like normal breast tissue versus one non-IDC-like normal breast tissue, showing their molecular differences.** (a-b) Percentile rank distribution of one IDC-like normal breast tissue for (a) Up-regulated genes and (b) Down-regulated genes, (c-d) Percentile rank distribution of one non-IDC-like normal breast tissue for (c) Up-regulated genes and (d) Down-regulated genes. Two plots were generated for each tissue. One plot was for up-regulated IDC genes at various fold-change cutoffs (Figure a or c) and the other plot was for down-regulated genes (Figure b or d). At each plot, the Y axis represented the tissue percentile rank and the X axis indicated the fold-change cutoff to determine the number of IDC genes (up- or down- regulated) which was displayed on the top of the plot (e.g., there were 528 up- regulated IDC genes with a fold greater than 2). Each plot had several boxplots with each boxplot to display the distribution of a tissue percentile rank at a specific fold cutoff. For example, the 1<sup>st</sup> boxplot in the Figure (a) was an IDC-like normal tissue percentile rank at a 2-fold change (fold>2). For each gene (from the 528 genes), we ranked this IDC-like normal tissue among all the histologically normal breast (HNB) tissues to obtain the percentile rank. So there were 528 percentile ranks (corresponding to the 528 genes) to indicate the position of this IDC-like normal tissue compared to the rest of the HNB tissues. The boxplot for these 528 percentile ranks (the 1<sup>st</sup> boxplot in Figure a) showed the median of percentile rank for this IDC-like normal tissue was beyond 90%. In summary, Figure (a) showed that the median of percentile rank was greater than 90% at all fold cutoffs for the up-regulated genes. On the other hand, the median of percentile rank was below 20% for the down-regulated genes in the Figure (b). The results showed this IDC-like normal tissue had higher expression (up or down) than the other normal tissues. In contrast, the non-IDC-like normal tissue yielded the median of percentile rank around 40% and 60% for up- and down- regulated genes, respectively (Figures c and d). The other IDC-like normal tissues are in the Figure B.

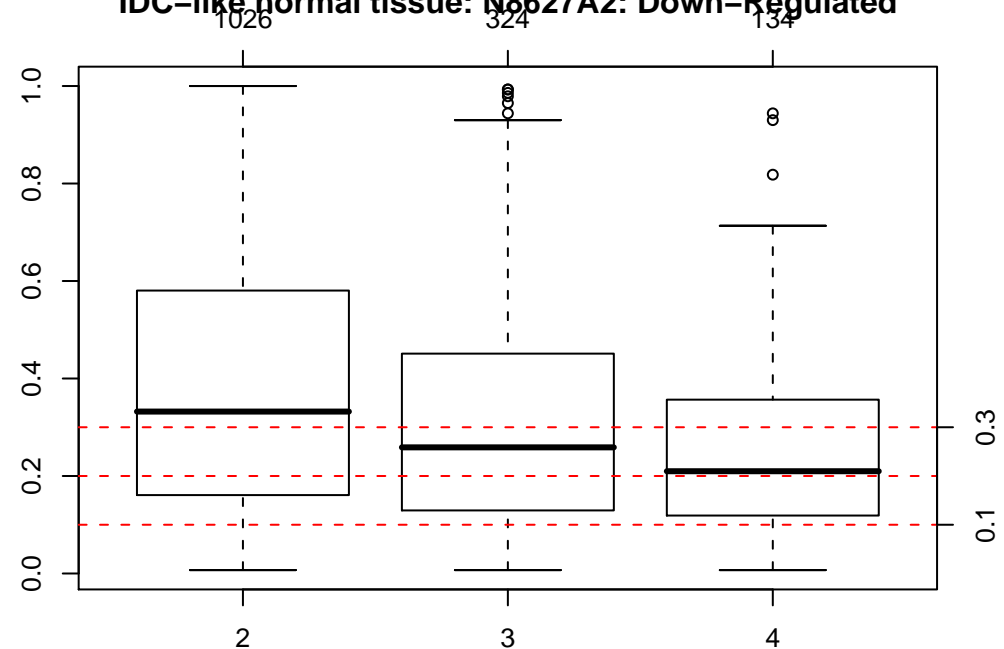


**Figure B: Percentile rank distribution of all IDC-like normal breast tissues.**

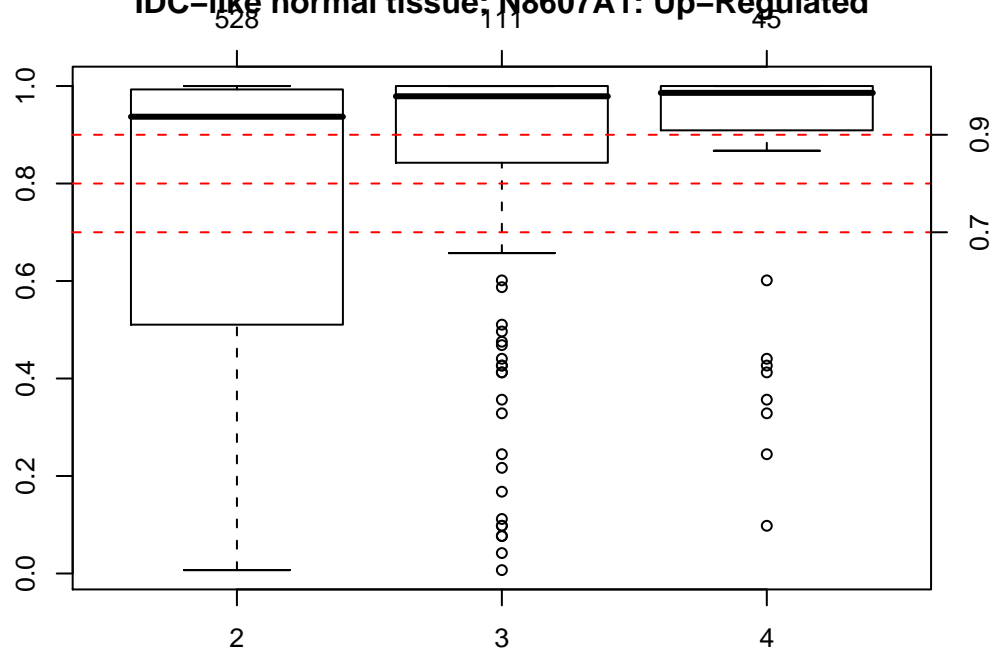
**IDC-like normal tissue: N8627A2: Up-Regulated**



