

UCL Gene Annotation group

Providing functional annotation to human genes and gene products

Ruth Lovering

Collaborative projects between UCL, Kings and EBI



RG/13/5/30112: 2013-2018

PARKINSON'S^{UK}
CHANGE ATTITUDES.
FIND A CURE.
JOIN US.

G-1307: 2014-2016

Co-grant holders

Philippa Talmud

Steve Humphries

Manuel Mayr

Rolf Apweiler

John Hardy

Biocurators

Paul Denny

Anna Melidoni

Nancy Campbell (maternity leave)

Milagros Rodríguez-López (maternity cover)



GENEONTOLOGY
Unifying Biology

1. GO annotation of cardiovascular (CV) relevant proteins
2. Capture **protein-protein** interactions (IntAct editing tool)
3. Annotation of microRNAs with experimentally validated data
 - **biological process** regulated
 - miR-33 regulates cholesterol homeostasis
 - miR-223 regulates FOXO1 expression
 - what **cells and tissues** these RNAs are expressed in
 - Manuel Mayr proteomics and miRNAs

- Only 1 hour
- Summary of different groups miR GO data
- Wiki page lists several areas needed for discussion
- Start with microRNA/miRNA annotation
 - Questionnaire
 - Annotation of differentially expressed miRNAs
 - Method limitations
 - Pre-miRs
 - Anti-miRs
 - Standardised analysis of differential expression
 - Field currently establishing consistent controls
 - Results very dependent on controls

Summary of MOD miRNA annotations [\[edit\]](#)

MOD/database	Brief summary of progress	Attending meeting
TAIR	Have been creating Arabidopsis miRNAs annotations for years. 199 miRNA genes so far, 228 annotations to 'gene silencing by miRNA' or kids, for the 73 experimental annotations, mostly IDA, IMP, IGI but also a few IEP	Tanya/Donghui
Wormbase	Mostly focused on miRNA annotations for the greatest hits of C. elegans miRNAs, namely let-7, lin-4, lsy-6, etc. Our BP annotations to the miRNAs have included the cellular level processes they control and also the organismal level processes that they, in turn, regulate, e.g. heterochronic development, neuronal fate specification, etc. We also have some MF annotations for their mRNA binding activities. Users have recently ask us about including these types of annotations in GO. We haven't yet considered making IEP annotations for miRNAs but I'm glad you raised that issue again.	Kimberley
GOA/UCL	Waiting for human microRNA identifiers, due Spring 2014	Ruth/Rachael
RGD	We currently take ISO annotations for rat orthologs from mouse and human miRNA-related genes. Our curation prioritization is disease-centric so as miRNAs come up as related to the disease category we are covering we will curate them. We also may do a focused effort to curate miRNAs and their targets at some point.	Jennifer
MGI	Have been creating mouse miRNAs annotations for years	
Dicty	not a priority at the moment	
ZFIN	currently has 124 microRNA genes annotated with 260 non-ND evidence annotations using 69 distinct GO terms.	Doug
PomBase	Pombe also has lots of antisense RNAs which do not get cleaved into siRNA but are presumed to regulate the gene they are antisense to	Val

Current stats: Number of annotations to "gene silencing by miRNA" or child terms;
UniProtKB(295)
TAIR(228)
ZFIN(144)
RGD(96)
MGI(84)
FB(76)
AspGD(15)
CGD(1)
dictyBase(1)

