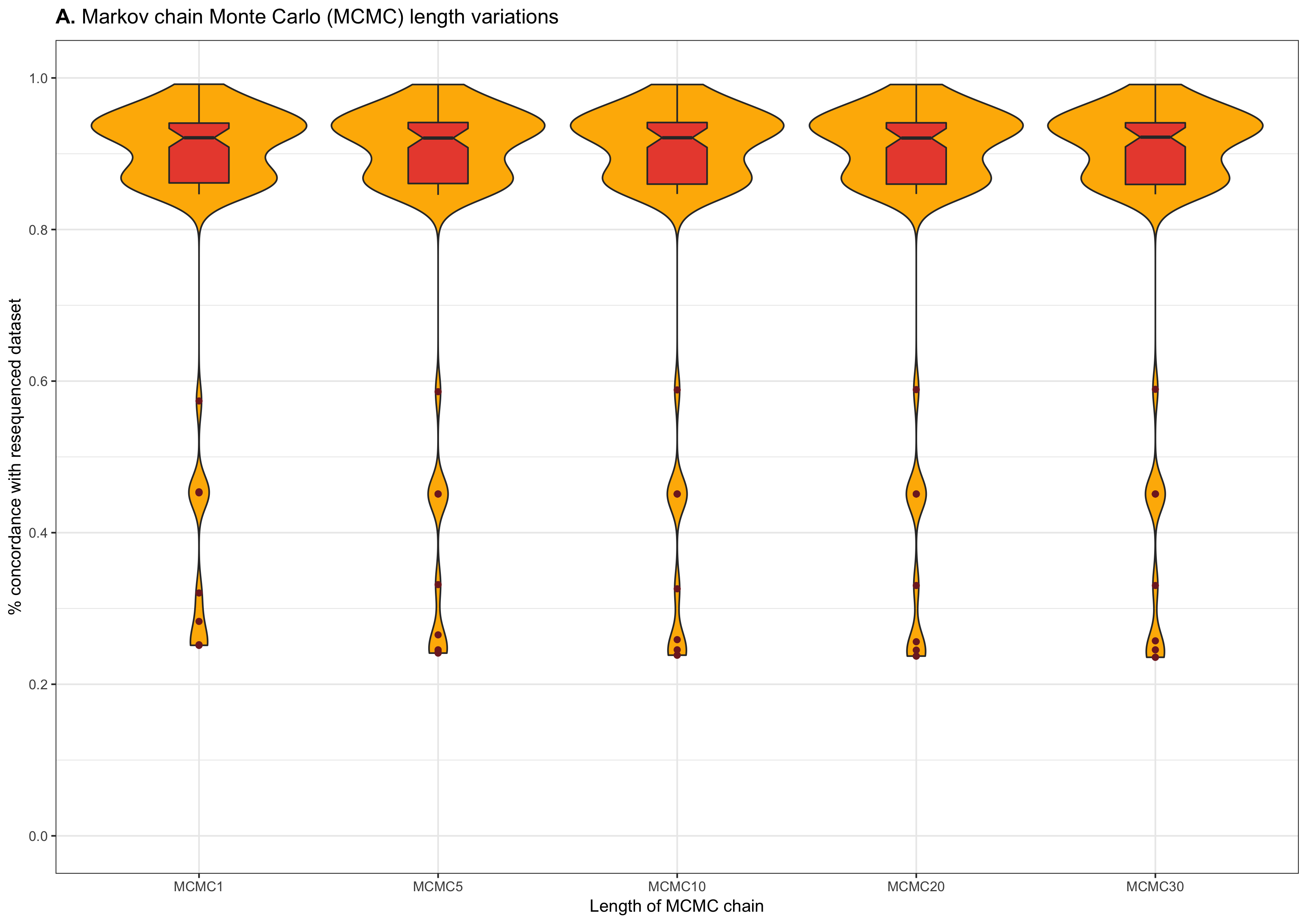
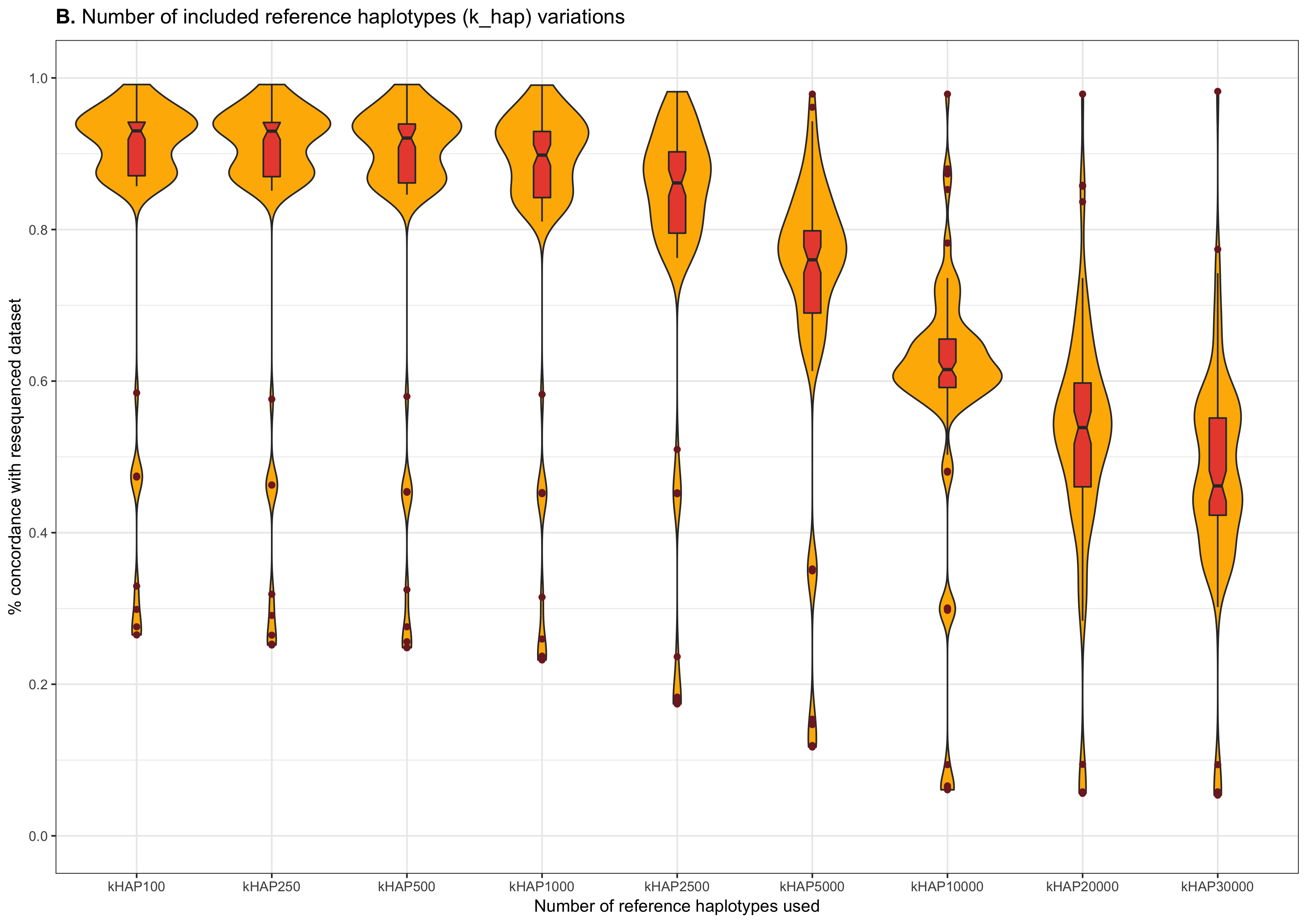
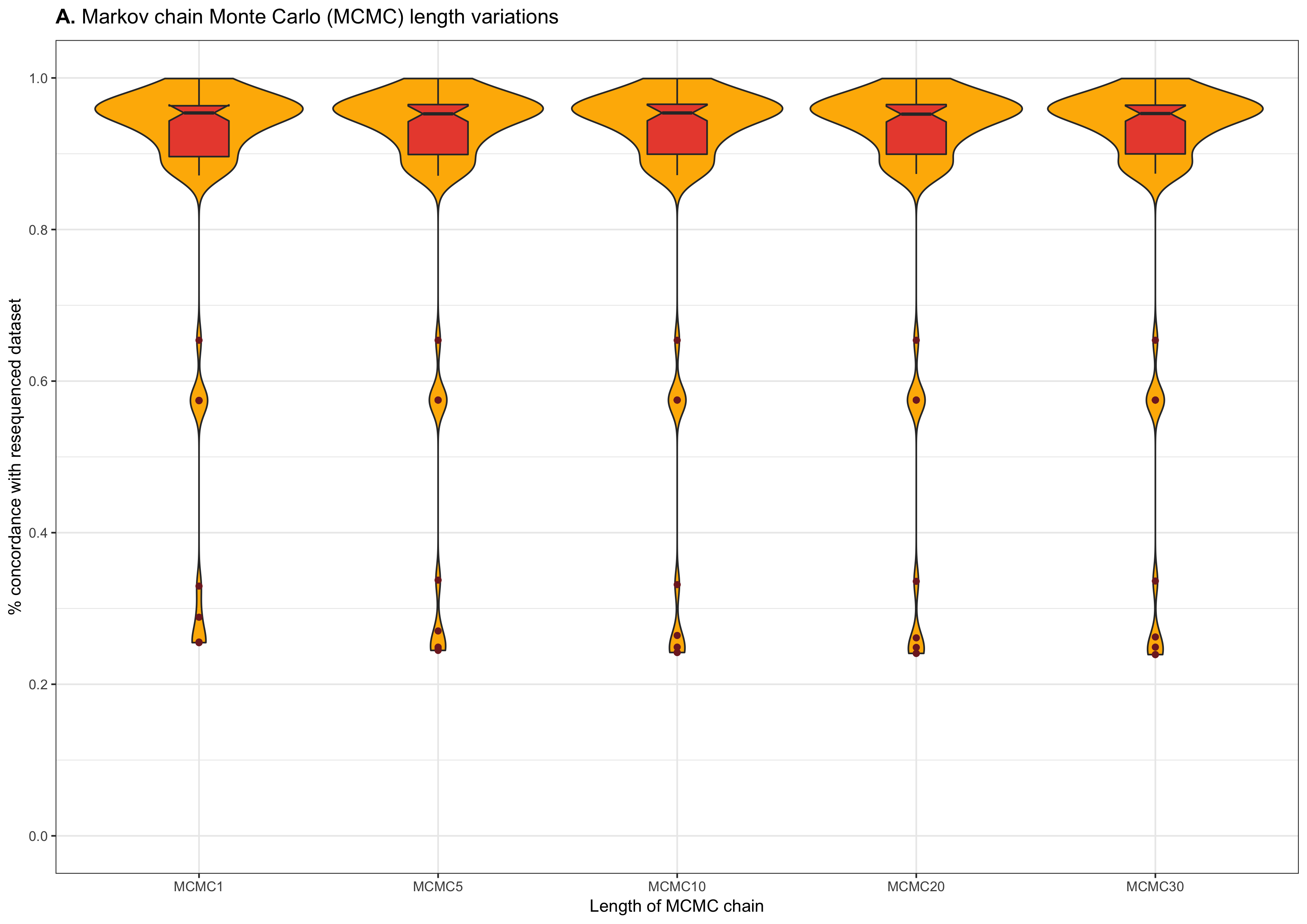
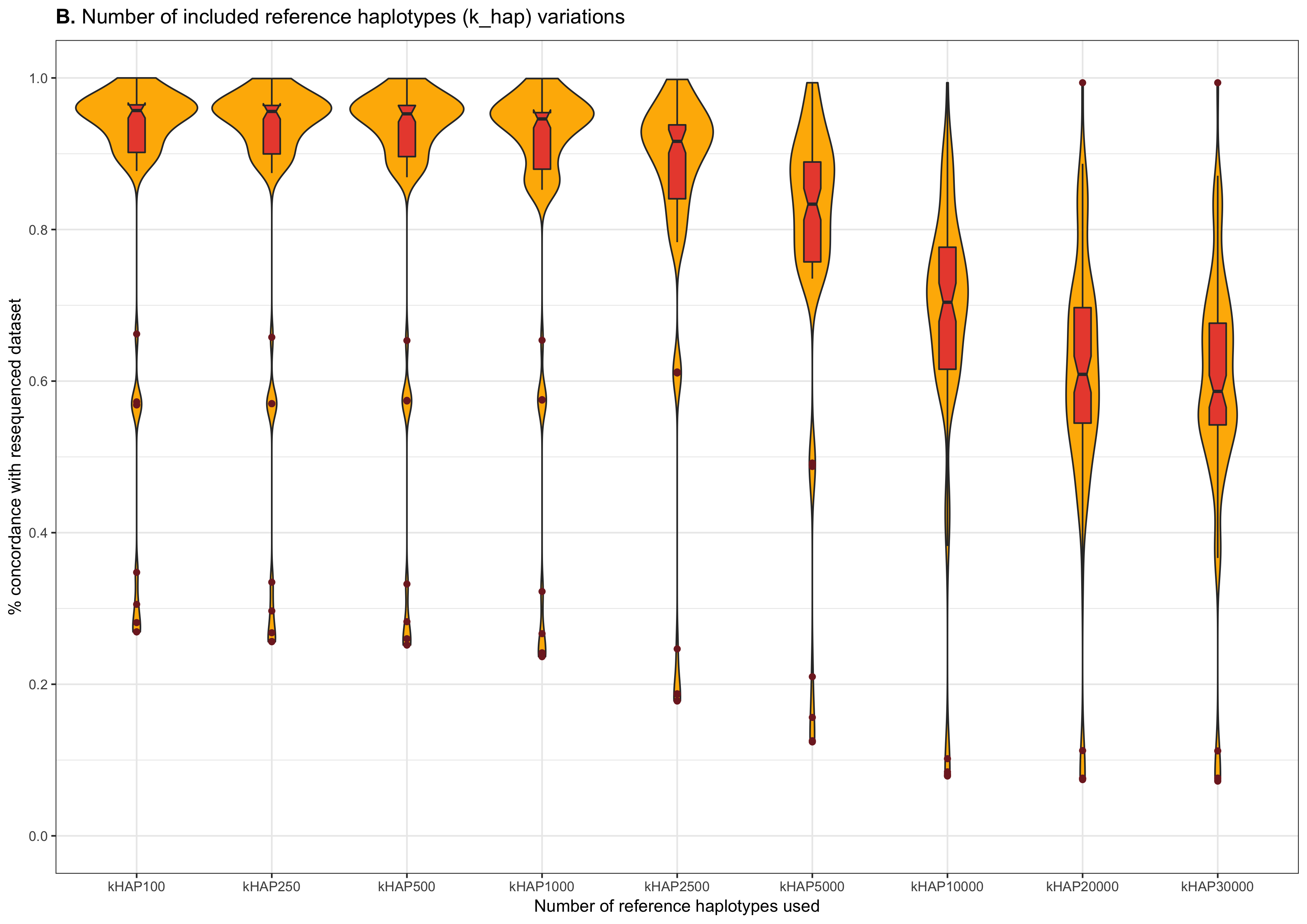
Since beginning the project just over a year ago, the following things have been achieved:

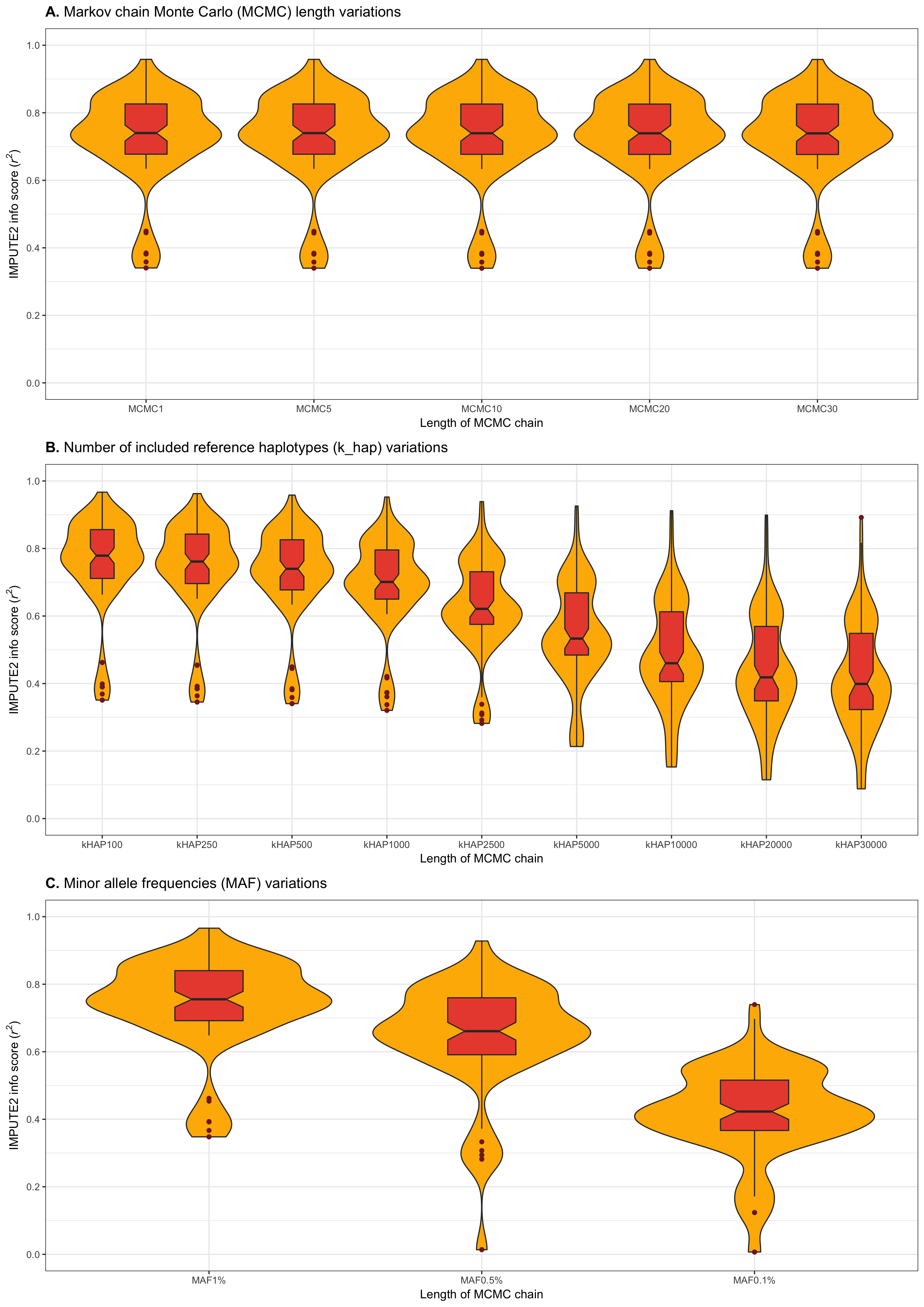
* A large reference panel of globally diverse complete human mitochondrial sequences has been constructed
  + n=44,299
  + When filtered to exclude low quality (>7 Ns), roughly ~36,000 sequences
  + Constructed by aligning to a reference alignment (as opposed to just a reference sequence). This was done to maintain consistency in gap placement.
  + This reference panel has been filtered to minor allele frequency (MAF) 1%, 0.5%, and 0.1% so we could test the effect of adding rarer variants on imputation performance
* 101 (or 103, depending on MAF) *in silico* microarrays were created to assess performance of imputation pipeline on SNVs called from commonly used microarrays
  + These were tested against different MAF, as well as varying the number of reference haplotypes IMPUTE2 considers for imputation (khap), and tested the effect of varying lengths of the Markov chain Monte Carlo length
  + We assess imputation accuracy by means of haplogroup concordance
  + HiMC was used to assign haplogroups.
    - MAF did not affect haplogroup concordance
    - MCMC did not affect haplogroup concordance
    - K\_hap did affect haplogroup performance. As you increase the number of considered haplotypes, the ability to correctly assign haplogroups decreases.
