

Statistical Modeling of Methylation Patterns in Ovarian Carcinomas

Michelle Lacey

Abstract

Changes in cytosine methylation at CpG nucleotides are observed in many cancers, but the biological mechanisms responsible for these changes are not yet fully understood. Previously developed stochastic models for cancer-related methylation change have either treated CpG sites independently or employed a context dependent approach to adjust model parameters according to regional methylation levels. However, our analyses of double-stranded methylation patterns in 0.2 kb regions of the tandem repeats Sat2 and NBL2 have detected small clusters of identically methylated sites in close proximity that could not be explained by random variation. These findings suggest a high degree of site-to-site dependence, and we have developed a neighboring sites model for methylation change as an alternative approach. We have compared the independent sites, context dependent, and neighboring sites models in their ability to generate simulated sequences statistically similar to our Sat2 and NBL2 carcinoma samples, and we demonstrate that the neighboring sites model is preferred in the majority of the cases considered.