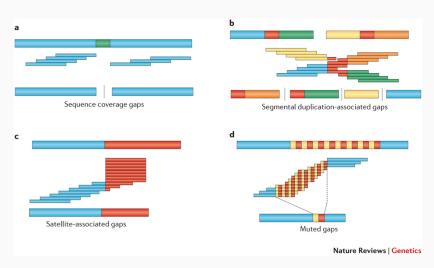


PROBLEMS





Source[Chaisson et al., 2015]

2

IMPROVING ILLUMINA SHORT READ ASSEMBLY



- 1. Within a taxanomic group, GC content is largely conserved.
- 2. Within a taxonomic group, genome size is largely conserved.
- 3. Bacterial genomes are dense.
- 4. Nucleotide order is not random.

POSSIBLE SOLUTION

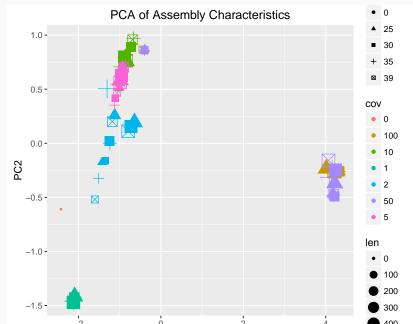




Figure: Bridge Reconstruction. Pink fragments are reads. Grey shows the gene of interest with interupted coverage. Orange fragemnt is a pseudoread generated from this situation under the hypothesis that the beige fragment exists but is underrepresented: that is, the bridge is there, but needing repair

4







Chaisson, M. J. P., Wilson, R. K., and Eichler, E. E. (2015). Genetic variation and the de novo assembly of human genomes. Nature Publishing Group, 16.



Chaisson, M. J. P., Wilson, R. K., and Eichler, E. E. (2015). Genetic variation and the de novo assembly of human genomes. Nature Publishing Group, 16.



Chaisson, M. J. P., Wilson, R. K., and Eichler, E. E. (2015). Genetic variation and the de novo assembly of human genomes. Nature Publishing Group, 16.

ACKNOWLEDGEMENTS





OÉ Gaillimh NUI Galway

NUIG Microbiology

- · Dr. Fiona Brennan
- · Dr. Florence Abram
- · Matthias Waibel

NUIG Maths

· Pilib



James Hutton Institute

· Leighton Pritchard

