

Viral Beacon Data freeze 20200701 - Variant files statistics

1 Summary stats

1. 65239 variant files

by source data type :

raw sequences = 6917

consensus sequences= 58322

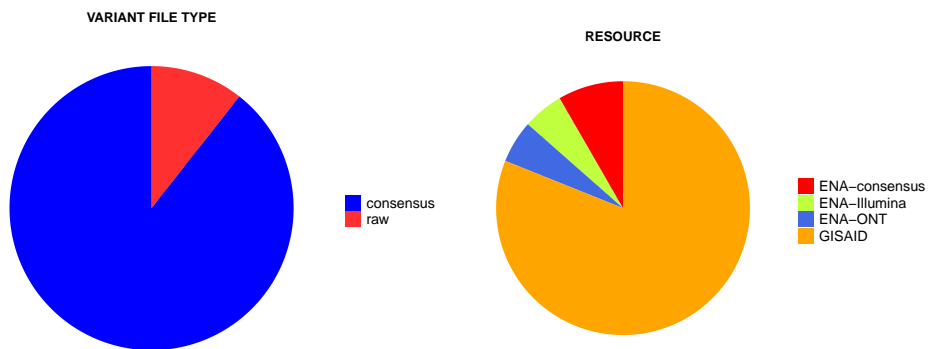


Figure 1: Distribution of variant files by source data type and resource

by resource & pipeline:

ENA-Illumina (Galaxy LoFreq): 3378 (intrahost variants)

ENA-ONT (Medaka): 3539 (consensus variants)

ENA-consensus: 5430 (consensus variants)

GISAID: 52892 (consensus variants)

