Ambry Genetics®
A TEMPUS COMPANY

P552

Use of Pangenome reference improves variant calling in clinical genome sequencing

Erica D Smith, Ermanno Florio, Julian Stone-Farhat, Matthew Schultz, Ali Alhafidh, David Brohawn, Marcus Mahar

BACKGROUND

- Choice of reference genome affects what variants we can call.
- **Linear reference genomes** hg19 and hg38 have enabled decades of genetic research, but they have gaps and bias.
- **Graph-aware reference** mapping allows alignment of more continuous stretches of DNA sequence[1,2].
- We directly compared the effect of using linear vs graph reference sequence in a **short read whole genome** clinical assay.

METHODS

- Three cohorts were analyzed: Genome In A Bottle (GIAB), the expanded 1000 Genomes Project (1kGP), and clinical specimens
- DNA was extracted from blood or cell lines using MagMAX DNA Multi-sample Ultra 2.0" Kit (ThermoFisher)
- DNA was sequenced on Illumina NovaSeqXPlus platform
- Reads were aligned with dragen v4.3 or 4.4[3] as noted
- Performance was compared between linear (called hg38 here) vs graph (called pangenome or pan here)
- Variant calling was evaluated in stratified bins defined by the GA4GH v3.1[4]

RESULTS I: SNV Sensitivity x depth in GIAB specimens

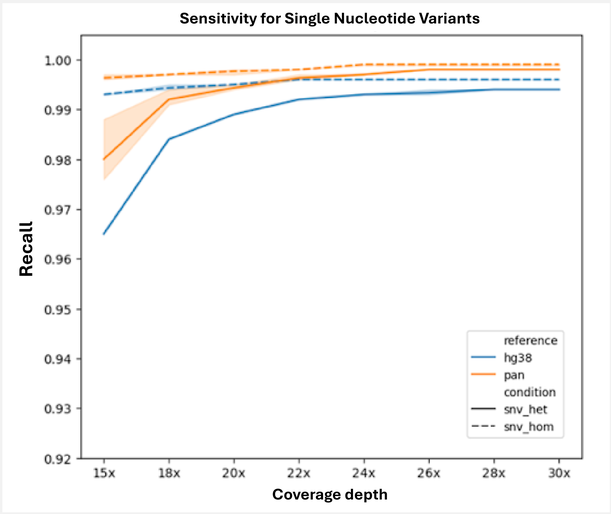

Figure 1. Sensitivity for heterozygous and homozygous SNVs were calculated using downsampling analysis on GIAB HG001-HG007. Specimens aligned using dragen v4.3.