  + IMPUTE2’s *r*2 metric was also used as a guide for imputation performance

Micro-haplogroups:



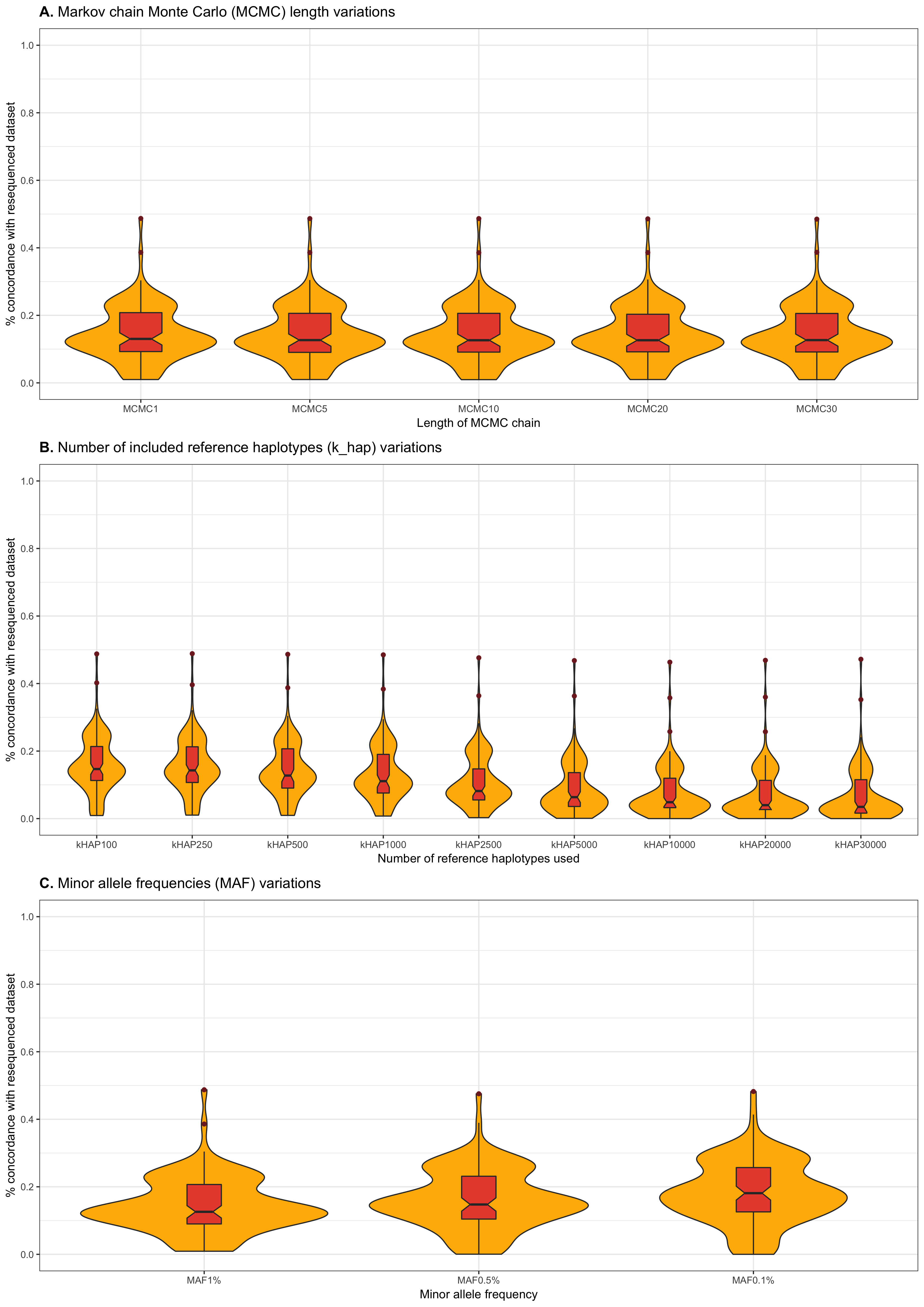
Macro-haplogroups:



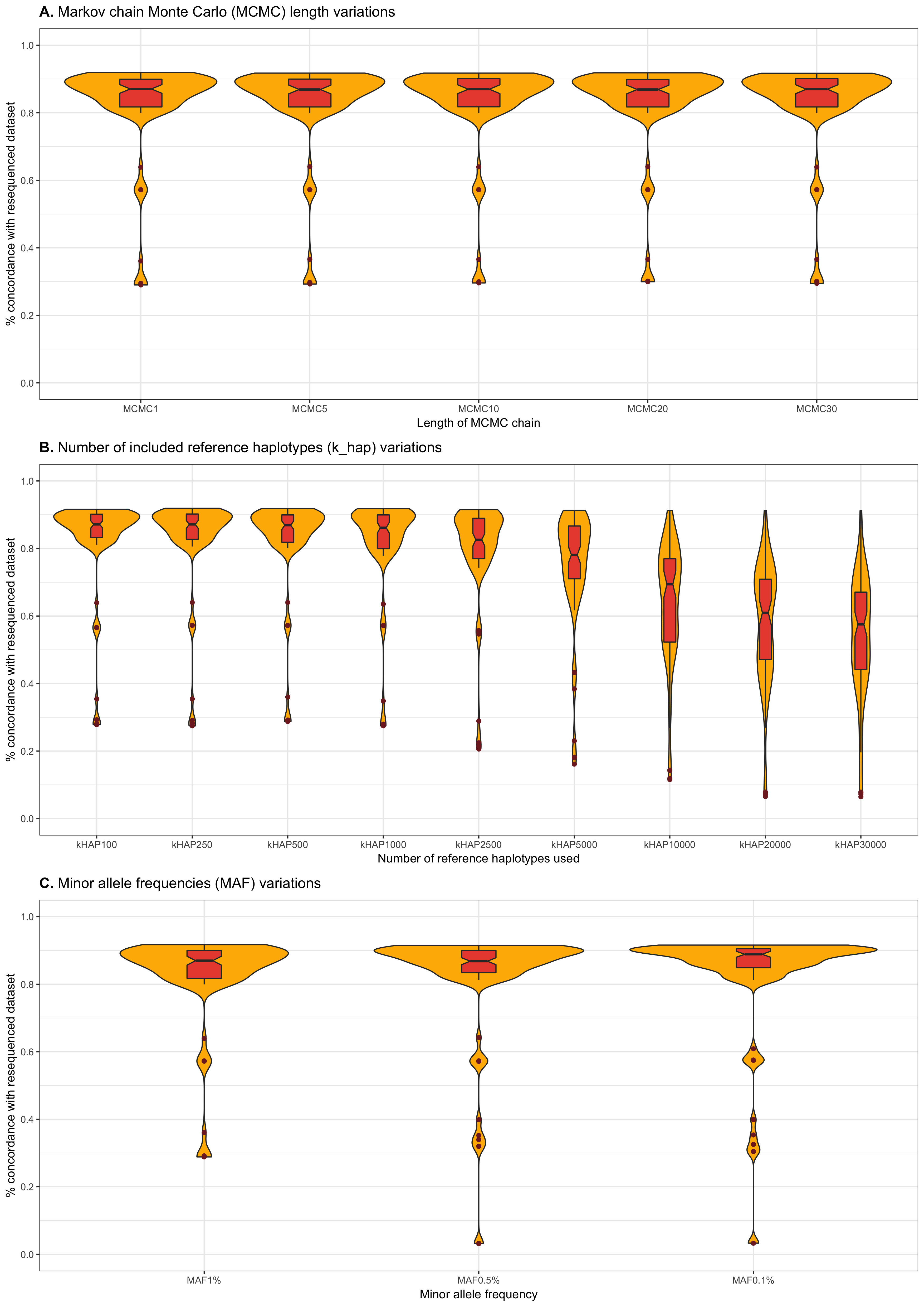
INFO SCORE:

* We also applied HaploGrep2.0 to assign haplogroups
  + Haplogroup concordance showed similar patterns when macro-haplogroups were considered. However, MAF did have a noticeable affect this time.