2 Stats variant files metadata by resources (ENA-consensus, ENA-Illumina, ENA-ONT, GISAID)

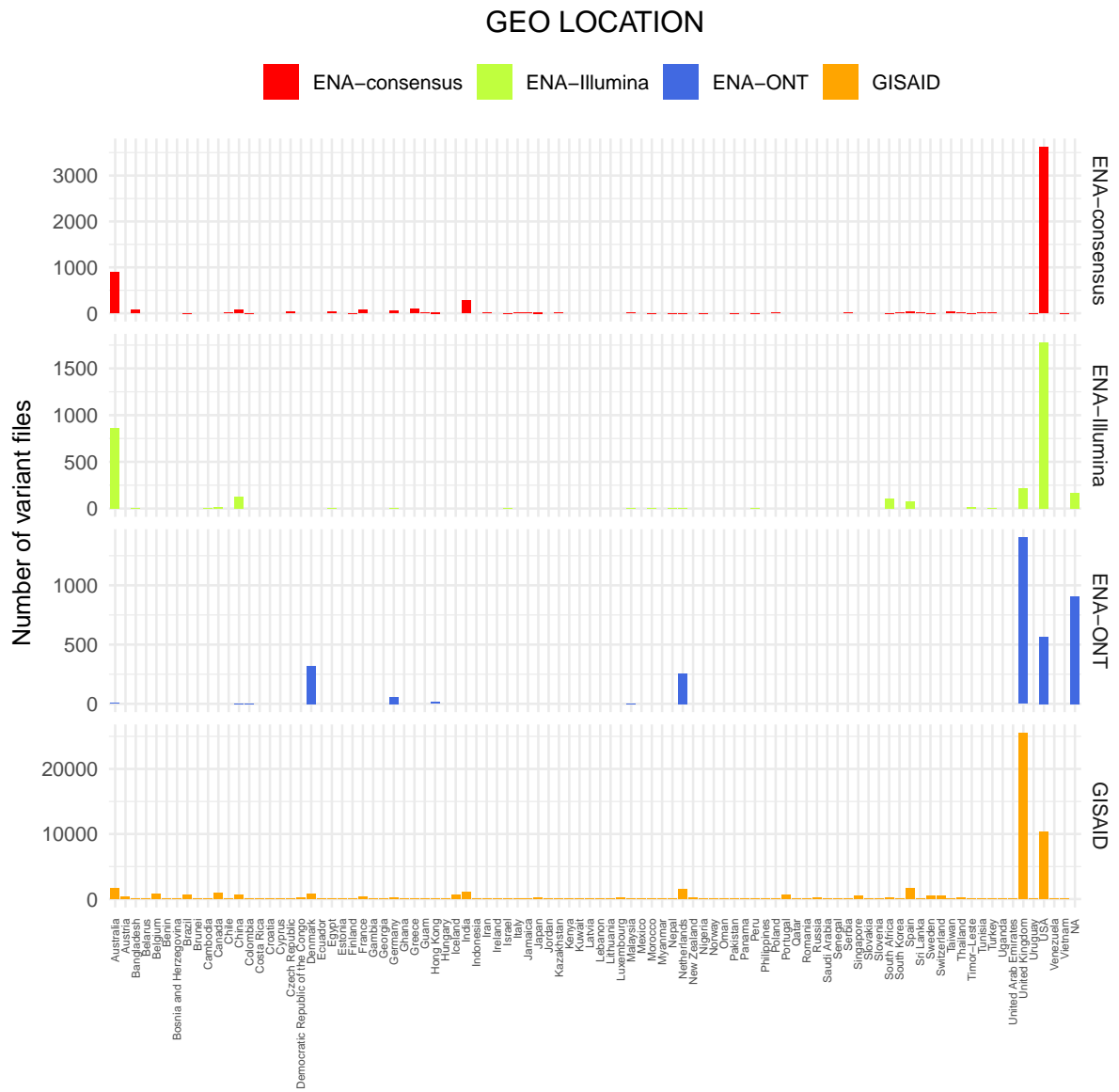


