

**Figure S3** Average distance to neighbouring samples using kNNi and LD-kNNi. To allow comparison, the  $d_l$  distance measure (Equation 3) is used in each case and both methods use k = 5. Exactly one is shown as a separate bar due its high frequency. The average distance to neighbouring samples is far lower for LD-kNNi than for kNNi.