Gene Ontology annotation: Worked examples

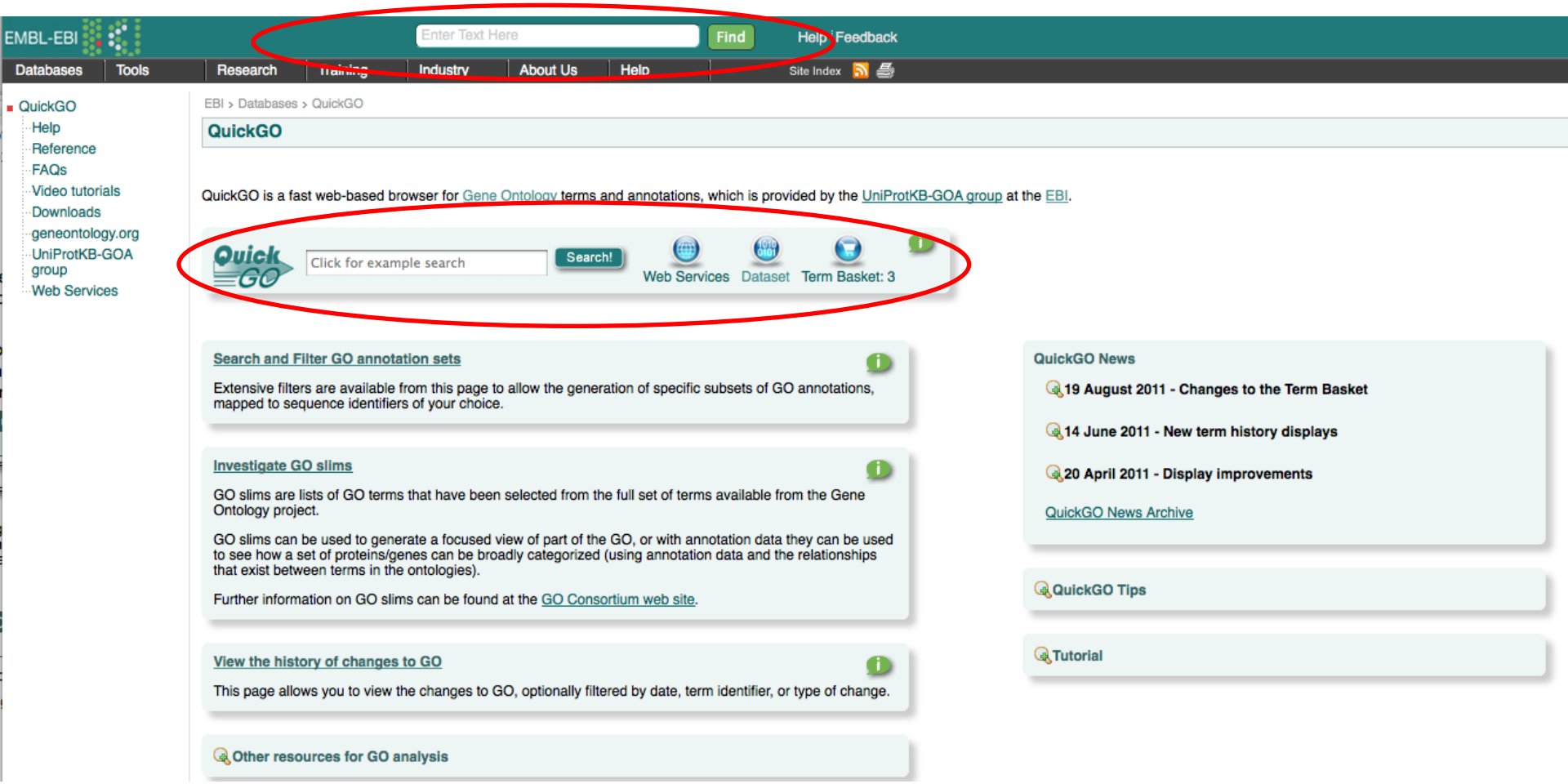
Prudence Mutowo
UniProt-GOA

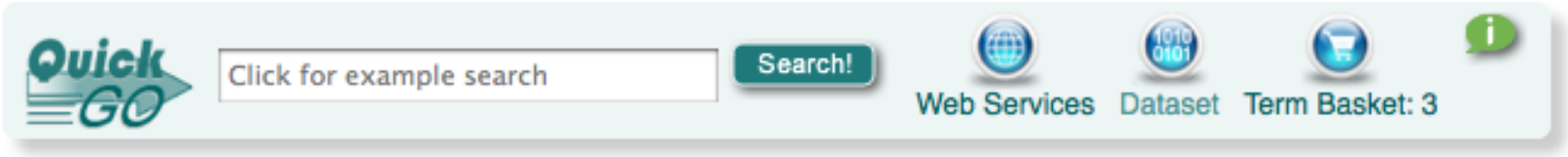
UniProt

EMBL-EBI

Talk Overview

- Intro to GO and GO terms
- Exercise
- Annotating to GO
- Accessing GO annotations
- Exercise
- Practical use of GO
- Exercise
- Precautions


EMBL-EBI
Enter Text Here
Find
Help Feedback
Databases
Tools
Research
Training
Industry
About Us
Help
Site Index
QuickGO
Help
Reference
FAQs
Video tutorials
Downloads
geneontology.org
UniProtKB-GOA group
Web Services
EBI > Databases > QuickGO
QuickGO
QuickGO is a fast web-based browser for Gene Ontology terms and annotations, which is provided by the UniProtKB-GOA group at the EBI.
Click for example search
Search!
Web Services
Dataset
Term Basket: 3
Search and Filter GO annotation sets
Extensive filters are available from this page to allow the generation of specific subsets of GO annotations, mapped to sequence identifiers of your choice.
Investigate GO slims
GO slims are lists of GO terms that have been selected from the full set of terms available from the Gene Ontology project.
GO slims can be used to generate a focused view of part of the GO, or with annotation data they can be used to see how a set of proteins/genes can be broadly categorized (using annotation data and the relationships that exist between terms in the ontologies).
Further information on GO slims can be found at the GO Consortium web site.
View the history of changes to GO
This page allows you to view the changes to GO, optionally filtered by date, term identifier, or type of change.
Other resources for GO analysis
QuickGO News
19 August 2011 - Changes to the Term Basket
14 June 2011 - New term history displays
20 April 2011 - Display improvements
QuickGO News Archive
QuickGO Tips
Tutorial



Search field
Global tool bar
Term basket

QuickGO
- Help
- Reference
- FAQs
- Video tutorials
- Downloads
- geneontology.org
- UniProtKB-GOA group
- Web Services

QuickGO

QuickGO is a fast web-based browser for Gene Ontology terms and annotations, which is provided by the UniProtKB-GOA group

Click for example search

Search!

Web Services Dataset Term Basket: 3

Search and Filter GO annotation sets

Extensive filters are available from this page to allow the generation of specific subsets of GO annotations, mapped to sequence identifiers of your choice.

Investigate GO slims

GO slims are lists of GO terms that have been selected from the full set of terms available from the Gene Ontology project.

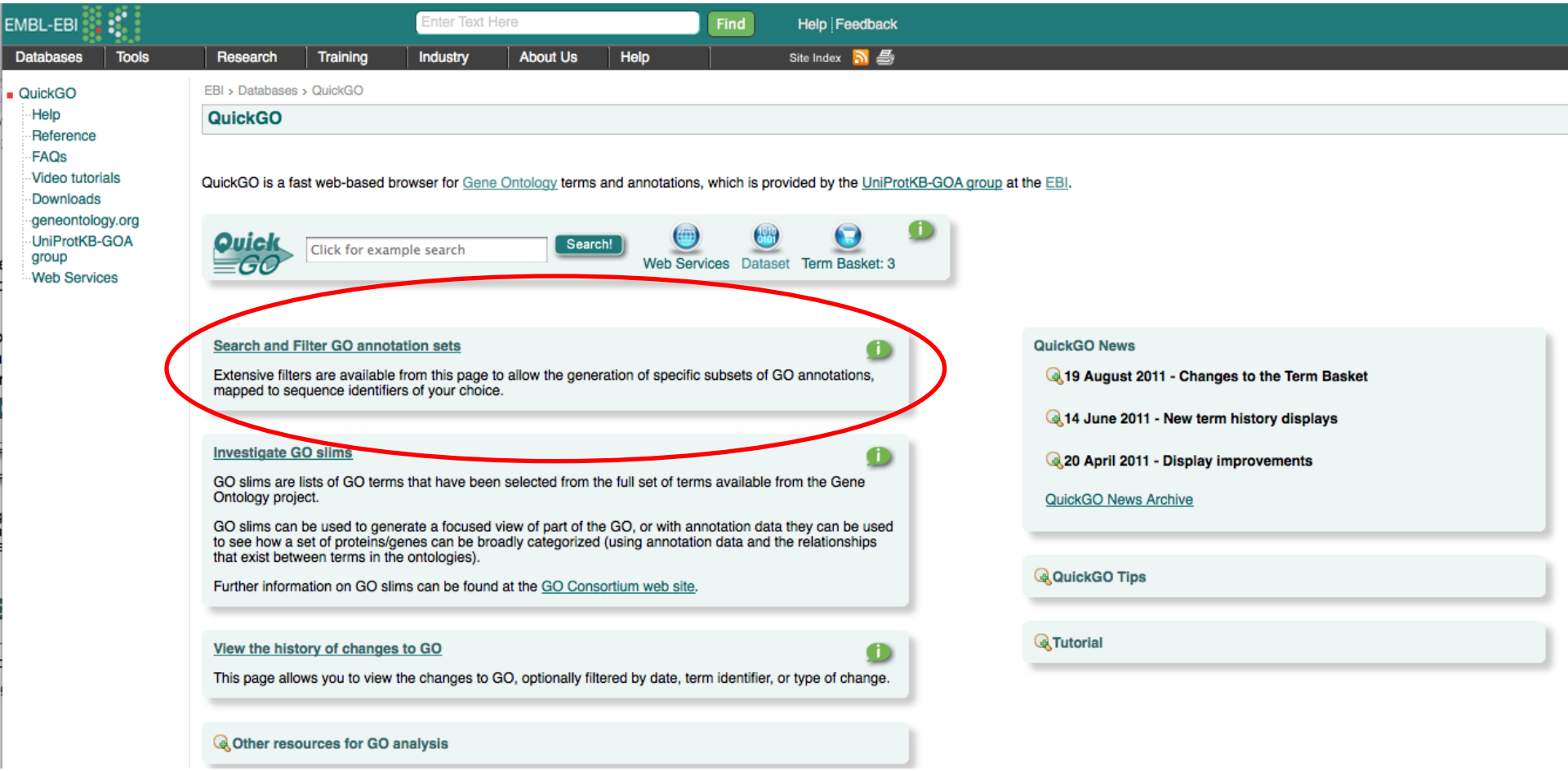
GO slims can be used to generate a focused view of part of the GO, or with annotation data they can be used to see how a set of proteins/genes can be broadly categorized (using annotation data and the relationships that exist between terms in the ontologies).