Figure 2: Distribution of variant files by country by resource

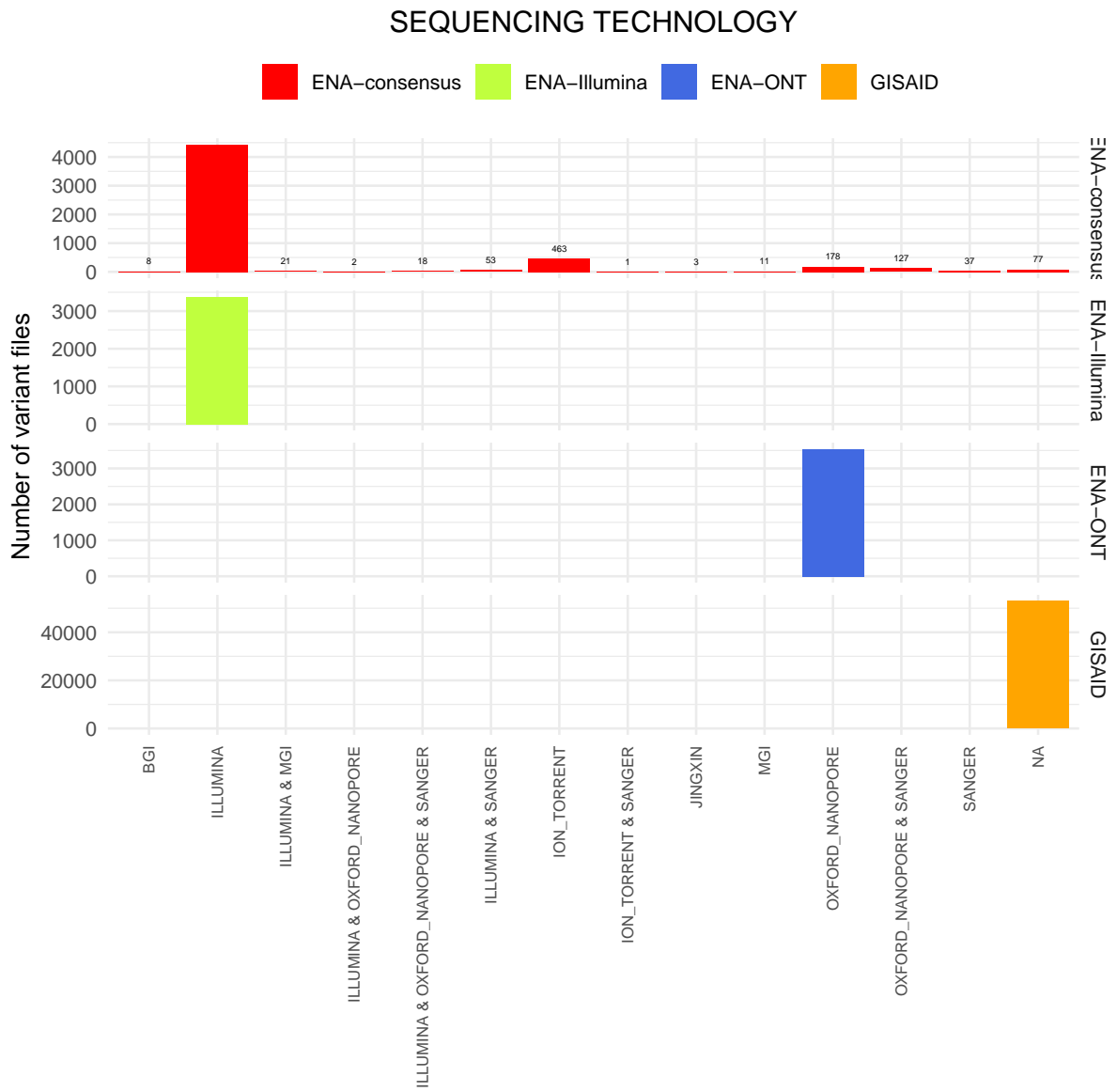


Figure 3: Distribution of variant files by sequencing technology by resource

