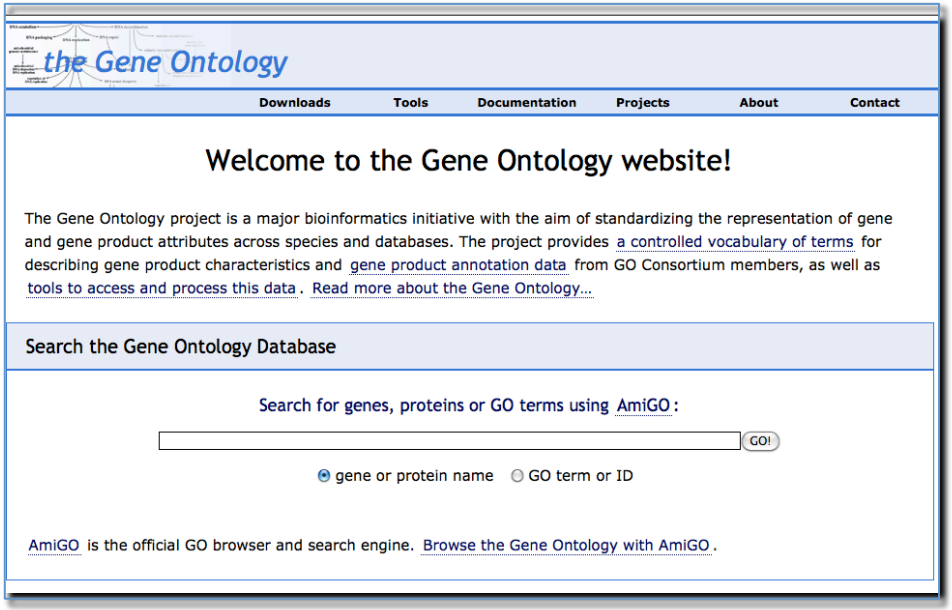
PANTHER Classification System

Gene Ontology

- The first ontology that was designed as a formal representation of biological knowledge
- Three knowledge domains:
  - molecular function
  - biological process
  - cellular component.



EGFR (MGI:95294)

Molecular function: GO:0005006: epidermal growth factor-activated receptor activity

Biological process: GO:0007173: epidermal growth factor receptor signaling pathway

Cellular component: GO:0016021 integral component of membrane

Ashburner, M., et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. **25**: p. 25-9.

The Gene Ontology Consortium (2012) The Gene Ontology: enhancements for 2011. *Nucleic Acids Res*. 40:D559

GO Annotation

- 4.1 million annotations to 758K genes
- 568K annotations were curated manually with experimental evidences.
- 121K to human genes (counted by IDs, so there are multiple IDs from different curation groups for the same gene).

PANTHER Classification System

Ontology

Molecular function

Ancestors of epidermal growth factor-activated receptor activity (GO:0005006)

subject	relation	object
epidermal growth factor-activated receptor activity	I is_a	transmembrane receptor protein tyrosine kinase activity (GO:0004714)
epidermal growth factor-activated receptor activity	I is_a (inferred)	molecular_function (GO:0003674)
epidermal growth factor-activated receptor activity	I is_a (inferred)	catalytic activity (GO:0003824)
epidermal growth factor-activated receptor activity	I is_a (inferred)	molecular transducer activity (GO:0060089)
epidermal growth factor-activated receptor activity	I is_a (inferred)	transferase activity (GO:0016740)
epidermal growth factor-activated receptor activity	I is_a (inferred)	receptor activity (GO:0004872)
epidermal growth factor-activated receptor activity	I is_a (inferred)	signal transducer activity (GO:0004871)
epidermal growth factor-activated receptor activity	I is_a (inferred)	transferase activity, transferring phosphorus-containing groups (GO:0016772)
epidermal growth factor-activated receptor activity	P part_of (inferred)	ERBB signaling pathway (GO:0038127)
epidermal growth factor-activated receptor activity	P part_of	epidermal growth factor receptor signaling pathway (GO:0007173)
epidermal growth factor-activated receptor activity	P part_of (inferred)	enzyme linked receptor protein signaling pathway (GO:0007167)
epidermal growth factor-activated receptor activity	P part_of (inferred)	cellular response to stimulus (GO:0051716)
epidermal growth factor-activated receptor activity	P part_of (inferred)	cellular protein modification process (GO:0006464)
epidermal growth factor-activated receptor activity	P part_of (inferred)	cellular protein metabolic process (GO:0044267)
epidermal growth factor-activated receptor activity	P part_of (inferred)	cellular process (GO:0009987)
epidermal growth factor-activated receptor activity	P part_of (inferred)	cellular metabolic process (GO:0044237)

