Introduction

This document provides an outline of the target validation service that Euretos offers. The approach leverages the Euretos Knowledge Platform, where over 200 life sciences data sources are integrated into a common data model and consulted simultaneously. By consulting all these sources at once, Euretos is able to provide unique value:

1. The analysis automatically includes a very *comprehensive body of knowledge* and is therefore *likely to include unknown information* relevant to the researcher
2. The required effort to do an analysis is significantly reduced making it *cost effective to undertake an analysis of this magnitude* and level of detail
3. The results can be delivered within the *fraction of the time* normally associated with these types of research projects.

The target assessment report assess the following type of information. Please note that the specific focus and direction of the investigation will determine which sections are emphasised:

1. Target overview, classification & functions
2. Target expression (RNA & protein)
    a. Healthy tissue
    b. Differential expression
    c. Upstream regulators inducing transcription
3. Target Interactions
    a. Target (de)activation
    b. Degradation & Secretion
    c. Other post translational modifications
4. Downstream effects
    a. Gene transcription
    b. Transport
    c. Protein complexes
    d. Pathways and molecular mechanisms
5. Co-expression analysis (optional)
    a. Co-expressed genes
    b. Co-expression regulation mechanisms
6. CONCLUSION (Phenotype analysis)
    a. Key associated phenotypes
    b. In depth analysis of selected phenotype

The phenotype analysis is the essential part of the service where the up and downstream genetic, proteomic and metabolomic interactions analysed in sections 2 to 5 are assessed in the context of the specific phenotype.

1 - Target overview, classification & functions

This section described the key aspects of the target including:

Gene name	[Target Name]
Synonyms (selection)	Selected synonyms of the target
Protein names (selection)	Protein(s) coded by the target
Classification	Protein functional classifications
Functions	Key molecular and cell functions
Expression	Main expression tissues and cell types
Pathway membership	Key associated pathways

Table 1. Overview of [Target Name]

2 - Target expression (RNA & Protein)

This section details the expression profile of the target and covers:

a. Healthy expression
b. Differential expression
c. Upstream factors inducing transcription
d. Conclusions

2a - Expression in healthy tissue (RNA & Protein)

Shows gene expression based on RNA-seq results (measured in FPKM: Fragments Per Kilobase of transcript per Million mapped reads) is analysed taking into account various sources in particular Fantom 5, GTEx, Human Protein Atlas (HPA), Illumina Body Map (IBM) and Evolution of Expression Levels (EEL).

Tissue or cell type	Fantom 5 Reference	GTEx Reference	HPA Reference	IBM Reference	EEL Reference
Tissue or cell type	FPKM	FPKM	FPKM	FPKM	FPKM

Table 2. RNA Expression of [Target Name]

Protein expression is also considered showing protein abundance measures available through the Human Proteome Map.

Tissue or cell type	Human Protein Map Reference
Tissue or cell type	Abundance

Table 3. Protein Expression of [Target Name]

Where relevant expression graphs are included in report.