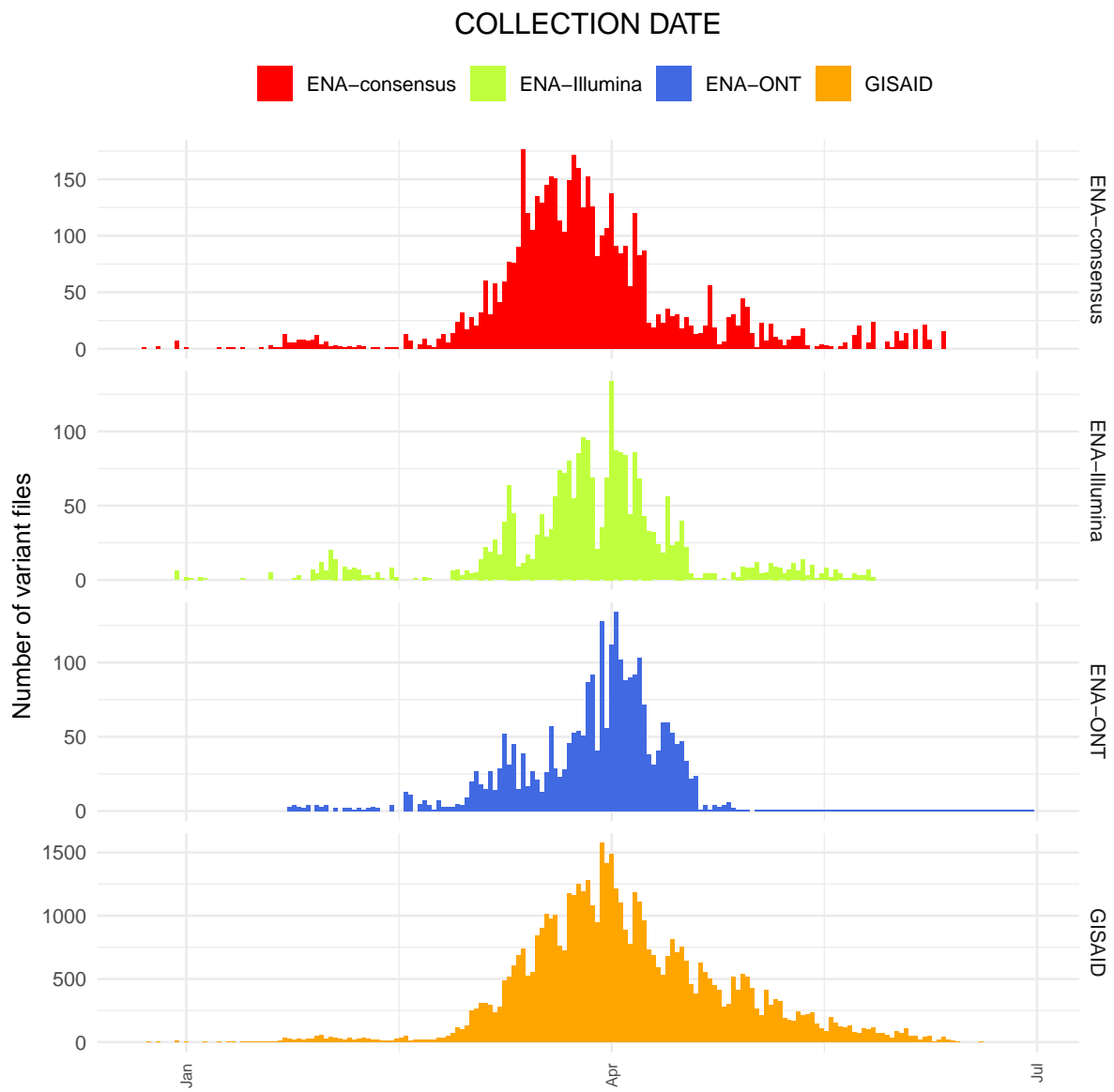


Figure 4: Distribution of variant files by collection date by resource

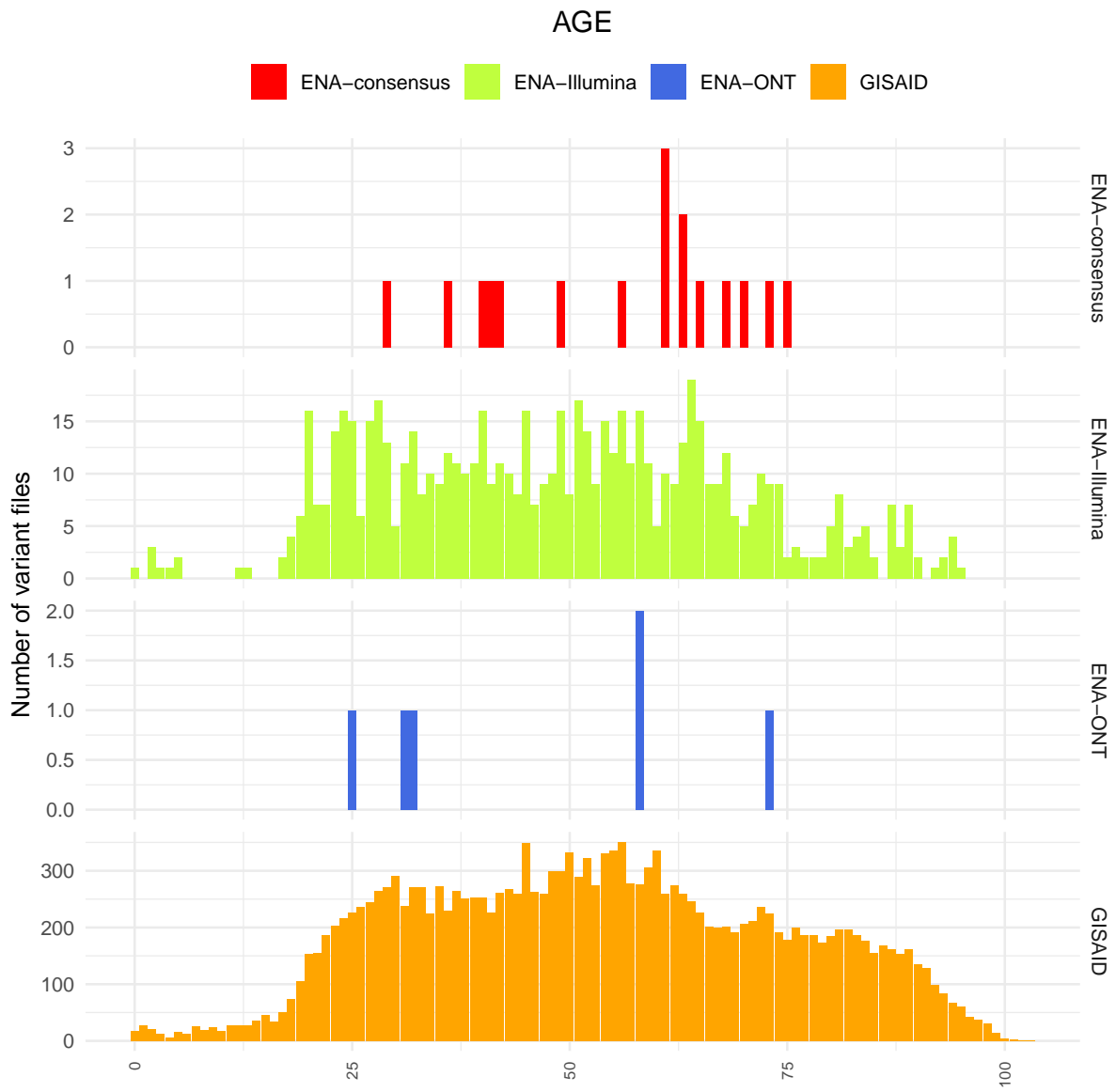