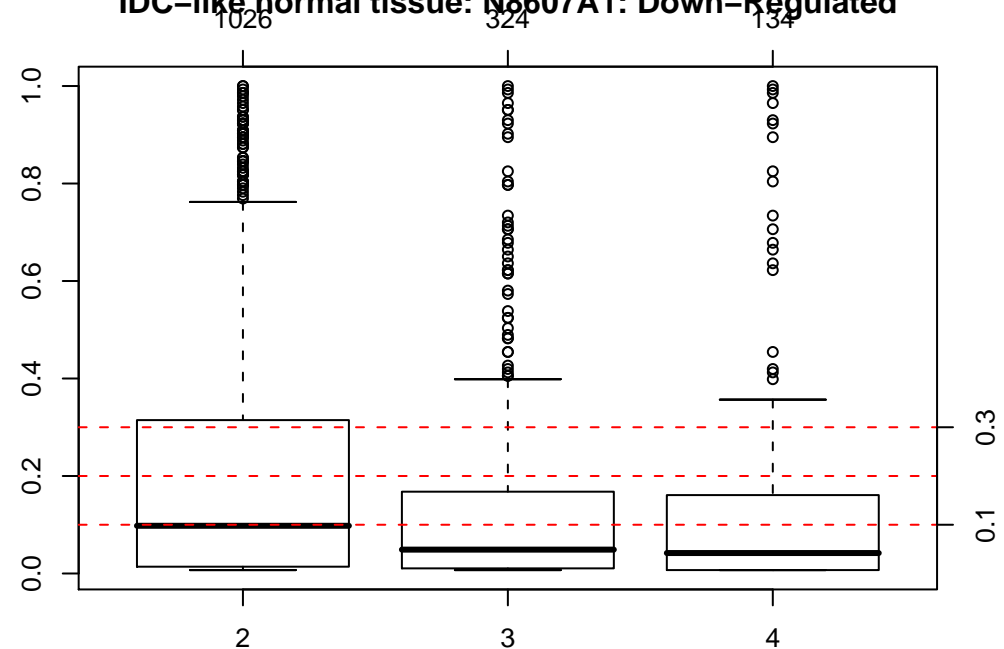
**IDC-like normal tissue: N8627A2: Down-Regulated**



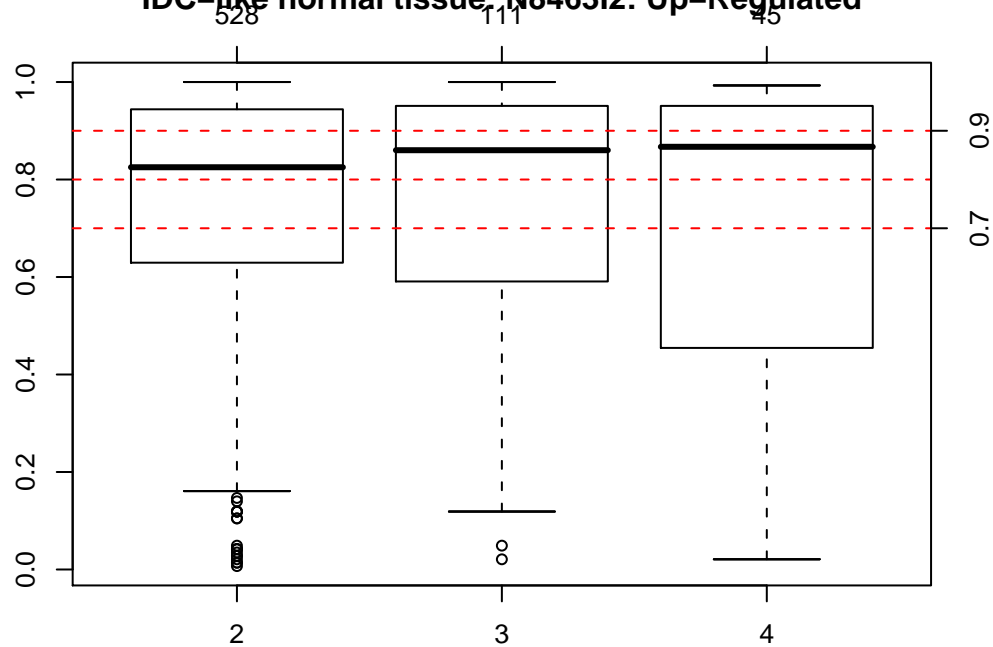
**IDC-like normal tissue: N8607A1: Up-Regulated**