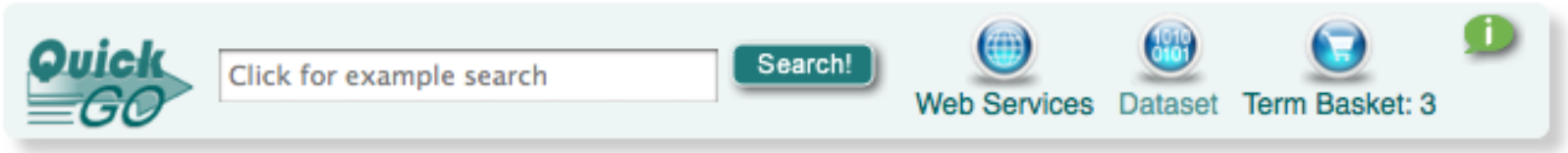
Further information on GO slims can be found at the GO Consortium web site.

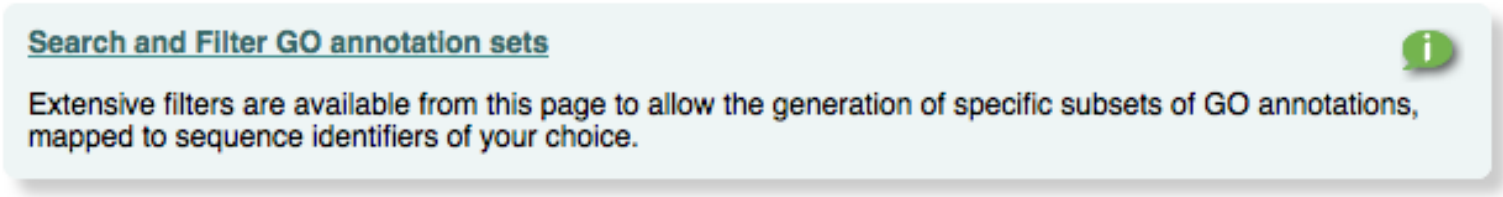
View the history of changes to GO

This page allows you to view the changes to GO, optionally filtered by date, term identifier, or type of change.

Other resources for GO analysis


EMBL-EBI
Enter Text Here
Find
Help | Feedback
Databases
Tools
Research
Training
Industry
About Us
Help
Site Index
QuickGO
Help
Reference
FAQs
Video tutorials
Downloads
geneontology.org
UniProtKB-GOA group
Web Services
EBI > Databases > QuickGO
QuickGO
QuickGO is a fast web-based browser for Gene Ontology terms and annotations, which is provided by the UniProtKB-GOA group at the EBI.
Click for example search
Search!
Web Services
Dataset
Term Basket: 3
Search and Filter GO annotation sets
Extensive filters are available from this page to allow the generation of specific subsets of GO annotations, mapped to sequence identifiers of your choice.
Investigate GO slims
GO slims are lists of GO terms that have been selected from the full set of terms available from the Gene Ontology project.
GO slims can be used to generate a focused view of part of the GO, or with annotation data they can be used to see how a set of proteins/genes can be broadly categorized (using annotation data and the relationships that exist between terms in the ontologies).
Further information on GO slims can be found at the GO Consortium web site.
View the history of changes to GO
This page allows you to view the changes to GO, optionally filtered by date, term identifier, or type of change.
Other resources for GO analysis
QuickGO News
19 August 2011 - Changes to the Term Basket
14 June 2011 - New term history displays
20 April 2011 - Display improvements
QuickGO News Archive
QuickGO Tips
Tutorial

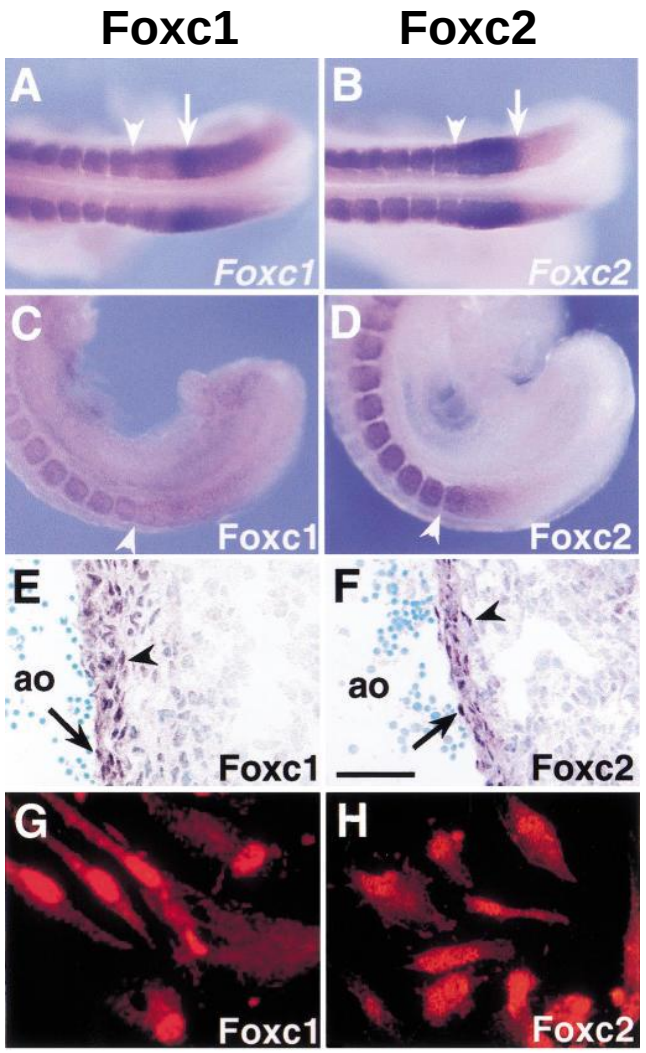




Search whole proteomes

- What information can GO capture?
- What information can GO not capture?
- How specific can the annotations be?
- How many annotations can be derived from a single paper?
- How many annotations can be made to a single gene product?
  - Repetitive annotations?
  - Conflicting annotations?

