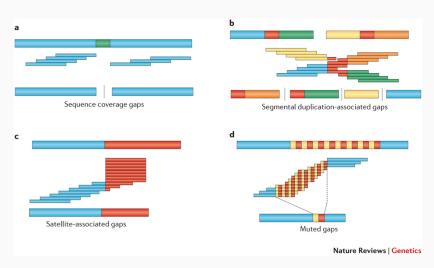


PROBLEMS



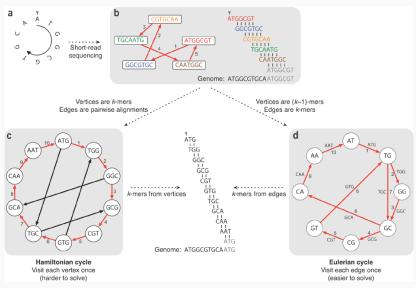


Source[Chaisson et al., 2015]

2

DE BRUIJN GRAPHS AND EULERIAN PATHS



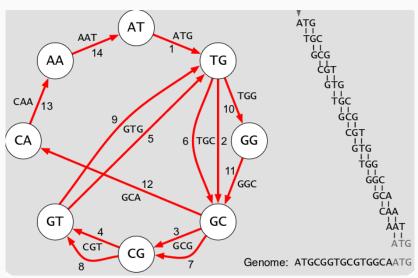


Source[Compeau et al., 2011]

3

EULERIAN PATHS WITH REPEATS





Source[Compeau et al., 2011]

ŀ

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

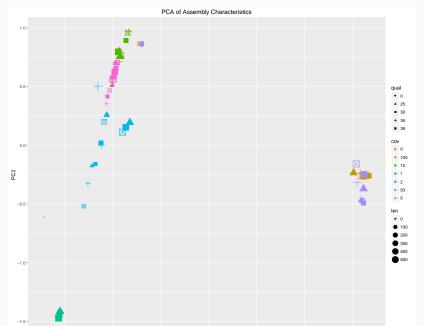


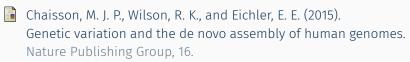


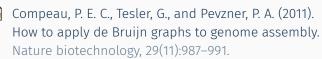
Figure: Bridge Reconstruction. Pink fragments are reads. Grey shows the gene of interest with interupted coverage. Orange fragemnt is a pseudoread generated from this situation under the hypothesis that the beige fragment exists but is underrepresented: that is, the bridge is there, but needing repair

INITIAL DATA









ACKNOWLEDGEMENTS





OÉ Gaillimh NUI Galway

NUIG Microbiology

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

NUIG Maths

· Pilib



James Hutton Institute

· Leighton Pritchard

