Supplemental Fig S2: Details of model design. (A) Three branches of convolutional blocks were used to deal with sequence features, chromatin features, and TF binding features, respectively, in which each convolutional block employs Conv2D of sape C*1*K to deal with cCRE's inputs independently. Then, three latent featrues output from three branches were concatenated along with the C channel and input to a branch of transformer blocks, in which each transformer block emplys the multi-head attention mechanism to learn the relationships between genes and their corresponding cCREs. At last, the middle row of outputs corresponding to the gene index was extracted, representing the relationships between a gene and its surrounding cCREs, and was then fed into the output module to predict the expression level of genes. (B) The interaction matrix, signifying if there exists interations between genes and their cCREs, can be directly fed into each transformer block by modifying the equation of computing attentions, or this matrix can be indirectly incorporated into the supervised training process by adding a KL Divergence loss to the training loss.