Polly^{TM}: A novel cloud-based platform for Metabolism labs

Abhishek Jha[1,2], Swetabh Pathak[1]; Sabu George[1]; Raghav Sehgal[1]; Shefali Lathwal[1]; Darren Dumlao[3]; Mary Piotrowski[3]; John Janiszewski[3]
[1]Elucidata, New Delhi, India; [2]Elucidata, Cambridge, Massachusetts; [3]Pfizer, Groton

INTRODUCTION

Investigating cellular metabolism to develop new therapeutics has proven to be a promising approach to drug discovery for both cancer and immunometabolism. Despite these successes, broader integration of metabolism studies in drug discovery programs has been limited due to computational and logistical challenges of translating raw data into metabolic insights. Noticeably, there is a lack of efficient, iterative, and scalable data analytics platform, limited methods to integrate different types of omics data, and inadequate capabilities to support targeted flux experiments. We present Polly^{TM}, a platform that combines high performance raw data processing pipelines, sophisticated data analysis algorithms, and data visualization dashboards on cloud. Polly^{TM} has been developed in collaboration with leading pharmaceutical companies and cell metabolism labs to increase their throughput.

METHODS

To demonstrate the high-throughput capability of Polly^{TM}, we used unbiased global metabolite profiling data, RNAseq data, and [$U^{13}C_6$] glucose and [1,2-$^{13}C_2$] glutamine labeled MRM data from primary human CD4+ T cells at 0 and 24 hrs of activation. We tried two workflows on Polly^{TM} for analysing the above mentioned data, essentially, Polly^{TM} IntOmix Workflow for integrative omics analysis and PollyPhi^{TM} Relative LC-MS/MS workflow for fluxomics analysis.

Polly IntOmix Workflow

Polly^{TM} IntOmix workflow helps in finding the most relevant metabolic module (a subset of the 9,000 reactions and 16,000 metabolites in KEGG) for a given biological context and thus is a critical step in understanding cellular phenotypes. There are three parts to it:

- Polly^{TM} MetScape workflow: Input is MZXML files, output is differential expression and pathway visualizations.
- Polly^{TM} RNAseq pipeline: Input is Fastq files, output is differential expression
- Polly^{TM} IntOmix: Takes differential expression files as input from the MetScape workflow and the RNAseq pipeline, and returns the most significantly changing subnetwork.

MetScape Workflow

RNASeq Pipline

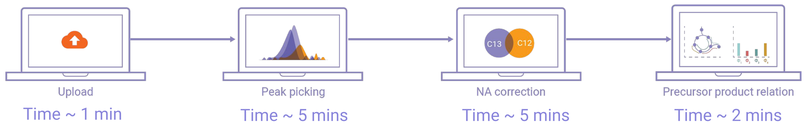
PollyPhi^{TM} Relative LC-MS/MS Workflow

Methods to catalog and quantify metabolites in metabolomic modules are informative but they do not present a complete view and can yield misleading conclusions in terms of the directionality and the rate of flow of the metabolites. PollyPhi^{TM} Relative LC-MS/MS workflow helps in solving this problem by obtaining key insights using isotopic information to measure the flow of atoms in a pathway. It has three steps:

- Peak curation using El-MAVEN[1]: typically takes <5 mins to curate an MS-MS dataset.
- Natural Abundance correction - Performed using Polly^{TM} IsoCorrect in < 5mins.
- Phi Calculations - The PollyPhi^{TM} Relative LC-MS/MS dashboard enables users to calculate upto 56 phis[2] related to TCA in <2 mins.

The whole workflow can run in less than 15 minutes.

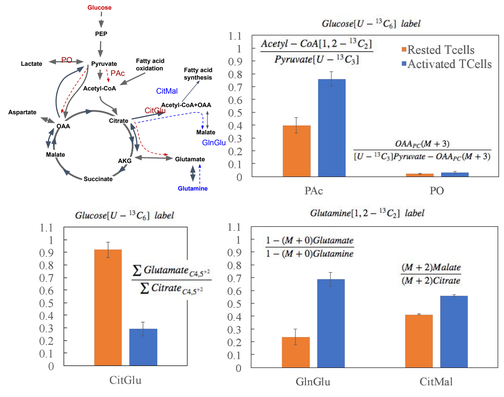
PollyPhi^{TM} Relative LC-MS/MS



RESULTS

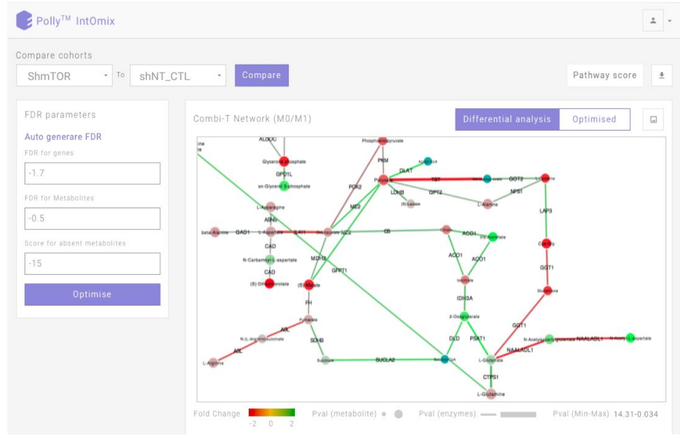
PollyPhi^{TM} Relative LC-MS/MS Workflow shows decreased fatty acid breakdown and increased influx of glutamate carbons into TCA cycle for activated T-cells

Labeled LC-MS/MS data for T-cells was used to calculate ratios of precursors and labeled products in the lower glycolysis and TCA cycle using the PollyPhi^{TM} Relative LC-MS/MS application. Going from fractional enrichment values to visualization of ratios, which allow relative comparisons of changing fluxes in the TCA cycle, for both glucose and glutamine labels took less than five minutes.



Polly^{TM} IntOmix Workflow shows mTOR knockdown cells compensate for reduced citric acid cycle intermediates using alternative carbon sources such as arginine

- Polly^{TM} MetScape processed 2GB of RAW metabolomic data with 24 samples in under 30 mins. The following steps were performed:
  - Peak picking performed on 2700 metabolites
  - Sample quality check, Principal component analysis
  - Differential expression analysis to identify 700 metabolites which were changing
- Polly^{TM} RNA seq Pipeline ran in 60 mins on transcriptomics data.
- Processed metabolomic data and transcriptional T cell data at 0 and 24 hrs of activation were processed through Polly^{TM} Intomix using CoMBI-T[3] in under 5 mins.



CONCLUSION

- Polly^{TM} IntOmix workflow can generate integrative omics insights from raw data in under 60 mins
- PollyPhi^{TM} can help in getting Relative Flux insights in under 30 mins
- Polly^{TM} in general can be used to generate metabolic insights very quickly thereby enabling scientists significantly

REFERENCES

1. El-MAVEN v0.3.2, https://elucidatainc.github.io/ElMaven/
2. Integrated, Step-Wise, Mass-Isotopomeric Flux Analysis of the TCA Cycle: Alves, Tiago C. et al. Cell Metabolism , 2015, Volume 22 , Issue 5 , 936 - 947
3. Network Integration of Parallel Metabolic and Transcriptional Data Reveals Metabolic Modules that Regulate Macrophage Polarization: Jha, Abhishek K. et al. Immunity, 2015, Volume 42 , Issue 3 , 419 - 430