



FIGURE S2. — Procedures Differentiating LCGI- and SCGI- promoters. We found that among genes harboring CpG islands in their promoters (CGI-promoters), a subset of genes with unusually long CpG islands in their promoters exhibit distinctively narrower gene expression breadths. This pattern was obvious when we used several different cutoff values to define ‘long-CpG island’ promoters (see below).

In Figure S2, we show results from human promoter analyses, using several cutoff values (5%, 10%, 12.5%, 15% in each bin). From the analyses of 20 bins (5% of the promoters in each bin; Figure S2A), we found that bins with mean $> \sim 2\text{Kb}$ showed reduced expression breadth compared to other bins. This value of $\sim 2\text{Kb}$ did not change when different bin sizes were used (10 bins, 8 bins, 6 bins, representing 10%, 12.5%, 17% cutoff: Figures S2B-D). We chose to use the cutoff value of 2kb, and defined CGI promoters with longer than 2kb CGI as ‘long-CGI promoters’. Promoters with CGI lengths shorter than 2kb are defined as ‘short-CGI promoters’.

The difference in expression breadths between SCGI and LCGI are highly significant ($P < 10^{-14}$, t -test). Using slightly different cutoff values near 2kb also provides highly significant results (for example, cutoff = 1600bps; $P < 10^{-10}$, 1800bps; $P < 10^{-12}$, 2200bps; $P < 10^{-10}$). However, the P -value was the smallest when 2kb was used, supporting that 2kb represents the value closest to the actual cutoff between LCGI and SCGI.