- Foxc1, Foxc2
- Look at the figure legend



Whole-mount *in situ* hybridization of 9.5 dpc embryos

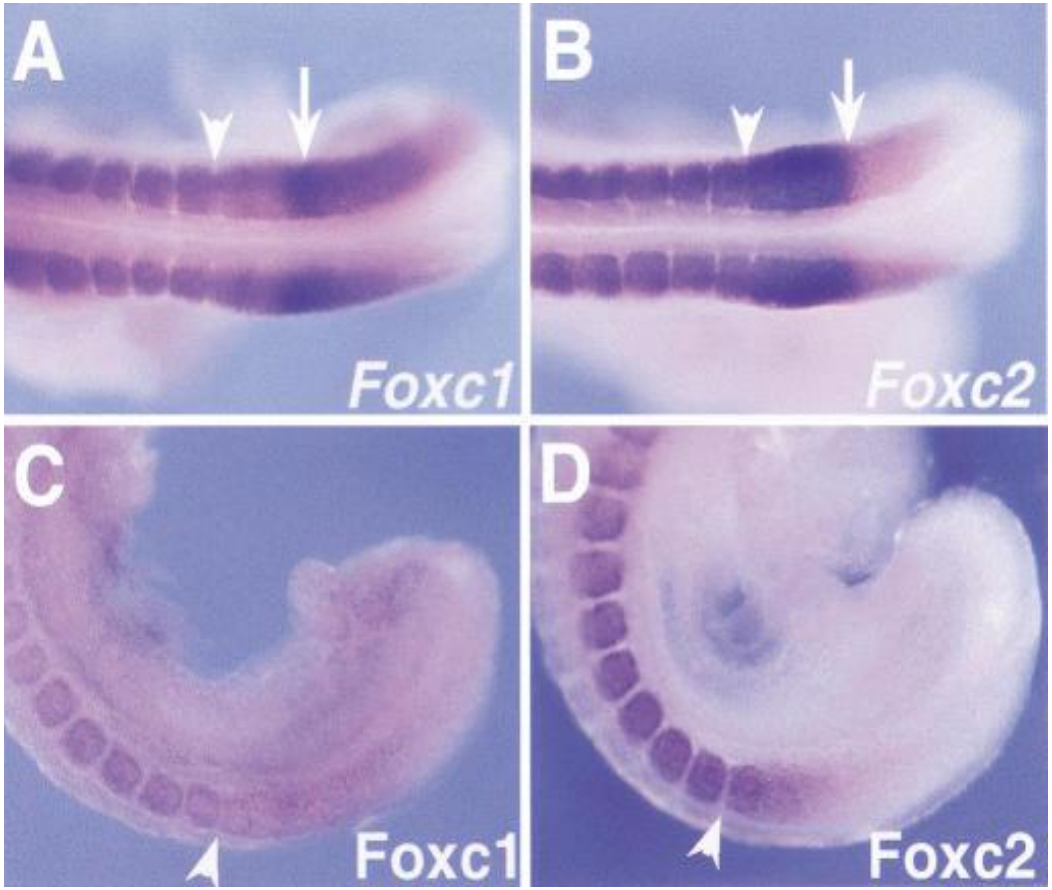
anti-digoxigenin antibodies

Transcripts in presomitic mesoderm and somites

Whole-mount immunohistochemistry of 9.5 dpc

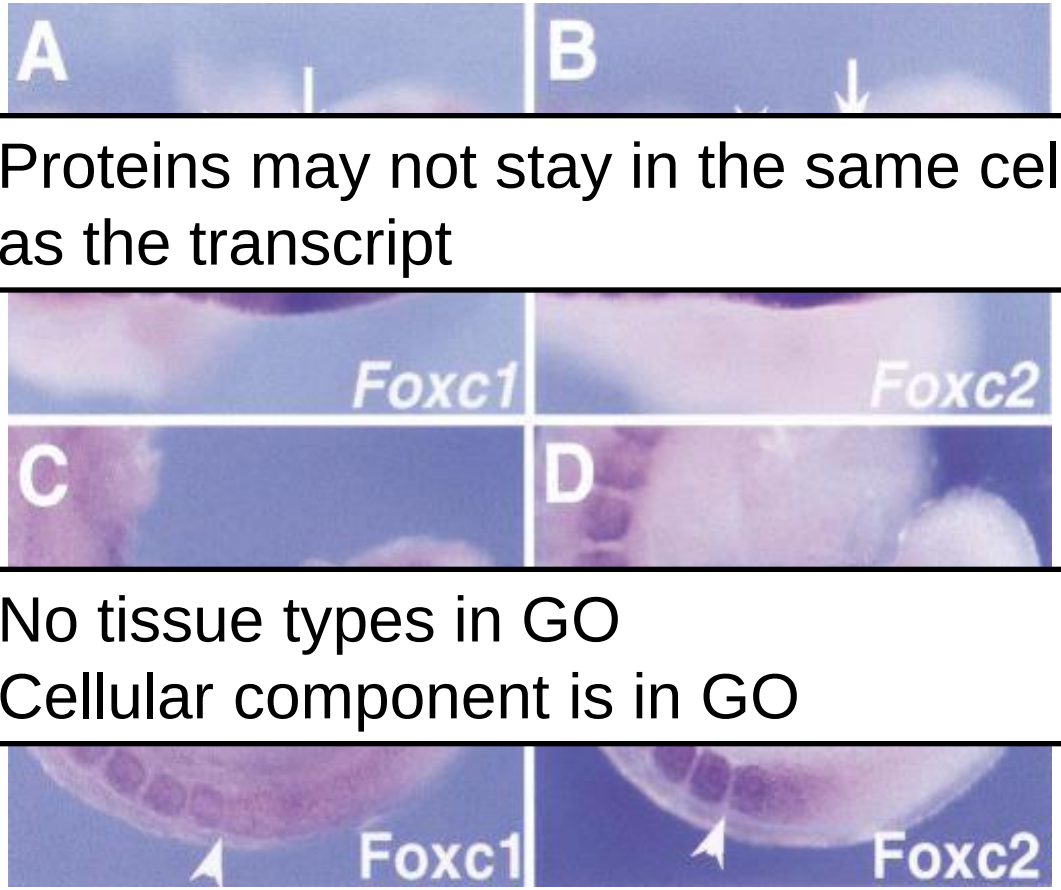
anti-Foxc1/c2 antibodies

Proteins in presomitic mesoderm



Whol mount *i situ*
hyb .5 dpc
embryc
Tran somitic
mesc m and mites

Whol mount
imm istry
of 9.5 c
Prote itic
mesc m

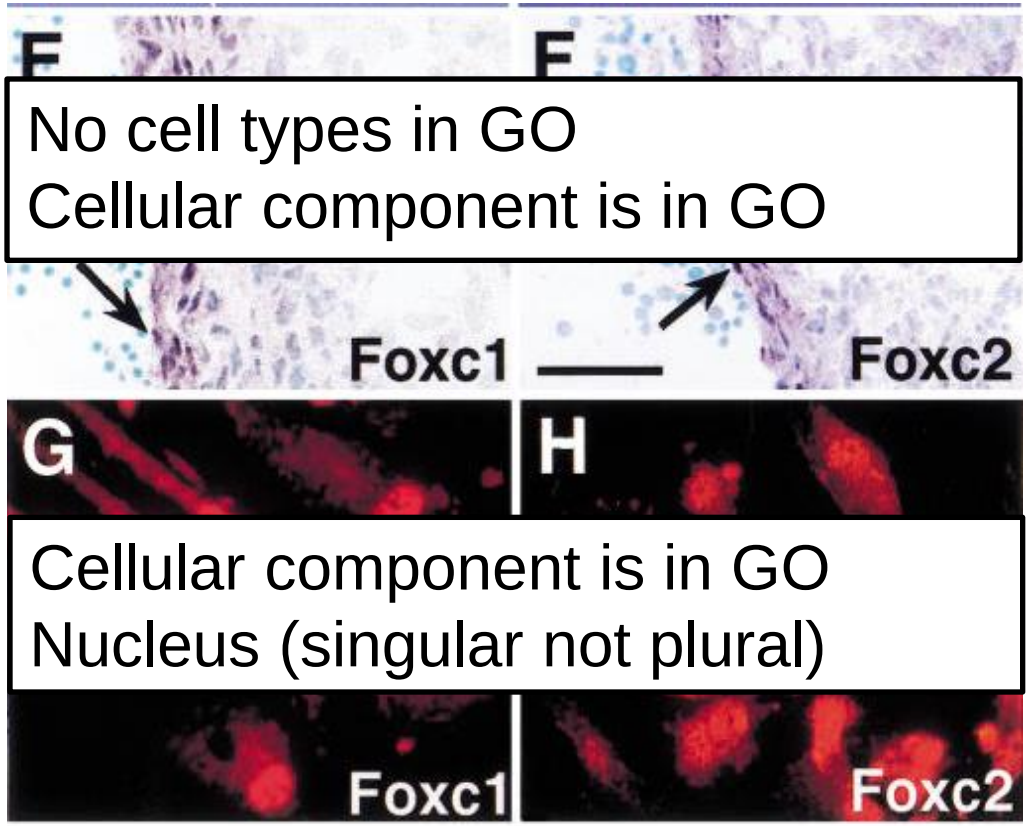


Proteins may not stay in the same cell as the transcript

No tissue types in GO
Cellular component is in GO

Antibody stained dorsal aor[illegible] 5 [illegible]mbryo)
Protei[illegible]ed in endo[illegible]mooth mus[illegible]cells

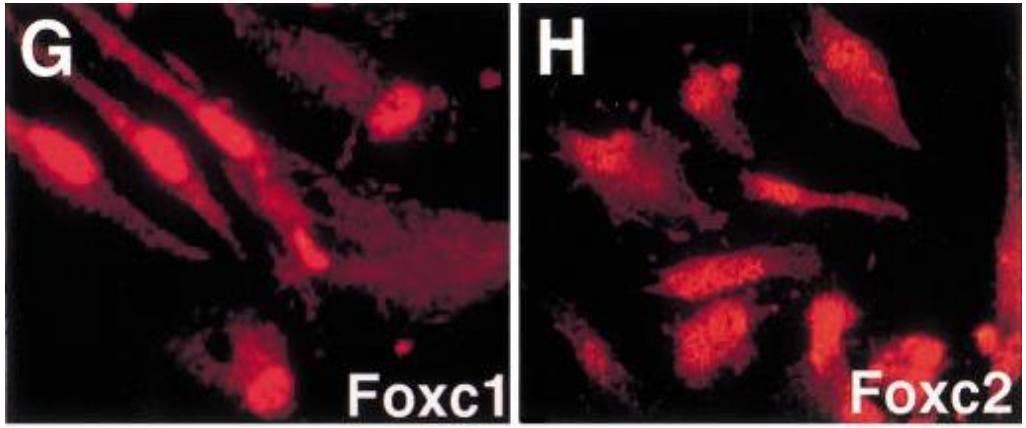
Immunostaining of human aortic smooth muscle cells
Nuclear localisation of proteins



Protein	GO term
Foxc1	GO:0005634 nucleus **(cellular component)**
Foxc2	GO:0005634 nucleus **(cellular component)**

Immunostaining of human aortic smooth muscle cells

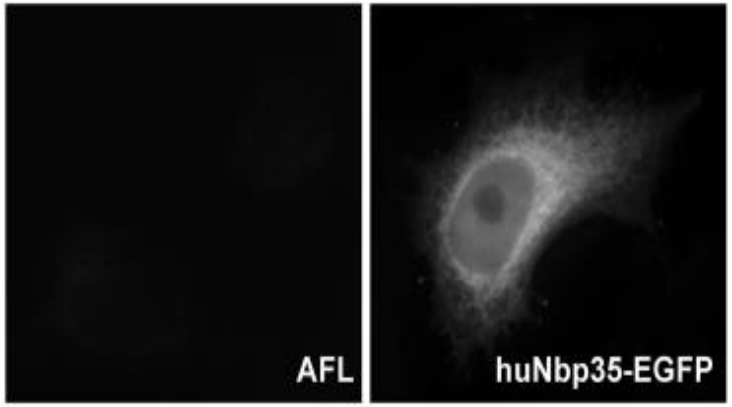
Nuclear localisation of proteins



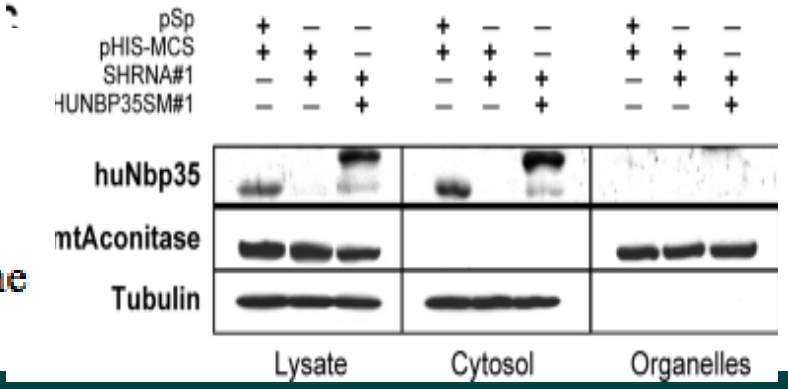
Protein	GO term	Supporting evidence
Foxc1	GO:0005634 nucleus	Figure 1G immunostaining shows nuclear location
Foxc2	GO:0005634 nucleus	Figure 1H immunostaining shows nuclear location

Figure legend:
FIG. 2.
Human Nbp35 is a cytosolic protein.