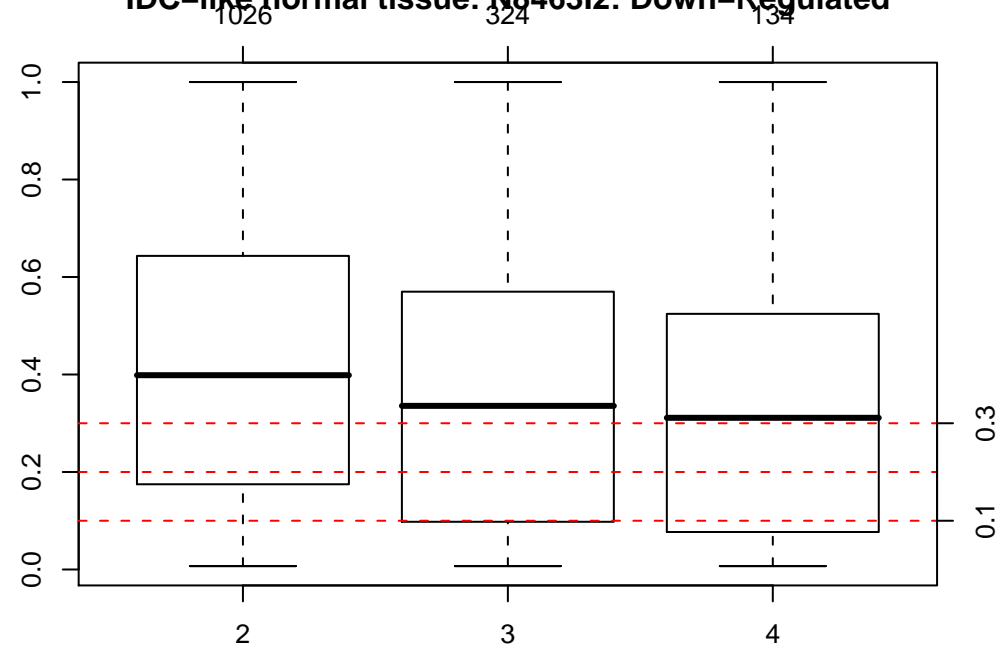
**IDC-like normal tissue: N8607A1: Down-Regulated**



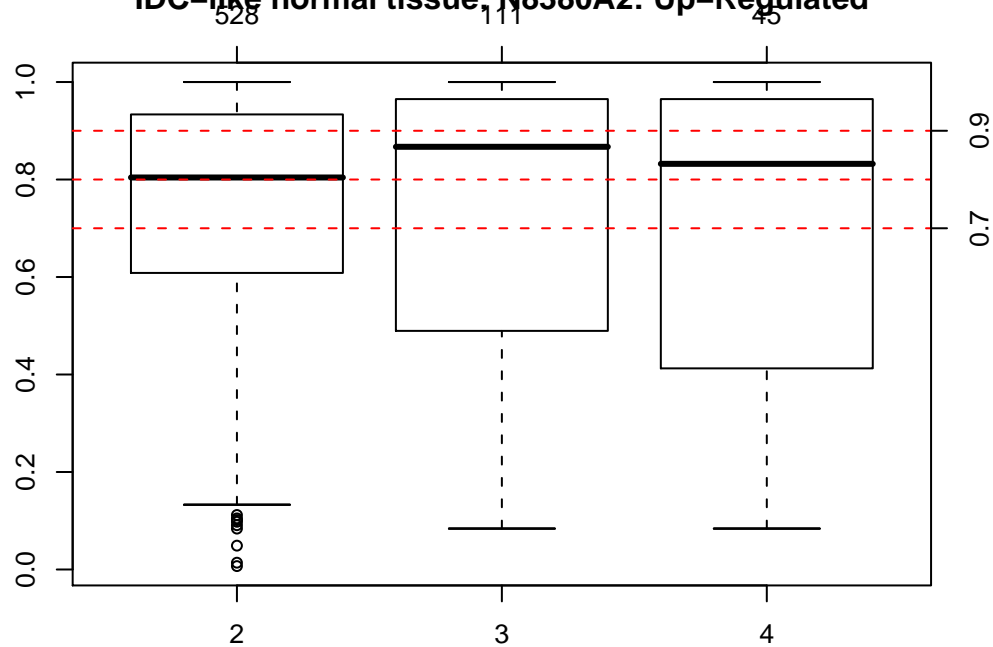
**IDC-like normal tissue: N8463I2: Up-Regulated**



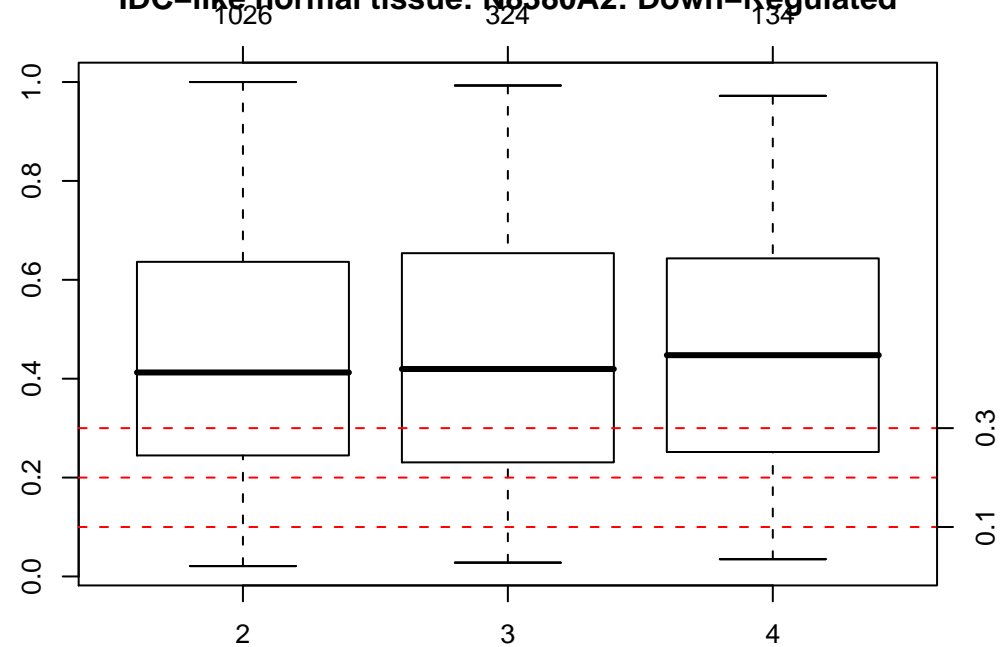
**IDC-like normal tissue: N8463I2: Down-Regulated**



