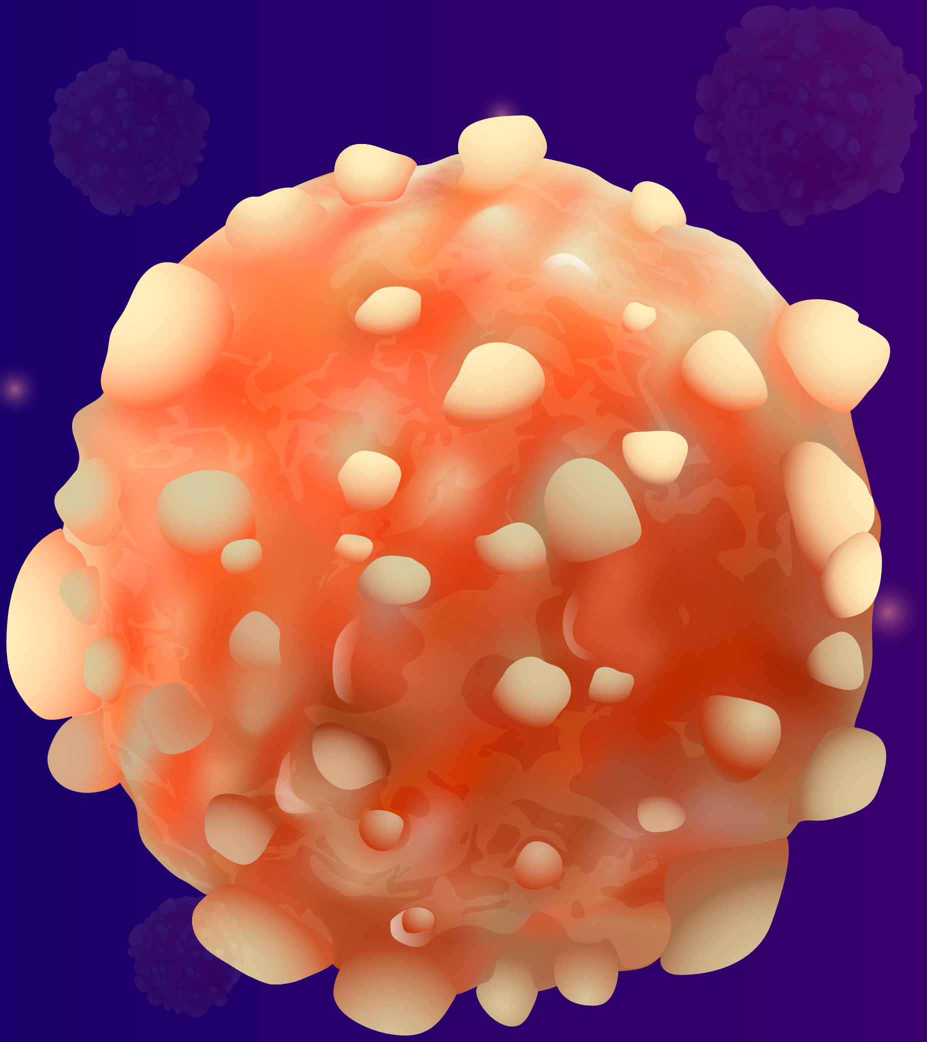
excelra

Where data means more

CASE STUDY

To perform tumor enriched target identification, prioritization and tissue restricted/lineage target identification and prioritization for Small cell lung cancer and Ovary carcinoma



Objective

- To collect datasets pertaining to small cell lung carcinoma (SCLC) and ovary cancer (OC) and along with control samples studies from public data sources to perform target identification and prioritization.

- Integrate public bulk RNA data and scRNA data available for both lung and ovary cancer, their respective normal samples and perform lineage study with normal vital organs samples

Input

- TCGA and GEO cancer data (raw counts) with appropriate parameters.

- TCGA and GEO cancer data (raw counts) with appropriate parameters.

Workflow (overall)

- Identification of accurate target datasets from public data sources.

- Integration, Batch effect removal & Normalization of the extracted data.

- Identification of differentially upregulated genes from comparison of cancer dataset with control and vital normal organs

- Identification of differentially upregulated genes from comparison of control datasets with vital normal organs for Ovary lineage study