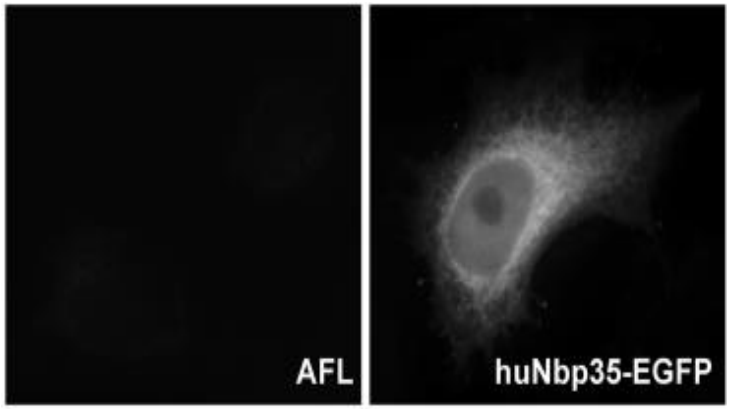
FIG. 2.
Human Nbp35 is a cytosolic protein. (A) EGFP fluorescence of a HeLa cell transiently transfected with a vector encoding a huNbp35-EGFP fusion protein (right) in comparison to the endogenous autofluorescence (AFL) of control cells (left). (B)

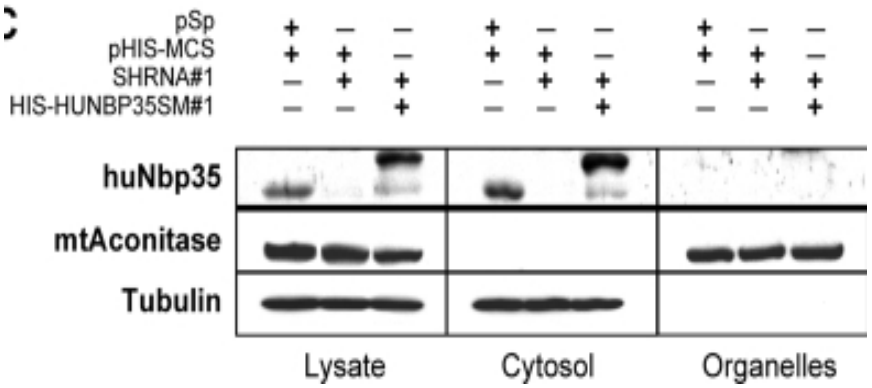


15,000 × *g*. Supernatant (cytosol) and pellet (organelles) fractions were analyzed by immunoblotting. HuNbp35 exclusively colocalizes with tubulin in the cytosolic fraction, but not with mitochondrial aconitase (mtAconitase) present in the membrane fraction.



Human Nbp35 is a cytosolic protein.

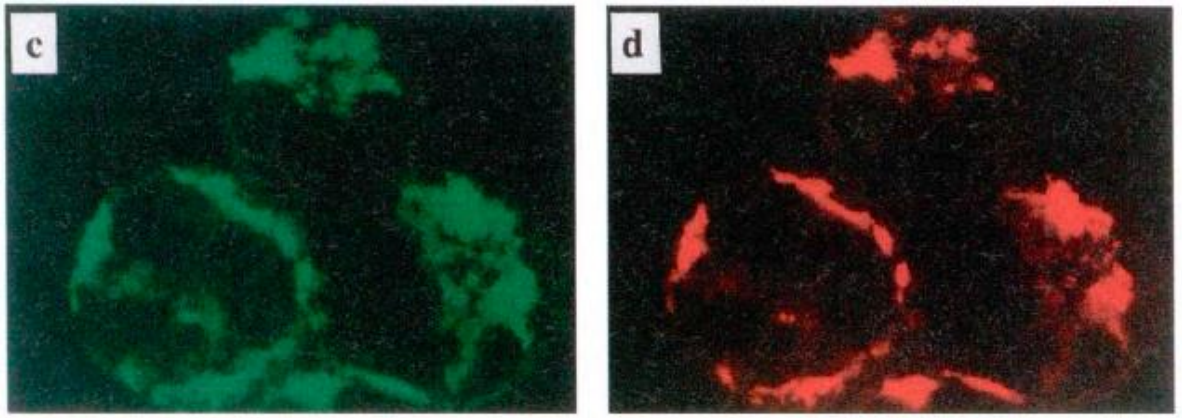




Protein	GO term	Supporting evidence
NUBP1	GO:0005829 Cytosol	Fig 2A Immunofluorescence and/or Fig 2C subcellular fractionation

Bcl-$X_{L/S}$: marker for outer mitochondrial membrane.

HAX1 annotation:



magnification view is shown in *b*. Daudi cells were double labeled with rabbit anti-HAX-1 polyclonal Ab (*c*) and goat anti-Bcl-$X_{L/S}$ polyclonal Ab (*d*) followed by secondary Abs coupled to fluorescein (green; *c*) and TRITC (red; *d*).

GO term	Supporting evidence
GO:0005739 mitochondrion Or GO:0005741 mitochondrial outer membrane	Fig 7a, b and d, co-localization of HAX1 and Bcl-$X_{L/S\ by}$ immunofluorescence confirms localisation of HAX1 to mitochondria

GO has 22 different evidence codes

IEA, Inferred from Electronic Annotation

IDA, Inferred from Direct Assay
IMP, Inferred from Mutant Phenotype
IPI, Inferred from Physical Interaction

} Experimental evidence, Methods & Results

TAS, Traceable author statement
NAS, Non-traceable author statement

} Abstract & Introduction

ISS, Inferred from Sequence Similarity

Users can quickly see how reliable the annotation is likely to be

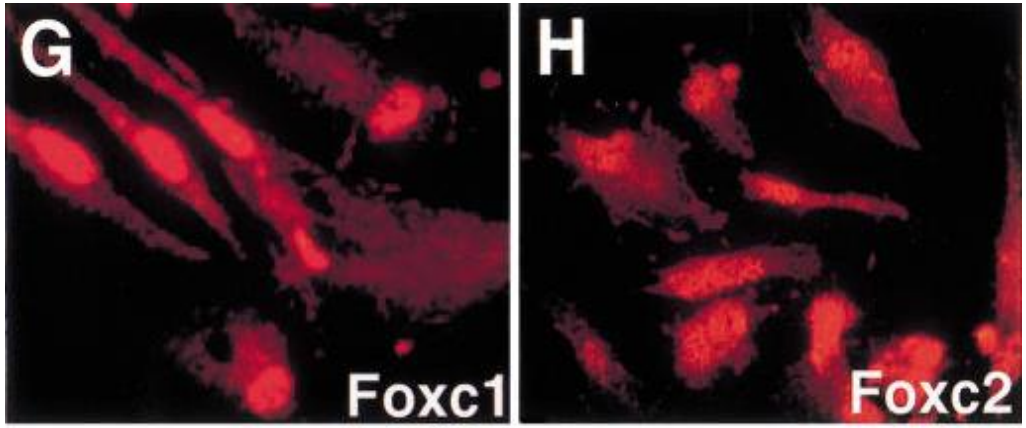
Eg. An author statement may suggest that the function of the protein has not been demonstrated in your species.

Users can filter on specific evidence types

Eg. only using manual annotations in their analysis

Immunostaining of human aortic smooth muscle cells

Nuclear localisation of proteins



Protein	GO term	Which evidence code?
Foxc1	GO:0005634 nucleus	IDA
Foxc2	GO:0005634 nucleus	IDA

IDA: inferred from Direct Assay

IMP: inferred from Mutant Phenotype

IDA Inferred from direct assay
IMP Inferred from mutant phenotype

Is the experiment demonstrating the normal function of the gene product?

eg immunofluorescence, transfection, enzyme assay

➢IDA

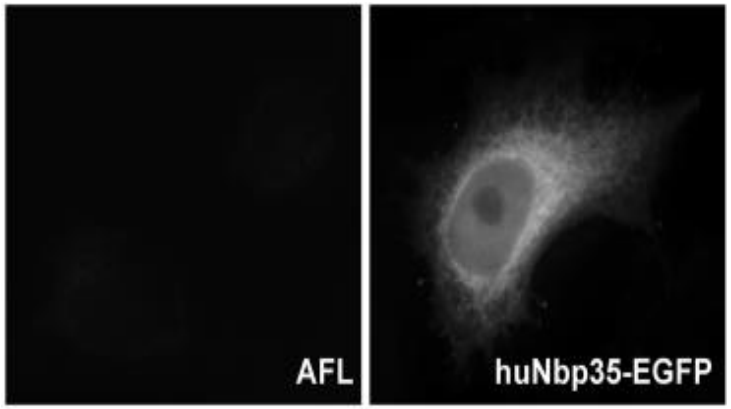
Is the abnormal expression of a gene or a mutation in a protein indicating something about the '*normal*' function?

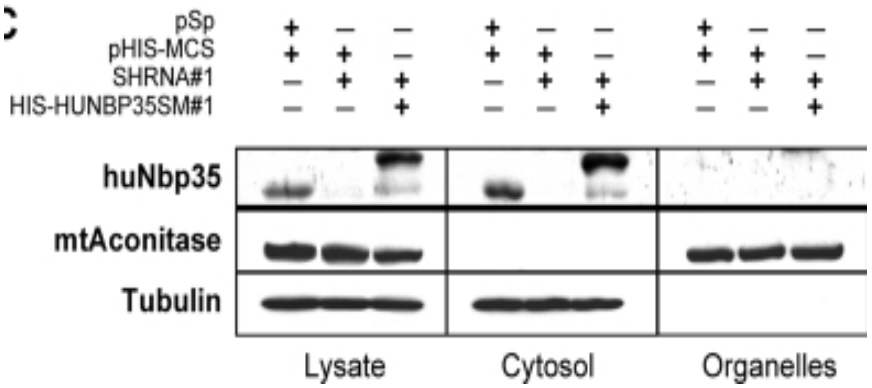
eg knockout mouse, siRNA, overexpression

➢IMP

Choose evidence codes for examples 2 and 3 – 10 minutes

Human Nbp35 is a cytosolic protein.





