Nature Reviews Genetics | Published online 29 Feb 2016

IN BRIEF

TECHNIQUE

Chicago HighRise for genome scaffolding

Putnam *et al.* report an *in vitro* method for generating long-range linkage data that improves the scaffolding of *de novo* assembled genomes. Their approach, called Chicago, requires only small amounts of high-molecular-weight DNA as starting material and uses reconstituted chromatin as a substrate to produce proximity ligation libraries. Using the HighRise (HiRiSE) software, genomic scaffolds can be generated. The authors show the utility of their method for human genome assembly and scaffolding. Moreover, Chicago libraries can be used to improve existing assemblies, which the authors illustrate by reassembling and scaffolding the genome of the American alligator, and for haplotype phasing.

ORIGINAL ARTICLE Putnam, N. H. *et al.* Chromosome-scale shotgun assembly using an *in vitro* method for long-range linkage. *Genome Res.* http://dx.doi.org//10.1101/gr.193474.115 (2016)

GENETIC SCREENS

CombiGEM–CRISPR: a creative combination

Researchers have generated a platform that uses combinatorial genetics en masse (CombiGEM) and the CRISPR–Cas9 system for the rapid assembly of barcoded combinatorial genetic libraries that can be tracked with high-throughput sequencing. The approach can be used to undertake pooled combinatorial genetic perturbations in human cells and to identify the synergistic actions of genes that control complex traits. These findings can then be leveraged to determine novel combinatorial drug therapies.

Wong *et al.* applied CombiGEM–CRISPR to create a library of 23,409 barcoded dual guide-RNA combinations. Gene pairs that inhibit ovarian cancer cell growth when targeted were identified by a high-throughput pooled screen and validated by CRISPR–Cas9-mediated knockout or RNA-interference-mediated knockdown assays. Small-molecule drug pairs directed against the identified gene pairs were shown to synergistically inhibit proliferation of ovarian cancer cells.

ORIGINAL ARTICLE Wong, A. S. L. *et al.* Multiplexed barcoded CRISPR-Cas9 screening enabled by CombiGEM. *Proc. Natl Acad. Sci. USA* http://dx.doi.org/10.1073/pnas.1517883113 (2016)

BIOINFORMATICS

Sequence searches blossom

A novel data mining algorithm, based on a newly developed indexing data structure called Sequence Bloom Trees (SBTs), efficiently searches large-scale short-read sequencing repositories for experiments containing a sequence of interest, at 162 times the speed of existing search methods. The authors validate their algorithm by building an SBT from 2,652 human RNA-seq short-read sequencing runs deposited in the NIH Sequence Read Archive (SRA). These runs represent the entire set of publicly available, human RNA-seq runs from blood, brain and breast tissues stored within the SRA at the time. Solomon and Kingsford searched the SBT for the expression of all 214,293 known human transcripts to identify tissue-specific transcripts. Results were obtained in 3.3 days, as opposed to an estimated 92 days with an alternative search method. The SBT index could be used to identify unknown long non-coding RNAs, for example, as it does not require prior knowledge of sequences of interest.

ORIGINAL ARTICLE Solomon, B. & Kingsford, C. Fast search of thousands of short-read sequencing experiments. *Nat. Biotechnol.* http://dx.doi.org/10.1038/nbt.3442 (2016)