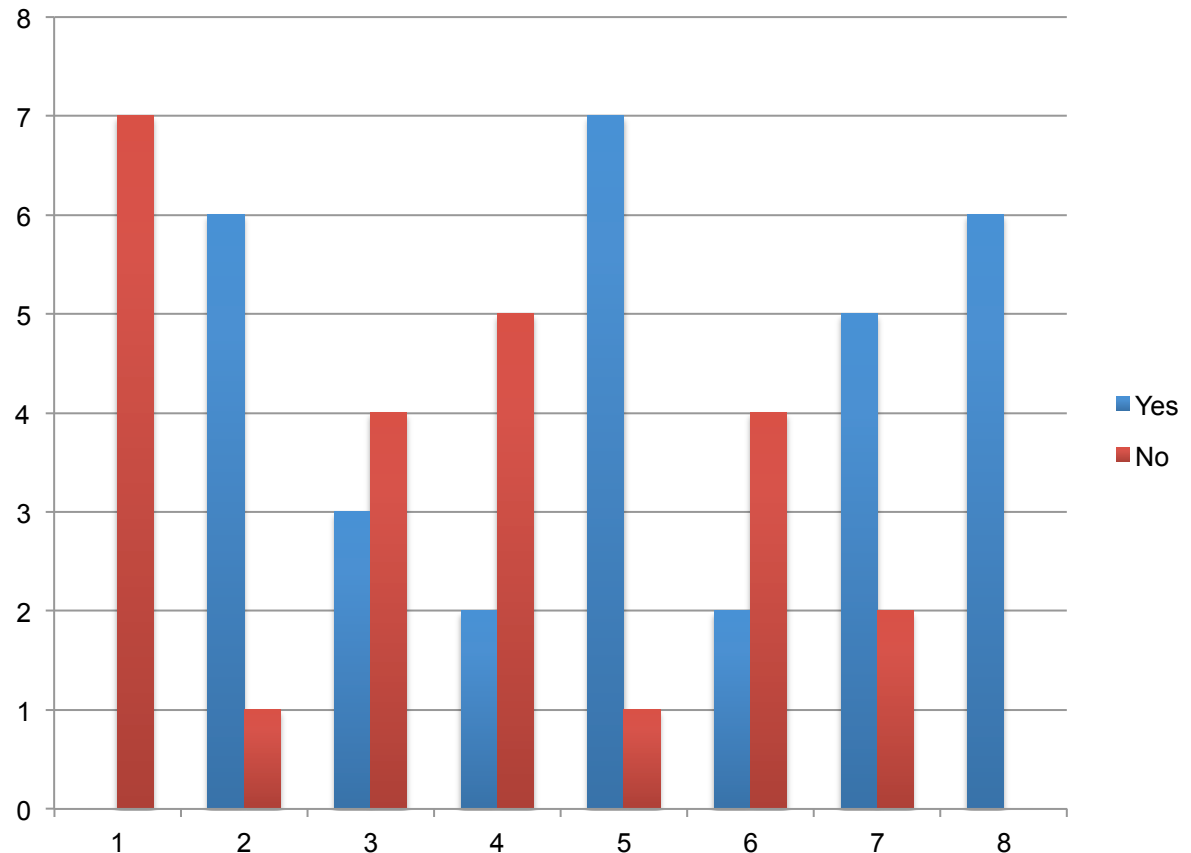
Inhibition of microRNA-29b reduces murine abdominal aortic aneurysm (AAA) development, PMID: [22269326](#)

All questions based on this paper

Some consistency in replies

Plenty to discuss

Can we reach an agreement today?



Murine model of experimental AAA: the porcine pancreatic elastase (PPE) infusion model in C57BL/6 mice

miR-29 in AAA, induced by PPE infusion in mice

(A) AAD (versus baseline in percentage) in PPE-induced (ELAST) AAA compared with that in sham-operated control mice (sham). (B) miR-29 family (miR-29a, miR-29b, miR-29c) expression in ELAST mice compared with that in sham-operated control mice. (C) ISH for miR-29b (purple chromogen) in control aortas (untreated), sham-operated mice, and ELAST mice 14 days after surgery (original magnification, $\times 200$). (D and E) mRNA expression levels in ELAST mice compared with those in sham-operated mice for (D) *Col1a1*, *Col3a1*, and *Col5a1* as well as (E) *Eln* and *Fbn1*. $n = 5-8$ for each treatment group and time point. Data are mean \pm SEM. * $P < 0.05$ versus sham.

