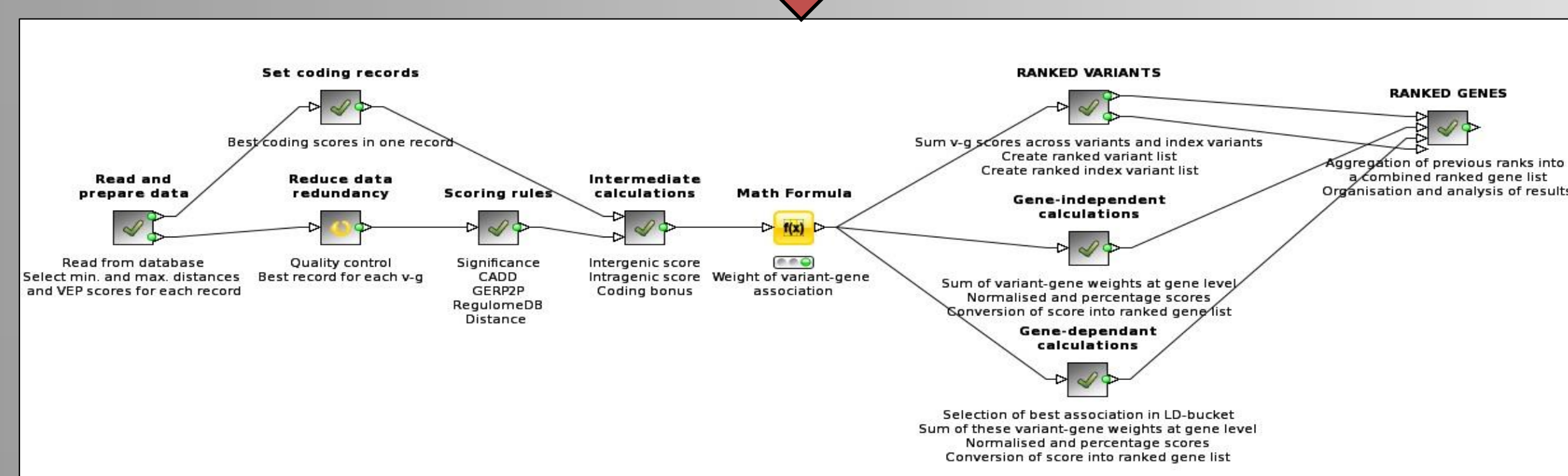


Background and Proposal

- GWAS data is difficult to interpret
- Novel scoring methodology allows ranking of variants and genes
- Used as input to an OPS workflow of investigative questions

1. In which pathways are the affected genes involved?
2. Which other genes are present in these pathways?
3. Are there any common genes amongst the pathways?
4. What is the differential expression of these genes when the indication is present/absent?
5. In which tissues does target expression occur?
6. Is the gene a valid and druggable target?
7. Are any existing drugs or small molecules known to interact with the target?