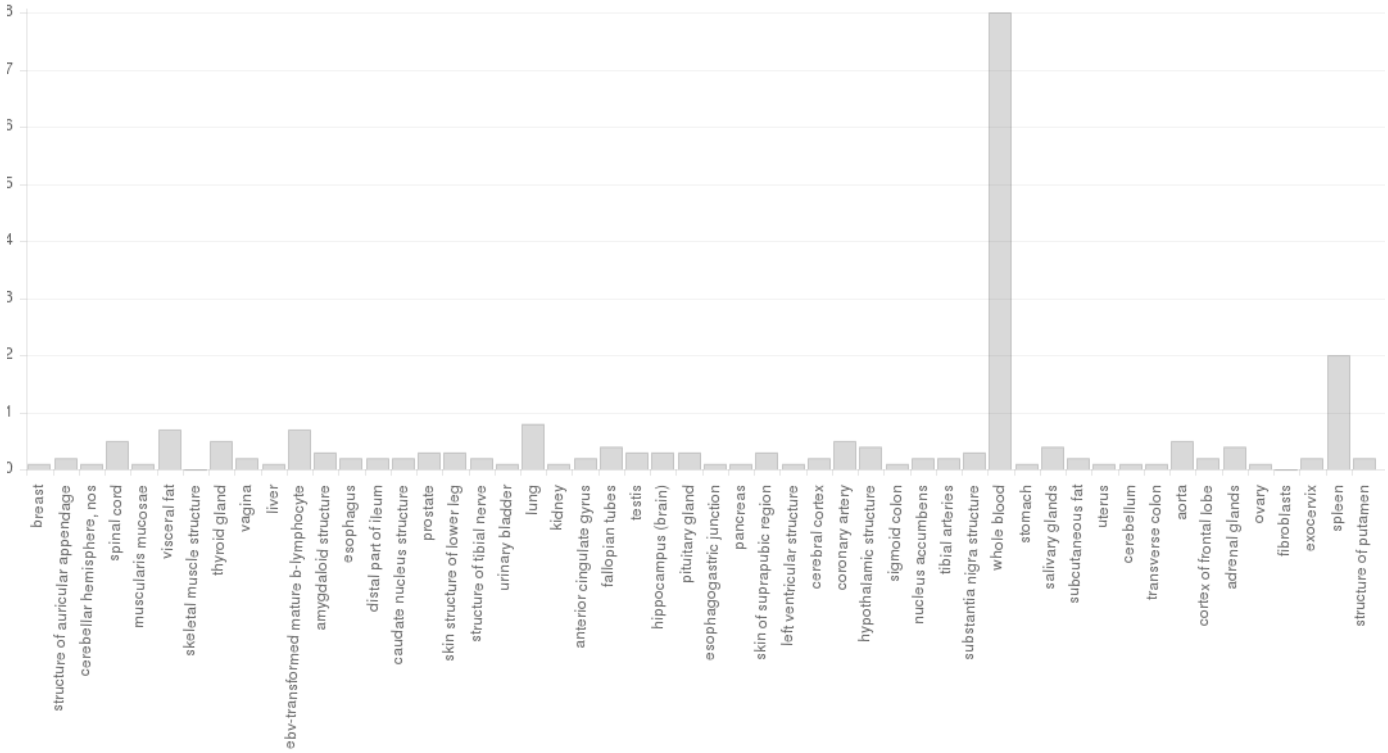


Figure 1. RNA expression profile for Target X according to GTEx. Y-axis: FPKM. X-axis: tissue type.

2b - Differential Expression

Known differential expression assays from sources such as GEO and Expression Atlas are analysed to determine the distinct conditions where the target is differentially regulated.

Tissue or cell type	Experimental conditions	Log2fold change	Assay reference
Tissue or cell location of the differential expression	Conditions such as disorder, compounds etc.	Value of log2fold change	Reference

Table 4. Differential expression for [Target Name]

2c - Upstream regulators inducing transcription
Identifies all upstream transcription regulators, which control the expression of the target:

Gene	Gene classification	Reference
Identifier of the upstream transcription regulator	Classes the transcription factor gene belongs to	Reference

Table 5. Upstream transcription regulators of [Target Name]

2d - Conclusions
Based on the RNA and protein expression profiles and the upstream transcription regulators identified above, overall conclusions are drawn from an expression perspective.

3 - Target interactions

This section details the interactions on the target and covers:

a. Activation & Deactivation through:
    i. Ligand engagement
    ii. Activation by post translational modifications
    iii. Physiological conditions
b. Degradation & Secretion
c. Other post translational modifications
d. conclusion

3a - Target (de)activation (ligand engagement or physiological conditions)
i - Ligand engagement - Identifies the ligands, physiological conditions that impact target activation and deactivation. Identifies target interactions where ligands (small molecules, antibodies & peptides) for which direct target engagement has been demonstrated are reported detailing the mechanism of action, if the experiment was in-vitro and/or in-vivo, and in which species the experiment took place.

Ligand	**Species**	**Type**	**Action**	**Type**	**Evidence**
Identifier of the ligand	Species involved	Small molecule, antibody or peptide	Agonist or Antagonist	In vitro or in vivo	Reference

Table 6. Ligand engagement with target

ii - Activation by post translational modifications - Identifies the post translational modifications that impact target activation and deactivation.

Gene	Gene classification	Type	Reference
Gene identifier	Classes the gene belongs to	Type of post translational modification	Reference

Table 7. Post translational modification factors of target

iii - Physiological conditions - Identifies known exogenous physical conditions that may cause target activation or deactivation such as temperature, pressure, etc.

3b - Degradation & Secretion
Identifies all genes that impact the degradation of the protein target through ubiquitination, proteasomal or lysosomal degradation, and secretion.

