MyInformatics® - Streamlining End-to-End NGS Analysis with Automated, Containerized Workflows for Faster Results

Jillian Burke[1], Julian D'Angelo[1], Travis Johnson[1], Caistarrin Mystical [1], Paulina Sanchez[2], Juan Sierra [1], Javier Velazquez-Muriel [1], Joshua Wemmer[1],
[1]Invivoscribe, San Diego, CA, [2]LabPMM, San Diego, CA

INTRODUCTION

- MyInformatics® is an NGS software system that streamlines the entire sample processing workflow for Invivoscribe's CAP/CLIA-certified MyAML® and MyMRD® Assays - from data creation and entry into a laboratory information management system (LIMS), through comprehensive bioinformatics analyses, to final variant reports.
- Processes such as quality control and sample tracking are automated through integration with the LIMS and web display components, enhancing data transparency.
- These seamlessly interlinked processes also eliminate many sources of human error and increase efficiency, resulting in decreased turnaround times.

METHODS

- The MyInformatics ecosystem consists of interconnected components that enable efficient data exchange. Each component undergoes rigorous validation before integration into the broader system, which is then subject to comprehensive system validation.
- The LIMS receives QC data from the bioinformatics pipeline, determines sample quality (pass/fail), and returns results to the pipeline, allowing for automated decision-making.
- The bioinformatics portion is fully automated with distinct subcomponents separated into Docker containers to create a modular design that facilitates quick updating of the software. Subcomponents include quality control (QC), variant calling, and annotation with relevant metadata. The workflow is run with NextFlow, allowing seamless pipeline orchestration.
- Pipeline outputs are transferred to the MyInformatics Annotation Portal (MAP) via LIMS, ensuring only high-quality samples with complete datasets are imported.
- Clinical Genomics scientists interpret and classify variants in MAP. To ensure accurate and actionable reporting, the final report aligns with clinical guidelines from Clinical Genome Resource, Cancer Genomics Consortium, Variant Interpretation for Cancer Consortium, and Association for Molecular Pathology.

