# A Tool for analyzing and interpreting DNA double-strand break sequencing

DNA double-strand breaks (DSBs) are the most genotoxic form of DNA damage. There is extensive evidence that DSBs are a major threat to genomic stability, responsible for chromosomal instability and rearrangements, and arise through the direct action of ionizing radiation or chemicals (including many chemotherapy drugs), as well as during replication. A better understanding of their formation and occurrence is paramount to effective prevention and treatment of the associated diseases, including cancer. Despite this importance, knowledge of how the genome breaks in response to various stressors is still sorely lacking. This deficit has been a direct result of the absence of two critical research tools. Until recently, there has been no technology to reliably and accurately detect DSBs throughout the genome, and still lacking are advanced statistical methods and tools for analysis and interpretation of these challenging data. Here we present the novel advanced statistical methods implemented as user-friendly software tools that overcome this final challenge.

Working with more than 100 samples from DSB-sequencing using these various methods, we developed the first dedicated software suite for analysis of DSB sequencing data, especially customized for this unusual, emerging data type. The paucity of DSBs in genome makes their analysis particularly challenging and results in weak signal. Therefore, normalization is crucial. Moreover, DSB sequencing leads to different patterns than, for example ChIP-Seq, so using software created for ChIP-Seq analysis, such as MACS, is inappropriate. Moreover, some standard practices such as removing duplicate reads, often leads to incorrect results in DSB sequencing. Therefore, we developed a dedicated pipeline for DSB sequencing data that starts from pre-processing the raw sequencing reads and ends with high-level data interpretation. Our approach is also unique in that it provides multiscale analysis, starting from base pairs resolution and gradually progressing up to mega base resolution, and automatically generates reports about trends across scales.
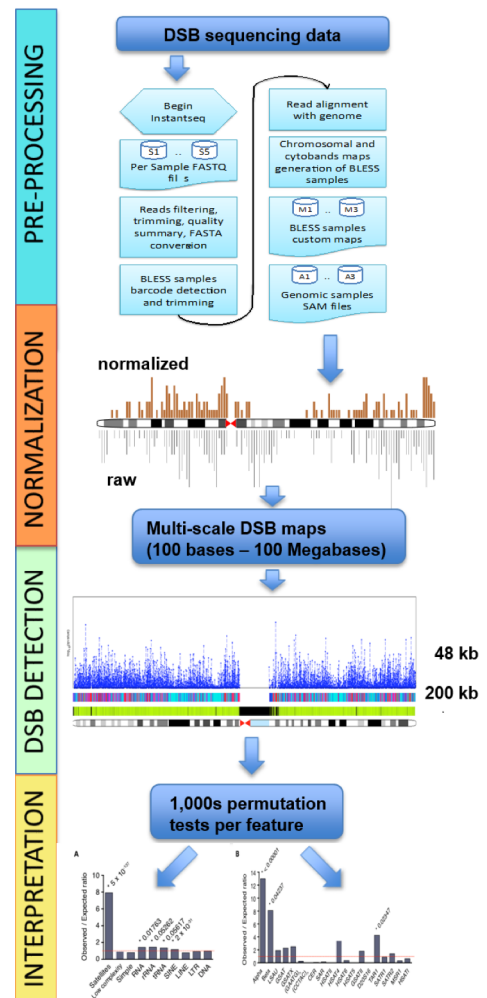
The software home page is http://breakome.utmb.edu/software.html, to download use **reviewer** for both user ID and password.