Gene	Gene classification	Type	Reference
Gene identifier	Classes the gene belongs to	Ubiquitination, lysosomal interaction, Secretion	Reference

Table 8. Degradation and secretion factors of [Target Name]

3c - Other post translational modifications
Identifies the genes that are involved in post translational modifications of the target through various mechanisms such as acetylation, glycosylation, amidation, hydroxylation and methylation.

Gene	Gene classification	Type	Reference
Gene identifier	Classes the gene belongs to	Type of post translational modification	Reference

Table 9. Post translational modification factors of [Target Name]

3d - Conclusions
Based on the interactions identified above, overall conclusions are drawn from a target interaction perspective.

4 - Downstream interactions

This section details the downstream effects of the target and covers:

a. Gene transcription
b. Transport
c. Protein complexes

d. Pathways and molecular mechanisms
e. Conclusion

4a - Gene transcription

Identifies the genes for which the target affects transcription.

Gene	Gene classification	Reference
Identifier of the transcripted gene	Classes the transcripted gene belongs to	Reference

Table 10. Genes for which the target is a transcription regulator

4b - Transport

Identifies to which extent the target impacts molecular transport of other proteins

Molecule	Process description	Reference
Identifier of the transported gene/protein	Classes the gene/protein belongs to	Reference

Table 11. Target involvement in molecular transport

4c - Protein complexes

Identifies the most important pathways and molecular mechanisms the target is involved in by analysing the protein complexes the target belongs to. Per protein complex the associated genes are identified. For these major genes an interaction analysis is done to identify potential interaction mechanisms based on Fisher Exact test:

Process	Involved protein complex genes	P-value
Pathway or molecular mechanism	Involved genes in protein complex	p-value (Fisher Exact test)

Table 12. Potential mechanisms based on protein complex association of the target

4d - Pathways and molecular mechanisms

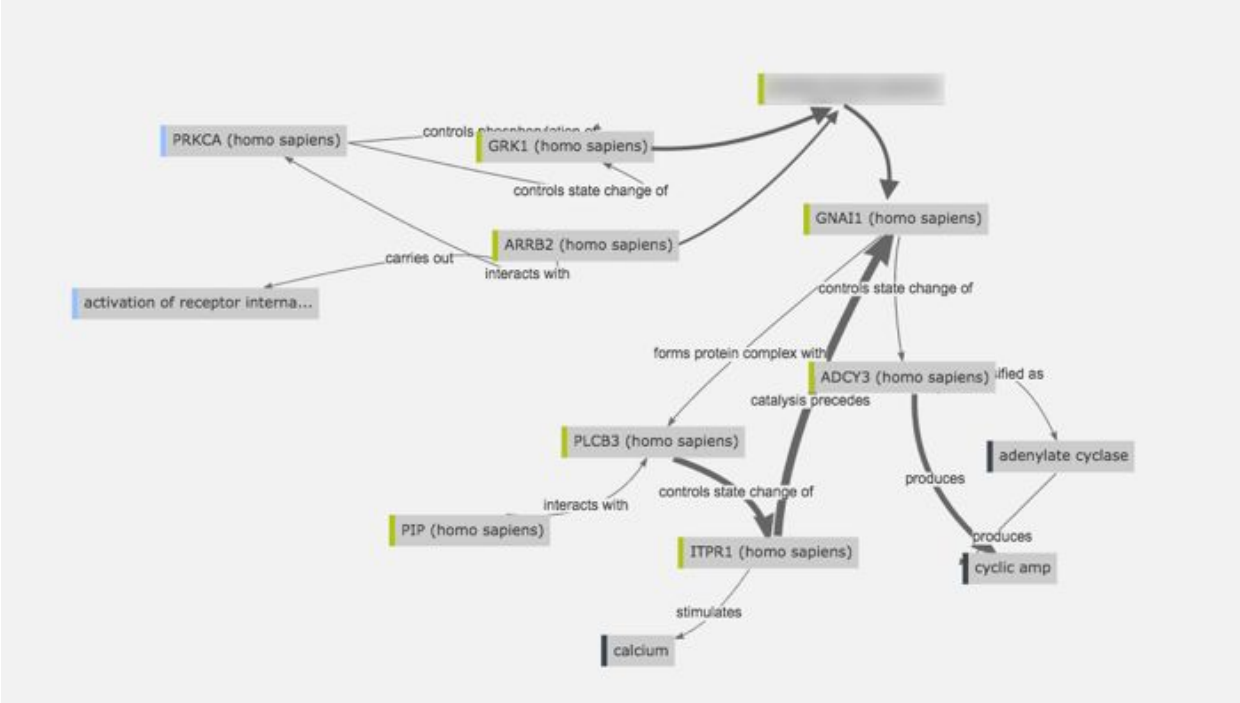
Identifies the most important pathways and molecular mechanisms the target is involved in by analysing the downstream interactions of the target using Fisher Exact test.