RESULTS

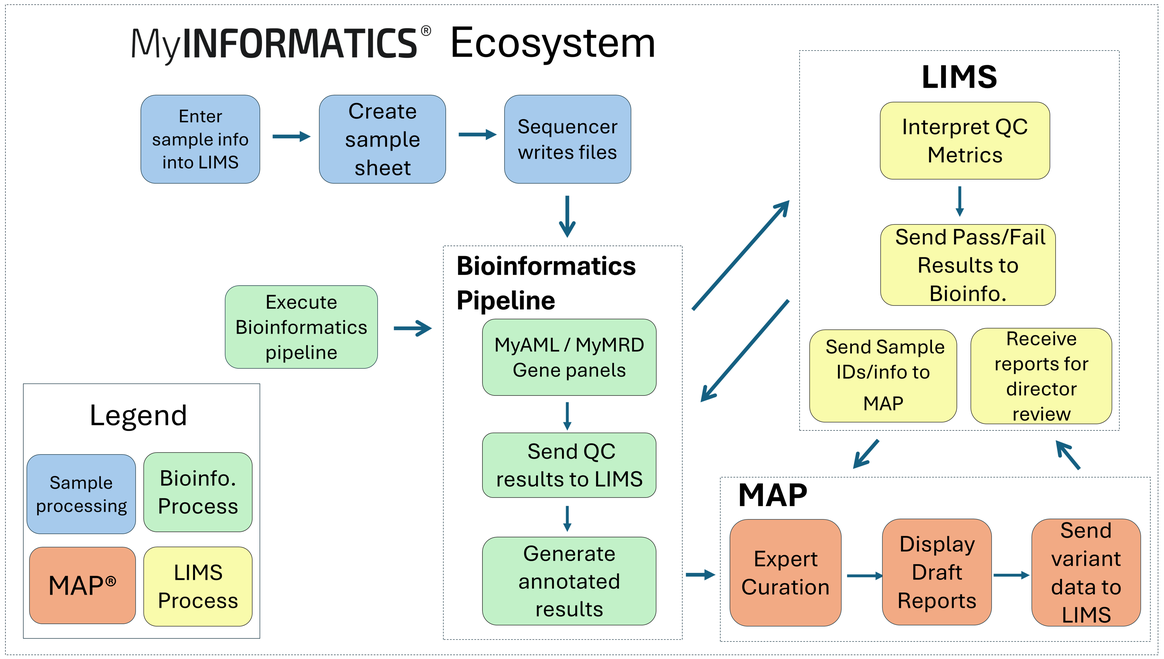

Fig 1. Overview of MyInformatics Ecosystem. The system begins with entering sample IDs into the LIMS and sample processing. After data generation, bioinformatics analysis is started. All bioinformatics steps following pipeline execution, and all subsequent data transfer steps, are automated. Sample quality information is sent to LIMS, which returns a pass/fail decision. Following analysis, the LIMS imports only the passing, high-quality samples into the MAP software, where they receive expert variant interpretation and classification before the reportable variants are sent to the LIMS for final report creation

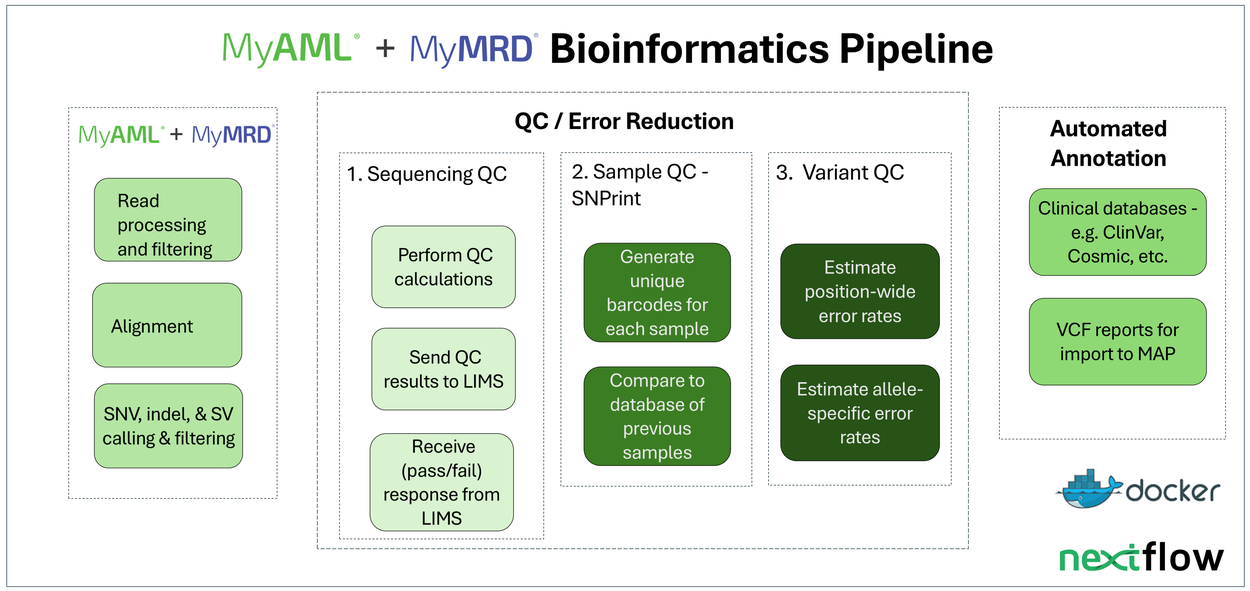

Fig 2. Overview of Bioinformatics Components. Subcomponents of the analysis pipeline are housed in unique Docker containers run in series using a NextFlow workflow. The pipeline contains multiple QC steps that ensure the quality of the sequencing run and individual samples, along with a proprietary error reduction step that reduces false positive variants. Predicted variants are populated with metadata from databases that will be passed automatically by the LIMS for curation in the MAP interface.

