

Case Study: iCore as an Informatics platform to execute flexible comparative analysis

Description

iNDx was approached with one of the mid-sized pharmaceutical company running its clinical trial to perform RNAseq analysis and downstream GSEA analysis for them. The client had previously approached another vendor for performing their analysis. The problem was that they had to periodically send in the sets of comparisons they were interested in and wait for the execution results which typically took a week to 10 days. This made it extremely restrictive for them to conduct an in-depth analysis and explore different analysis combinations based on their parameters of interest. They were looking for an extended informatics team who could not just provide service-based analysis for a set of sample sets but also provide a solution to conduct their analysis flexibly and efficiently. They were also interested in conducting GSEA analysis for their specific genesets.

Problem

To benefit maximize the benefit from conducting RNAseq assay, different sample sets must be compared. The grouping of samples can be determined by various factors. Some examples can be

- To determine the changes of gene-expression before and after treatment administration, timepoint based analysis may be required.
- To compare the patients belonging to a certain arm with Dosage X with another arm with dosage XX
- To compare different arms getting different treatment administration
- To compare the patients showing positive response till week 16 to non-reponders
- To compare patients with different types of solid tumors(e.g: Head and Neck Vs Melanoma)

As the trial progresses, the response groups and patient parameters in terms of their response and their reaction to a certain drug or adverse events can be factors for comparison. The study also must accommodate the new phases and additional samples from those phases and compare them with the previous data. So, its almost not possible to come up with a complete panel of analysis to be conducted and the present landscape with vendors offering the service-based analysis incurs a lot of turn-around time and becomes extremely restrictive for scientists to conduct a flexible research based on various parameters of interest.

On the other hand of the spectrum, there are big informatics vendors who provide very expensive solutions with a lot of unnecessary features which must be bought as a single system which does not cater to the specific problem at hand. iNDx addresses this problem by providing a platform which is flexible, modular and extensible to accommodate this.

Solution Features and Workflow

iNDx provided a solution which tackled the pain points of the customers by providing the following salient features-

The execution of the analysis was moved into a cloud-based kubernetes platform where the user is now able to run the analysis of interest with a click of the button. This on-demand analysis takes about 5-10 minutes to execute and produce the necessary results – a drastic change from a wait time of a few weeks.

- The samples are presented to the user along with all the necessary parameters with an option to filter the samples and group them in accordance to criteria of interest. The samples of interest can be selected and groups created and saved for which can then be used for running multiple analysis.

Define Group

Filter

<input type="checkbox"/>	Filter	BARCODE	PUID	ARM	CHORT GROUP	TIMEPOINT	DISEASE NAME	DISEASE DETAIL	RESPONSE GROUP
<input type="checkbox"/>	Sample barcode								
<input type="checkbox"/>	PUID								
<input type="checkbox"/>	ARM								
<input type="checkbox"/>	Cohort Group								
<input type="checkbox"/>	Timepoint		007-001	A	Monotherapy 40mg	A TT	OVA	Ovarian	
<input type="checkbox"/>	Disease name		001-001	A	Monotherapy 10mg	A_C1D1	PANC	Pancreatic	
<input type="checkbox"/>	Disease detail								
<input type="checkbox"/>	Response group								
<input type="checkbox"/>		K00876CD01	001-003	A	Monotherapy 10mg	A_C1D1	OVA	Ovarian	
<input type="checkbox"/>		K00877CD01	001-004	A	Monotherapy 10mg	A_C1D1	END	Endometrial	
<input type="checkbox"/>		K00908CD01	001-001	A	Monotherapy	A_C1D8	PANC	Pancreatic	

- The GSEA analysis is also closely knit with the RNASeq differential analysis, where the user can add any number of geneset and check for the enrichment of the new geneset.

DEFINE GENESET

GeneList

GENEDB

Description

Geneset name

- The analysis execution can be monitored through a monitoring system which provides the status and logs of the executions.
- Once the results are available, they are available in a central place which is our Discovery Dashboard. The history of all executions is available for the user to choose and view.



Note:

The solution was also created in an extremely flexible fashion so that, any new data types and assays can be easily added based on user requirements. This would only mean plugging in the additional pipelines which can be started after choosing groups and as the new assay results are available, new cards are generated on the Discovery dashboard with the relevant results.