**Figure 2** Galaxy QC outputs for a representative sample. Read quality across the sequence was acceptable on average but G/C content was abnormally high at the first 9bp of the 5’ end (lower left box). Trimming of sequences (two boxes on right) removed this bias and shortened sequence was used for analysis.