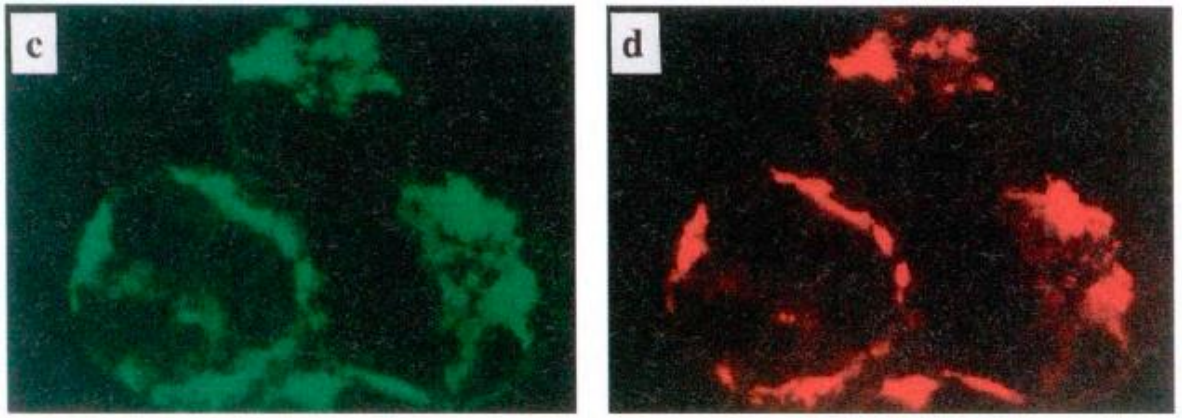
Protein	GO term	Supporting evidence
NUBP1	GO:0005829 Cytosol	Fig 2A Immunofluorescence and/or Fig 2C subcellular fractionation

- Immunofluorescence and subcellular fractionation cellular component annotations
- IDA – inferred from direct assay

Use IDA

eg	Protein	GO term	Evidence code
1	Foxc1	GO:0005634 nucleus	IDA
1	Foxc2	GO:0005634 nucleus	IDA
2	Nbp35	GO:0005829 cytosol	IDA

Bcl-$X_{L/S}$: marker for outer mitochondrial membrane.



magnification view is shown in *b*. Daudi cells were double labeled with rabbit anti-HAX-1 polyclonal Ab (*c*) and goat anti-Bcl-$X_{L/S}$ polyclonal Ab (*d*) followed by secondary Abs coupled to fluorescein (green; *c*) and TRITC (red; *d*).

HAX1 annotation:

GO term	Supporting evidence
GO:0005739 mitochondrion Or GO:0005741 mitochondrial outer membrane	Fig 7a, b and d, co-localization of HAX1 and Bcl-$X_{L/S\ by}$ immunofluorescence confirms localisation of HAX1 to mitochondria

- Immunofluorescence and subcellular fractionation cellular component annotations
- IDA – inferred from direct assay
- EXP – inferred from EXPeriment

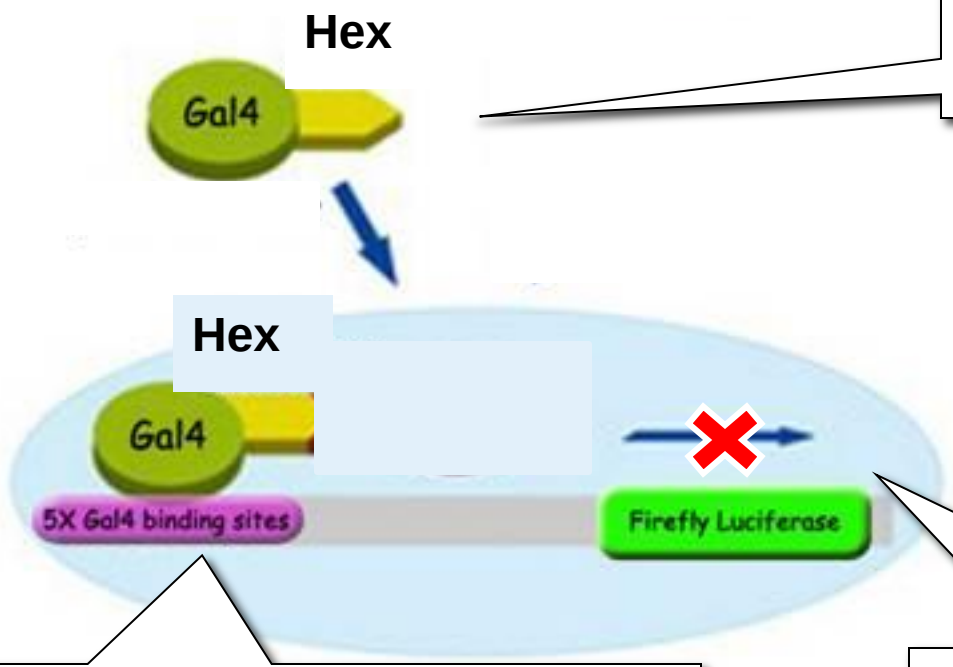
eg	Protein	GO term	Evidence code
1	Foxc1	GO:0005634 nucleus	IDA
1	Foxc2	GO:0005634 nucleus	IDA
2	Nbp35	GO:0005829 cytosol	IDA
3	HAX1	GO:0005739 mitochondrion	IDA

Table 1 Transcriptional activity of Hex

The indicated amounts of the expression plasmid were transfected into HepG2 cells with pRL-SV and reporter plasmid 5 × GAL4–GL3Control. The luciferase activity was normalized with respect to the activity of the *Renilla* luciferase. Data are expressed as the percentage of luciferase activity relative to the value of cells transfected with 0.2 μg of pSG424 and are means ± S.E.M.

Amount of pSG–Hex(6–271) plasmid (μg)	Relative luciferase activity (%)
0	100
0.001	98.7 ± 4.4
0.005	66.9 ± 3.7
0.01	58.5 ± 3.0
0.025	39.3 ± 2.7
0.05	30.1 ± 2.1
0.1	21.3 ± 1.1
0.2	18.9 ± 1.4