Process	Involved interacting genes	P-value
Pathway or molecular mechanism	Genes involved from interacting genes with target	p-value (Fisher Exact test)

Table 13. Potential gene regulating mechanisms based on co-expressed with [Target Name]

4e - Conclusion

Based on the identified downstream interactions common denominators are identified and highlighted which help to identify potential underlying molecular mechanisms. The most important interactions are visualised:



5 - Target co-expression (optional)

This section identifies co-expressed genes by analysing a large set of experiments and identifies the potential pathways and regulatory mechanisms which are likely driving the expression patterns.

5a - Co-expressed targets

Identifies all genes that are co-expressed with the target from known RNA sequencing experiments to identify highly correlated genes.

Gene (co-expressed with [Target Name])	Classifications
Gene identifier of co-expressed gene	Classifications of the co-expressed gene

Table 14. Co-expressed genes with [Target Name]

5b - Co-expression regulation mechanisms

Based on the identified co-expressed genes potential underlying regulation mechanisms, driven by tissue and cell functions, (signaling) pathways or other physiological processes are identified. For each mechanism a p-value is calculated based on Fisher Exact test.

Physiological process	Involved co-expressed genes	P-value
Potential gene regulating mechanism	Involved genes from co-expression list	p-value (Fisher Exact test)

Table 15. Potential gene regulating mechanisms based on co-expression with [Target Name]

6 - CONCLUSION (Phenotype analysis)

This section is the concluding section of the target assessment report and analyses the target from a phenotype perspective. Depending on the type of investigation either a list of phenotypes is requested (usually in the context of drug repurposing within a disease area) or a specific phenotype is assessed (usually for new drug development).

6a - Key associated phenotypes (usually in drug repurposing cases)
Based on the target profile established in the previous section a list of most highly associated phenotypes is created. For each of the key phenotypes the relevant aspects of the target profile are mentioned. For each phenotype a p-value is calculated based on Fisher Exact test.

Phenotype	Associated aspects	P-value
Potential associated phenotype	Involved aspects of target profile	p-value (Fisher Exact test)

Table 16. Key associated phenotypes based on target profile [Target Name]
For selected phenotypes a detailed analysis can then be undertaken as a separate activity.

6b - In depth analysis of selected phenotype (usually in new drug development)

For a selected phenotype, determined by the client, a detailed analysis will be undertaken. The result of this section is a detailed ‘article style’ analysis where, all relevant up and downstream genetic, proteomic and metabolomic interactions analysed in sections 2 to 5 are assessed in the context of the specific phenotype.

IN ORDER TO BEST DEMONSTRATE THE TYPE OF OUTPUT USUALLY PROVIDED IN THIS SECTION, BELOW AN ANONYMISED EXAMPLE IS PROVIDED
TARGET is highly expressed in a subset of CELL TYPE, and its activation by the neurosteroid MOLECULE or by noxious heat evokes PHENOTYPE in mice (reference). TARGET-deficient

mice failed to develop SUB-PHENOTYPE, suggesting that the channel may be sensitized in the context of inflamed tissue (reference). These findings together with the functional analyses (Pathways, Cellular function, Phenotypes, see Appendix 2) highlights a concrete link to PHENOTYPE, with the pathway described in (reference) appearing in the top 10 pathways.

The key genes in this pathway form the AP-1 complex and EGR / MAPK related proteins which are activated by Calcium influx after MOLECULE activation of TARGET, and which leads to induction and trafficking of NMDA receptors to the cell surface. Delayed-onset potentiation by MOLECULE occurs via a noncanonical, G protein–coupled, and Ca2+-dependent mechanism that is independent of NMDAR ion channel activation (reference). At the same time, this mechanism leads to the inhibition of GABA receptor (reference). Together this points to a key signaling role of TARGET in glutamate related synaptic potentiation.