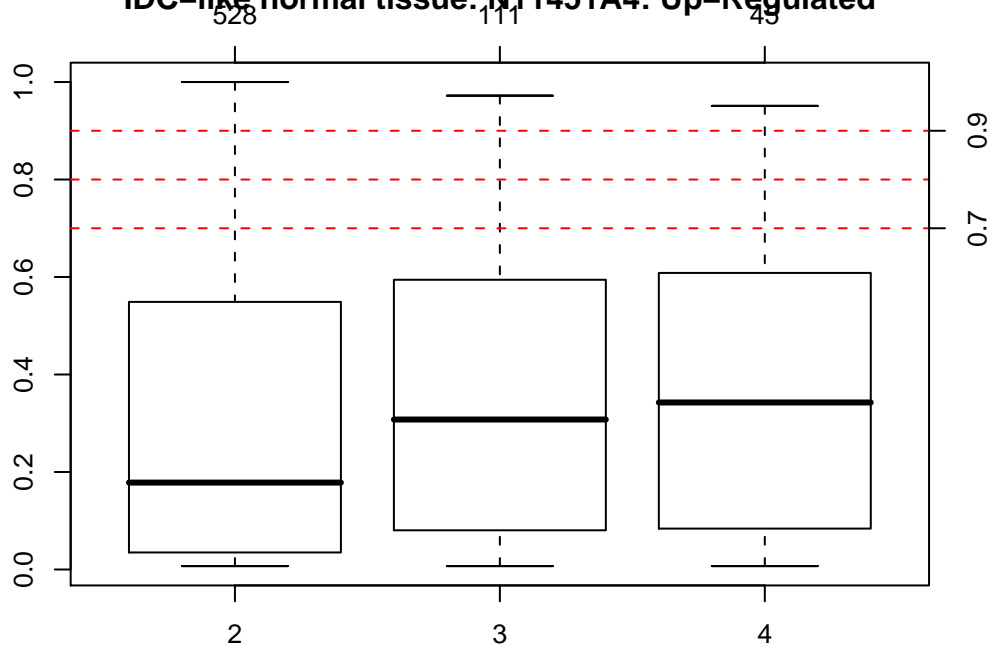
**IDC-like normal tissue: N8380A2: Up-Regulated**



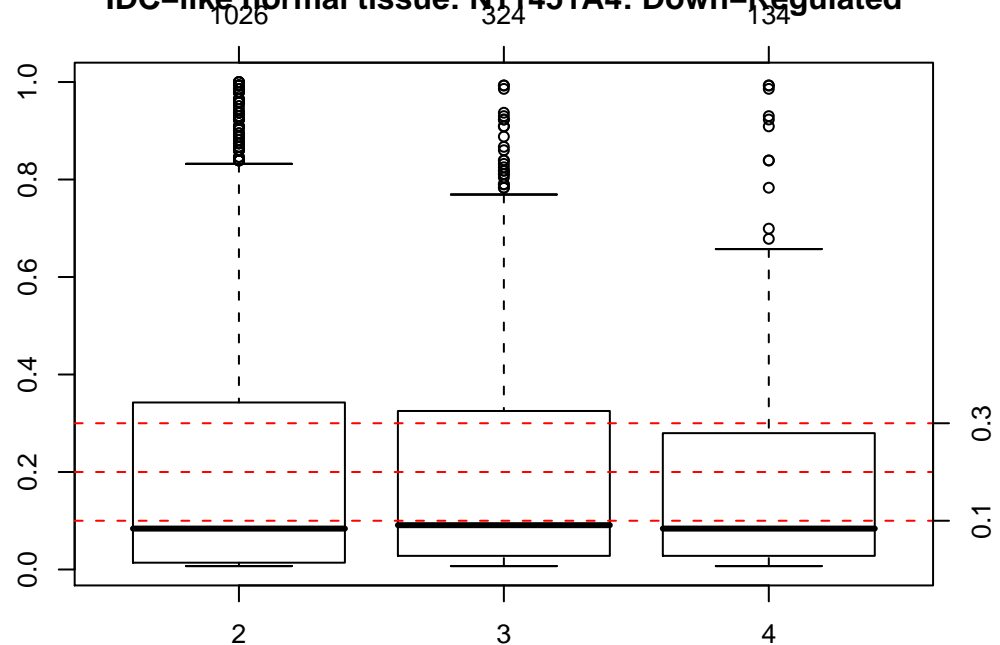
**IDC-like normal tissue: N8380A2: Down-Regulated**



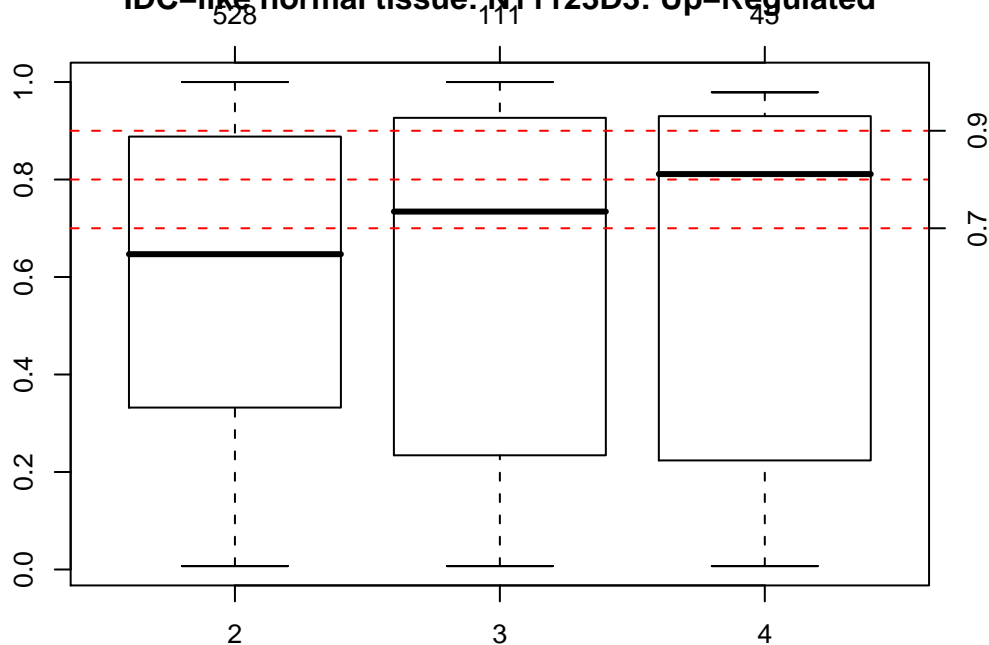
**IDC-like normal tissue: N11451A4: Up-Regulated**



