Best practices for SNV and methylation calling from bisulfite sequencing data

Guillermo Barturen[1,2], José L. Oliver[1,2] and Michael Hackenberg[1,2]

[1]Dpto. de Genética, Facultad de Ciencias, Universidad de Granada, 18071-Granada, Spain
[2]Lab. de Bioinformática, Inst. de Biotecnología, Centro de Investigación Biomédica, 18100-Granada, Spain

Abstract. The advent of high-throughput sequencing techniques, together with bisulfite treatment of the DNA, allows for genome methylation profiling at a single cytosine resolution. Some modern tools, as *MethylExtract* [1], are also able to detect variation (SNVs) using the same bisulfite sequencing library, which is crucial for many downstream analyses, for example when deciphering the impact of sequence variation on differential methylation. However, many error sources do exist like sequencing errors, bisulfite failures, clonal reads and single nucleotide variants. We evaluate here the impact of all these potential error sources, and provide a list of best practices to overcome them.

References

1. Barturen, G. Rueda, A. Oliver, J.L. and Hackenberg, M. MethylExtract: High-Quality methylation maps and SNV calling from whole genome bisulfite sequencing data (submitted).