Fermentation Measurements and KPIs

This page lists the high-level records that are stored in LIMS and way they are related. the The goal of this page is to describe the fermentation data stored in LIMS and how one might access that data.

See Also

- Airflow Manual: Troubleshooting Fermentation Excel File Errors (how to handle errors in KPI calculations)
- Fermentation KPIs (initial project implementation)

Creation of the records

Experiment and Fermentation Run records are created often before anything is physically done in the lab. When the Fermentation Run is started, a sample is specified and is transferred into the virtual plate record representing the tank. If the model is being run via Polaris, the plate may be added to a cohort.

When samples are taken, they are tracked using Polaris, that creates the test tube "wells" record with the third party barcode, performs a liquid transfer to the sample well from the tank, and creates several extended attributes on the well record. For ambr runs, the liquid transfers are processed by the Hydra software.

During the run, several run-level measurements are collected either into Polaris or the Excel-based Universal Template.

After the run is completed, if a Universal Template was used, it is placed in Dropbox and is picked up by ZTL/Airflow (see Airflow Manual: Troubleshooting Fermentation Excel File Errors). The DAG writes measurements and runs the KPI calculation process, uploading measurements as Measurement Batches and the KPIs as Analytic Summaries.

Views

There are several fermentation views that exist in both the LIMS replica and Snowflake.

View Name	Description	Status
`fermentation_runs_vw`	An all-in-one view for data available at the fermentation run level of granularity.	STABLE
`fermentation_run_cohorts_vw`	Runs associated with cohorts. This association is made by the fermentation plate having the "main_tank" or "seed_tank" workflow item.	STABLE
`fermentation_run_tanks_vw`	The tank details for each run (includes tank_name and tank_vendor columns). One row for each fermentation run.	STABLE
`fermentation_run_strains_vw`	The fermentation_run_id, sample_id, and strain_id for each run (strain_id is pulled from the fermentation plate). One row for each fermentation run.	STABLE
`fermentation_run_quality_vw`	The latest run quality score for each fermentation run that has a quality score. One row for each fermentation run (if a score exists).	STABLE
`fermentation_run_measurements_vw`	Run level measurements for each run.	STABLE
`fermentation_run_data_status_vw`	The status of LIMS data for each fermentation run (used for notification emails).	STABLE
`fermentation_samples_vw`	Wells (both plate and cryovials) containing liquid samples taken from the tank during the run.	STABLE
`fermentation_sample_final_wells_vw`	Downstream wells containing liquid transferred to them from the sample wells.	STABLE
`fermentation_sample_measurements_vw`	Sample level measurements for each run.	STABLE
fermentation_sample_single_well_measurements_vw	Sample level measurements for cryovial wells.	STABLE
`fermentation_sample_plate_based_measurements_vw`	Sample/AC level measurements for plate-based sampling (e.g. AMBR).	STABLE
`fermentation_sample_ac_measurements_vw`	Analytical Chemistry measurements from the samples taken during the run.	STABLE KNOWN ISSUE: Sample wells must have a third_party_barcode to show up in this view.

`fermentation_sample_hplc_measurements_vw`
`fermentation_sample_gcms_measurements_vw`
`fermentation_sample_lcms_measurements_vw`
	Faster measurement-specific versions of the fermentation_sample_ac_measurements_vw view.	STABLE

`fermentation_all_kpis_vw`	KPIs for each run where it was successfully calculated.	STABLE

View Name	Description	Status
`fermentation_all_kpis_norm_vw`	Normalized KPIs from FermPipe. See the fermpipe_normalization Analytic Processes records in LIMS and the "fermpipe__*" DAGS in ZTL.	STABLE
`fermentation_hydra_plates_vw`	Record of which plates were handled or seen by Hydra on the AMBR 250s	STABLE

High Level Records

This section contains an high-level explanation of how each record is related to each other in LIMS.

Fermentation Sample Measurements

Sample measurements are stored under a measurement batch with `batch_type IN ('fermentationSample','fermentationSamplePolaris')`. For the 'fermentationSample' batch, the data are taken from the "Sample Data Output" sheet of the template file.