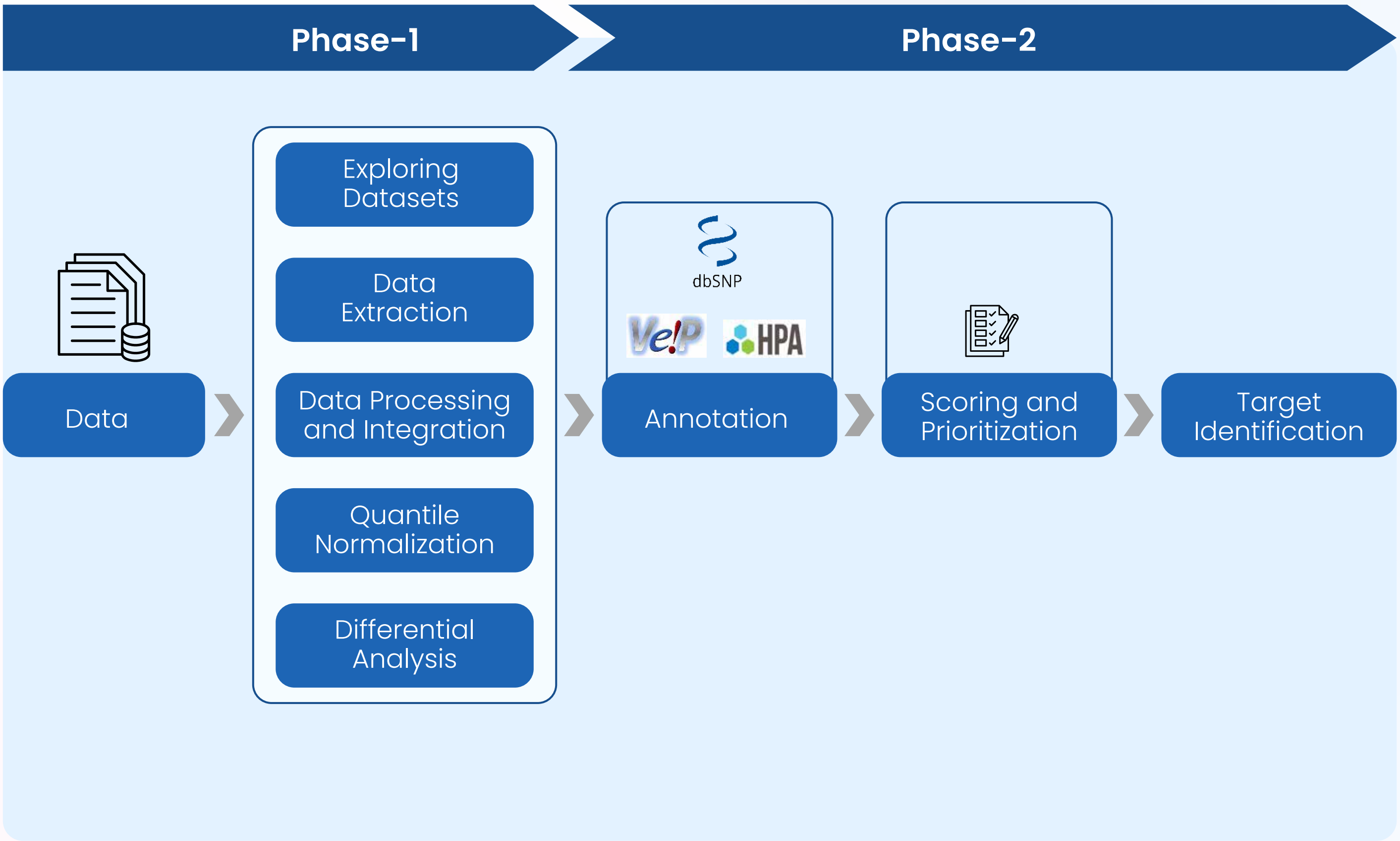
- Identification of tumor specific genes from scRNA data for lung and comparison with scRNA vital organs data for Lung lineage study

- Annotation of the genes using multiple databases and prioritization of targets/genes

Deliverables

- Comprehensive list of genes/targets from tumor enrichment and lineage analysis
- List of prioritized targets for objectives stated above
- Detailed report & necessary files

Approach



Phase-1

Phase-2

Data

Exploring Datasets

Data Extraction

Data Processing and Integration

Quantile Normalization

Differential Analysis

Annotation

Scoring and Prioritization

Target Identification

Results

Curated Data

Annotated Data

dbSNP

Ve!P

HPA

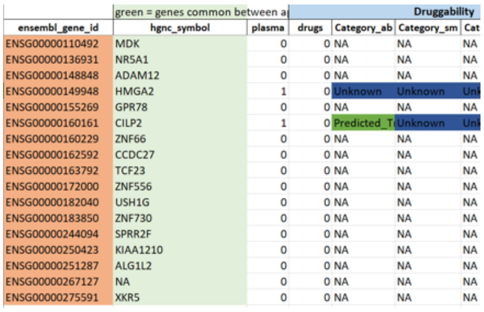
Gene Buckets

Clinical Connections

Pathway Information

Drug Information

Gene Buckets



	green = genes common between a		Druggability			
ensembl_gene_id	hgnc_symbol	plasma	drugs	Category_ab	Category_sm	Cat
ENSG00000110492	MDK	0	0	NA	NA	NA
ENSG00000136931	NR5A1	0	0	NA	NA	NA
ENSG00000148848	ADAM12	0	0	NA	NA	NA
ENSG00000149948	HMGA2	1	0	Unknown	Unknown	Unk
ENSG00000155269	GPR78	0	0	NA	NA	NA
ENSG00000160161	CILP2	1	0	Predicted_T	Unknown	Unk
ENSG00000160229	ZNF66	0	0	NA	NA	NA
ENSG00000162592	CCDC27	0	0	NA	NA	NA
ENSG00000163792	TCF23	0	0	NA	NA	NA
ENSG00000172000	ZNF556	0	0	NA	NA	NA
ENSG00000182040	USH1G	0	0	NA	NA	NA
ENSG00000183850	ZNF730	0	0	NA	NA	NA
ENSG00000244094	SPRR2F	0	0	NA	NA	NA
ENSG00000250423	KIAA1210	0	0	NA	NA	NA
ENSG00000251287	ALG1L2	0	0	NA	NA	NA
ENSG00000267127	NA	0	0	NA	NA	NA
ENSG00000275591	XKR5	0	0	NA	NA	NA

Where data means more

excelra

SAN FRANCISCO • BOSTON • LONDON • GHENT/GENT
UTRECHT • HYDERABAD

Connect with our experts: marketing@excelra.com

www.excelra.com