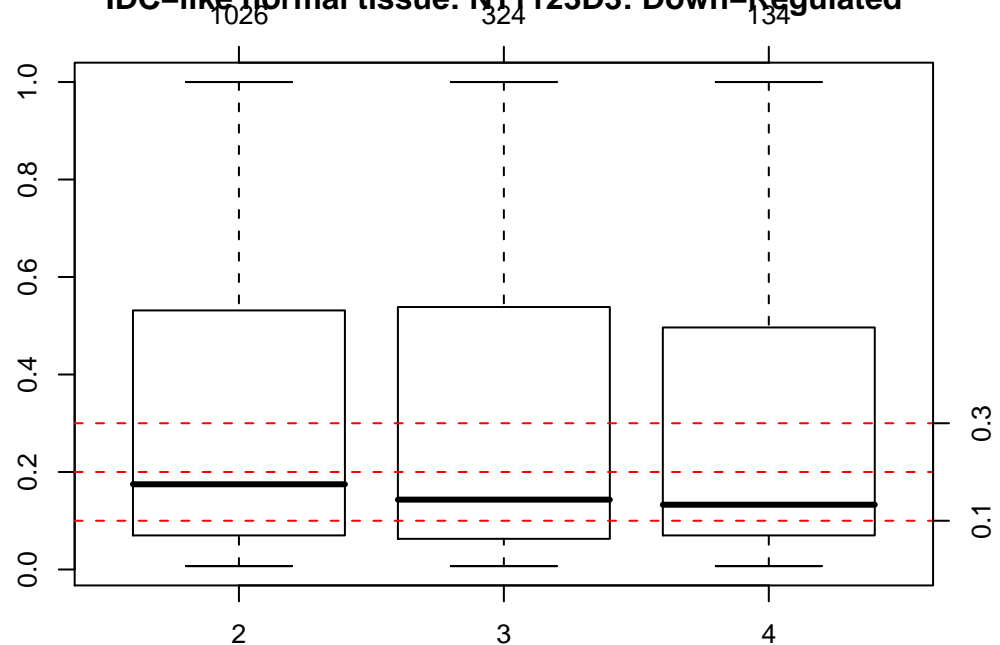
**IDC-like normal tissue: N11451A4: Down-Regulated**



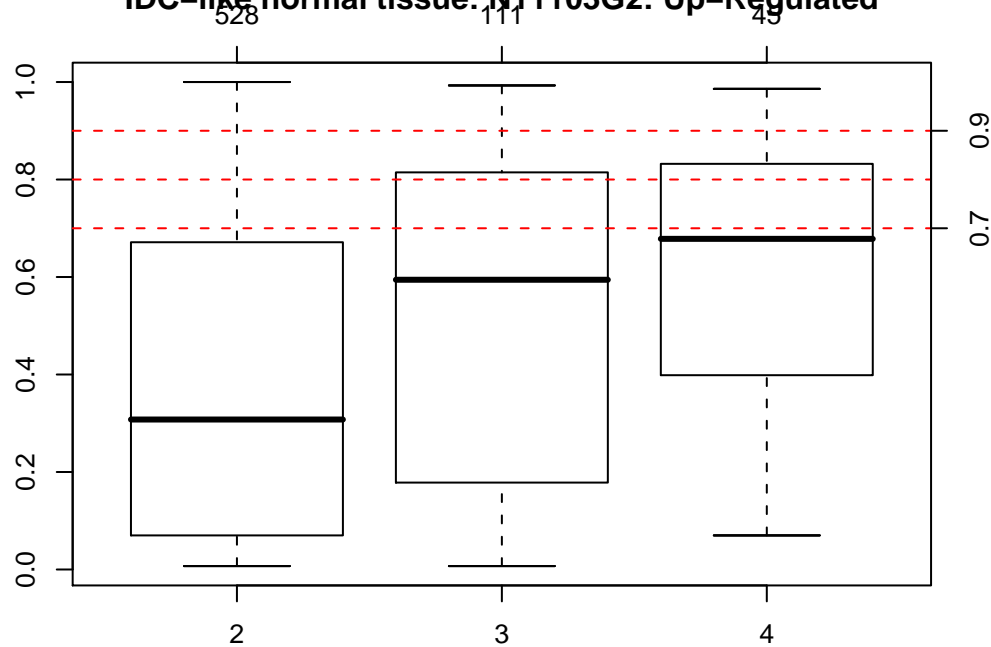
**IDC-like normal tissue: N11123D3: Up-Regulated**