The "plateId", "fermentation sample type" and "fermentation sample tag" columns are important and are used to index the data later. The template should pre-populate these, so they shouldn't get messed up.

The value in the "Retain Barcode" column is used to look up the sample well in LIMS via the `third_party_barcode` field. If the well is not uploaded to LIMS for some reason, the upload process cannot associate the measurements with the right sample. This is a very common problem with the upload process.

Measurement columns are expected to have their units in square brackets (e.g. "sample volume mL") except for certain unitless measurements (e.g. "pH").

Fermentation Run Measurements

Run measurements are stored under a measurement batch with `batch_type IN ('fermentationRun','polarisFermentationRun')`. For the 'fermentationRun' batch, the data are taken from the "Run Level Output" sheet of the template file.

Run measurements are divided into multiple components, and each component is given a tuple in the measurement batch. Typically, this relationship is shown on the "Run Level Input" sheet of the template file. It's helpful to keep that in sync so that program-specific terminology can be mapped back to the measurements stored in LIMS.

Tuple 0 is for run-level measurements (e.g. tare weight of the fermentation tank, final weight of the tank, length of fermentation).

The components are dependent on the program. If you have a mixed component, it should possibly be split up into several sub-components in the later tuples (e.g. tuple=2 for the mixture, tuple=3 for the amount of dextrose that was added, tuple=4 for the amount of water that was added).

The "component name" measurement is important and is used to index the table later. By convention, tuple=0 should have a value of 'None' for both "component name" and "feed type".

Fermentation KPIs

Fermentation KPIs are stored as AnalyticSummary records. In order for the KPIs to be run a new AnalyticProcess with modelName="fermentation_kpis", modelVersion="1.0" must exist for the program (See the existing records in LIMS). If the AnalyticProcess is missing, no KPIs will be calculated. The KPI code lives in the fermentation-kpis repo.

If a value is needed to calculate a KPI, it must be included in one of the measurement batches above.

The fermentation KPI process happens after the sample and run measurements are stored. The measurements are pulled from LIMS and indexed according to certain columns (e.g. "component name" for run measurements, and type/tag for samples).

If a necessary measurement is missing, the whole KPI job will fail and nothing will be uploaded.

NOTE: KPIs were previously created as measurement batches. Some of these records still exist in LIMS.

Entity Details

- Experiment
  - Experiments can be identified as fermentation experiments by the "phase" attribute (i.e. `phase='fermentation'`).
- Fermentation Run
  - A fermentation run is a combination of three things:
    - The fermentation experiment.
    - An identifier for the physical hardware that is used (i.e. the "fermentation tank").
    - A `run_type` attribute that specifies whether the run is in a "seed", "seed2", or "main" stage.
    - A `sample_id` referring to a Sample, and in turn a Strain, that was used for the run.
      - NOTE: The "fermentation_runs.sample_id" column is empty in the database. The sample in the well of the plate (via the "plate_id") is the definitive record. If necessary to query the strain for all runs, use the "fermentation_strains_vw" view.
- Fermentation Plate
  - A single well plate used during a fermentation run to represent the fermentation vessel.
  - When samples are taken during the fermentation run, a liquid transfer is performed from the fermentation plate into a sample well.
- Sample Wells
  - Sample test tubes with contents created via a liquid transfer from the fermentation vessel.
  - They are placed in a "cryobox" plate during the sample process and then subsequently removed from the plate. As a result, these wells have no associated "plate_id" as a result (they are "orphaned wells").
  - These wells have a `third_party_barcode` set on them that is used to retrieve them during the upload of measurement batches.
- Sample Measurement Batch
  - A measurement batch with `batch_type='fermentationSample'` contains Measurement records associated with each sample well.
  - NOTE: Each Measurement record can be associated with a Fermentation Run via the `plate_id` field and the "fermentation sample tag" Measurement's `categorical_value` (to identify "seed" vs "main").