Figure 1

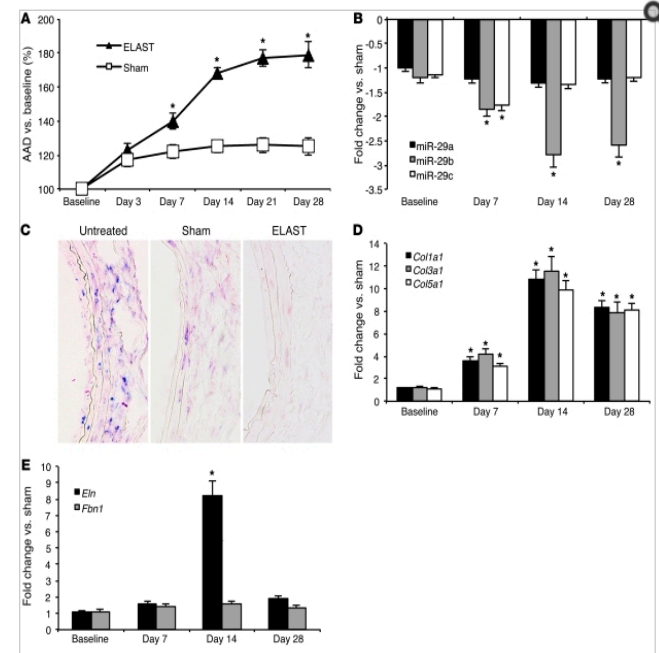


Figure 1: Murine model of experimental abdominal aortic aneurysm: the porcine pancreatic elastase (PPE) infusion model in C57BL/6 mice

Possible annotation porcine elastase with GO:0030198 extracellular matrix organization IDA

Would you annotate figure 1?

A Does anyone think we should annotate this?

2 comments

Not familiar enough with this model and system to make a confident decision one way or the other.

Fair comment

Although the proposed annotation is a reasonable assumption, the data in fig. 1 show that treatment with PPE results in an increased abdominal aortic diameter and reduced miR29 expression and increased Eln, Col1, Col3 and Col5 expression. I can't tell from that data if the changes in gene expression cause the increased aortic diameter or vice versa. I wouldn't annotate anything.

Good point

Human aortic smooth muscle cells (hASMCs)
Human aortic adventitial fibroblasts (hAFBs)

Cells were treated with human TGF- β 1

- Known regulator of miR-29b
- Profibrotic stimulus
- Treatment decreased miR-29b expression in hAFBs but not in hASMCs

Suggested annotation:

Human TGF- β 1 GO:2000628 regulation of miRNA metabolic process IDA

C16: has_regulation_target human miR-29b, occurs_in CL:0002547 fibroblast of the aortic adventitia

Would you include the cell specific information in the annotation extension field associated with your annotation of TGF- β 1

Suggested annotation:

Human TGF- β 1 GO:2000628 regulation of miRNA metabolic process IDA

C16: has_regulation_target human miR-29b, occurs_in CL:0002547 fibroblast of the aortic adventitia

Would you include the cell specific information in the annotation extension field associated with your annotation of TGF- β 1

Pretty much all agreed that they would include the cell specific data

Yes: 6 No: 1

4 responses

It seemed the cell type information was critical in this experiment. Is any more known about the mechanism of regulation of miR-29b by TGF-beta1?

probably...since they showed a specific effect in fibroblast vs muscle cell.

I think these cells will be exposed in vivo to TGF- β 1

I don't think the term "regulation of miRNA metabolic process" describes correctly the role of TGFB1 here. Also, in GO "expression" is not related to "metabolic processes".

Suggested annotation:

Human TGF- β 1 GO:2000628 regulation of miRNA metabolic process IDA

C16: has_regulation_target human miR-29b, occurs_in CL:0002547 fibroblast of the aortic adventitia

Would you include the cell specific information in the annotation extension field associated with your annotation of TGF- β 1

Pretty much all agreed that they would include the cell specific data

Yes: 6 No: 1

Does anyone want to discuss this further?

the role of TGFB1 here. Also, in GO "expression" is not related to "metabolic processes".

Note that miRNAs are the final product, unlike mRNAs which are then translated to proteins.

Is miRNA metabolic process in appropriate term?

Should there be a relationship between expression and metabolic process?

Should there be a GO term specific for the miRNAs?

Eg 'miRNA activity involved in gene silencing'

Parent terms:

- has_part GO:0003729 mRNA binding
- part_of GO:0035195 gene silencing by miRNA
- is_a molecular_function.
- Definition: Interacting selectively and non-covalently with an RNA sequence in order to modulate translation.