The more Hex transfected into the cell, the lower the luciferase activity



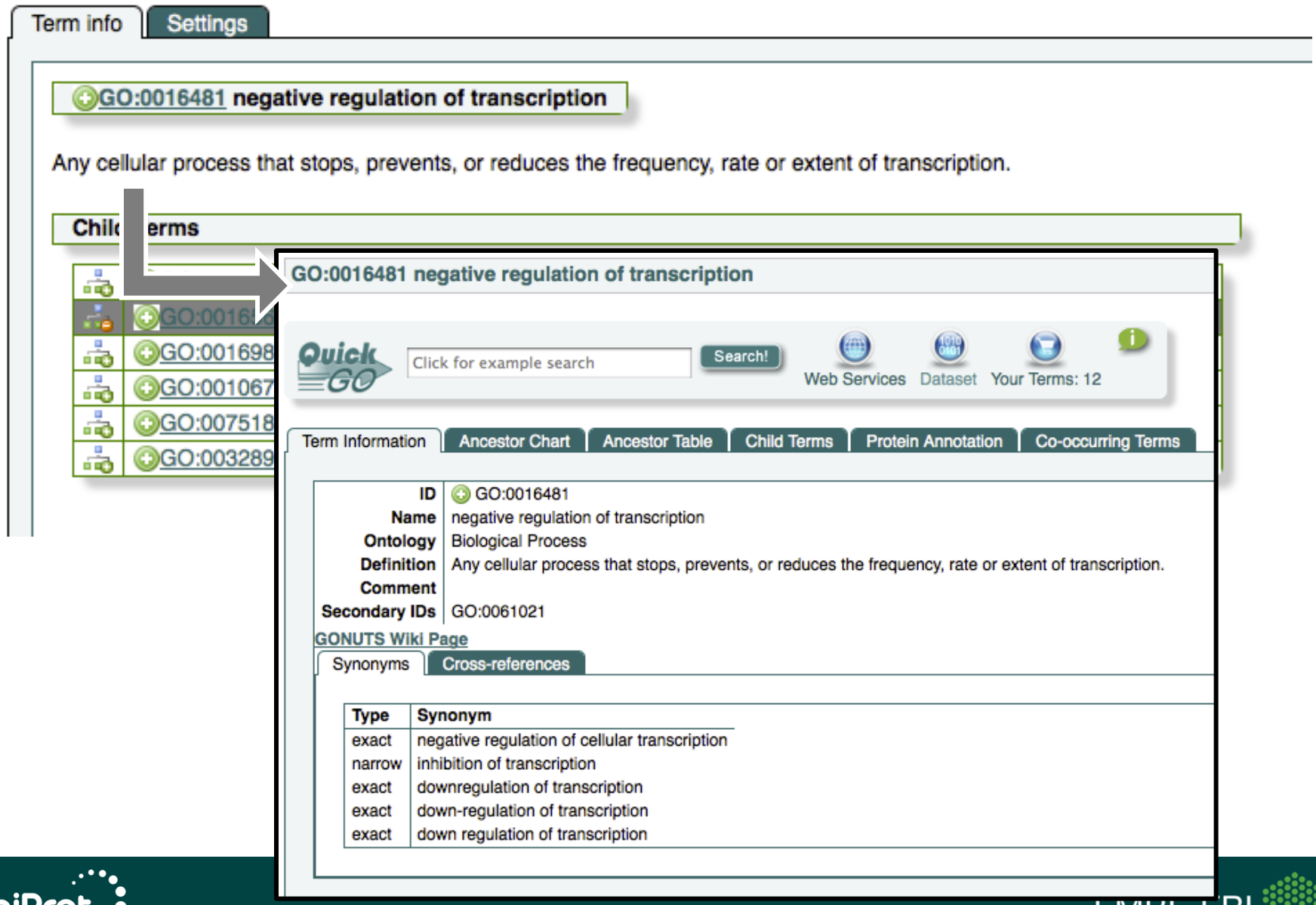
Cells synthesis GAL4-Hex fusion protein

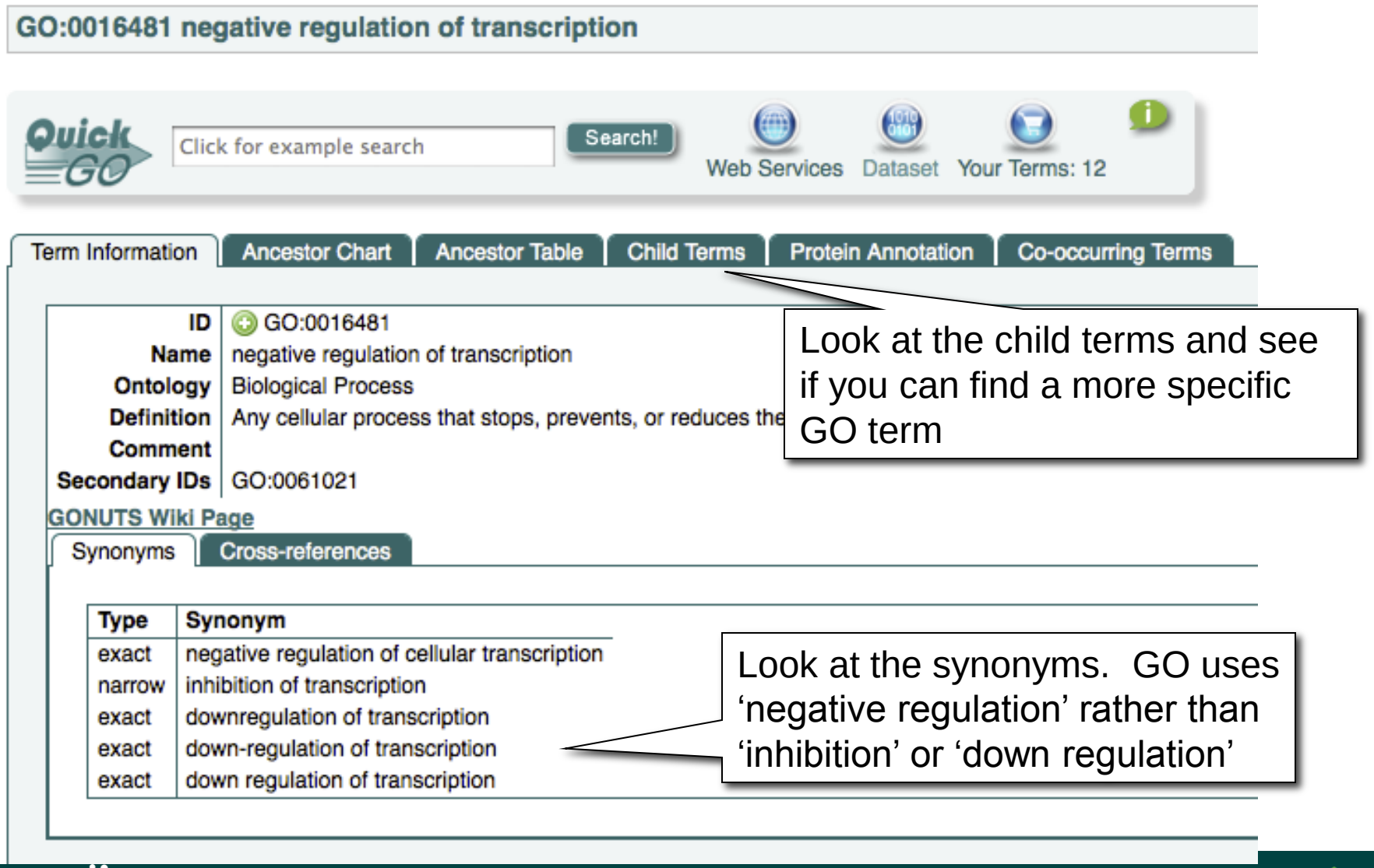
Amount of pSG–Hex(6–271) plasmid (μg)	Relative luciferase activity (%)
0	100
0.001	98.7 ± 4.4
0.005	66.9 ± 3.7
0.01	58.5 ± 3.0
0.025	39.3 ± 2.7
0.05	30.1 ± 2.1
	21.3 ± 1.1
	18.9 ± 1.4

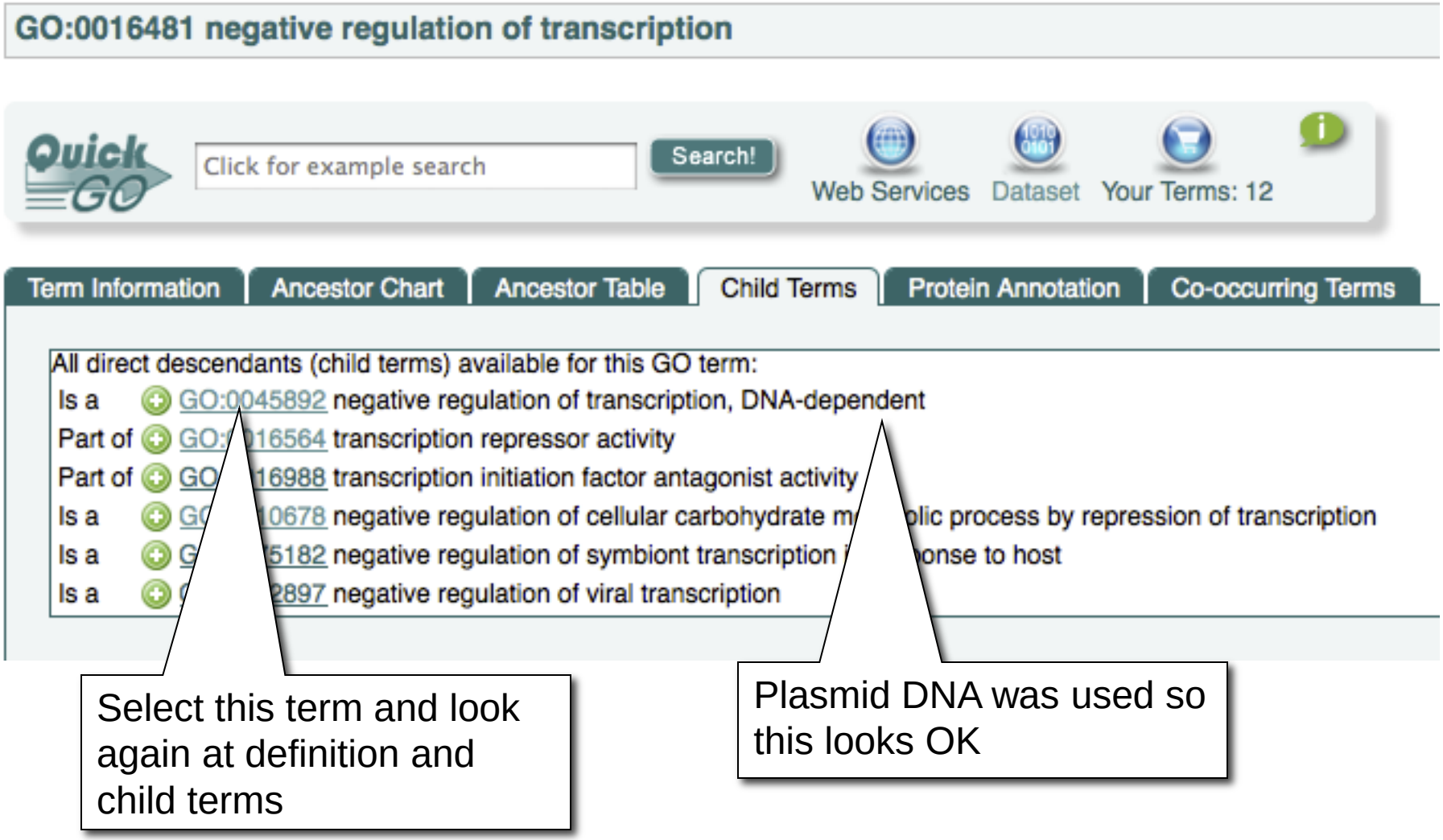
Five copies of the GAL4-binding site upstream of promoter

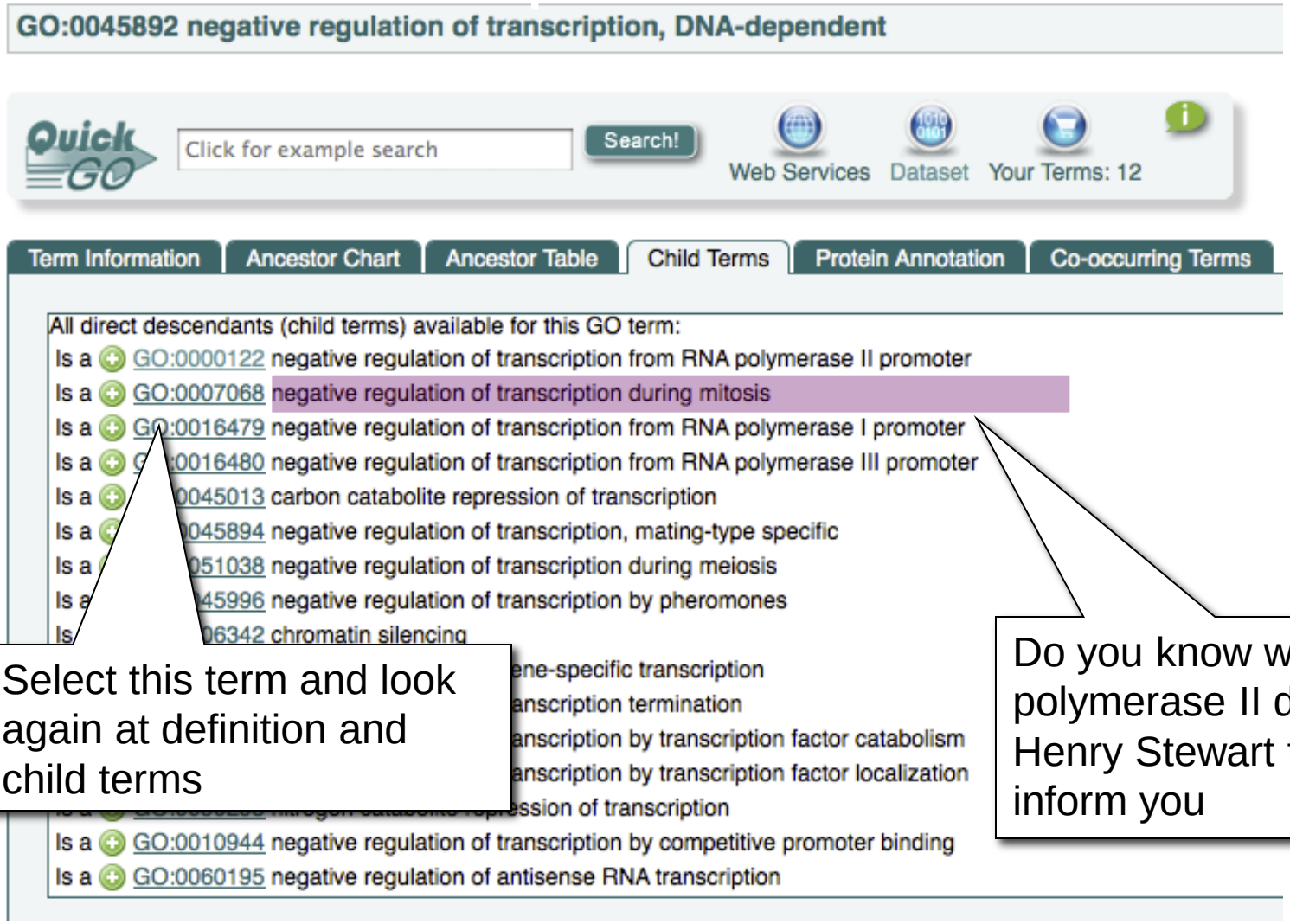
Luciferase transcription inhibited, no luciferase activity

These results show that GAL4-Hex fusion protein represses luciferase expression by binding specifically to the GAL4-binding sites, suggesting that Hex functions as a transcriptional repressor.


