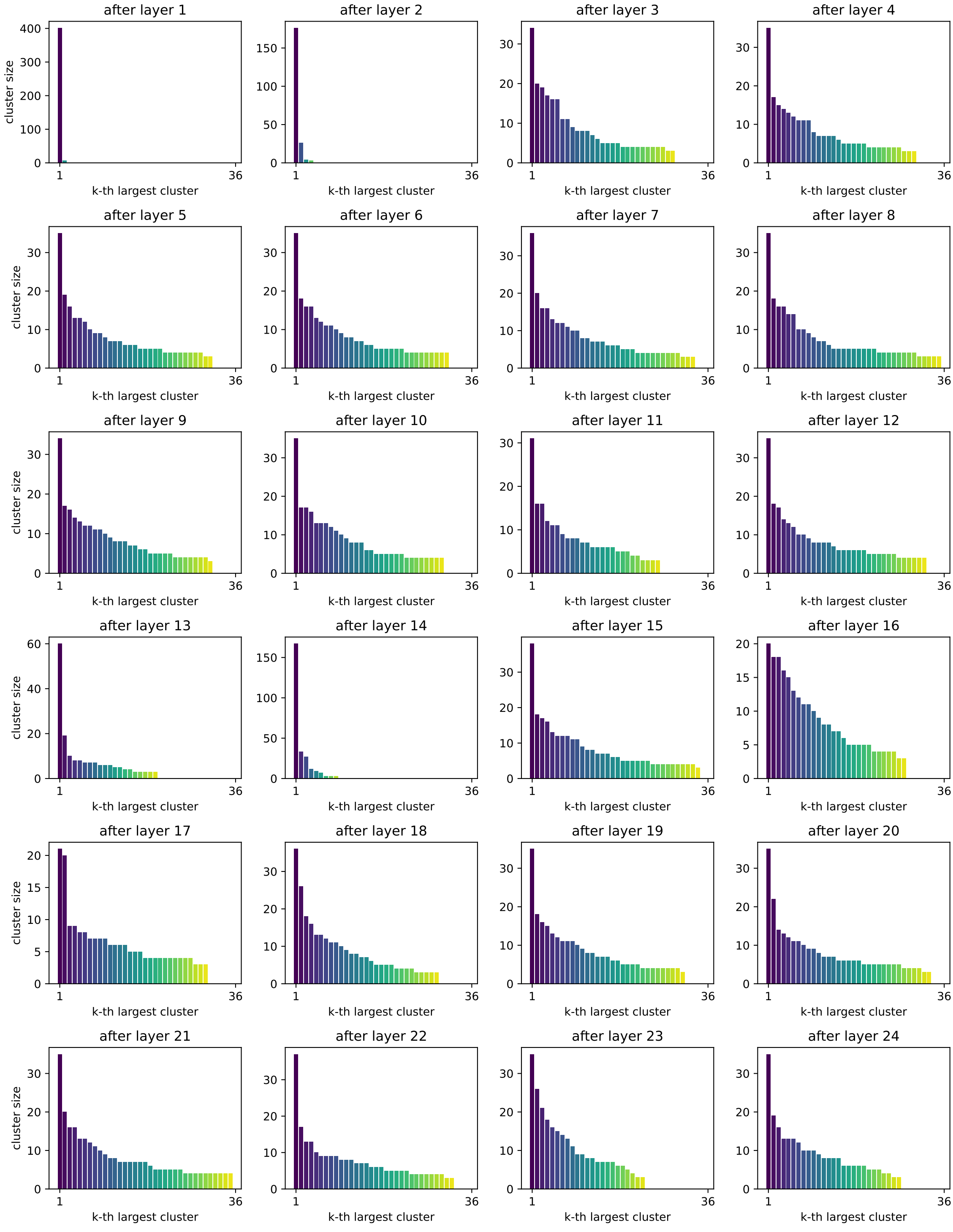HDBSCAN token cluster sizes
(wikitext_bert-large-uncased_params-trained_dense-off_layers-24)