Note that proteins are annotated to the term GO:0035195 gene silencing by miRNA

No obvious agreement

Yes: 3 No: 4

Should there be a GO term specific for the miRNAs?

Eg 'miRNA activity involved in gene silencing'

No obvious agreement

Yes: 3 No: 4

Does anyone want to discuss this further?

I guess I can see the use of an MF term such as this, but I'm not sure that 'miRNA activity' is sufficiently descriptive enough. What about something like 'mRNA binding involved in post-transcriptional gene silencing'?

Not sure if I should answer yes or no: I suppose all miRNAs would have this annotation?

Based on the definition of GO:0035195 'gene silencing by miRNA', it doesn't seem like proteins should be annotated with that term...I thought that WAS the term to annotate miRNA activity with... :)

Note that GO:0035195 'gene silencing by miRNA' is a biological process

To apply the GO term 'GO:0035195 gene silencing by miRNA' to a miRNA **would there have to be evidence in this paper that this miRNA bound to the target mRNA?**

Most people felt that there wouldn't have to be evidence in this paper that the miRNA bound the target mRNA

Yes: 2 No: 5

To apply the GO term 'GO:0035195 gene silencing by miRNA' to a miRNA would there have to be evidence in this paper that this miRNA bound to the target mRNA?

Yes: 2 No: 5

3 comments

If we are already annotating proteins to this term, doesn't that imply that the answer to this question is 'no'?

The question is about miRNAs, proteins involved in this process wouldn't be expected to bind the target.

Depends on the evidence code you use. For an IDA yes, otherwise no.

So for an anti-miR experiment no evidence is required for target bound? Whereas, for a pre-miR expt evidence of target being bound is required?

Question 4

To apply the GO term 'GO:0035195 gene silencing by miRNA' to a miRNA **would there have to be evidence in this paper that this miRNA bound to the target mRNA?**

Yes: 2 No: 5

3 comments...continued

I would annotate based on what the authors state.

If the authors claim the miRNA silences expression of gene 1, I would make the annotation even if they don't show the direct physical interaction.

So this is a No vote

Maybe the gene they show is silenced is actually regulated indirectly by the miRNA silencing expression of some other gene...does that matter?

I think this is a No vote (as it still relies on the miR silencing a gene)

Maybe if we put the silenced gene in column 16

So this is a No vote if nothing in C16

And Yes vote if target added to C16

...despite that I would still annotate the miRNA to 'gene silencing by miRNA'.... **So this is still a No vote**

Question 4

To apply the GO term 'GO:0035195 gene silencing by miRNA' to a miRNA **would there have to be evidence in this paper that this miRNA bound to the target mRNA?**

Yes: 2 No: 5

3 comments...last comments continued

**So this is a No vote if nothing in C16
and Yes vote if target added to C16**

Maybe I wouldn't put the regulated gene in column 16 though unless I thought it was direct regulation?

How do we decide whether or not this is direct regulation?

This is question 9

Does anyone want to discuss anything else in this figure?

Figure 3B Primary aortic fibroblasts were treated with human TGF- β 1

- Known regulator of miR-29b
- Profibrotic stimulus

Expression levels of miR-29b target genes (COL1A1, COL3A1, ELN, FBN1) in Tgf- β -stimulated anti-29b- and pre-29b-transfected hAFBs were measured.

Results

COL1A1 and COL3A1 were significantly up-regulated with Tgf- β -stimulation

COL1A1 and COL3A1 further up-regulated with anti-29b treatment

COL1A1 and COL3A1 down-regulated with pre-29b

If the anti-29b data was not available the experimental data involving the addition of pre-29b could be annotated as IDA (or IMP) as:

Human miR-29b: GO:0035195 gene silencing by miRNA IDA C16: has_
regulation_target human COL1A1 [and COL3A1] happens_during GO:0071560
cellular response to transforming growth factor beta stimulus.

As this experiment involves the addition of an miR which is known to be expressed in these cells would you also include the C16 statement: occurs_in CL:0002547 fibroblast of the aortic adventitia?

Figure 3B

Expression levels of miR-29b target genes (COL1A1, COL3A1, ELN, FBN1) in Tgf- β -stimulated pre-29b-transfected hAFBs were measured.

