4.2.0-1 Release Notes

Table of contents

Several packages have been updated to GATK 4.3. Support for read lengths of up to 500 base pairs has been added to haplotypecaller and mutectcaller.

Several performance improvements.

A number of bugs have been fixed.

New Tools

Added markdup, a tool to locate and tag duplicate reads in a BAM or SAM file, where duplicate reads are defined as originating from a single fragment of DNA.

Improvements

deepvariant:

- Major performance improvements.

collectmultiplemetrics:

- Supports read lengths of up to 500 base pairs.

haplotypecaller:

- Updated to GATK 4.3.

- Supports read lengths of up to 500 base pairs.

- Adds support for the `--htvc-bam-output` option.

- `--batch` is deprecated.

mutectcaller:

- Updated to GATK 4.3.

- Supports read lengths of up to 500 base pairs.

- Adds support for the `--run-partition` option.

bammetrics

- Updated to GATK 4.3.

bqsr

- Clarified error messages.

Improvements spanning multiple tools

- Added more error checking when writing BAM files using `--gpuwrite` .

- Improved performance for BWA alignment, particularly for reads > 250 bases in length.

- Added GPUDirect Storage (GDS) support for fq2bam (BWA-MEM + GATK) and associated pipelines.

- Added a low memory mode for `--gpuwrite` (affects sorting).

Bug Fixes

- deepvariant Could only handle 32 alt alleles per candidate; can now handle up to 64.

- bamsort Bamsort could fail for certain coordinate sorts.

For further information see the Parabricks datasheet.