Nature Reviews Genetics | AOP, published online 8 December 2009; doi:10.1038/nrg2726

BIOINFORMATICS

SNPs while you wait

Computational power is struggling to keep pace with the amount of genome variation data that are being collected. One solution is offered by an open access software tool, Crossbow, that allows billions of bases of sequence to be compared and SNP-genotyped accurately and cheaply in a few hours.

Currently the only way to accelerate genome analysis time is to run many computers in parallel. This approach involves costly resources that are available only in dedicated sequencing centres; parallelization sometimes demands custom-made software, which can prevent peers from repeating the analyses. Crossbow overcomes these issues by using 'cloud computing', an increasingly popular approach in which computing space and central processing unit cycles are 'rented' from idle server capacity on the web, allowing large-scale analyses to be carried out on a virtual computer cluster from any online terminal.

Crossbow combines the speed of an existing alignment program (Bowtie) and the accuracy of the SNP-caller SOAPsnp. The authors tested its performance in aligning and calling SNPs from 2.7 billion reads that were sequenced from a Han Chinese male (corresponding to 38-fold genome coverage). The calling accuracy for known and novel SNPs compared favourably (98% concordance) to the published set of SNPs obtained using an Illumina genotyping assay. Moreover, the entire analysis, which cost only US$85, was completed in 3 hours using a 320-core cluster rented from Amazon.com.

The time and cost associated with Crossbow, and the fact that it requires standard software and hardware, makes it widely available to perform new analyses, replicate existing ones or extend it to analyse RNA–seq and ChIP–seq data.

Tanita Casci

ORIGINAL RESEARCH PAPER Langmead, B. *et al.* Searching for SNPs with cloud computing. *Genome Biol.* **10**, R134 (2009)