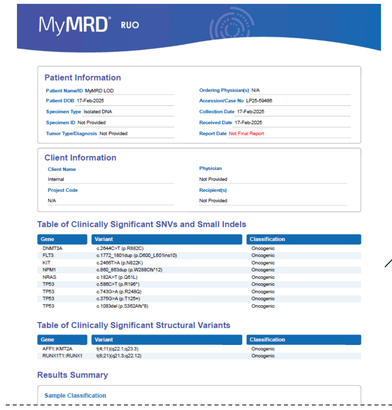

Figure 3a. The MyInformatics system improves efficiency in data reporting by eliminating manual data transfer. In this example report for the MyMRD assay, reported variants are automatically imported into MAP. After expert interpretation and classification, reports are sent directly to the LIMS for release by the clinical lab director. This automation reduces labor and accelerates sample processing.

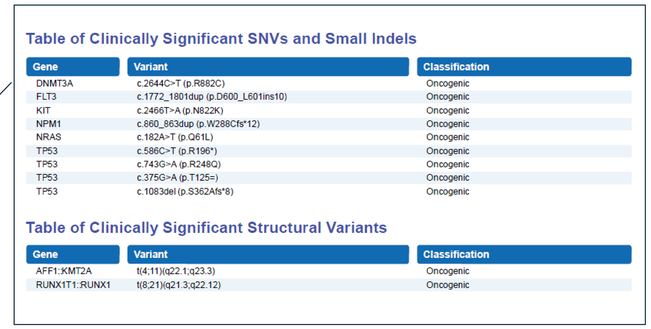

Table of Clinically Significant SNVs and Small Indels

Gene	Variant	Classification
DNMT3A	c.2644C>T (p.R882C)	Oncogenic
FLT3	c.1772_1801dup (p.D600_L601ins10)	Oncogenic
KIT	c.2466T>A (p.N822K)	Oncogenic
NPM1	c.860_863dup (p.W288Cfs*12)	Oncogenic
NRAS	c.182A>T (p.Q61L)	Oncogenic
TP53	c.586C>T (p.R196*)	Oncogenic
TP53	c.743G>A (p.R248Q)	Oncogenic
TP53	c.375G>A (p.T125=)	Oncogenic
TP53	c.1083del (p.S362Afs*8)	Oncogenic

Table of Clinically Significant Structural Variants

Gene	Variant	Classification
AFF1:KMT2A	t(4;11)(q22.1;q23.3)	Oncogenic
RUNX1T1:RUNX1	t(8;21)(q21.3;q22.12)	Oncogenic

Figure 3b. End-to-end sample processing reduces turnaround time in clinical reporting. This section highlights a table from **Fig. 3a,** showing predicted genomic variants and their clinical classifications. Metadata from clinical databases is automatically transferred and displayed for curators, streamlining the classification process and reducing the overall workflow time.

CONCLUSIONS

- MyInformatics connects all key components of a modern NGS assay analysis workflow, from the sequencing instrument to the annotated sample reports.
- Containerization makes deployment of the software in different environments easy.
- Automation through connecting multiple software components reduces overall turnaround time, improves data transparency, and limits errors.
- MyInformatics enables consistency in results by standardizing assay and analytical workflows across LabPMM, Invivoscribe's global network of reference laboratories.

REFERENCES

1.Horak P et al. Standards for classifying somatic variant pathogenicity in cancer. Genet Med. 2022;24(5):986-998. PMID: 35101336.
2. Vogelstein B et al. Cancer genome landscapes. Science. 2013;339(6127):1546-1558. PMID: 23539594.
3. Hegde M et al. Reporting incidental findings in genomic sequencing. J Mol Diagn. 2015;17(2):107-117. PMID: 25684271.

invivoscribe®