How to get the most out of your RNA-seq data? Precision insights is your answer

Tamara Krsmanovic[1] *, Katica Jankovic[1], Toni Franjkic[1], Bendetta Leoni[1], Gordana Apic[1], Robert B. Russell[2]

[1] Metisox, Cambridge, UK; [2] University of Heidelberg, Cell Networks, Heidelberg, Germany, *tamara.krsmanovic@metisox.com

OBJECTIVES

RNA-sequencing (RNA-seq) has a wide variety of applications, but no single analysis pipeline can be used in all cases (1). We have developed Precision Insights protocol that allows to: gain a detailed and personalized understanding of genetics, gain molecular and clinical factors to improve disease diagnoses, treatment, and overall healthcare outcomes, make informed decisions, optimize processes, and develop products that are tailored towards specific needs and goals.

The aim of the study was to validate the Precision Insights protocol by exploring the RNA-seq data (GSE52778) (2) of airway smooth muscle cells (ASM) treated with dexamethasone.

METHODS

Precision Insights protocol (Figure 1) includes open-source tools for raw data processing, differential expression analysis and gene enrichment analysis and use of our proprietary database and software for data interpretation and visualization.

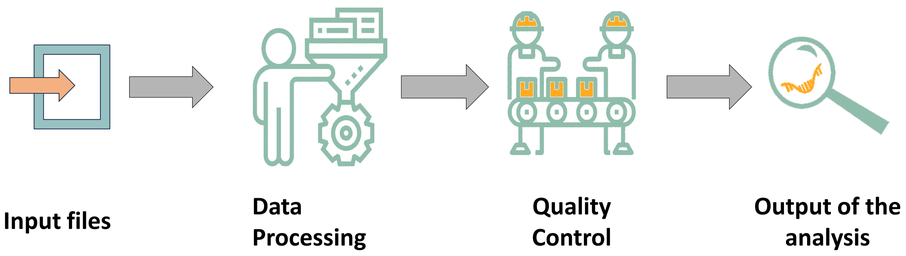

Figure 1. RNA-seq analysis workflow.

We generated our proprietary database to include internally curated and various publicly available data on proteins, biologically-active chemicals, their interactions, pathways and pathologies (Figure 2). We catalogued 4 million references, supporting each database entry, with hyperlinked interactions to appropriate PubMed articles as support, while proteins and chemicals were hyperlinked to EntrezProtein and PubChem, respectively. In order to support data integration and capture different levels of histopathological observations, we developed toxicology ontology with over 2,500 toxicity endpoints. The software used to query this database allows for enrichment of experimental data set with known protein and chemical interactions and effective data visualization.

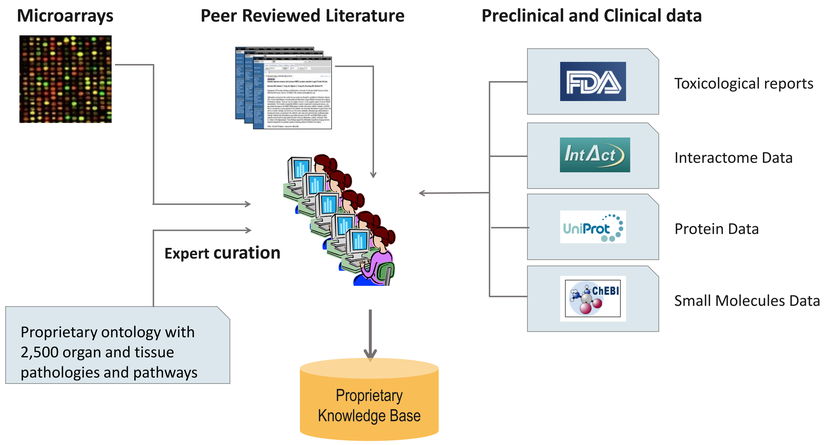

Figure 2. Manually curated database. Data was extracted from many diverse sources. Database contains complete human interactome, 2,500 organ and tissue pathologies, 500,000 synonyms/ontologies, 4 million linked articles (PubMed, FDA, clinical, patents), >10 million relationships, 50,000 biologically active chemicals.

RESULTS

Initial raw data processing of 8 samples (4 treated with dexamethasone and 4 with control vehicle) and differential expression analysis (cut off: logFC>±2, p_{adj}<0.05) identified 188 differentially expressed genes (Figure 3). Out of these, 103 were upregulated and 85 were downregulated in the treatment group. Functional profiling of this set of differentially expressed genes revealed overrepresented Gene Ontology and Reactome pathways that are potentially perturbed by dexamethasone treatment (Table 1).

