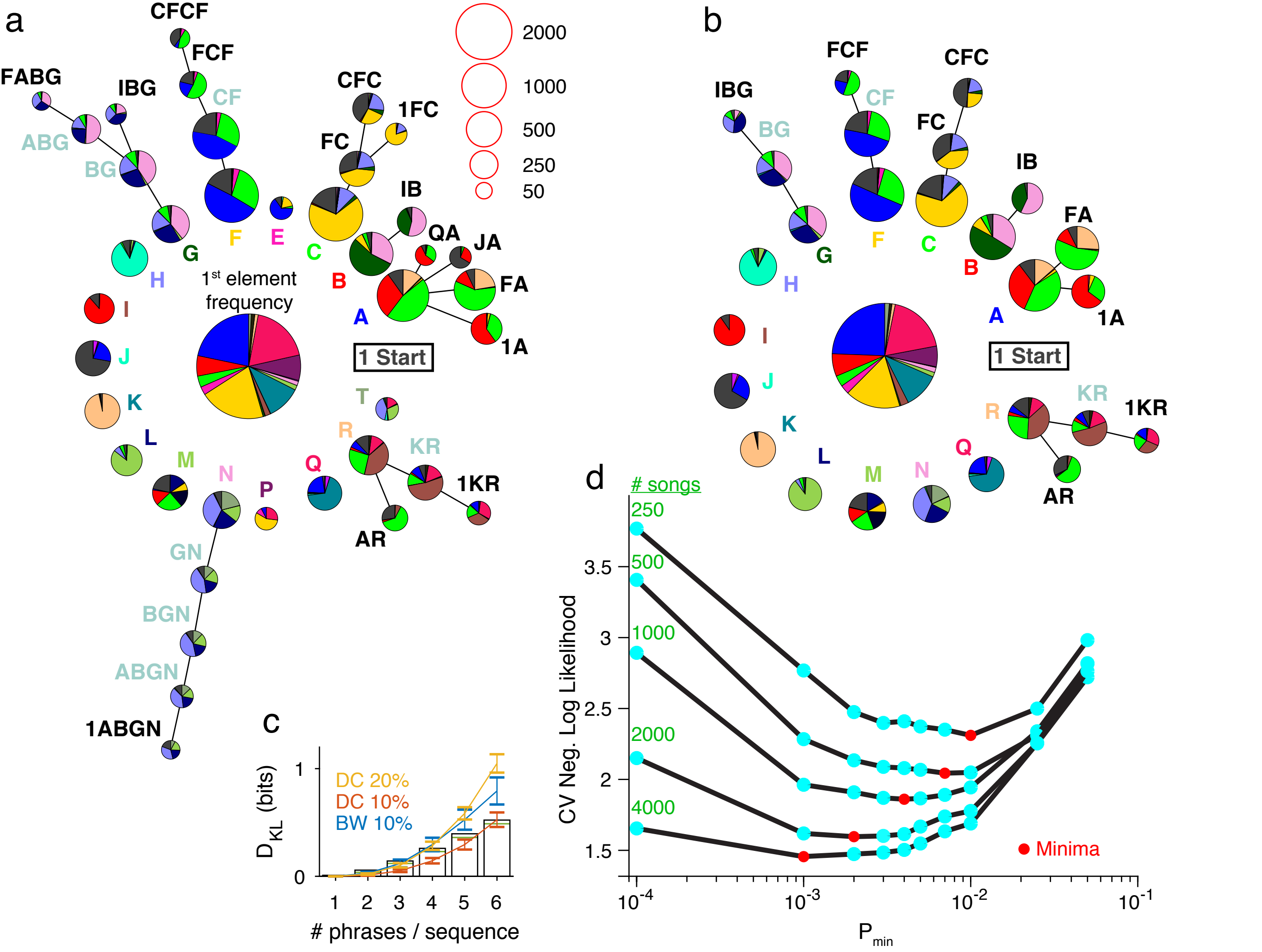

Figure 4 | **TweetyNet reproduces song syntax structure in greater precision using datasets more than 5 times larger than previously explored. a.** Long-range order found in ~2000 domestic canary songs. Letters and colors label different phrase types. Each branch terminating in a given phrase type indicates the extent to which song history impacts transition probabilities following that phrase. Each node corresponds to a phrase sequence, annotated in its title, and shows a pie chart representing the outgoing transition probabilities from that sequence. The nodes are scaled according to their frequency (legend). Nodes that can be grouped together (chunked as a sequence) without significantly reducing the power of the model are labeled with blue text. **b.** Repeats a for a subset of 500 songs. **c.** Bootstrap estimation of distribution reliability is calculated for phrase sequences of length 1-6 (x-axis). Error bars show the mean and standard deviation of the distance (D_{KL}, y-axis) between the sequence distribution of the full set of songs and the same set after ommiting a random 10 or 20 percent in red and blue. Bars and green lines show the 0.95 quantile and median of the sample ditribution (methods). Yellow error bars shows the same estimation for a previously published Belgian Waterslager canary dataset that is 5 times smaller. **d.** Ten-fold cross validation is used in selection of the minimal node probability for the PSTs (x-axis). Dots show the mean negative log-likelihood of test set data estimated by PSTs in 10 repetitions (methods). Curves are calculated for datasets that are sub sampled from about 4000 songs. Red dots shot minimal values - the optimum for building the PST.