Current method for creating background strain mosaics:

• Each marker is clustered in the xy-intensity space using CC founder and F1 replicates and DO samples. The number of clusters vary by marker, and the majority of markers have 3 clusters -- 2 homozygous and 1 heterozygous.

• At every marker, each MMRRC background strain is assigned a most likely cluster based on the probability of its xy-intensity value belonging to each cluster.

• At every marker, the target strain's xy-intensity value is compared with each of the two background strains' clusters to obtain an emission probability. The closer the target's xy-intensity is to a background strain's cluster (in terms of Mahalanobis distance), the higher the probability of that background strain emitting the observed target intensity. This emission probability, along with a transition probability, are fed into a hidden Markov model to find the mosaic of background strains at each marker that most likely resulted in the sequence of intensities observed across the genome.

• The 'unknown' background case is currently modeled as a constant, smaller-than-average emission probability for each marker. The probability of transitioning into the 'unknown' background from a known background are lower than probabilities for other types of transitions.

Reading the results:

•Noninformative markers occur where the two background strains have the same Illumina genotype call, or one of the backgrounds or the target strain has an 'N' call.

We only consider concordant and discordant calls among informative markers:
• For homozygous regions:
  ◦ Concordant calls are calls where the the Illumina genotype call for the target and the assigned background strain are the same.
  ◦ Discordant calls are where the Illumina genotype call for the target equals that of the other background strain, and neither are N calls.

• For heterozygous regions:
  ◦ Concordant calls are calls where the target's Illumina genotype call is 'H'.
  ◦ Discordant calls are where two background strains have differing genotype calls and the target strain is not called 'H'.