Term info
Settings
GO:0016481 negative regulation of transcription
Any cellular process that stops, prevents, or reduces the frequency, rate or extent of transcription.
Chil erms
GO:0016
GO:001698
GO:001067
GO:007518
GO:003289
GO:0016481 negative regulation of transcription
QuickGO
Click for example search
Search!
Web Services
Dataset
Your Terms: 12
Term Information
Ancestor Chart
Ancestor Table
Child Terms
Protein Annotation
Co-occurring Terms
ID
GO:0016481
Name
negative regulation of transcription
Ontology
Biological Process
Definition
Any cellular process that stops, prevents, or reduces the frequency, rate or extent of transcription.
Comment
Secondary IDs
GO:0061021
GONUTS Wiki Page
Synonyms
Cross-references
Type
Synonym
exact
negative regulation of cellular transcription
narrow
inhibition of transcription
exact
downregulation of transcription
exact
down-regulation of transcription
exact
down regulation of transcription


GO:0016481 negative regulation of transcription
QuickGO
Click for example search
Search!
Web Services
Dataset
Your Terms: 12
Term Information
Ancestor Chart
Ancestor Table
Child Terms
Protein Annotation
Co-occurring Terms
ID GO:0016481
Name negative regulation of transcription
Ontology Biological Process
Definition Any cellular process that stops, prevents, or reduces the
Comment
Secondary IDs GO:0061021
GONUTS Wiki Page
Synonyms
Cross-references
Type Synonym
exact negative regulation of cellular transcription
narrow inhibition of transcription
exact downregulation of transcription
exact down-regulation of transcription
exact down regulation of transcription
Look at the child terms and see if you can find a more specific GO term
Look at the synonyms. GO uses 'negative regulation' rather than 'inhibition' or 'down regulation'


GO:0016481 negative regulation of transcription
QuickGO
Click for example search
Search!
Web Services
Dataset
Your Terms: 12
Term Information
Ancestor Chart
Ancestor Table
Child Terms
Protein Annotation
Co-occurring Terms
All direct descendants (child terms) available for this GO term:
Is a GO:0045892 negative regulation of transcription, DNA-dependent
Part of GO:0016564 transcription repressor activity
Part of GO:0016988 transcription initiation factor antagonist activity
Is a GO:0010678 negative regulation of cellular carbohydrate m olic process by repression of transcription
Is a G 5182 negative regulation of symbiont transcription onse to host
Is a 2897 negative regulation of viral transcription
Select this term and look again at definition and child terms
Plasmid DNA was used so this looks OK


GO:0045892 negative regulation of transcription, DNA-dependent
QuickGO
Click for example search
Search!
Web Services
Dataset
Your Terms: 12
Term Information
Ancestor Chart
Ancestor Table
Child Terms
Protein Annotation
Co-occurring Terms
All direct descendants (child terms) available for this GO term:
Is a GO:0000122 negative regulation of transcription from RNA polymerase II promoter
Is a GO:0007068 negative regulation of transcription during mitosis
Is a GO:0016479 negative regulation of transcription from RNA polymerase I promoter
Is a GO:0016480 negative regulation of transcription from RNA polymerase III promoter
Is a GO:0045013 carbon catabolite repression of transcription
Is a GO:0045894 negative regulation of transcription, mating-type specific
Is a GO:0051038 negative regulation of transcription during meiosis
Is a GO:0045996 negative regulation of transcription by pheromones
Is a GO:0006342 chromatin silencing
Is a GO:0010944 negative regulation of transcription by competitive promoter binding
Is a GO:0060195 negative regulation of antisense RNA transcription
Select this term and look again at definition and child terms
Do you know what RNA polymerase II does? The Henry Stewart talks will inform you

GO:0000122 negative regulation of transcription from RNA polymerase II promoter

QuickGO Search! Web Services Dataset Term Basket: 3

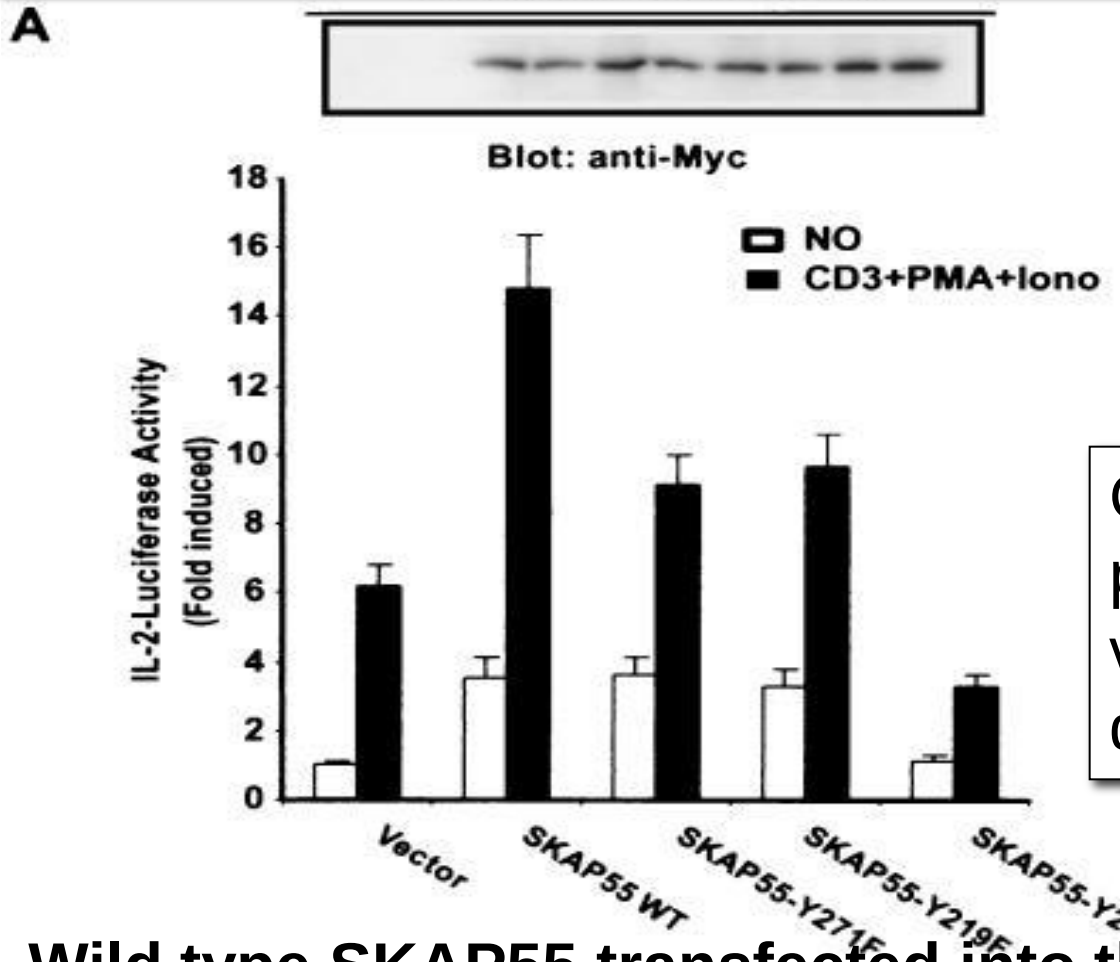
Term Information | Ancestor Chart | Child Terms | Protein Annotation | Co-occurring Terms | Change Log

This table lists all terms that are direct descendants (child terms) of GO:0000122:

Relationship To GO:0000122	Child Term	Child Term Name
Part of	GO:0001106	RNA polymerase II transcription corepressor activity
Is a	GO:0017055	negative regulation of RNA polymerase II transcriptional preinitiation complex assembly
Part of	GO:0001078	RNA polymerase II core promoter proximal region sequence-specific DNA binding transcription factor activity involved in negative regulation of transcription
Is a	GO:0007070	negative regulation of transcription from RNA polymerase II promoter during mitosis
Is a	GO:0010674	negative regulation of transcription from RNA polymerase II promoter during meiosis
Is a	GO:0071930	negative regulation of transcription involved in G1/S phase of mitotic cell cycle
Is a	GO:0072362	regulation of glycolysis by negative regulation of transcription from RNA polymerase II promoter
Is a	GO:0072368	regulation of lipid transport by negative regulation of transcription from RNA polymerase II promoter

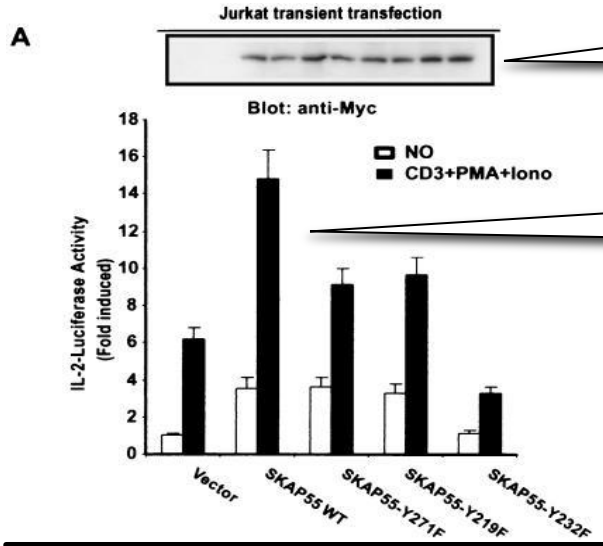
GO term	Supporting evidence	Evidence code
GO:0000122 negative regulation of transcription from RNA polymerase II promoter **(Biological Process)**	Table 1, luciferase assay shows Hex suppresses GAL4 mediated transcription	IDA
GO:0003714 transcription corepressor activity **(Molecular Function)**	Table 1, luciferase assay shows Hex suppresses GAL4 mediated transcription	IDA

What can we get from experiments using ‘mutant’ proteins?