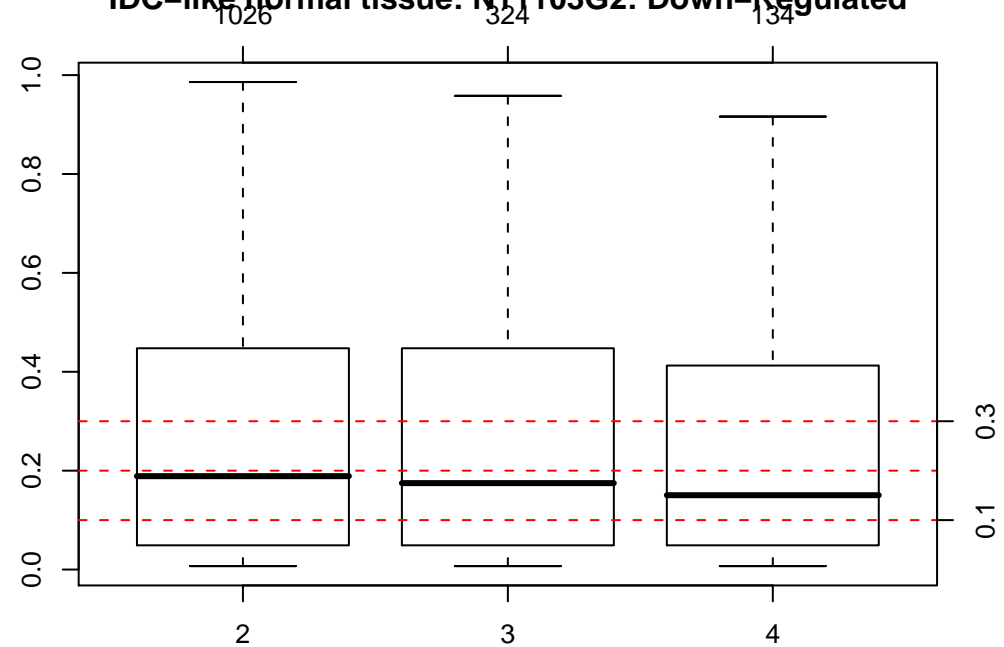
**IDC-like normal tissue: N11123D3: Down-Regulated**



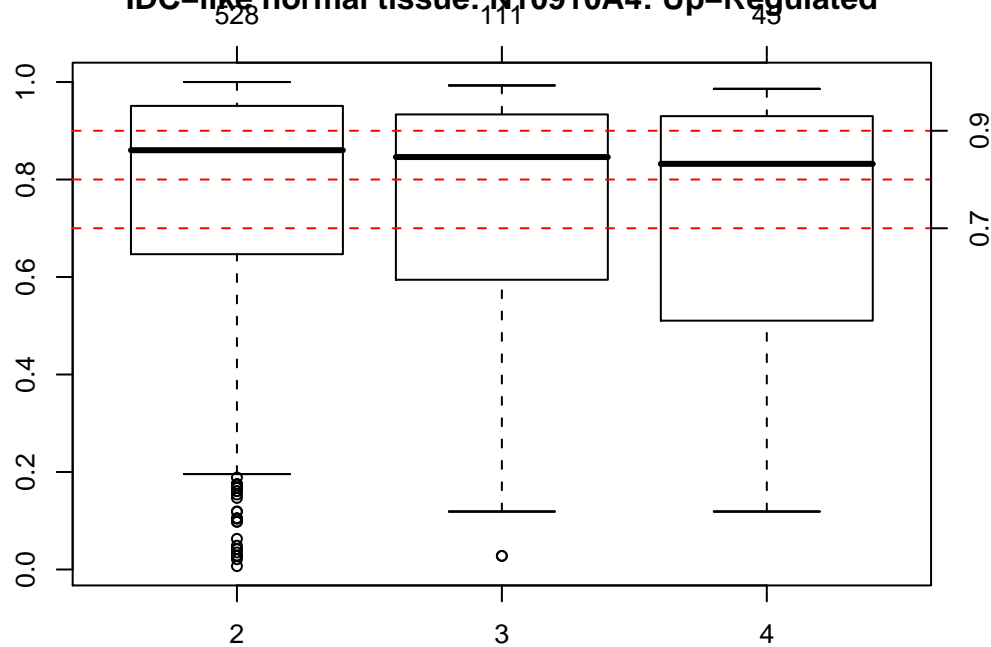
**IDC-like normal tissue: N11103G2: Up-Regulated**



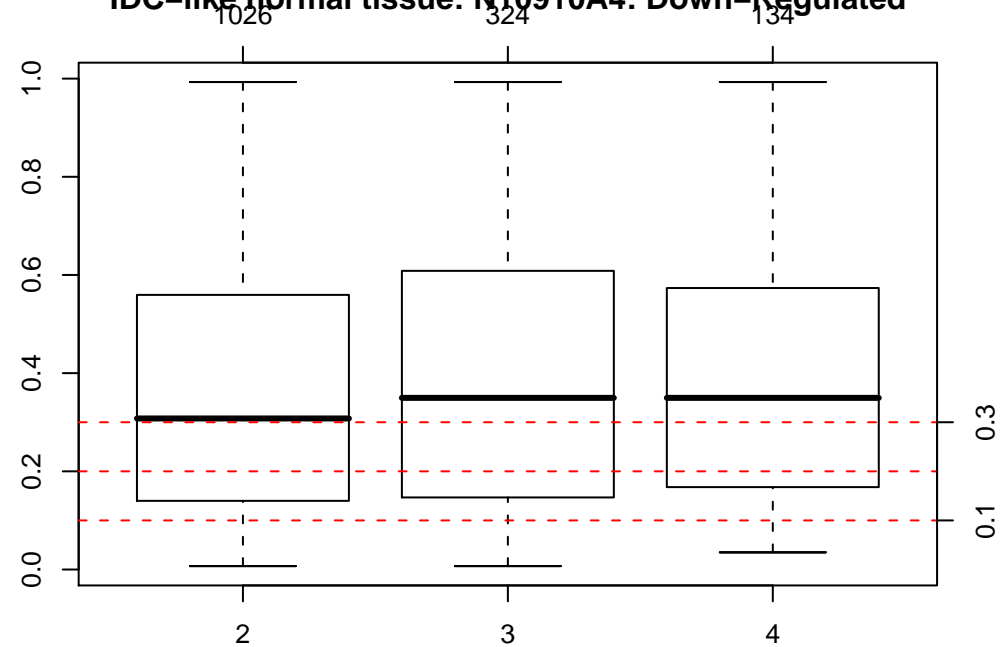
**IDC-like normal tissue: N11103G2: Down-Regulated**



