Indels : compare GATK with DINDEL; investigate the source of 1bp insertion false positives

GATK's genotype likelihoods calculation is inspired by and based on Dindel but GATK is based on a heuristic model requiring a minimal number of consistent indel alleles present in the reads(This computation is used to derive the prior probability each of the possible 10 diploid genotypes, using Bayesian formulation).

where

Let h is the probability that an aligned read was sampled from the chromosome with the reference allele and is the probability of observing a read with the allele given the genotype .

Reference:

Aaron McKenna, et al The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data, Genome Res. July 2010)