

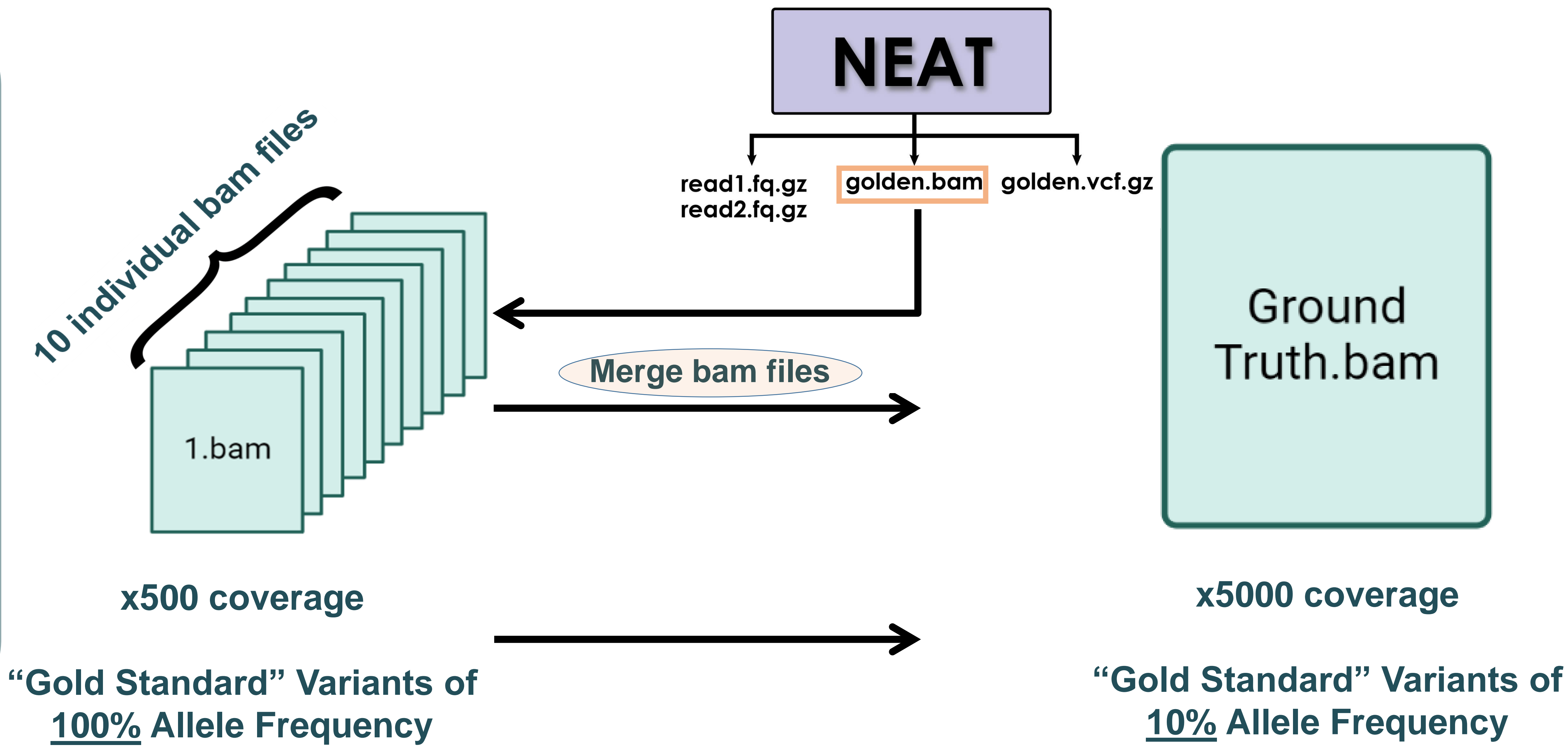
Synthetic Genomics Data Generation and Evaluation for the Use Case of Benchmarking Somatic Variant Calling Algorithms

Styliani-Christina Fragkouli^{1, 2}, Nikos Pechlivanis¹, Andreas Agathangelidis², Fotis Psomopoulos¹

¹Institute of Applied Biosciences, Centre of Research and Technology Hellas, Thessaloniki, Greece
²Department of Biology, National and Kapodistrian University of Athens, Athens 10679, GR

1 Synthetic «Ground Truth» Dataset Generation

- Highlights
- Generation of **synthetic genomics** data based on *TP53* gene
 - Define «Ground Truth» SNPs and INDELs in order to **benchmark** somatic variant callers
 - Investigate the impact of variant callers in variants at **low frequencies**



2 Benchmarking GATK-Mutect2