Biological process

PANTHER Classification System

MGI:95294

GO annotation

Found entities

Total: 12; showing 1-10 **Results count** 10

Gene/Product	Gene/Product name	Qualifier	Direct annotation	Annotation extension	Source	Taxon	Evidence	Evidence with	PANTHER family	Isoform	Reference
Egfr	epidermal growth factor receptor		signal transducer activity		MGI	Mus musculus	IDA		tyrosine-protein kinase receptor pthr24416	VEGA:OTTMUSP00000005385	MGI:MGI:2182234 PMID:11940581
Egfr	epidermal growth factor receptor		protein binding		MGI	Mus musculus	IPI	RefSeq:NP_851419	tyrosine-protein kinase receptor pthr24416	VEGA:OTTMUSP00000005385	MGI:MGI:3531070 PMID:15695332
Egfr	epidermal growth factor receptor		protein binding		MGI	Mus musculus	IPI	RefSeq:NP_851419	tyrosine-protein kinase receptor pthr24416	VEGA:OTTMUSP00000005385	MGI:MGI:3587490 PMID:15728722
Egfr	epidermal growth factor receptor		kinase activity		MGI	Mus musculus	IDA		tyrosine-protein kinase receptor pthr24416	VEGA:OTTMUSP00000005385	MGI:MGI:2664443 PMID:12808090
Egfr	epidermal growth factor receptor		protein binding		MGI	Mus musculus	IPI	UniProtKB:P01133	tyrosine-protein kinase receptor pthr24416	VEGA:OTTMUSP00000005385	MGI:MGI:4948998 PMID:21439278
Egfr	epidermal growth factor receptor		epidermal growth factor-activated receptor activity		MGI	Mus musculus	IDA		tyrosine-protein kinase receptor pthr24416	VEGA:OTTMUSP00000005385	MGI:MGI:3032880 PMID:14712229

Gene product

Molecular function

PANTHER Classification System

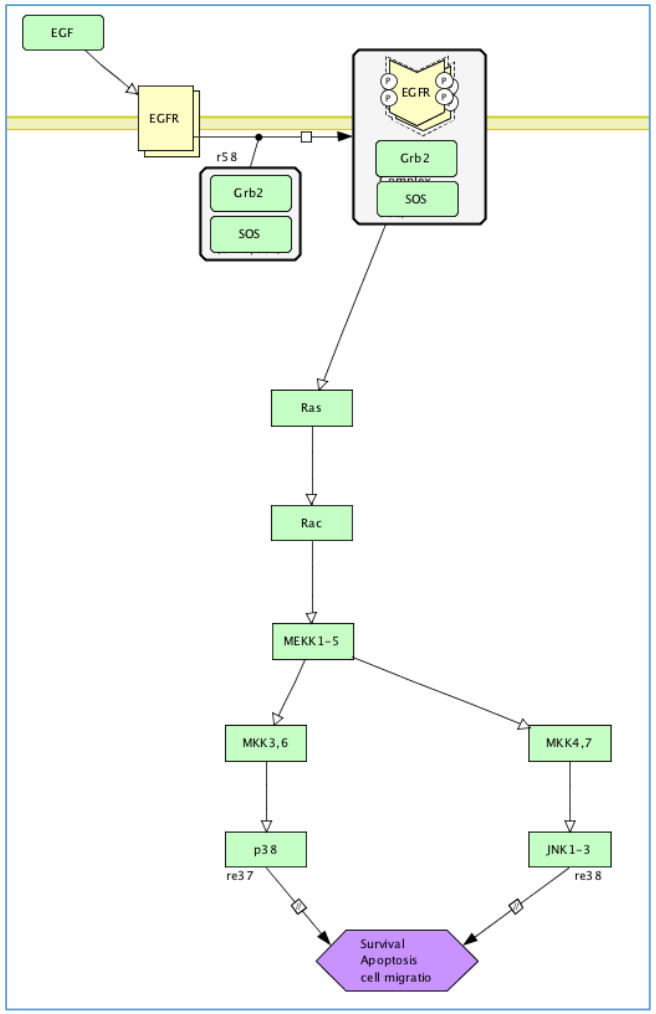
GO Relationships

Children of epidermal growth factor-activated receptor activity (GO:0005006)

subject	relation	object
gurken-activated receptor activity (GO:0008313)	is_a	epidermal growth factor-activated receptor activity
negative regulation of epidermal growth factor-activated receptor activity (GO:0007175)	negatively_regulates	epidermal growth factor-activated receptor activity
positive regulation of epidermal growth factor-activated receptor activity (GO:0045741)	positively_regulates	epidermal growth factor-activated receptor activity
regulation of epidermal growth factor-activated receptor activity (GO:0007176)	regulates	epidermal growth factor-activated receptor activity

