WiGiTs v2.0 Release Notes

Overview

WiGits 2.0 has these key improvements:

- New component ESVEE replaces GRIDSS and GRIPSS for SV calling and filtering
- New component vCHORD calculates HRD for panel samples based on PURPLE Circos images
- SAGE context building, support counting, jitter penalty and filtering models overhauled. Parameterisation is now depth independent. Also, BQR is calculated by consensus type
- Biallelic likelihood calculated in PURPLE. New somatic fitting added for panel / tumor only mode
- Microsatellite quality recalibration added to REDUX
- Hotspot or Clinvar pathogenic variants are conditionally rescued from PON filtering

For detailed tool and other updates see:

WiGits pipeline v2.0 docs

Resource File Updates

Tool / Stage	Description
Pave	Removed 58 variants from hg19 WGS PON and 57 variants from hg38 WGS PON
Ref genome	Masked ALTs added to hg38. Decoys added to hg38 False duplications masked (GRC recommendation)
Ensembl	hg38 updated to Ensembl v111
Known fusions	Exon ranges for hg38 updated to match transcript changes in v111. 18 fusions added. 2 fusions removed.
Driver gene panel	62 genes added. 16 genes reporting adjusted. 5 genes removed
MSI Jitter sites	Used by Redux to produce MSI Jitter params for Sage
Sage	3 somatic hotspots added. 4 somatic hotspots removed.
Peach	New resource files for Java rewrite of Peach.
Virusbreakend	Removed bacteria/archaea/vector sequences in database file to reduce size

Tool Updates

See tool releases and READ-ME for full details.

Tool	Version	Release Notes and Key Changes
Amber	4.1	Min depth bug fix for T/N panels
BamTools	1.3	BAM metrics produces read flag counts to replace flagstat
Cider	1.0.3	
Cobalt	2.0	
Cuppa	2.3	Full support for multi-sample mode

		Improvements to visualization
Esvee	1.0	Added new tool, replaces Gridss and Gripss
HealthChecker	3.6	
Linx	2.0	Highest JCN fusion is prioritised for reporting if otherwise equal Copy number change inferred for germline SV Plot issues fixed
Orange	3.7.1	
Pave	1.7	PONGnomad filter disabled for hotspot or clinvar pathogenic variants if GND_FREQ < 0.01. Conditionally recover artefact-prone hotspot/pathogenic variants with high VAF
Peach	2.0	Rewritten in java
Purple	4.1	Biallelic likelihood model for SNV/INDEL replaces boolean Somatic fitting now tested in tumor only mode Somatic fitting performs additional consistency checks in tumor + normal mode
Redux	1.1	Microsatellite recalibration model added Consensus qual is based on median instead of maximum input qual Reads with alignment score < 30 are unmapped Collapsing of orientation added, depending on UMI type
Sage	4.0	v4 SAGE is a major release, and includes an overhaul of the read context, read counting, quality and filtering models. The new model is depth independent and does not require any overrides for either shallow or deeper depth samples.
Teal	1.3.2	Consensus reads not counted by TEAL
Virus Interpreter	1.7	Blacklist HIV in output
vChord	1.0	Added new tool

Configuration Changes

Tool	Release Notes and Key Changes
Pave	'-force_pathogenic_pass' is deprecated
Peach	Configuration arguments entirely different from non-Java version.
Sage	'-germline' flag used for germline runs '-jitter_param_dir' specified for sample-specific jitter Config overriding filter behaviour (min tumor qual, min tumor vaf, or germline filter thresholds) is no longer required
VirusInterpreter	Added -virus_blacklisting_db_tsv

Functional Changes

Topic	Release Notes and Key Changes
ESVEE:SV calling	Esvee replaces both GRIDSS and GRIPSS. Esvee allows variants to be called with only 4 fragments support (compared to typically 7 for GRIDSS). ESVEE reduces FP arising from low mapping to low entropy sequencing, but allow improved sensitivity at low mappability sites. Assembly of LINE insertions is improved.
SAGE:Small variant calling	Probablistic model in SAGE allows depth independent parameterisation Read counting sensitivity is improved due to counting of partial core support BQR by consensus type allows DUPLEX fragments to be counted preferentially
PURPLE:Panel fits	Targeted somatic fitting removes common error of fitting ploidy 2 and purity 100% for low aneuploidy sample

PURPLE:Biallelic likelihood	New model considers both variant AF and regional copy number status to estimate biallelic likelihood
Linx fusions	The gene AL121790.1 has been dropped from Ensembl and so has the known fusion pair AL121790.1 - ETV1

Technical Changes

Tool	Performance Characteristics
Sage	Performance improvements, COLO829T somatic runtime = 42 mins (32 CPUs)
Esvee	Performance improvements, COLO829T somatic runtime = 25mins (32 CPUs)
Redux	Performance and memory improvements, COLO829T runtime = 65 mins (32 CPUs), COLO829R = 20 mins