  + Haplogroup concordance when downstream, or specific, micro-haplogroups showed that imputation doesn't work well for correctly assigning all layers of the haplogroup tag

Micro-haplogroups:

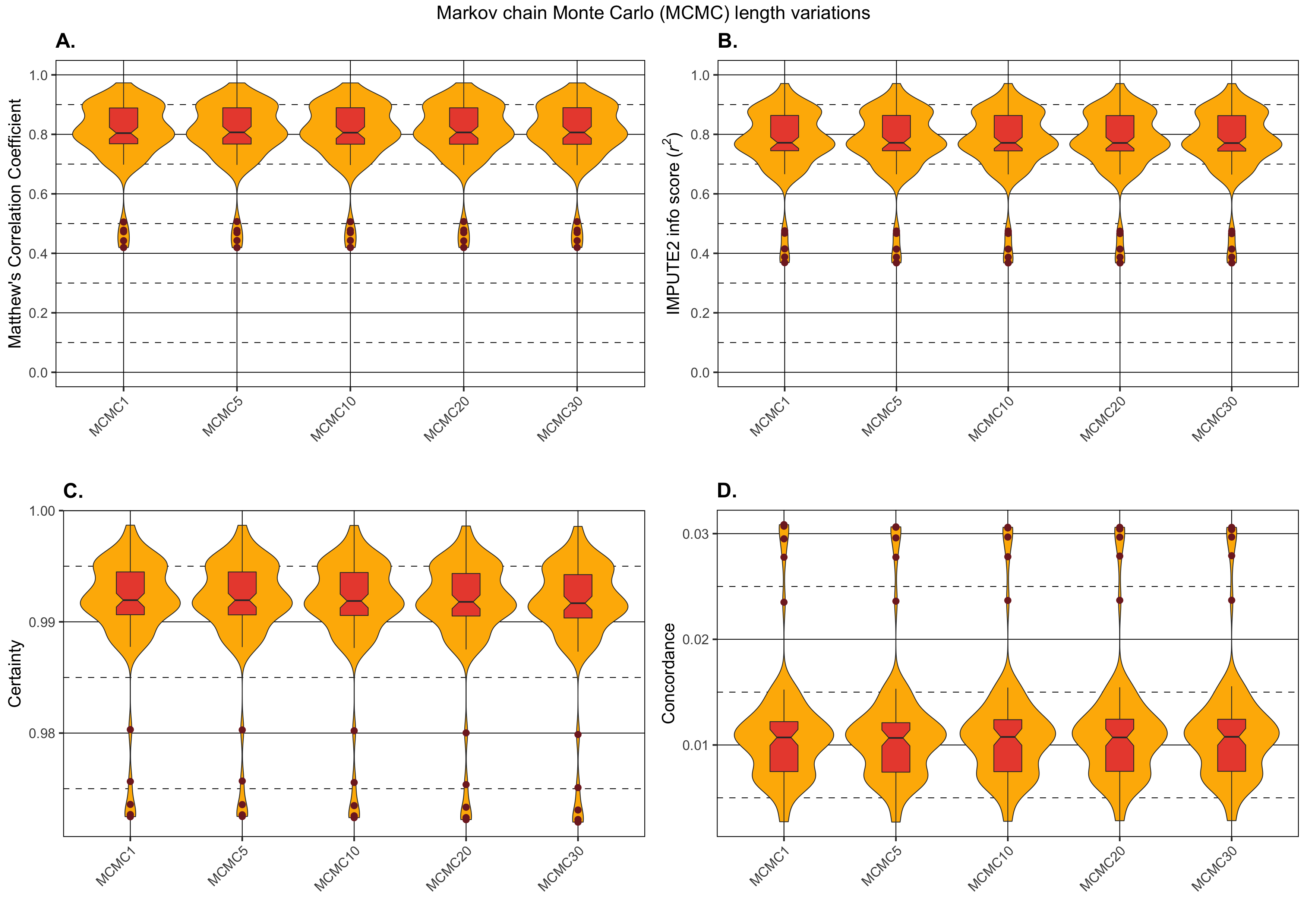
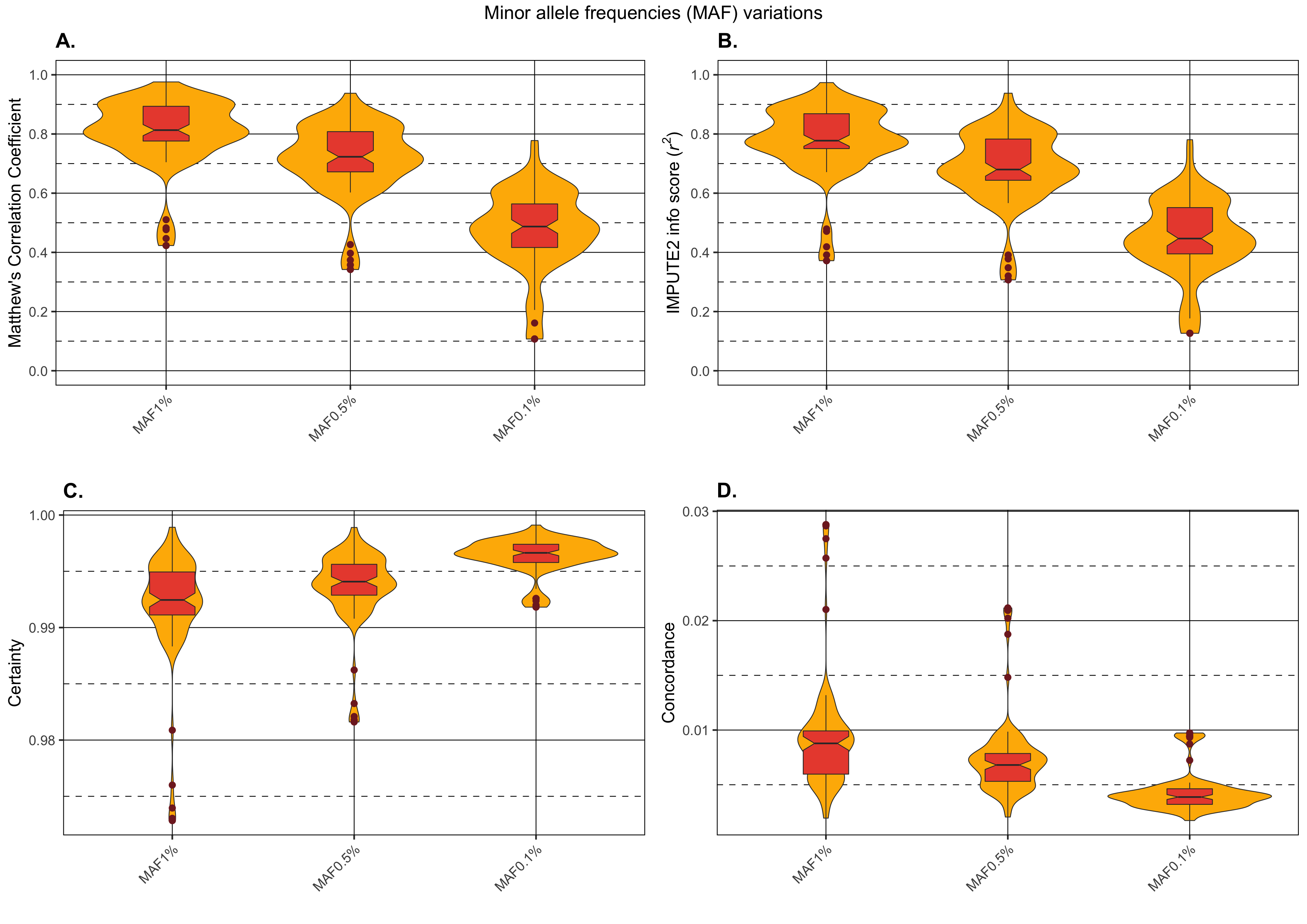
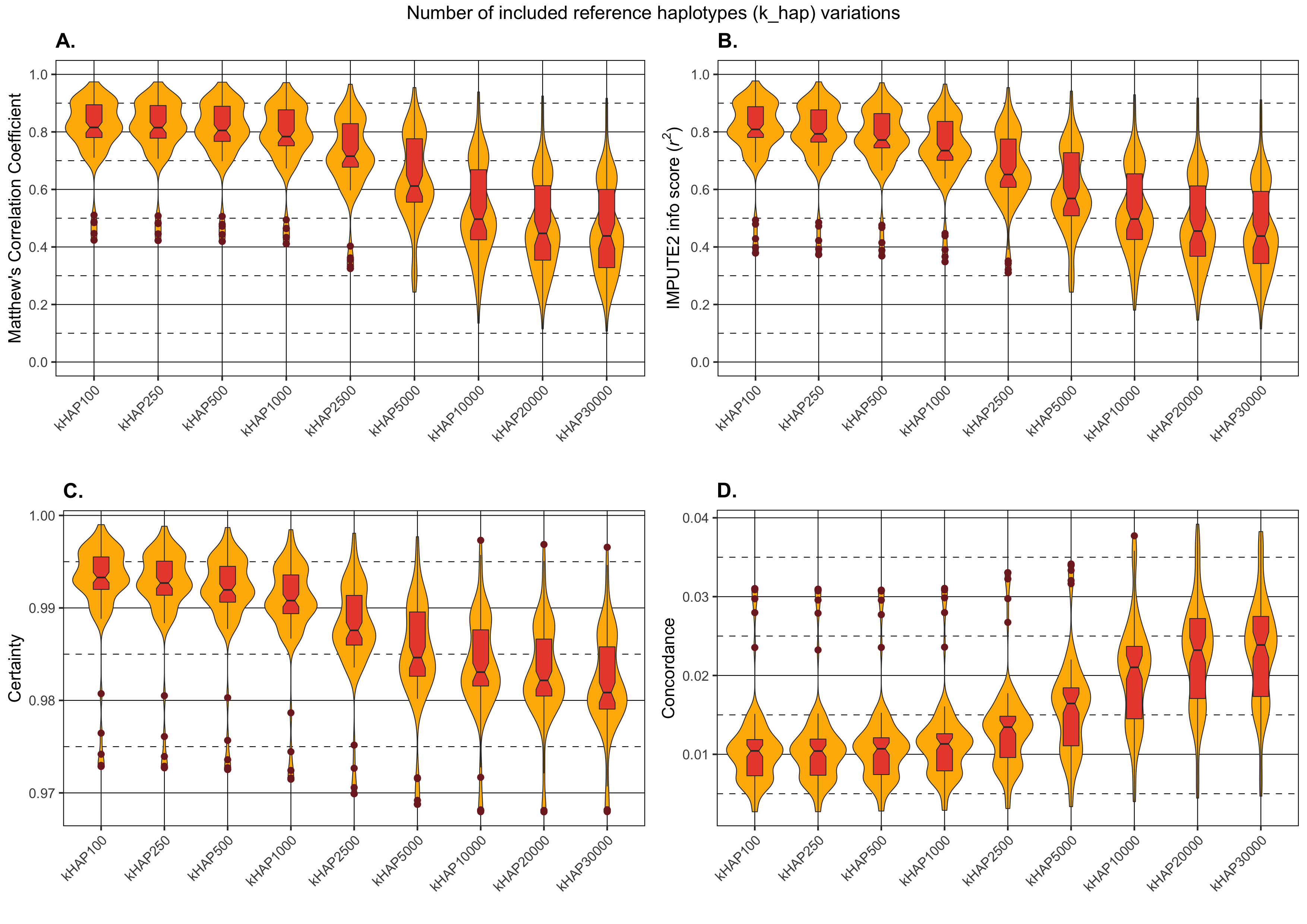


Macro-haplogroups:



* Recently, we have used the Matthew’s correlation coefficient (MCC) to determine genotype concordance
  + Shows same general trend as HaploGrep concordance measure
    - Decreasing MAF shows decrease in imputation performance
    - Increasing k\_hap shows decrease in imputation performance
    - MCMC is same
* Have applied the MCC concordance measure to the ADNI data, however at the moment we are only observing concordance of MCC = 0.13, which is not good
  + We are re-running the ADNI data
  + However, this may be a true result
  + Looking at different MAF values on imputation for ADNI at current (Devashi Paliwal)

Matthew’s Correlation Coefficient



* Paper all but written, just need to add in stuff about genotype concordance/MCC
* Will update the reference panel to include new sequences