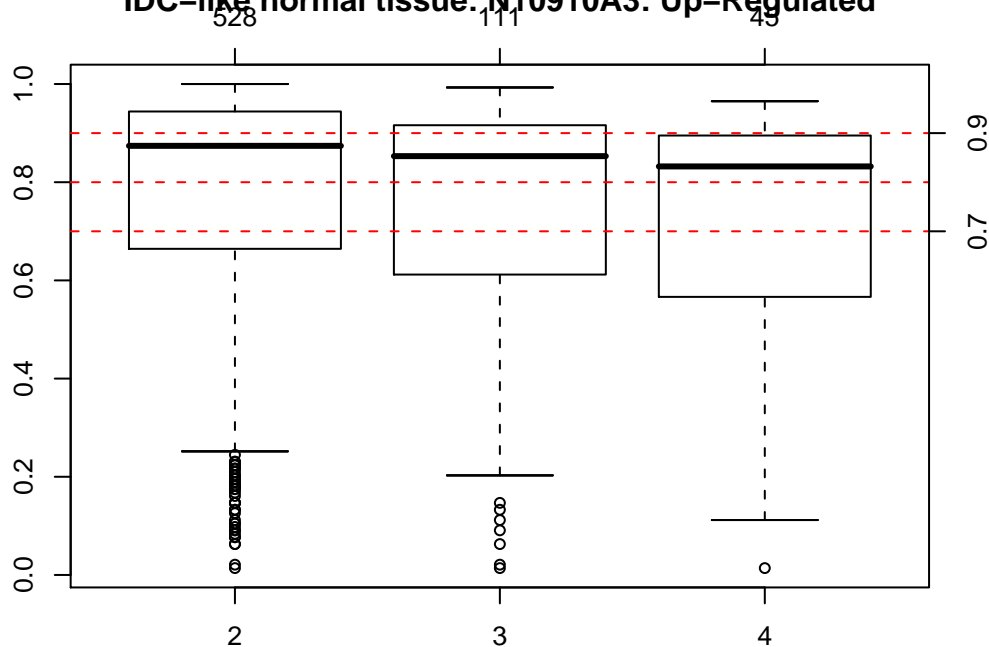
**IDC-like normal tissue: N10910A4: Up-Regulated**



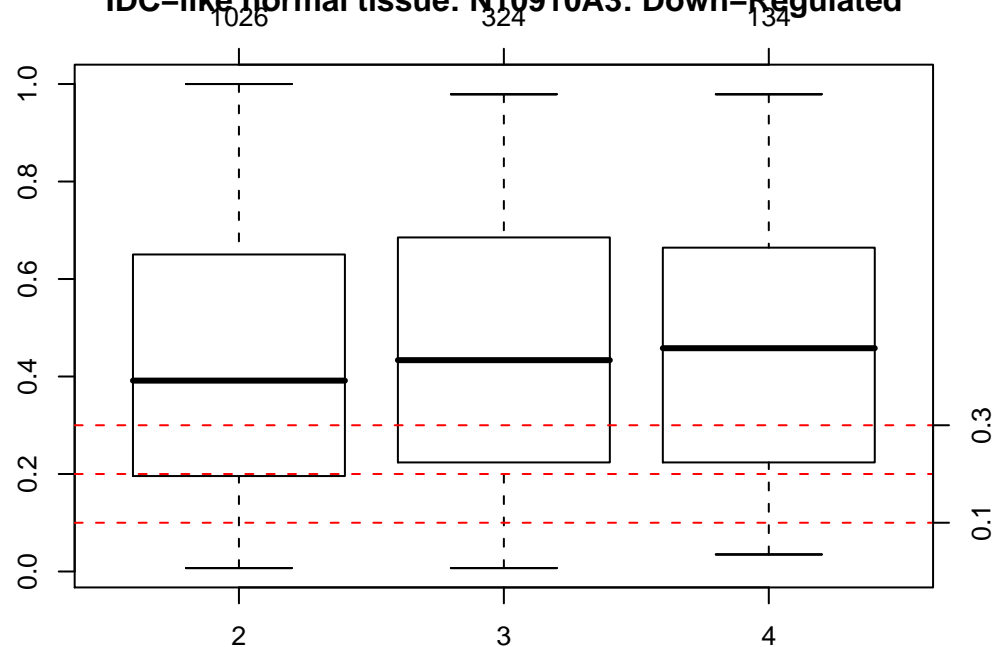
**IDC-like normal tissue: N10910A4: Down-Regulated**



