

Table 3

Advantages and Disadvantages of DNA Sequence Storage Strategies

Data Storage Strategy	Advantages	Disadvantages
Saving as a plain text file	Easy to retrieve and analyze the sequence	Large file size (3 GB/genome)
Saving as a compressed file	Smaller file size (roughly 1.5 to 2 GB/genome)	Time required (can be hours) to perform compression and decompression before data analysis
Saving only differences between the genome and the reference genome and compressing the file	Very small file size (4 MB/genome)	Time required to rebuild the needed DNA sequence; need to update all the files when the reference sequence is changed