GO does not annotate protein variants/mutation details

Wild type SKAP55 transfected into the cell, increases the luciferase activity

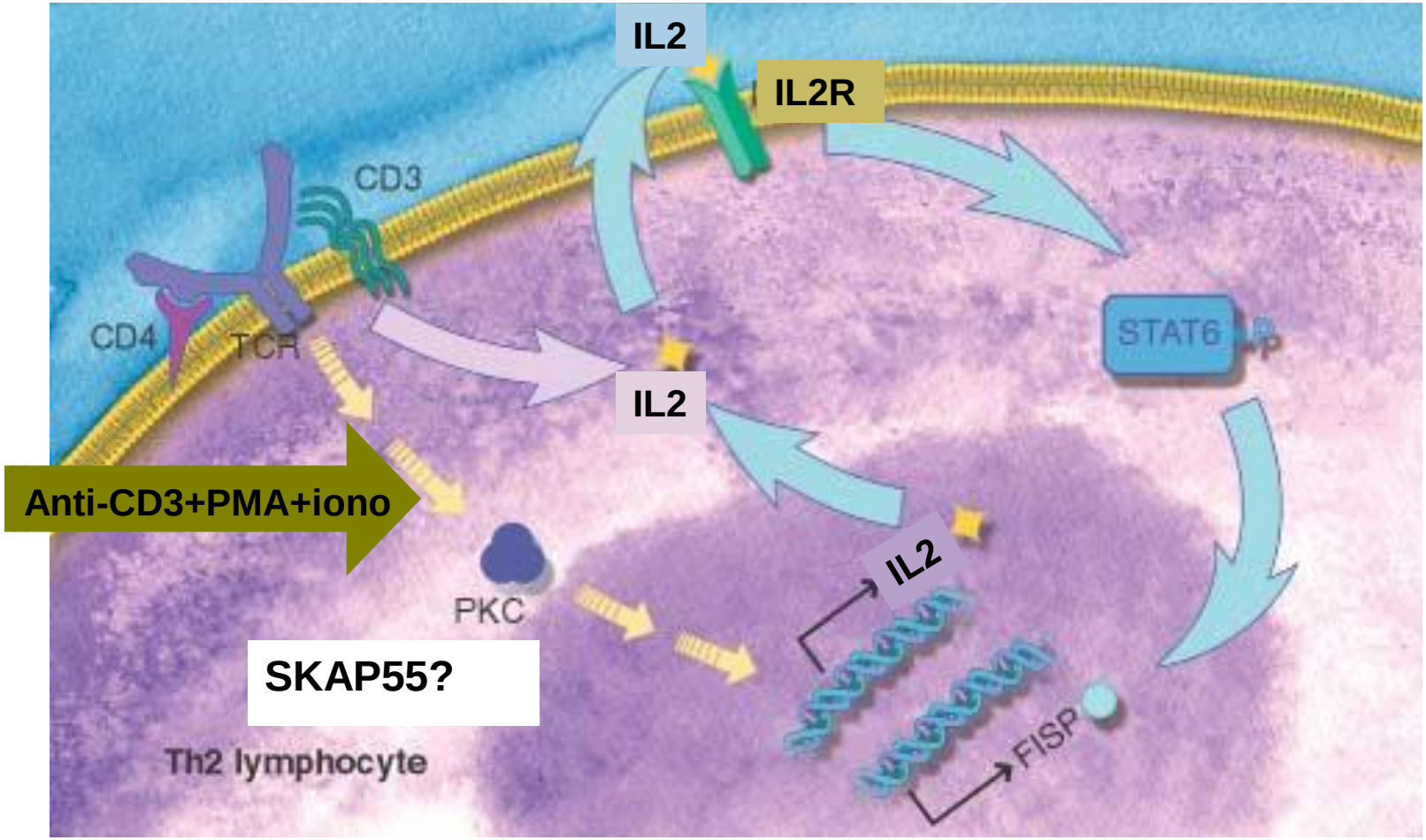


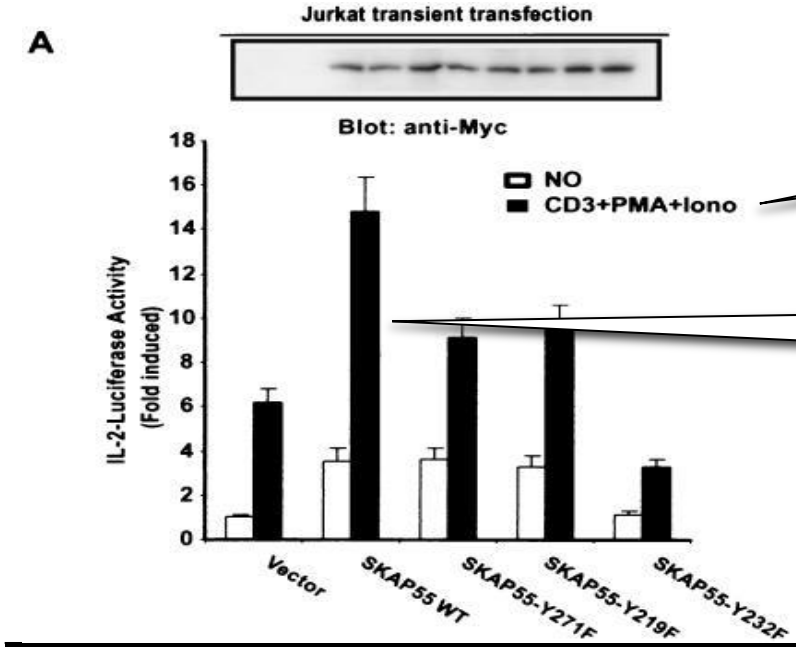
Cells synthesis Myc-SKAP55 fusion protein

IL-2-Luciferase transcription enhanced by SKAP55

This experiment is different from the previous one because the SKAP55 is not linked to a domain that binds DNA

Could SKAP55 affect a different part of the signaling pathway?


IL2
IL2R
CD3
CD4
TCR
STAT6
P
IL2
Anti-CD3+PMA+iono
PKC
IL2
SKAP55?
FISP
Th2 lymphocyte



Mimic TCR activation

SKAP55 expression increases transcription from IL2 promoter

No *Molecular Function* annotation

GO term (Biological Process)	Supporting evidence	Ev. code
GO:0045944 positive regulation of transcription from RNA polymerase II promoter	Figure 7A luciferase assay shows Wild type SKAP55 increases transcription from IL-2 promoter	IDA IMP
GO:0050852 T cell receptor signaling pathway	Figure 7A. anti-CD3 antibody plus PMA and ionomycin mimic stimulation of the T cell receptor (TCR). Over expression of SKAP55 increases the response to TCR stimulation	IDA IMP

Which protein to annotate?

Figure 1 from Kim M.L. *et al* PLoS One. 2010 Oct 15;5(10):e15371

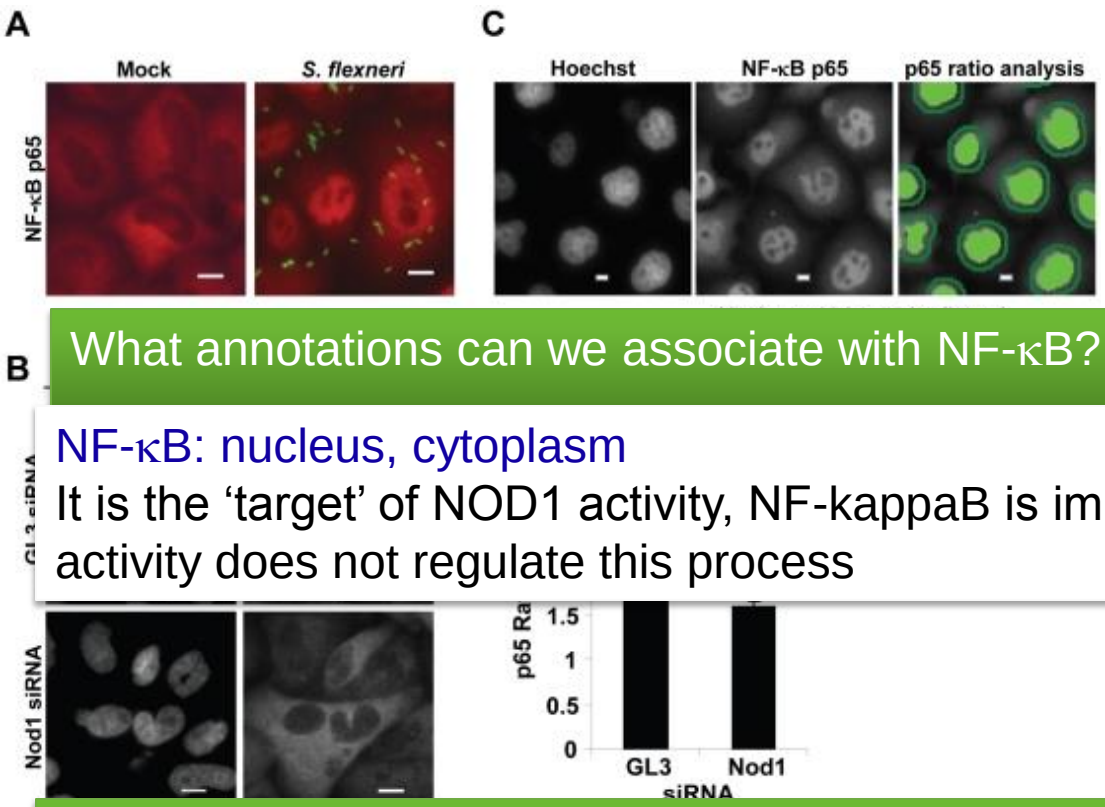


Figure A

NF-κB translocates to nucleus in response to *S. flexneri* infection.

What annotations can we associate with NF-κB?

NF-κB: nucleus, cytoplasm

It is the 'target' of NOD1 activity, NF-kappaB is imported to the nucleus, it's activity does not regulate this process

NOD1 protein, inhibits NF-κB translocation to nucleus in response to *S. flexneri*

What annotations can we associate with NOD1?

NOD1: positive regulation of NF-kappaB import into nucleus

S. flexneri infection.

Protein	GO term	Supporting evidence	Ev. code
p65 NFKB	GO:0005737 cytoplasm	Figure 1A immunoflorescence shows cytoplasmic location of NFkB in uninfected cells	IDA
p65 NFKB	GO:0005634 nucleus	Figure 1A immunoflorescence shows nuclear location of NFkB in infected cells	IDA
Nod1	GO:0042346 positive regulation of NF-kappaB import into nucleus	Figures 1B,C,D show depletion of NOD1 by RNAi, inhibits NFkB translocation into the nucleus in infected cells	IMP

Note that different evidence codes are used for a single paper