EGFR signaling

- GO annotation is gene-centric.
- The current ontology is not able to capture the relationship of GO terms between two different entities.



What is LEGO Project?

- LEGO is a new development under the Gene Ontology project.
- It is an extension of the existing GO by capturing relationships among various GO terms during the curation of genes.
- LEGO will enable curators to use the GO to express rich biological statements from the literature
  - Maximize biological knowledge captured by curator
  - Represent complex biology in an accurate, computable manner
  - Prevent
    - "kludges" that use existing terms misleadingly or inconsistently
    - Combinatorial explosion of GO terms

LEGO is an *extension* of GO

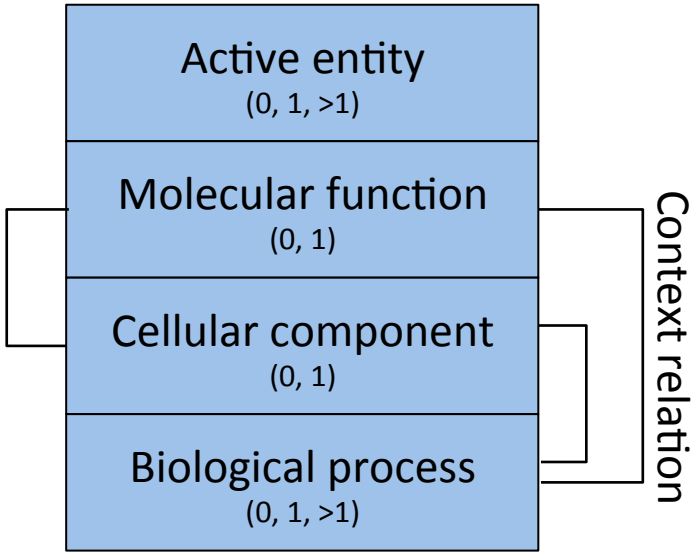
- Currently, GPs have separate MF, CC, BP annotations
  - In LEGO, a particular GP executes a particular MF in a particular CC as part of* a particular BP
  - LEGO is backwards compatible with current annotations
    - Current annotations are incomplete
      - E.g. an MF annotation states that a particular GP executes a particular MF in some CC as part of some BP

