

Guidelines for DNA-binding transcription factor annotation in eukaryotes

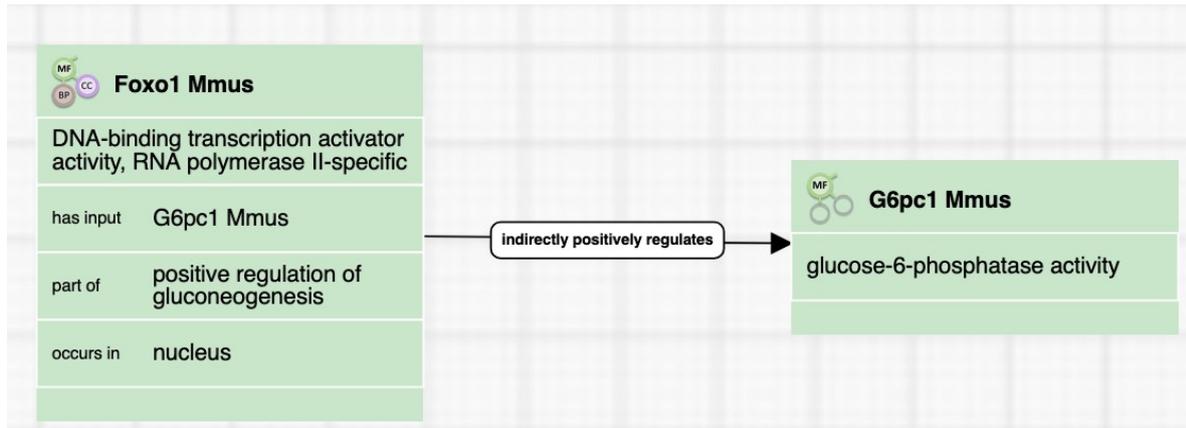
Pathway Editor

DNA-binding transcription factor activity - Single transcription target

The activity unit for a eukaryotic DNA-binding transcription factor is:

- **MF:** 'enables' a child of DNA binding transcription factor activity, RNA polymerase, II-specific ([GO:0000981](#)):
 - DNA-binding transcription activator activity, RNA polymerase, II-specific ([GO:0001228](#))
 - DNA-binding transcription repressor activity, RNA polymerase II-specific ([GO:0001227](#))
- **Context:**
 - The relation between the DNA-binding transcription factor activity and the gene it regulates is 'has input'
 - **BP:** 'part of' **regulation of the BP** in which the target participates (if known).
 - **CC:** 'occurs in' nucleus ([GO:0005634](#))
 - The causal relation between the transcription factor activity and the activity of its target gene is: 'indirectly positively regulates' or 'indirectly negatively regulates', since there are many steps between the activation of transcription and the activity of the target protein, including the production of a messenger RNA that is translated into a protein, i. e the regulator does not directly interact with the protein it regulates.

