**Summary**

The abundance of repeat elements in genomes can impede the assembly of a single sequence. The tool scaffold_builder was designed to generate scaffolds (super contigs of sequences joined by N-bases) using the homology provided by a closely related reference sequence.

**Methods**

Scaffold_builder is an advanced wrapper for Nucmer, written in Python. The Figure below illustrates how scaffold_builder resolves several situations that may arise when mapping contigs to the reference genome.

**Results**

The application was evaluated using simulated pyrosequencing reads of the three bacterial genomes, and two newly sequenced genomes. As shown in the Table below, scaffold_builder decreases the number of contigs by ~62% while increasing their average length by ~200%.

**Conclusions**

Scaffold_builder helps to create longer sequences during genome assembly. It allows the user to combine the strengths of de novo assembly with the structure provided by a closely related reference.

**Web-based version and Code:**

http://edwards.sdsu.edu/scaffold_builder