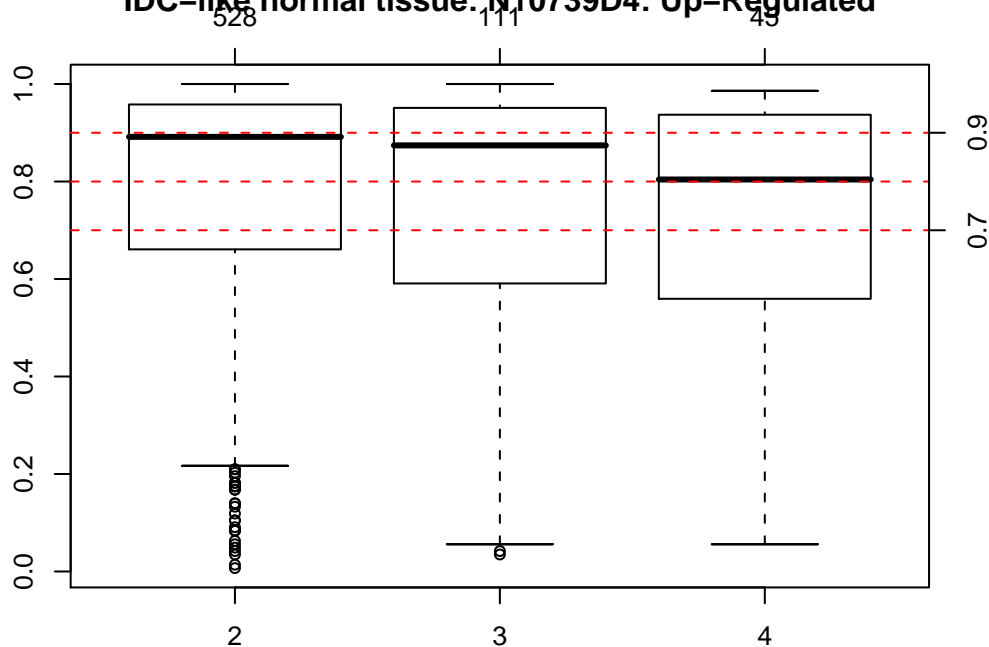
**IDC-like normal tissue: N10910A3: Up-Regulated**



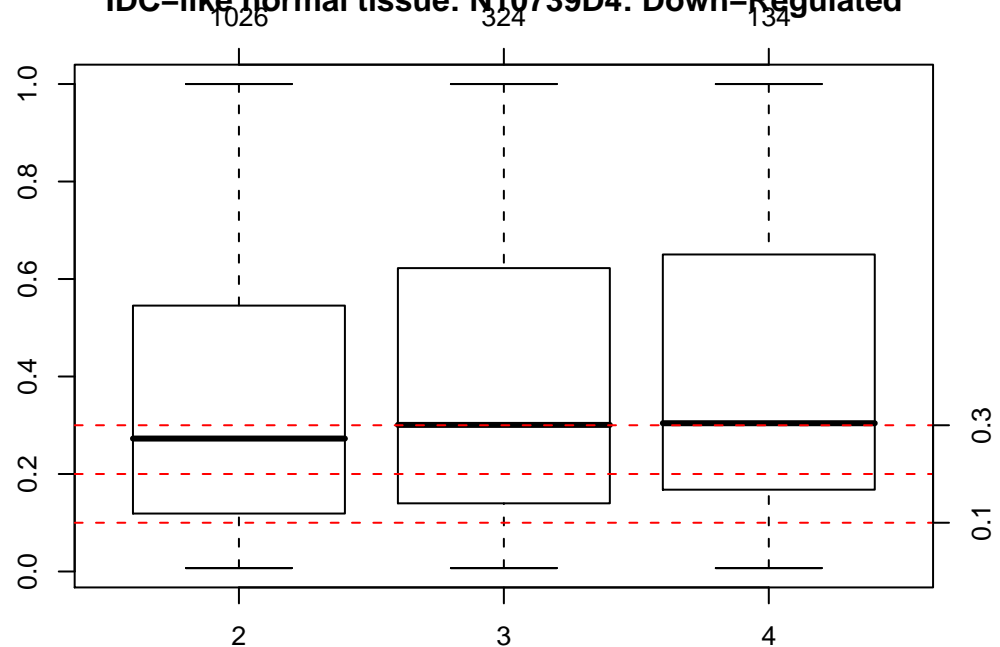
**IDC-like normal tissue: N10910A3: Down-Regulated**



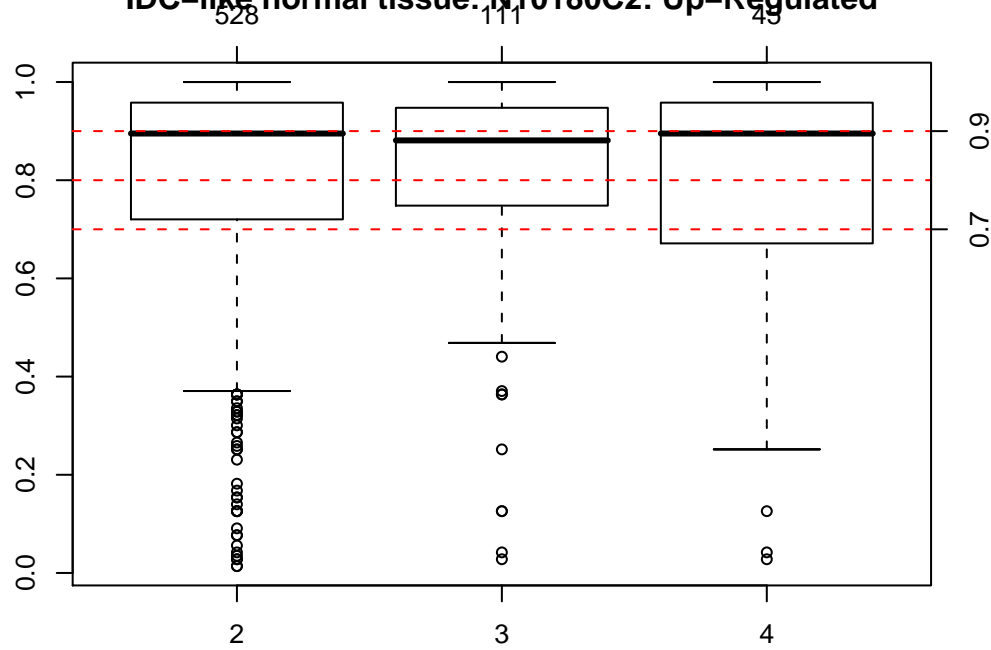
**IDC-like normal tissue: N10739D4: Up-Regulated**



**IDC-like normal tissue: N10739D4: Down-Regulated**



IDC-like normal tissue: N10180C2: Up-Regulated



IDC-like normal tissue: N10180C2: Down-Regulated