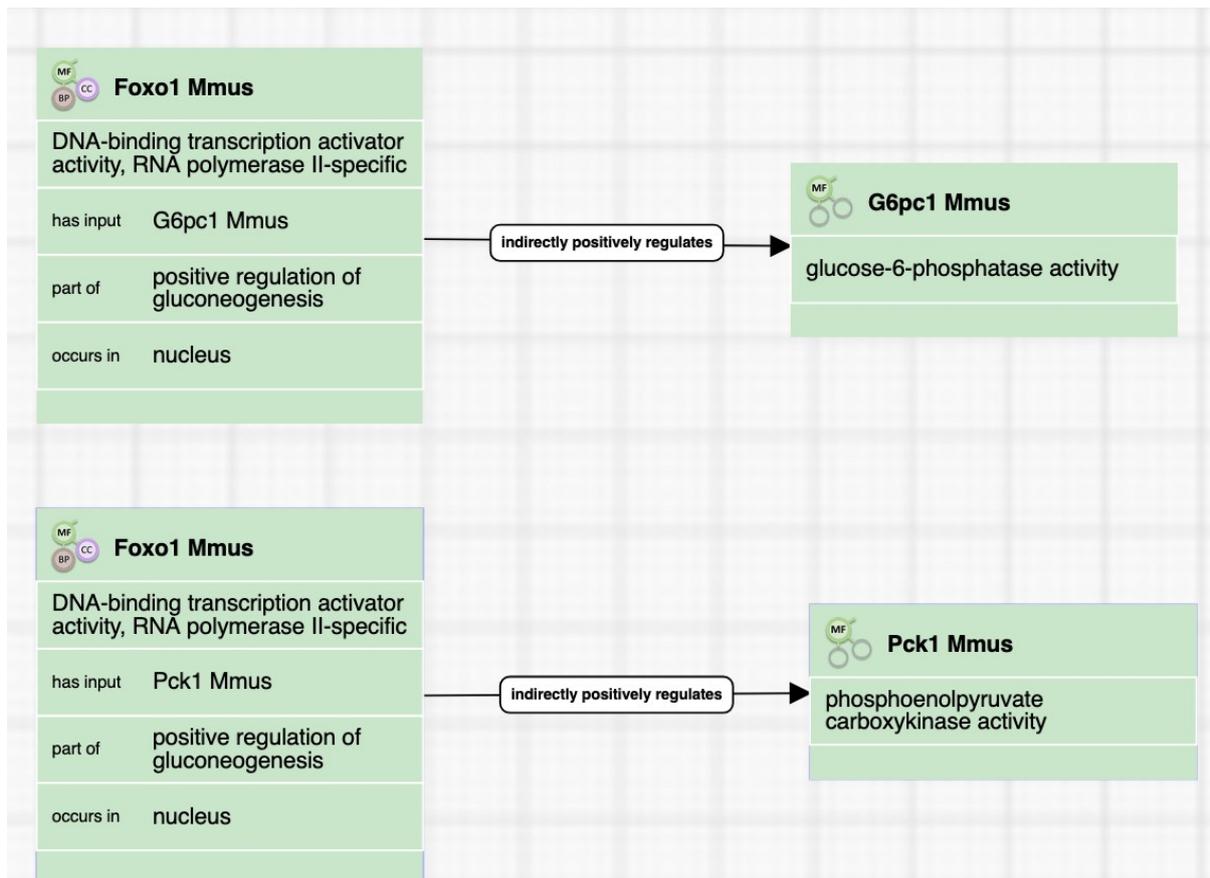
Example single target: [FOXO3 regulation of G6PC1](#)



DNA-binding transcription factor activity - Multiple transcription targets

In cases where transcription factor regulates multiple target genes, a separate activity unit is captured for each transcriptional target.

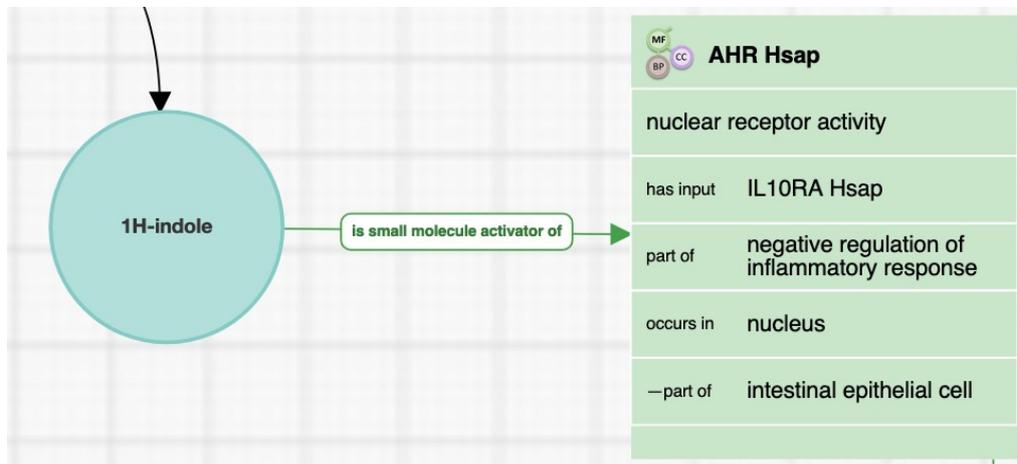
Example multiple targets: [FOXO3 regulation of G6PC1 and Pck1](#)



Nuclear receptors and ligand-activated transcription factors

- Nuclear receptors are positively regulated by a ligand, usually a small molecule (ChEBI).
- The activity unit for a nuclear receptor is:
 - **MF**: nuclear receptor activity ([GO:0004879](https://www.ebi.ac.uk/ontology/ontologies/go/terms/GO:0004879)) (a child of transcription factor activity)
 - **Context**: the causal relation between the small molecule and the nuclear receptor is 'is small activator of'.
 - Other data are captured the same way as for [other transcription factors](#).