In addition, Zinc ion's appear to play a key role in relation the PHENOTYPE mechanism, which are also channeled through TARGET. Zinc is an endogenous modulator of excitatory neurotransmission *in vivo* and identify a new mechanism in PHENOTYPE processing that relies on NR2A NMDA receptors (reference). So evidence points to TARGET as a signaling channel strengthening the SUB-PHENOTYPE, in CELL TYPE, with increased action potential through NMDA receptors in increased SUB-PHENOTYPE.

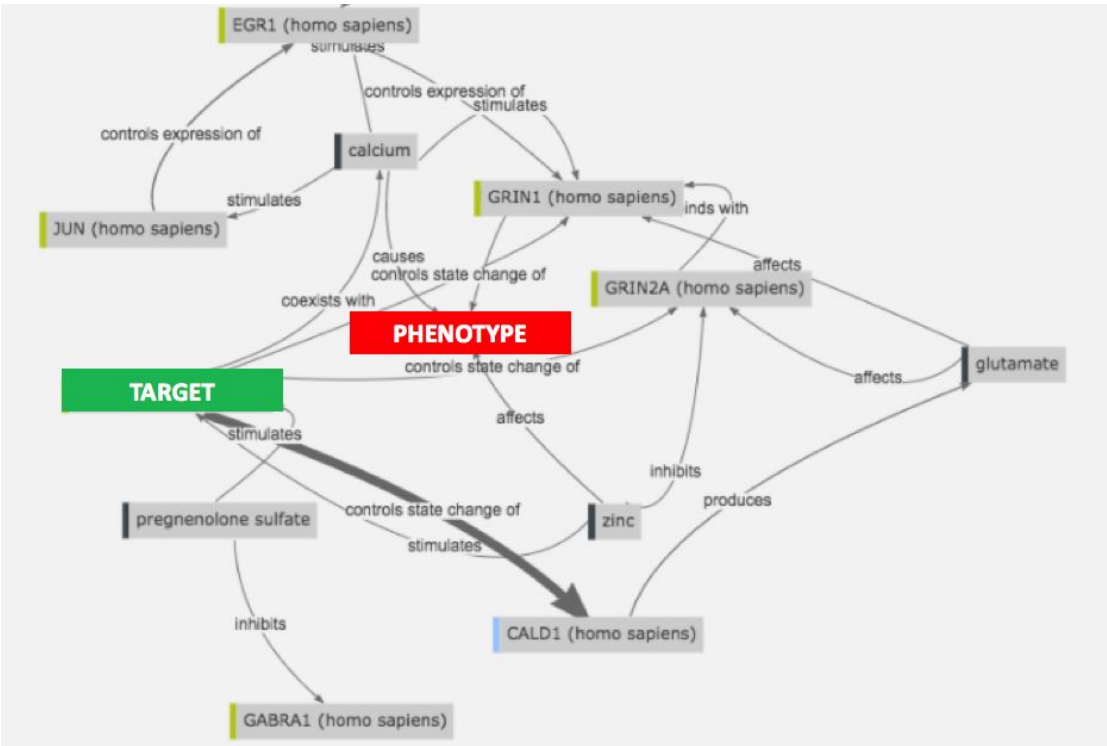


Figure x. TARGET in the context of glutamate driven potentiation via NMDA induction of PHENOTYPE

Finally, the role of MOLECULE activated TARGET in glutamatergic synapses is emphasized by it's induction of glutamate release in vivo. (reference).

However, the involvement of TARGET in MAPK signaling and pathways related to tumorigenicity and cancer might be side effect concern. TARGET promotes the growth of clear cell renal cell carcinoma (reference) and have been reported to show significantly increased expression in many other tumors (reference). Inhibition of TARGET makes TARGET an actionable target for experimental therapeutics for the treatment of clear cell renal cell carcinoma which would actually be a beneficial effect, but the exact mechanism needs further study.

MOLECULE is synthesized by ENZYME1 and ENZYME2 enzymes, which are upregulated in various tumors, such as colon cancer (4 fold in all colon cancer patients (The Cancer Genome Atlas), pancreatic cancer, breast cancer, lung cancer, thyroid- , and prostate cancers (reference). Activation of TARGET in tumor environment by MOLECULE is therefore not inconceivable and presents a potential opportunity for cancer related PHENOTYPE.