GENIE 18.0-public release notes

Sage Bionetworks

July 17, 2025

Release notes

These release notes will encompass the changes made in the 18 consortium series and general center information.

- 20565 samples have been added since the last public release.
- New panel(s) since the last public release:
  - UHN-HM-V1
  - UHN-HM-V2
  - UMIAMI-CA-HYBRID-WES
  - DFCI-F1-AB1
  - DFCI-F1-CF2
  - DFCI-F1-CF3
  - GRCC-2030859
  - GRCC-2040388
  - GRCC-2040390
  - GRCC-2058737
  - GRCC-2061930
  - GRCC-2063107
  - GRCC-2067719
  - GRCC-2067720
  - GRCC-2105390
  - PROV-MPNC
  - PROV-AMLC
  - PROV-MPNT
- New center:
  - UMIAMI

Data Concerns/Issues

These are the known data issues for this release. **Note**: There could be more undiscovered issues.

- The number of expected genes in certain gene panels may be incorrect due to sites submitting invalid gene symbols, coordinates, or incorrect assay information.
- **CRUK**
  - Confirmed genomic data for **CRUK-TS** samples contain no artifacts.
- **CHOP**
  - Fusion panel includes genes that are not on the other panels/pipelines that generate the SNV and indel calls. No BED file for fusion panel.
  - Confirmed genomic data for **CHOP-STNGS** samples contain no artifacts.
- **MSK**
  - Confirmed genomic data for MSK HEME panel samples contain no artifacts.
- **DFCI**
  - Expected discrepancies in expected and submitted gene count because of regions that are targeted for calling structural rearrangements.

- **UHN**
  - Confirmed genomic data for **UHN-54-V1** samples contain no artifacts.
  - UHN has some duplicate bed regions in their processed data. This is a processing issue, no action is needed from sites.
- **YALE**
  - Intentionally reports only amplifications in copy number data due to internal policy.
  - **YALE-OCP-V2** panel's expected gene count is 134 because one intron is included. Gene panels are created with only exons hence the difference in actual vs expected gene counts.
- Foundation Medicine genomic regions discrepancy
  - **DUKE/WAKE** use the Foundation Medicine T5A, T7, DX1, R2D2 bait sets, but some of the bed files uploaded don't seem to match the expected gene count per panel.
- Not all variants have variant counts (t_depth, t_alt_count, t_ref_count).
- Duplicated variants listed from **VICC** are expected due to providing Tempus samples.
- Genome Nexus related issues:
  - Duplicated variants