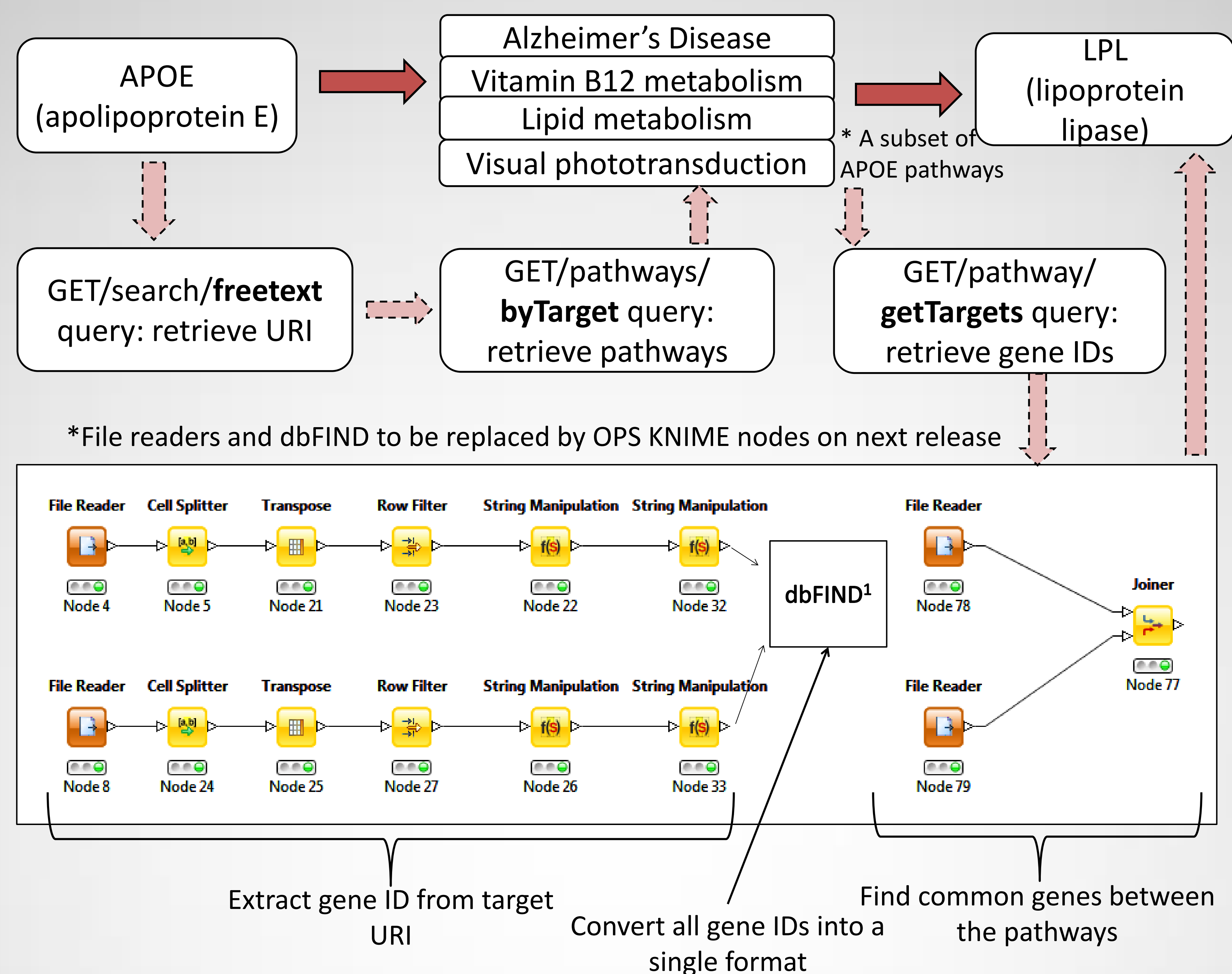
GWAS variants



List of possible targets

Preliminary analysis and literature curation – decide which targets to investigate further

Answering Q1-3: Gene-Pathway-Target



Answering Q4: A hypothetical use case for RNAseq data in Open PHACTS

Disease: Absent	Variant: Absent	Basic gene expression level (tissue):
Disease: Absent	Variant: Present	Change in gene expression (tissue):
Disease: Present	Variant: Absent	Change in gene expression (tissue):
Disease: Present	Variant: Present	Change in gene expression (tissue):

GET/target/classifications
ForCompound query: retrieve compounds that interact with a given target

Effect of compound 1 on gene expression:

Effect of compound 2 on gene expression:

Effect of compound 3 on gene expression:

Experimental validation likely required

Answering Q5-7: Knowledge mining

Existing drugs and molecules for target (LPL):

GET/target/classificationsBy
Compounds query: retrieve compounds interacting with the target

Carbonate ester, CETP inhibitor, trifluoromethyl benzenes...

GET/target query: retrieve existing drugs for target

Tyloxapol
AST-120

Tissues with target expression:

GET/tissue/byProtein query: retrieve tissues for a target

Adipose
Adrenal gland
Amygdala
Aorta
Bone marrow
Brain
Caudate nucleus
Cerebellum
Corpus callosum...

SUGGESTED ADDITION

Tissues involved in disease:

GET/tissue/byDisease query: retrieve tissues for a disease

Adrenal gland
Brain
Cerebrospinal fluid
Entorhinal cortex
Hippocampus
Muscle
Salivary gland
Subcutaneous tissue...

Benefits to life science research

- Open PHACTS facilitates recognition of indirect links between data
- KNIME workflows allow fast retrieval of information
- Quickly identify potential targets
- Gather excellent supporting evidence for each target from a variety of validated sources

References

¹Mudunuri,U., Che,A., Yi,M. and Stephens,R.M. (2009) bioDBnet: the biological database network. *Bioinformatics*, **25**, 555-556.