

# Pipeline v5.34 Release Notes

## Overview

5.34 has these key improvements:

- Teal (telomere length) added to pipeline
- Cider (CDR3 identification and enumeration) added to pipeline
- Cobalt now counts average coverage instead of number of reads starting in each 1kb window.
- Cuppa implemented with logistic regression model
- Various improvements in Purple, Pave, Sage and Amber, primarily related to target regions
- Adopted BWA-MEM2 to improve alignment speed
- Sambamba Mark Duplicates replaced by MarkDups component which includes logic to un-map poor alignments in problematic regions, and UMI aware duplicate collapsing and consensus read logic
- Visualisations added to Sage

## Resource File Updates

Tool / Stage	Description
Amber	new sites, now using the same sites AmberGermlineSites.37.tsv.gz, AmberGermlineSites.38.tsv.gz
MarkDups	Unmap regions file
Gripss	v38 PON updated from v37 lift-over
Pave	HMF panel target regions PON updated
Sage	Hotspots removed KMT2C and a range from AR chrX
Driver gene panel	Germline reportability changes for CDC73, TERT, TMEM127
HMF targeted panel	PON and Cobalt normalisation updated

## Tool Updates

Tool	Version	Release Notes and Key Changes
Amber	4.0	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/amber-v4.0">https://github.com/hartwigmedical/hmftools/releases/tag/amber-v4.0</a> Minor changes to min depth in tumor and targeted panel modes Restrict sites to target regions BED in panel mode
BamTools	1.2.1	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/bam-tools-v1.2.1">https://github.com/hartwigmedical/hmftools/releases/tag/bam-tools-v1.2.1</a> Write target region statistics file including dual-strand counts
BWA-MEM2	2.2.1	Now used for alignment instead of BWA
Cider	1.0	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/cider-v1.0.3">https://github.com/hartwigmedical/hmftools/releases/tag/cider-v1.0.3</a> Initial version
Cobalt	1.16	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/cobalt-v1.16">https://github.com/hartwigmedical/hmftools/releases/tag/cobalt-v1.16</a> Use mean read depth instead of read counts
Cuppa	2.0	Uses logistic regression model
Gripss	2.4	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/gripss-v2.4">https://github.com/hartwigmedical/hmftools/releases/tag/gripss-v2.4</a> Auto-set target panel settings Targeted mode: added QualPerAD filter, SGLs no longer called
Lilac	1.6	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/lilac-v1.6">https://github.com/hartwigmedical/hmftools/releases/tag/lilac-v1.6</a>
Linx	1.25	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/linx-v1.25">https://github.com/hartwigmedical/hmftools/releases/tag/linx-v1.25</a>
MarkDups	1.1	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/mark-dups-v1.1.2">https://github.com/hartwigmedical/hmftools/releases/tag/mark-dups-v1.1.2</a>

		Use unmap regions file to unmap problematic reads Performance improvements
Pave	1.6	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/pave-v1.6">https://github.com/hartwigmedical/hmftools/releases/tag/pave-v1.6</a> Minor HGVS coding and protein fixes Improved phased-variant logic Multi-threaded
Purple	4.0	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/purple-v4.0">https://github.com/hartwigmedical/hmftools/releases/tag/purple-v4.0</a> Mask IG/TCR, centromeric regions & don't fit short arms Chromosome X amplifications use 1.5x ploidy for males Don't smooth large germline deletions in diploid normalisation Fit in 0.5% intervals for purity <20% and lower min to 7% Run HRD detection routine <b>in targeted mode</b> <b>Ploidy penalty model customised for targeted mode</b> Only show diploid regions for input PNG
Sage	3.4	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/sage-v3.4">https://github.com/hartwigmedical/hmftools/releases/tag/sage-v3.4</a> Sync overlapping fragments on by default Track and use average map-quality, max read-edge distance for filtering New indel-deduplication logic HTML visualisations (see guide below) Improved read and fragment strand bias filters Targeted mode specific settings and logic
SvPrep	1.2.3	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/sv-prep-v1.2.3">https://github.com/hartwigmedical/hmftools/releases/tag/sv-prep-v1.2.3</a>
Teal	1.2.1	<a href="https://github.com/hartwigmedical/hmftools/releases/tag/teal-v1.2.1">https://github.com/hartwigmedical/hmftools/releases/tag/teal-v1.2.1</a> Initial version