LEGO formalism

Annoton – the annotation unit in LEGO

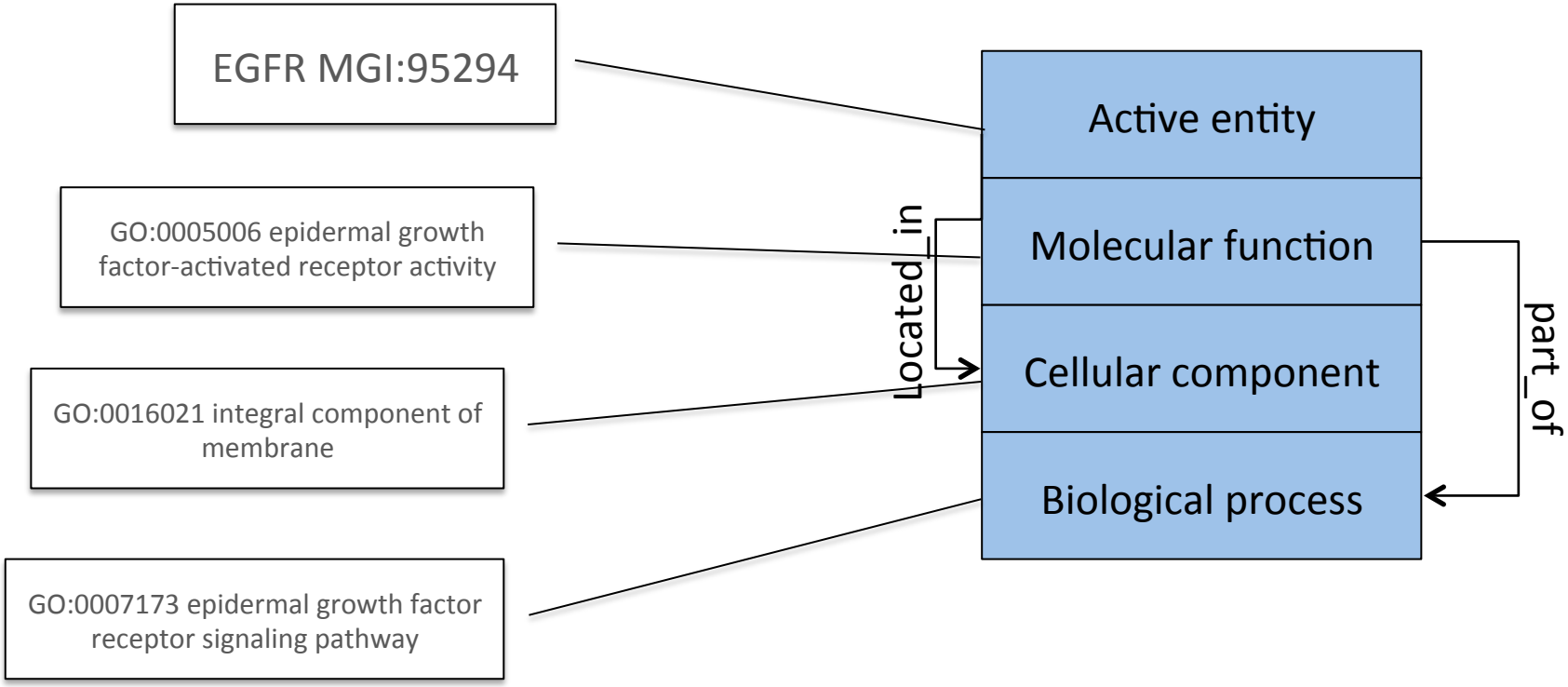
Molecular Activity

- A molecular activity is defined as the molecular function that an entity is capable of performing, via a particular biochemical mechanism, in a specific cellular location, as part of a biological process.