Figure 5: Distribution of variant files by host age by resource

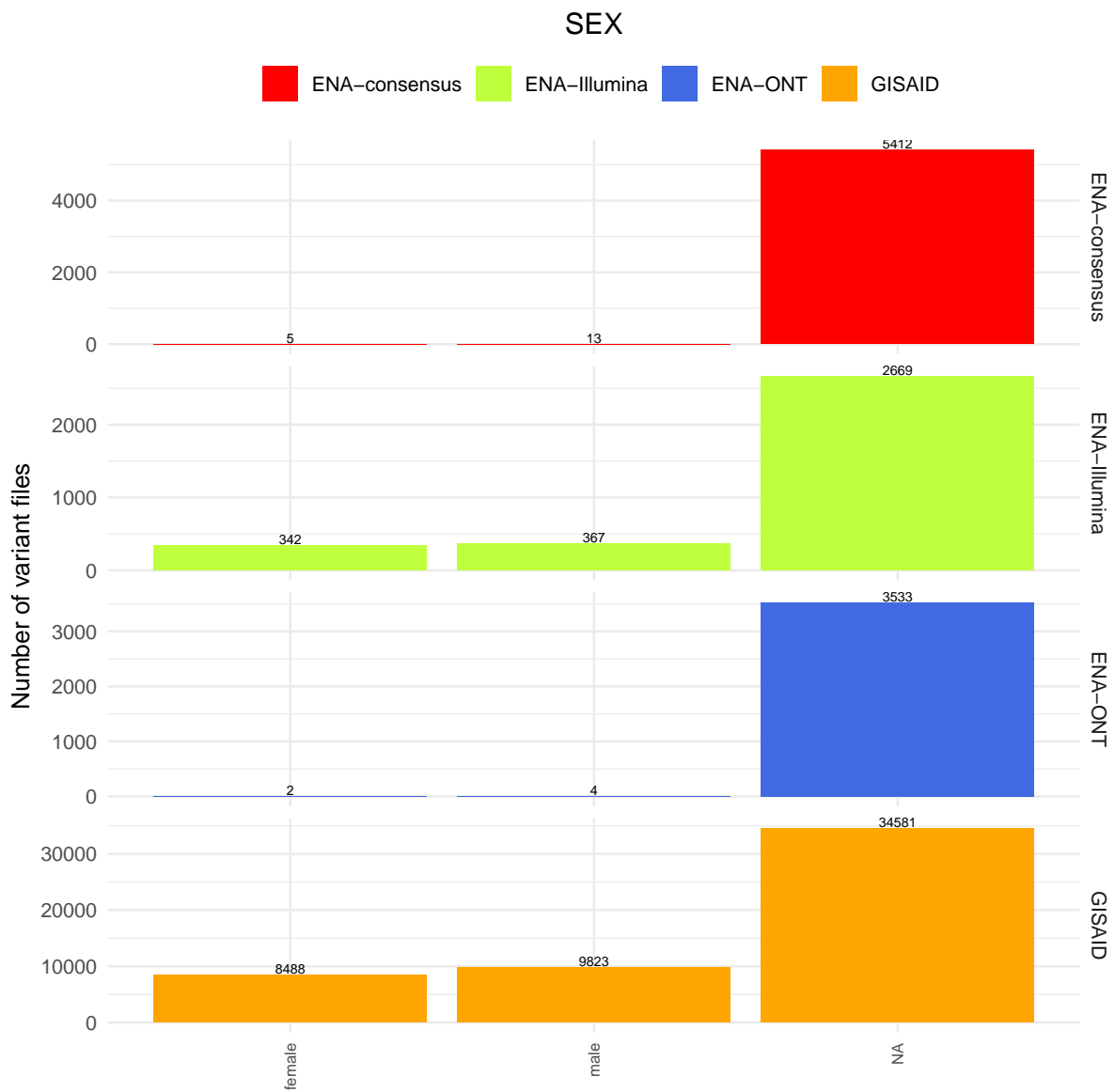


Figure 6: Distribution