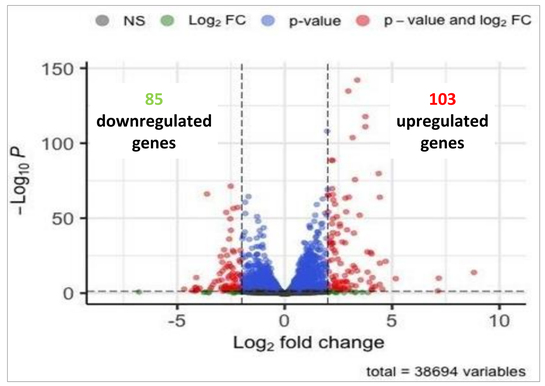

Figure 3. Differential expression analysis (enhanced Volcano plot).

Table 1. Functional profiling: GO overrepresentation test, Reactome overrepresented pathways, and pathologies identified with our proprietary database.

Enriched pathways		Pathologies
GO	**Reactome**	**Proprietary database**
Cellular response to heavy metal ions	Immune and neuronal system	Cancers (glioma, Retinoblastoma)
Regulation of peptide hormone secretion	Signal transduction	Neurodegenerative diseases (Alzheimer's disease)
Regulation of insulin secretion and response to insulin	Cellular response to stimuli	
Regulation of T cell proliferation and inflammatory response	Metabolism of proteins	Inflammatory diseases (liver cirrhosis, colitis, osteoarthritis, diabetic nephropathy)
Neurogenesis	Gene expression and disease	

We used our proprietary database and software to analyze the effects of dexamethasone on the respiratory system and identified recurring pathologies for this set of genes (Table 1). Literature evidence indicated that dexamethasone is used for symptom relief in various cancers treatments, or to treat many inflammatory and autoimmune disorders (3, 4).

Detailed interrogation of the ASM hyperplasia (one of the recognized phenotypes of asthma) network from our proprietary database (Figure 4), combined with differentially expressed genes from the experimental dataset, revealed that the mechanism behind ASM hyperplasia is complex and that dexamethasone treatment influences different aspects of it.

Dexamethasone downregulated the pro-inflammatory genes, as well as upregulated anti-inflammatory genes, effectively reducing the inflammation (Figure 4). Dexamethasone downregulated NOV, which in recent years emerged as a critical regulator in a variety of immune-related diseases. Dexamethasone upregulated KLF5 and IGFBP2, consequently ameliorating airway smooth muscle cell hypertrophy.

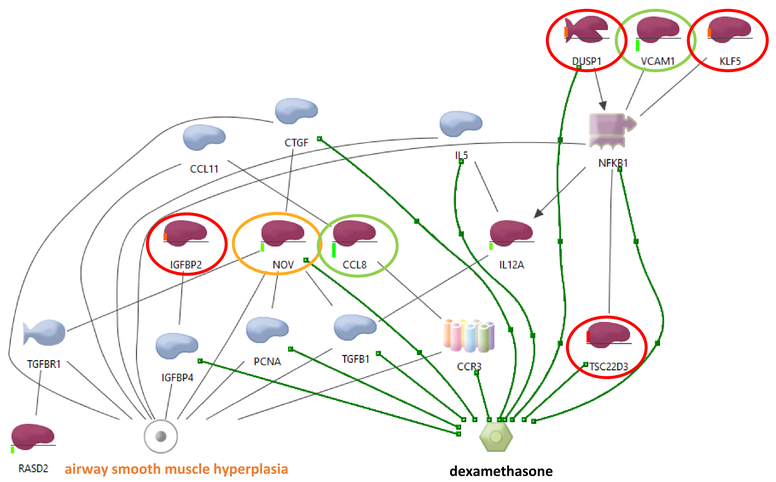

Figure 4. A segment of airway smooth muscle hyperplasia pathway showing potential mechanism of action of dexamethasone. NOV-nephroblastoma overexpressed gene; pro-inflammatory genes: VCAM1 and CCL8; anti-inflammatory genes: DUSP1 and TSC22D3; KLF-plays a role in regeneration and immune response; IGFBP2-inhibits IGF-mediated growth and developmental rates. Differentially expressed genes (DEGs) from the experimental dataset are labeled purple. Red circles indicate upregulated genes, green circles indicate downregulated genes, orange circle indicates direct link of the DEG to the airway smooth muscle hyperplasia.

CONCLUSION

- In this example, we showed that our Precision Insights protocol can help in deciphering the drug mechanism of action and in deeper analysis of specific aspects of the mechanism. It also helps in identifying potential adverse effects associated with drug target manipulation.

REFERENCES

1. Conesa *et al.*, Genome Biol. **2016** Jan 26:17:13. **2. Himes *et al.*,** PLoS One. **2014** Jun 13;9(6):e99625. **3. Noreen *et al.*,** Eur J Pharmacol. **2021** Mar 5:894:173854. **4. Cook *et al.*,** Oncoimmunology. **2015** Sep 16;5(3):e1066062.