PANTHER Classification System

Molecular Activity
-EGFR

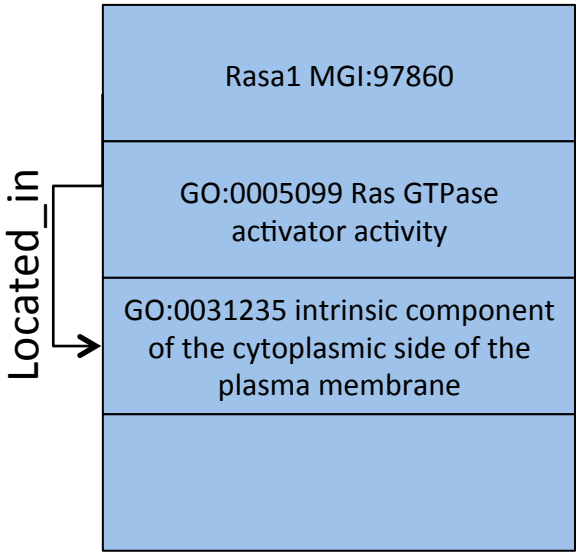


PANTHER Classification System

Molecular Activity -Rasa1

MGI:97860

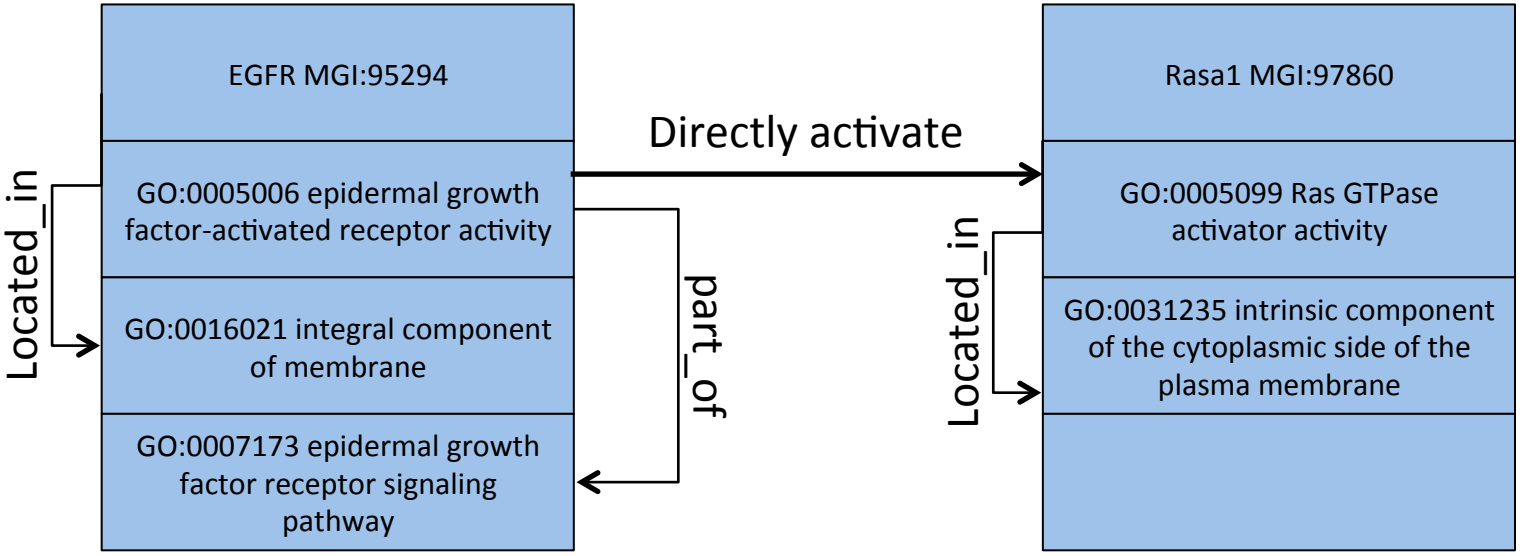
Rasa1	RAS p21 protein activator 1		protein binding		MGI	Mus musculus	IPI
Rasa1	RAS p21 protein activator 1		protein binding		MGI	Mus musculus	IPI
Rasa1	RAS p21 protein activator 1		Ras GTPase activator activity		MGI	Mus musculus	IBA
Rasa1	RAS p21 protein activator 1		positive regulation of Ras GTPase activity		MGI	Mus musculus	IBA
Rasa1	RAS p21 protein activator 1		negative regulation of Ras protein signal transduction		MGI	Mus musculus	IBA
Rasa1	RAS p21 protein activator 1		plasma membrane	part_of bone marrow part_of macrophage	MGI	Mus musculus	IDA
Rasa1	RAS p21 protein activator 1		intrinsic component of the cytoplasmic side of the plasma membrane		MGI	Mus musculus	IBA
Rasa1	RAS p21 protein activator 1		cytoplasm		MGI	Mus musculus	IBA



Effect Relation

- The *effect relation* describes the effects exerted by one *molecular activity* unit to the other.
  - Directly activate
  - Directly inhibit
  - Positively influence
  - Negatively influence
  - Regulate
  - Upstream

LEGO Model



Epidermal growth factor-activated receptor activity from mouse EGFR, which is an integral component of membrane and is involved in epidermal growth factor receptor signaling pathway, directly activates the Ras GTPase activator activity of mouse Rasa1 that is located in the intrinsic component of the cytoplasmic side of the plasma membrane.

SBGN-AF can be used as graphical representations of LEGO Models

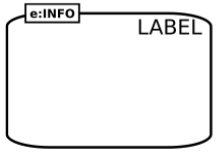
SBGN

LEGO

Unit of information

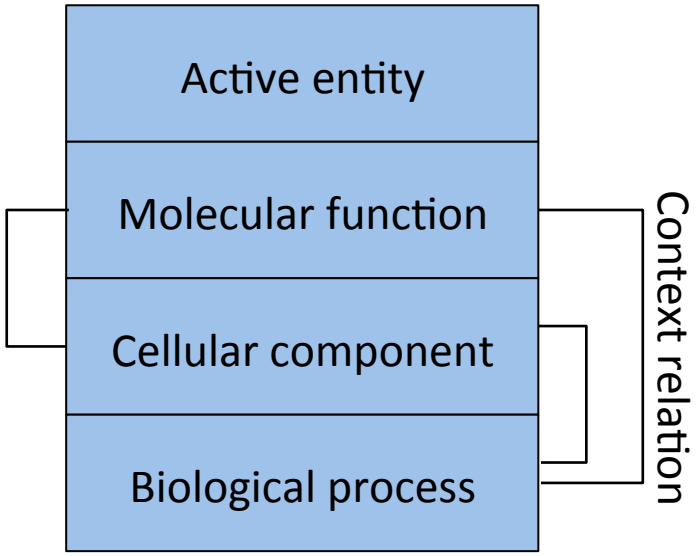
Biological activity

Compartment

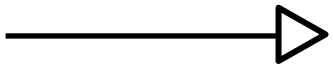
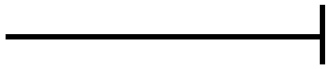