of variant files by host sex by resource

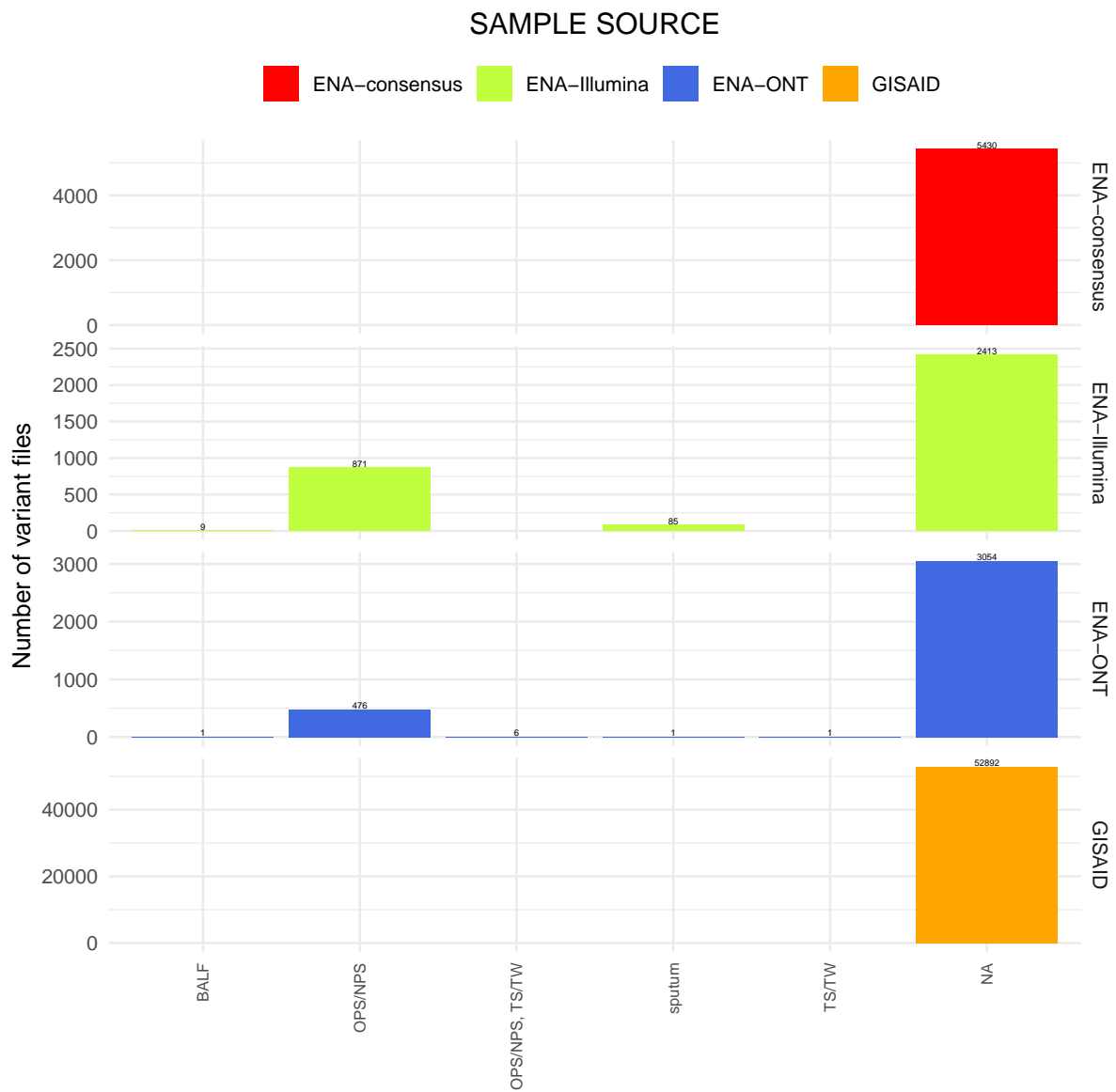


Figure 7: Distribution of variant files by sample source by resource