Box 1 'Rate limiters' of next-generation DNA sequencing experiments

Until recently, the act of acquiring high-throughput molecular data (for example, DNA sequencing or DNA microarrays) was the primary cost associated with many experiments in genomics (for example, genome assembly, expression analysis and so on). For some projects, the emergence of next-generation sequencing has simply increased ambitions, such that DNA sequencing costs remain dominant. However, for most experiments, other 'rate limiters' are an increasing fraction of the overall cost and effort. These include the following: first, the cost of generating, acquiring and/or storing samples; second, the costs of constructing and indexing fragment libraries; third, the costs of building and maintaining infrastructure for large-scale data analysis, storage, exchange and deposition to public repositories; fourth, the time and labor costs of executing on both routine (for example, read mapping) and specialized (for example, data interpretation) tasks in large-scale data analysis; fifth, the costs of training personnel for the experimental and analytical skill sets associated with next-generation DNA sequencing; sixth, the costs associated with transient or persistent mismatches between the local capacity and local demand for next-generation DNA sequencing; and finally, for clinical samples, the costs associated with phenotyping subjects, obtaining consent from subjects and complying with regulations for working with human subjects.