## Configuration Changes

Tool	Release Notes and Key Changes
Amber	Targeted panel: remove -tumor_only_min_depth 80 add -target_regions_bed target_regions_definition BED file
Cobalt	Targeted panel: -pcf_gamma 15 -> 50
Gripss	Targeted panel: remove -hard_min_tumor_qual 200 remove -min_qual_break_point 1000 remove -min_qual_break_end 1000
Lilac	In tumor-only mode (including for targeted region), now set the tumor BAM file with -tumor_bam instead of -reference_bam
Pave	v38: remove -gnomad_load_chr_on_demand Set threads, recommend 8 and increase max heap size to 30GB Targeted panel: force_pathogenic_pass
Purple	Targeted panel: remove -min_diploid_tumor_ratio_count 3 remove -min_diploid_tumor_ratio_count_centromere 3 remove -ploidy_penalty_factor 0.6 remove -ploidy_penalty_standard_deviation 0.15 (ie only show now specify the 3 target regions files)
Sage	Change -output -> -output_vcf Targeted panel:

	add -high_depth_mode change -hard_min_tumor_vaf 0.002 change -hotspot_min_tumor_vaf 0.015 change -panel_min_tumor_qual 150
--	---

## Functional Changes

Topic	Release Notes and Key Changes
Variant calling	Sage is more sensitive to calling low VAF variants in a targeted panel. The rate of false-positives is also reduced for both panel and WGS runs
Cobalt	Targeted panel - use of read depth in 1000-base regions gives smoother copy number profiling
MarkDups	Formation of a consensus read gives more precise base and base qualities
Purple	Expected improvements to fits in which scenarios? LST/TAI/LOH based HRD prediction added for targeted mode. Currently tuned for HMF panel only

## Technical Changes

Tool	Performance Characteristics
BWA-MEM2	Roughly 40% improvement in alignment time
MarkDups	MarkDups runs in about 2/3 the time of Sambamba MarkDuplicates
Pave	Multi-threading reduces the run time by approximately 2/3

## Sage Visualisation Guide

### Variant overview

7:152079940 G > T TIER = LOW\_CONFIDENCE QUAL = 288 NO REPEAT MED = 73 FILTER = PASS CONTEXT = GGTTTAACTATAAATACATCATCTCT

SAMPLE	QUAL	AD	AF	DP	FULL	PARTIAL	CORE	REALIGNED	CORE	PARTIAL	ALT	REF	ABQ	AMQ	ANM	SB	RSB	JIT
COL0829v003T	288	33	0.079	419	31	2	0	0	0	396	24	35	41	3.9	0.48	0.34	0	0
COL0829v003R	11	1	0.009	109	1	0	0	0	0	0	0	108	6	37	47	4.0	1.00	0.00

Sample level statistics  
(see VCF headers)

COL0829v003T (tumor)

Type Male MAPQ FinalQ modBQ modMQ rawBQ Orientation



SAGE qual calculations per fragment (see readme)

Final contribution to variant qual score for fragment

Read name: A00624.8.HKX.HDSXX.1.1331.17815.1078  
Alignment: 7:152079940-152079958, 7:152079984-152079958  
Cigar: 75S76M, 73M76S  
Insert size: 75  
Orientation: R1F2  
MapQ: 60, 60  
NM: 1, 1  
Dup count: 0

Consensus fragment reads  
(left mouse click)

Fragments are grouped by support type  
and ordered by (strand, map quality)