Example: [Model for nuclear receptor annotation](#)



Form Editor

DNA-binding transcription factor activity

- **MF:** 'enables' a child of DNA binding transcription factor activity, RNA polymerase, II-specific ([GO:0000981](#)):
 - DNA-binding transcription activator activity, RNA polymerase, II-specific ([GO:0001228](#))
 - DNA-binding transcription repressor activity, RNA polymerase II-specific ([GO:0001227](#))
- **Context:**
 - The relation between the DNA-binding transcription factor activity and the gene it regulates is 'has input'. A single input is captured per activity unit.
 - **regulation of transcription may be 'part of' a larger BP**, specifically, regulation of the BP in which the target participates (if known).
 - **CC:** 'occurs in' nucleus ([GO:0005634](#))

Example DNA binding transcription factor activity: [FOXO3 regulation of G6PC1](#)

RELATIONSHIP	TERM	ASP	EXT	EVIDENCE	REFERENCE
	DNA-binding transcription activator activity, RNA polymerase II-specific GO:0001228	F			
	Foxo1 Mmus MGI:MGI:1890077			direct assay evidence used in manual assertion ECO:0000314	PMID:1234566
has input	G6pc1 Mmus MGI:MGI:95607		ext.	direct assay evidence used in manual assertion ECO:0000314	PMID:1234566
part of	positive regulation of transcription by RNA polymerase II GO:0045944	P		direct assay evidence used in manual assertion ECO:0000314	PMID:1234566
part of	positive regulation of glycolytic process GO:0045821	P	ext.	author statement supported by traceable reference used in manual assertion ECO:0000304	PMID:1234566
occurs in	nucleus GO:0005634	C		direct assay evidence used in manual assertion ECO:0000314	PMID:1234566

Nuclear receptors and ligand-activated transcription factors

Example: [Model for nuclear receptor annotation guidelines](#)

The annotations are the same as for DNA binding transcription factor activity, except using the more precise MF nuclear receptor activity ([GO:0004879](#)).

RELATIONSHIP	TERM	ASP	EXT	EVIDENCE	REFERENCE	W
	nuclear receptor activity GO:0004879		F			
	AHR Hsap UniProtKB:P35869			direct assay evidence used in manual assertion ECO:0000314	PMID:28602820	
has input	IL10RA Hsap UniProtKB:Q13651		ext	curator inference used in manual assertion ECO:0000305	PMID:29454749	
part of	negative regulation of inflammatory response GO:0050728		P	direct assay evidence used in manual assertion ECO:0000314	PMID:28602820	
occurs in	nucleus GO:0005634		C	direct assay evidence used in manual assertion ECO:0000314	PMID:28602820	
part of	intestinal epithelial cell CL:0002563		ext	direct assay evidence used in manual assertion ECO:0000314	PMID:29454749	

Differences between GO-CAM and standard annotation of a DNA-binding transcription factor activity

In standard annotation (captured with the Noctua Form or Protein2GO), relations between molecular functions are not captured, so there is no relation between the DNA binding transcription factor and the MF of its transcriptional target.

For nuclear receptors, the relation between the small molecule activator and the transcription factor is not captured.

Open questions

- FORM: For nuclear receptors, the relation between the small molecule activator and the transcription factor is not captured: can we add the relation in the Form?

Future features

Chromosomal coordinates of the promoter/enhancer/loop anchor binding site of a DNA-binding transcription factor will be captured as 'has input'. For for the human genome, the syntax is: hg38_chr6:12334566-12335555* if we want to capture the chromosomal region

* <https://eu.idtdna.com/pages/support/faqs/how-are-genomic-coordinates-defined>

Review information

Review date: 2023-07-20

Reviewed by: Cristina Casals, Pascale Gaudet, Patrick Masson