Results

COL1A1 and COL3A1 down-regulated with pre-29b

Annotation:

Human miR-29b: GO:0035195 gene silencing by miRNA IDA C16: has_
regulation_target human COL1A1 [and COL3A1] happens_during GO:0071560
cellular response to transforming growth factor beta stimulus.

As this experiment involves the addition of an miR which is known to be expressed in these cells would you also include the C16 statement: occurs_in CL:0002547 fibroblast of the aortic adventitia?

Pretty much all agreed that they would include the cell specific data

Yes: 7 No: 1

3 comments – in agreement

Figure 3B

Annotation:

Human miR-29b: GO:0035195 gene silencing by miRNA IDA C16: has_
regulation_target human COL1A1 [and COL3A1] happens_during GO:0071560
cellular response to transforming growth factor beta stimulus.

As this experiment involves the addition of an miR which is known to be expressed
in these cells would you also include the C16 statement: occurs_in CL:0002547
fibroblast of the aortic adventitia?

Pretty much all agreed that they would include the cell specific data

Yes: 7 No: 1

Does anyone want to discuss anything else about this Q?

We just need to be careful to distinguish between the experimental system and the real biological context. In this case it seems to be relevant biologically.

probably...I guess I would rely on what the authors conclude about their assay. If the cells are a good model of aortic fibroblasts then yes...if the authors are unclear about that..then I probably wouldn't include the cell in column 16.

Following on from the previous question:

Figure 3B assuming the addition of pre-29b was annotated as: Human miR-29b: GO:0035195 gene silencing by miRNA IDA C16: has_ regulation_target human COL1A1 [and COL3A1] happens_during GO:0071560 cellular response to transforming growth factor beta stimulus.

If the miR was not known to be expressed in fibroblast of the aortic adventitia cells would you also include the C16 statement: occurs_in CL: 0002547 fibroblast of the aortic adventitia?

Most people felt that they wouldn't have added this cell specific data if the miR was not known to normally be present in these cells

Yes: 2 No: 4

1 comment

Following on from the previous question:

If the miR was not known to be expressed in fibroblast of the aortic adventitia cells would you also include the C16 statement: occurs_in CL: 0002547 fibroblast of the aortic adventitia?

Most people felt that they wouldn't have added this cell specific data if the miR was not known to normally be present in these cells

Yes: 2 No: 4

1 comment

I think this question goes to the heart of whether GO annotation reflects what the

Does anyone want to discuss anything else about this Q?

expanded ECO term to indicate that the experiment reflects heterologous expression would be useful?

Let me reword...if the miR was "known not" to be expressed in fibroblast of the aortic adventitia cells...then I would not include that in the annotation. I rely heavily on trying to capture the authors intent in cases like this. We don't have time to do gene expression research for such an annotation, so we just have to rely on the authors statements and try to be cautious about how specific we annotate.

Figure 5D-F demonstrates a change in MMP activity
MMPs are matrix metalloprotease associated with ECM degradation

Results (F)

Mmp2 and Mmp9 expression up-regulated with pre-29b

Mmp2 and Mmp9 expression down-regulated with anti-29b

Note that this is not gene silencing of MMPs, but the reverse! Ie there is likely to be other genes being regulated before these.

Annotation suggestion:

Mouse miR-29b GO:0090091 positive regulation of extracellular matrix disassembly IMP

Do you think this interpretation is too downstream from the observed increase in MMP activity?

Most people felt that this was too far downstream

Yes: 5 No: 2

Question 7

Figure 5D-F demonstrates a change in MMP activity associated with miR-29b levels. Annotation suggestion:

Mouse miR-29b GO:0090091 positive regulation of extracellular matrix disassembly IMP

Do you think this interpretation is too downstream from the observed increase in MMP activity?

Most people felt that this was too far downstream

Yes: 5 No: 2

3 Comments

Generally no, this type of annotation seems to provide some information about the organismal level processes affected by miR-29b. If we don't make these types of annotations, are we limiting the information that is available wrt miR biology?

Although I understand that we would like to capture this type of information.

Question 7

Figure 5D-F demonstrates a change in MMP activity associated with miR-29b levels. Annotation suggestion:

Mouse miR-29b GO:0090091 positive regulation of extracellular matrix disassembly IMP

Do you think this interpretation is too downstream from the observed increase in MMP activity?

