

ART-illumina (simulator)

Introduction: ART is a set of simulation tools to generate synthetic next-generation sequencing reads. ART simulates sequencing reads by mimicking real sequencing process with empirical error models or quality profiles summarized from large recalibrated sequencing data. ART can also simulate reads using user own read error model or quality profiles. ART supports simulation of single-end, paired-end/mate-pair reads of three major commercial next-generation sequencing platforms: Illumina's Solexa, Roche's 454 and Applied Biosystems' SOLiD.

Input: Genome reference sequence in fasta format.

Output: Simulated pair-end reads in fastq format.

Test inputs:

input_ref (fasta/fna/fas/fa format)

imcas:/test_kira/Lactobacillus.fasta

Test arguments:

args_f (the fold of read coverage to be simulated or number of reads/read pairs generated for each amplicon, integer, [20]) 20

args_l (the length of reads to be simulated, integer, [150]) 150

args_m (the mean size of DNA/RNA fragments for paired-end simulations, integer, [200]) 200

args_s: (the standard deviation of DNA/RNA fragment size for paired-end simulations, integer, [10]) 10

Test example:

*.fastq 41.27MB

*.fastq 41.27MB