

# Supplementary File 1

## Effect of file split size and number of cores per file

Read Length = 75 bp

### DATA RANGE

File size (number of reads):  $10^6$ ,  $2 \times 10^6$ ,  $3 \times 10^6$ ,  $4 \times 10^6$ ,  $5 \times 10^6$ ,  $10^7$ ,  $1.5 \times 10^7$ ,  $2 \times 10^7$ ,  $3 \times 10^7$ ,  $5 \times 10^7$ ,  $10^8$

Segment size (number of reads):  $10^5$ ,  $2 \times 10^5$ ,  $5 \times 10^5$ ,  $10^6$ ,  $2 \times 10^6$ ,  $5 \times 10^6$ ,  $10^7$ ,  $3 \times 10^7$ ,  $5 \times 10^7$ ,  $10^8$

Number of cores: 10, 16, 20, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140











