

Additional File 5

A)

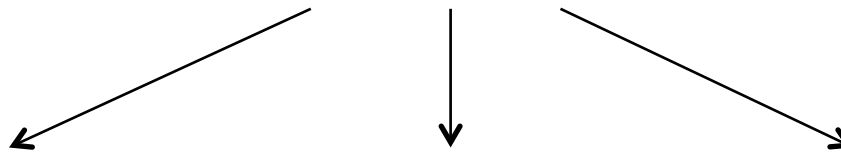
for low GC (35% GC) / GC50 (50% GC) / high GC (65% GC)

3x "references"	3x "references"
low divergence	high divergence
3x diploid "samples"	3x diploid "samples"

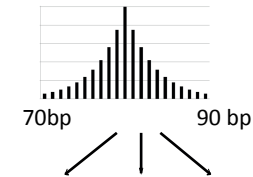
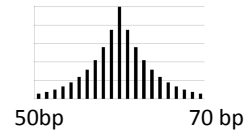
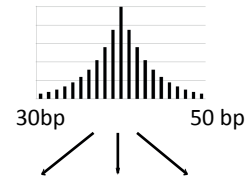


B)

600x coverage depth read pool created from "sample" sequences
trim reads to size distributions



C)



D)

low fragmentation bias and low
damage frequency at read ends
for low damage read pools

no fragmentation bias or
damage at read ends for
no damage read pools

high fragmentation bias and high
damage frequency at read ends
for high damage read pools



randomly draw reads to coverage

E)

0.1x 0.5x 1x 2x 4x 8x 16x

align reads from each final pool against original "reference".
Compare called variants to actual variants