AI-Augmented Quality Control for Ion Torrent Sequencing Workflows Using Automated Report Extraction, Longitudinal Analysis, and LLM-Driven Interpretation

Casiana Fernandez-Bango, Emilio Margolles, Ozzarah Ferrer, Syduet Gonzalez, Hugo Kaneku

Immunology and Histocompatibility Laboratory · University of Miami · Miami, FL

Why is this needed?

Ion Torrent S5 QC review is largely manual and based on individual PDF reports. This makes it difficult to identify longitudinal trends and compare performance across instruments, chips, and reagent lot numbers.

Objective

Develop an automated, AI-augmented framework to extract, standardize, analyze, and interpret Ion Torrent S5 QC data for longitudinal performance monitoring.

Methods

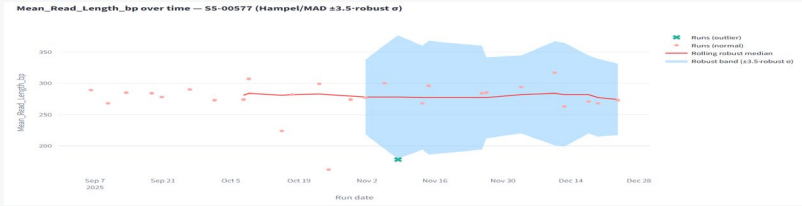
The pipeline parses unstructured Ion Torrent S5 sequencer-generated reports and converts QC metrics into a validated, standardized dataset for downstream analysis. The dataset is used to generate longitudinal visualizations, including read-length trends, instrument-level comparisons, and stability metrics. Outputs can be exported for documentation or reviewed through an interactive Streamlit interface. An integrated large language model (LLM) generates standardized interpretive summaries, identifies potential QC concerns, proposes technical causes, and recommends corrective actions.

REPORTS →EXTRACT →STRUCTURE →ANALYZE →INTERPRET

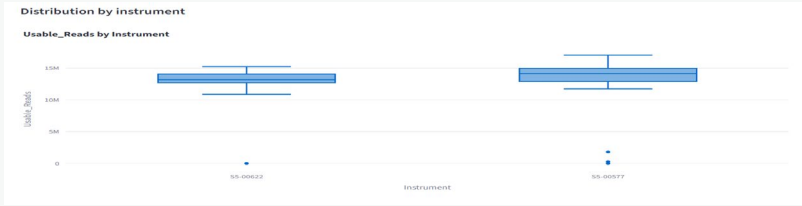
PDFs · Parsing · QC Dataset · Statistics · AI Insights

Results

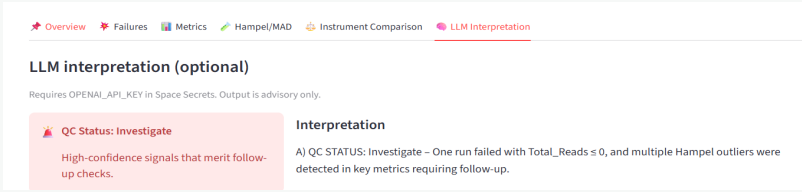
AUTOMATED LONGITUDINAL SURVEILLANCE: *Robust statistics identify deviations that may be difficult to recognize during individual-run review.*



AUTOMATED INSTRUMENT COMPARISON: *QC metrics are continuously stratified and compared across sequencing instruments.*



AI-ASSISTED INTERPRETATION: Statistical signals → standardized interpretation → actionable QC guidance



Laboratory Operational Impact

- **AUTOMATES**
  Transforms static PDF reports into structured QC data.
- **MONITORS**
  Enables longitudinal surveillance across runs, instruments, and consumables.
- **STANDARDIZES**
  Supports consistent statistical and AI-assisted interpretation.
- **SUPPORTS ACTION**
  Highlights QC concerns and supports targeted investigation.

Conclusions

This proof-of-concept demonstrates that automated extraction, robust longitudinal analysis, and LLM-assisted interpretation can provide a scalable framework for sequencing QC monitoring. The approach reduces subjectivity, improves consistency, and supports enhanced operational oversight in transplant immunology laboratories.

Live Demo

Hugging Face

Public Source Code

GitHub