RESULTS II: Performance calling variants in 43 specimens from 1kGP

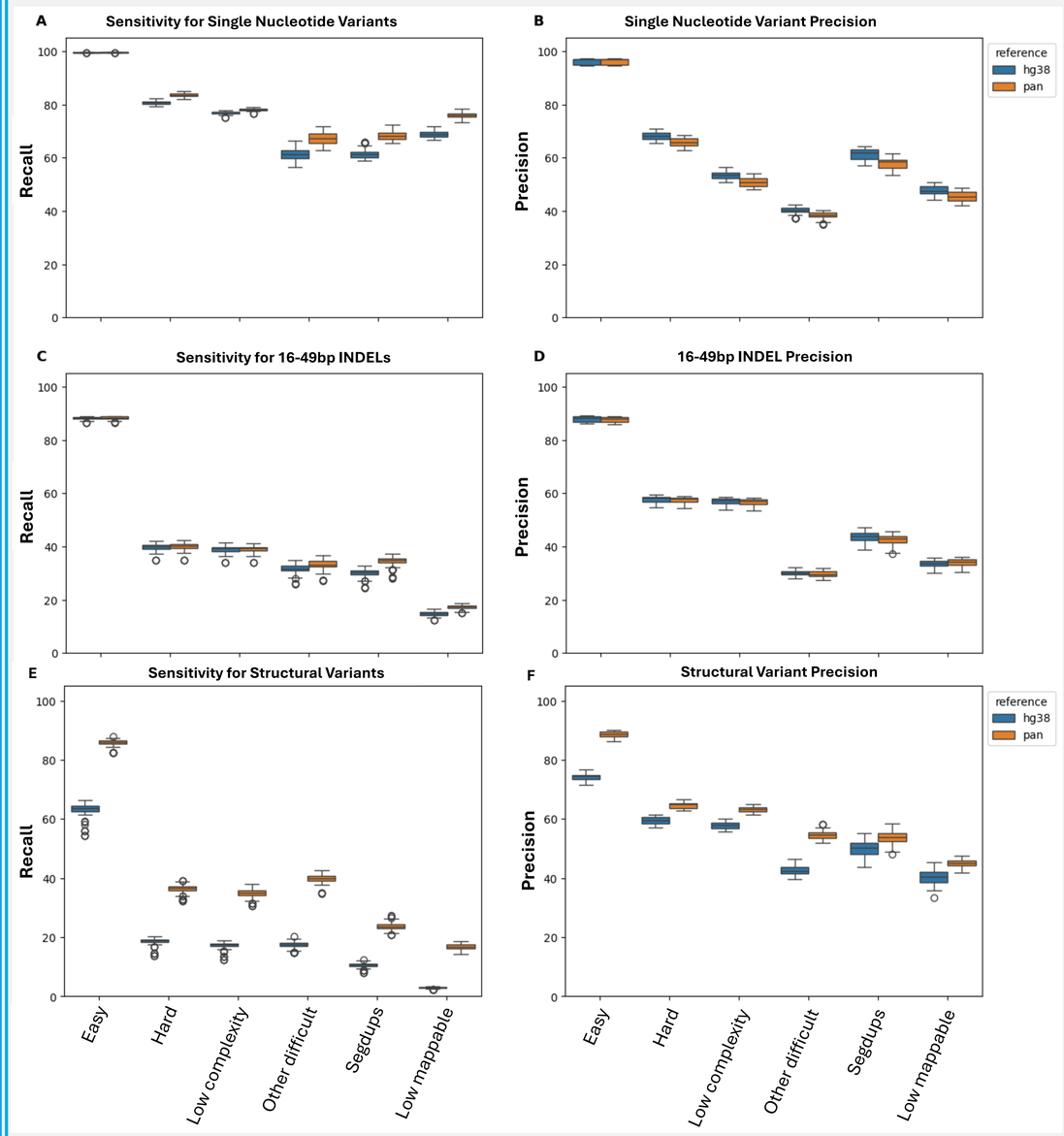

Figure 2. Sensitivity and precision compared to a published benchmark[1]. SNVs **A** and **B,** and small indels in **C** and **D,** SVs **E** and **F**. Variants were binned by how challenging they are to call as defined by the GA4GH Consortium[4]. Specimens aligned using dragen 4.4.

ACRONYMS

SNV – Single Nucleotide Variant
Indel – insertions and deletions <50bp
SV – Structural Variant, >50bp
GA4GH - Global Alliance for Genomics and Health

RESULTS III: selected SV sensitivity in clinical specimens

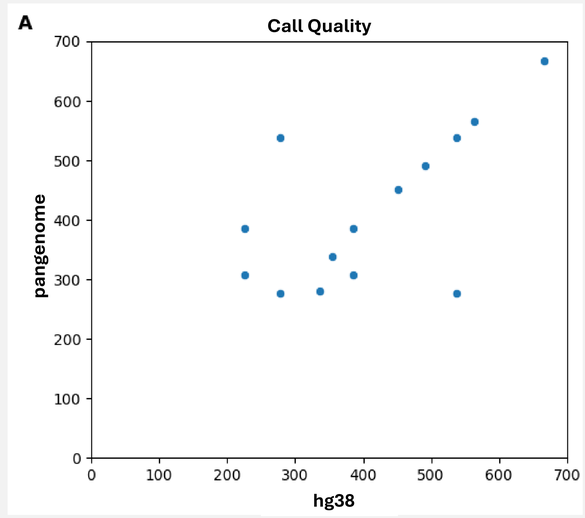

Figure 3. Strength of 14 diagnostic SVs (mostly ALU insertions) previously detected on multigene panel sequencing. Call Quality reflects confidence in the identity of the individual base pair called. Specimens aligned with dragen 4.4.

THE CHALLENGE

The benefits of using pangenome reference in research settings has been demonstrated. Here we define the specific benefits and limitations in clinical diagnostic workflows.

TAKE HOME POINTS

- Pangenome improved cohort-wide SNV sensitivity at low coverage depth (Fig 1) but not full depth (Fig 2 A,B)
- Pangenome increased cohort-wide sensitivity and precision for SVs (Fig 2 E,F) overall.
- For 14 SVs previously reported on clinical assays, both hg38 and pangenome were sufficient (Fig 3).
- Performance gains of pangenome may be greatest for complex variants and at low coverage depth.

REFERENCES

1. Liao, *et al* 2023 Nature. PMID: 37165242
2. Nyaga, *et al* 2025 Front Genet. PMID: 41050061
3. Behera, *et al* 2024 Nat Biotechnol PMID: 39455800
4. Krusche, *et al* 2019, Nature Biotechnology PMID: 30858580