    - Each measurement has the following:
        - The `plate_id` refers to the Fermentation Plate.
        - The `well_id` refers to the Sample Well.
        - The `aliquot_id` refers to the Aliquot that was transferred from the Fermentation Plate. It should contain a Sample that refers to a specific Strain. This should match the Strain associated with the Fermentation Run. *Note: in some old data, this might not be the case.*
- Run Measurement Batch
    - A measurement batch with `batch_type='fermentationRun'` contains measurements for the run as a whole.
    - For each Measurement:
        - The `plate_id` for each measurement refers to the Fermentation Plate.
        - The `aliquot_id` and `well_id` fields will be `null`.
- KPI Measurement Batch
    - DEPRECATED: These have been migrated to AnalyticSummary records. They exist for older runs, but should generally be ignored.
    - A measurement batch with batch_type='fermentationKPI' contains the KPIs for each run.
    - The `plate_id` for each measurement refers to the Fermentation Plate.
    - The `aliquot_id` and `well_id` fields will be null for each of these measurements.
- KPI Analytic Process
    - An Analytic Process record with `model_name='fermentation_kpis'` and `model_version='1.0'` controls whether KPIs are run for a given program.
    - This record is created manually and has:
        - The `program_id` refers to the Program
        - The `restricted_output` determines whether KPIs should be secured with the "Ferm KPI" ZAuth class.
        - The `program_parameters` value is unused and is set to '{}'.
- KPI Analytic Process Run
    - After the Run and Sample Measurement Batches are uploaded, the KPI process is started.
    - An Analytic Process Run is created every time the KPIs are calculated.
    - Each Analytic Process Run has multiple Analytic Summary records. Each summary contains a KPI, or a calculation made to create a KPI. Each Analytic Summary has:
        - The `target_id` refers to the Fermentation Run
        - The `name` should be a unique name for each Analytic Process Run (not enforced).
        - The `value` contains the raw value of the KPI.
        - The `units` contains the units for the KPI.
- QC Process
    - One or more manual QC Process with `purpose='run quality'` may exist.
    - Each QC Process has several Quality Score records, each with:
        - The zid refers to the Fermentation Run.
        - The `pass` will be `0` or `1`
        - The `score` is a value that indicates the failure mode of the run (use or inspect the fermentation_run_quality_vw view for the interpretation).
        - The `reason_for_score`, and `comment` fields may or may not be filled out.
- Extended Attributes
    - There may be extended attributes for various records
    - Fermentation Run
        - `'is operations'` indicates that the run is a TEST run, not a DEV or EQUIV run.
        - `'tank type'` indicates the fementor type (e.g. BioFlo or DasGip)
        - `'project name'` indicates the program
        - `'fermentation process version'` indicates the process version (program specific meaning)
        - `'fermentation process details'` describe the process conditions
    - Sample Well
        - A 'sample source' record may exist with values such as ({{'Harvest', 'Head Plate', 'Sampling Port')

Sample measurement names

These are the measurement names used for Delta sample fermentation runs.

- absorbance
- dextrose
- fermentation sample tag
- fermentation sample type
- gravimetric product titer
- NH3
- offline pH
- pH
- product removed
- production fermenter post weight
- production fermenter pre weight
- sample volume
- volumetric product titer

Run measurement names

These are the measurement names used for Delta fermentation run measurements.

- antifoam final weight
- antifoam initial weight
- base final weight

- base initial weight
- component description
- component name
- feed number
- feed type
- final overflow weight
- final weight
- harvest final weight
- initial weight
- length of production fermentor
- length of seed fermentor
- lot
- post autoclave weight
- process deviation
- production fermentor final volume
- production fermentor final weight
- production inoculation final weight
- production inoculation initial weight
- tare weight
- Weight

KPI names (measurement or analytic summary)

These will be the same name of the Analytic Summary records after the migration.

- dextrose
- final harvest volume
- final od
- final tank mass
- final tank volume
- final titer
- length of production fermentor
- max od
- productivity
- total product
- yield