Submap

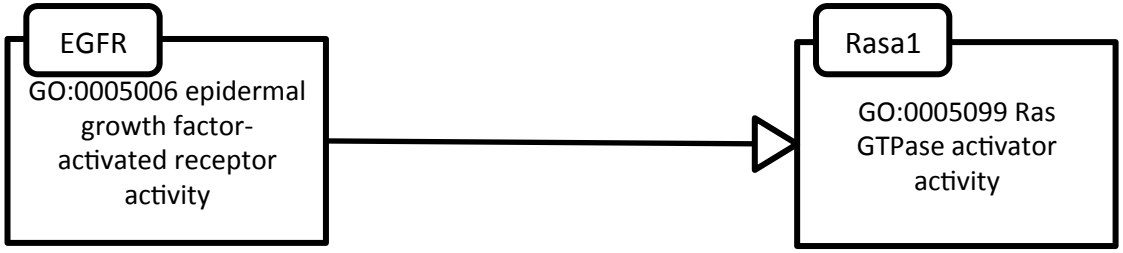
Phenotype (when MF=0)



Representing LEGO Model in SBGN-AF (cont.)

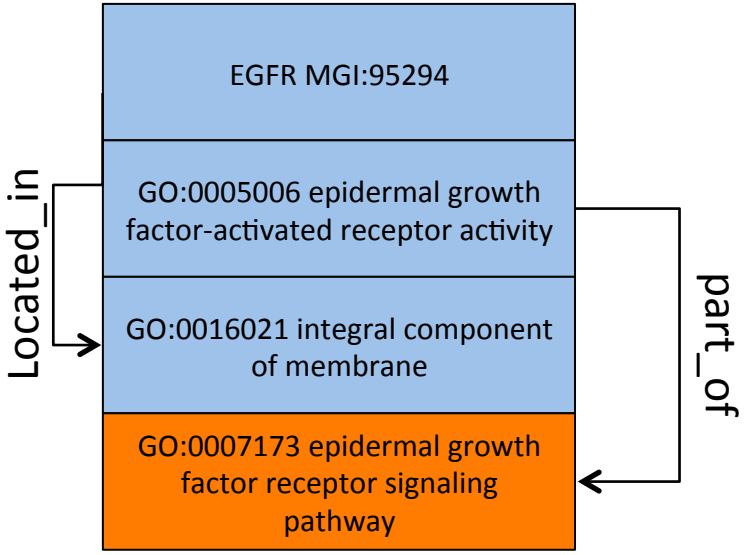
SBGN		LEGO
Positive influence		Directly activate Positively influence
Negative influence		Directly inhibit Negatively influence
Unknown influence		Regulate Upstream

SBGN-AF Representation
-activation of Rasa1 by EGFR



Issues

- Context relation between MF and BP.
  - The BP annotation here is to show the context relationship between the MF and BP.
  - The current spec does not require that the BP annotations be consistent throughout the model.
  - Submap may not be an option because it will make the map look really bad



Issues

- Context relation between MF and BP.
  - The BP annotation here is to show the context relationship between the MF and BP.
  - The current spec does not require that the BP annotations be consistent throughout the model.
  - Submap may not be an option because it will make the map look really bad
- Effect relations
  - "Directly …." have implications of direct interaction between the affect and effect proteins. SBGN-AF does not differentiate whether the influence is direct or indirect.
  - Upstream has no regulation or influence implication.

LEGO annotation tool

- Launch page
  - http://go-genkisugi.rhcloud.com
- Examples
  - Single paper curation
    - http://www.ncbi.nlm.nih.gov/pubmed/?term=17996703
    - http://go-genkisugi.rhcloud.com/seed/model/gomodel:goa_human-5323da180000002
  - Seeding
    - http://go-genkisugi.rhcloud.com/seed/model/gomodel:pombase-GO-0051306

LEGO team

- Lawrence Berkeley Lab
  - Seth Carbon
  - Chris Mungall
  - Heiko Dietze
  - Suzi Lewis
- USC
  - Anushya Muruganujan
  - Huaiyu Mi
  - Paul Thomas