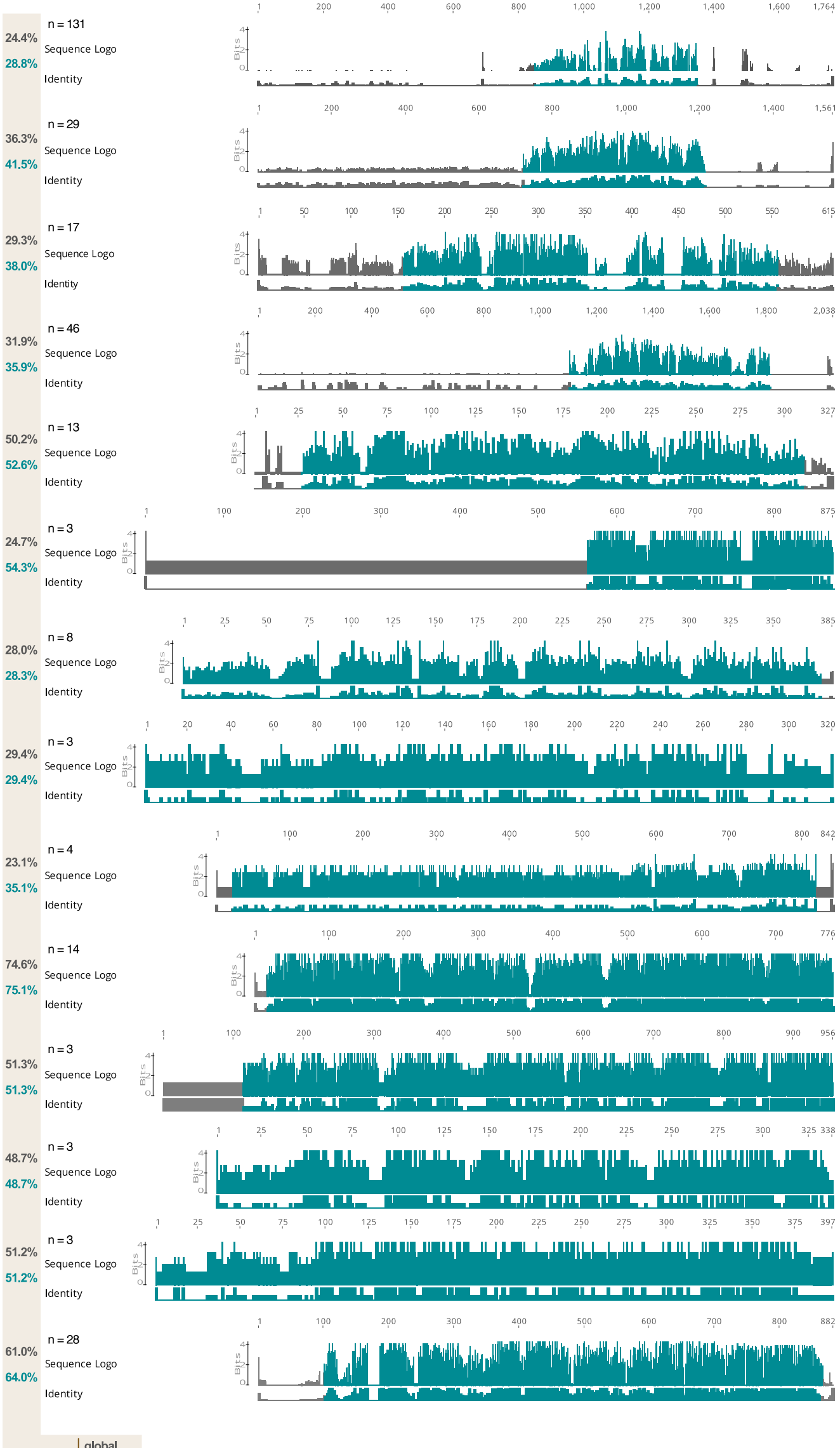

Figure S2. Alignment of the CE10 sequence clusters. All CE10 protein sequences of a clusters (1 to 14) were aligned using MAFFT with the G-INS-I settings, resulting in 14 alignments. The number of sequences in a cluster ("n =") are given on the left. The sequence fingerprint ("sequence logo") and a sliding window-based plot of the identity of these 14 alignments are depicted on the right. In the highlighted text on left, the global identity is given in grey font and the local identity is given in teal font; the teal colour of the alignment plots highlights the conserved region used for calculating the local identity. Note that sometimes the entire sequence was used for calculating the local identity due to its high conservation throughout the alignment.