Most people felt that this was too far downstream

Yes: 5 No: 2

Does anyone want to discuss anything else about this Q?

Depends...I try to answer that question in general by asking myself "would authors expect to get this gene (miR-29b in this case) back if they search for genes involved in regulation of extracellular matrix assembly/disassembly. I think in this case the answer would be 'yes'...so I would probably make the annotation even though the effect of miR-29b on extracellular matrix disassembly is one step removed from the primary function of the gene (regulating gene expression of MMPs). As an alternative, I would want to annotate to something like 'gene silencing by miRNA' and use the MMP genes as targets in column 16.

The custom-made LNA-anti-miR-29b 5'-3' sequence aligns 100% with both human sequences

Would you expect both miRs to be annotated based on the anti-29b experiments?

(presumably supported with the IGI evidence code and the WITH field including the other miR sequence)

All agreed that they would annotate both miRs

Yes: 6 No: 0

3 Comments

Question 8

The custom-made LNA-anti-miR-29b 5'-3' sequence aligns 100% with both human sequences

Would you expect both miRs to be annotated based on the anti-29b experiments?

All agreed that they would annotate both miRs

Yes: 6 No: 0

3 Comments

Does anyone want to discuss anything else about this Q?

This is a really good question for any sequence-based GO annotation.

Do curators routinely check the targets of sequence-specific reagents and annotate to all expected targets?

In this case, it seems reasonable to annotate to both miRs, **but is that consistent with the authors' intents?**

again...depends a little on what the authors tell me. We probably wouldn't be BLASTing the reagents to see what they hit, but if the authors told me they designed it to target both mir-29b genes and offered no information regarding which of them may be expressed in the cells of interest, I would annotate both genes by IGI as you suggest.

When associating 'GO:0035195 gene silencing by miRNA' with a miRNA, based on an experiment using an anti-miR to demonstrate up-regulation of specific mRNA levels in the absence of the miR,

How do we decide whether or not an mRNA is a direct target of a miRNA?

I wouldn't apply this GO term just based on up-regulation of specific mRNAs (discussed in Q4)

1

The majority of people (Q4) would create

miR-# 'GO:0035195 gene silencing by miRNA' IMP based on an experiment using an anti-miR which showed up-regulation of specific mRNA levels.

When would you add the target mRNA to C16?

Note I didn't explain that you could choose multiple options

I would apply this GO term and would add the regulated mRNAs as direct targets in C16 field.....	
based on this experiment alone	0
if there was some evidence that they likely to be targets, such as predicted targets	1
if there was some evidence that they likely to be targets, such as known targets (as stated by author)	2
if they had been shown in the paper to be directly base-pairing with each mRNA listed	3
if they had been shown in the paper (or in another paper) to be directly base-pairing with each mRNA listed	3
Total Respondents: 8	10

The majority of people (Q4) would create

miR-# 'GO:0035195 gene silencing by miRNA' IMP based on an experiment using an anti-miR which showed up-regulation of specific mRNA levels.

When would you add the target mRNA to C16?

Last option: if they had been shown in the paper (or in another paper) to be directly base-pairing with each mRNA listed

Does anyone want to discuss anything else about this Q?

I think I'd opt for the last choice in the list, but what about the possibility of just using has_input?

... and use the appropriate evidence code.

Suggestion is

Use has_direct_input if mRNA shown to be bound

Use has_input if binding not demonstrated in paper

What is meant by using correct evidence codes??

PMID:21993888 In vitro studies using vascular SMCs (smooth muscle cells) of the renin lineage cells permanently labeled with CFP (a marker of cells of the renin lineage) that express renin upon stimulation with cAMP.

Does anyone want to discuss anything else about this Q?

22 miRNAs were up-regulated and 26 were down-regulated.

Annotation of 48 different miRNAs: cellular response to inorganic substance IEP (these have the IDA evidence code in MGI).

Which evidence code should be applied to these 48 miRs?

Should there be different rules for the annotation of expression data describing miRs v mRNAs?

Should this data be included in GO?

How reliable is this data?

Should we limit these annotations to those which apply a minimum standard use of controls