BioPAX: An Introduction

Emek Demir
Computational Biology
Memorial Sloan Kettering Cancer Center

High throughput era

Goals:

Causality Analysis
Experiment Selection
Visualization

Requirements:
High Coverage
Integration

Complete Listing of All Pathguide Resources

Pathguide contains information about **291** biological pathway resources. Click on a link to go to the resource home page or 'Details' for a description page. Databases that are free and those supporting BioPAX, CellML, PSI-MI or SBML standards are respectively indicated.

If you know of a pathway resource that is not listed here, or have other questions or comments, please send us an e-mail.

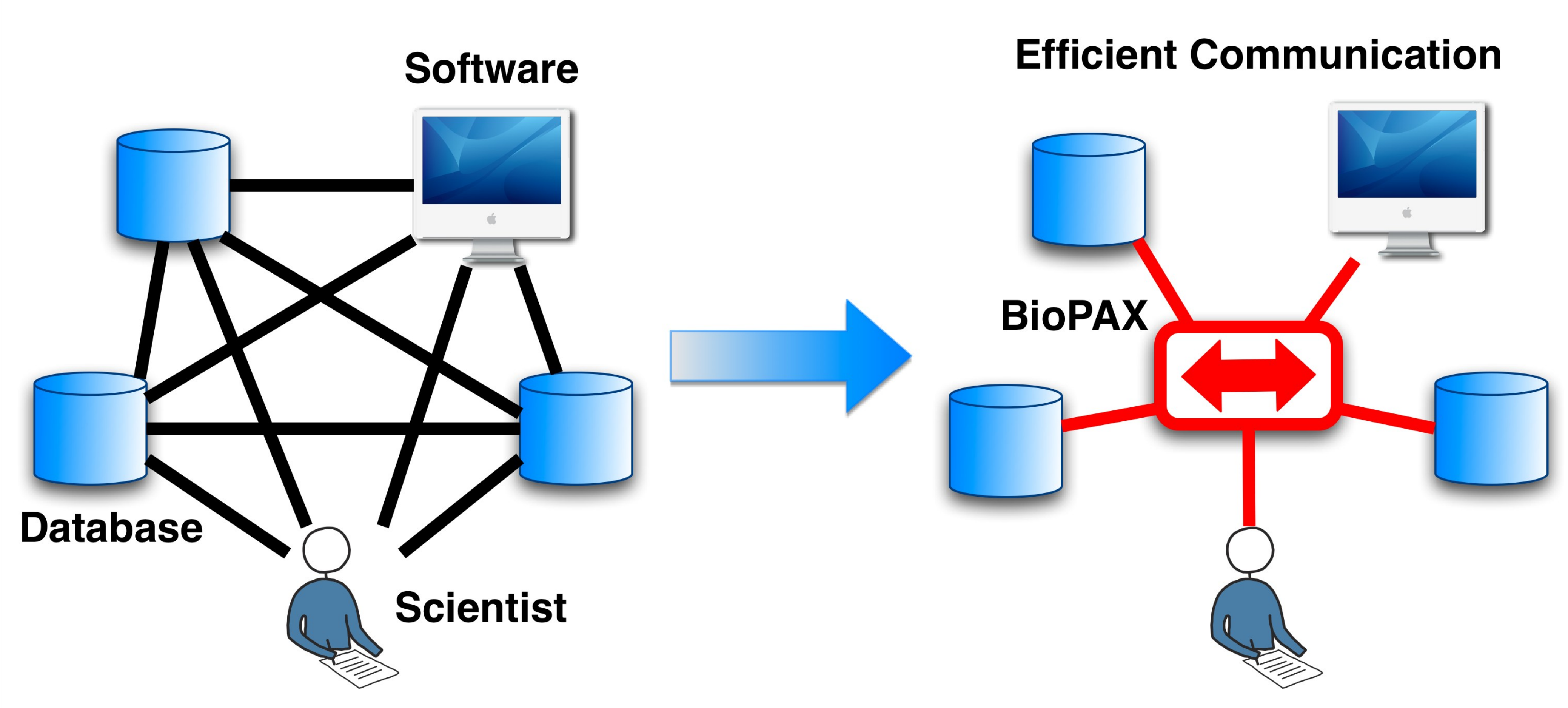
News

Major update
All resources were recently reviewed and many new ones were added

Get the Stats
Detailed Pathguide resource statistics now available

Protein-Protein Interactions

Database Name (Order: alphabetically | by web popularity)	Full Record	Availability	Standards
3DID - 3D interacting domains	Details	Free	
ABCdb - Archaea and Bacteria ABC transporter database	Details	Free	
AfCS - Alliance for Cellular Signaling Molecule Pages Database	Details	Free	
AllFuse - Functional Associations of Proteins in Complete Genomes	Details	X	
aMAZE - Protein Function and Biochemical Pathways Project	Details	Free	
ASEdb - Alanine Scanning Energetics Database	Details	Free	
ASPD - Artificial Selected Proteins/Peptides Database	Details	Free	
BID - Binding Interface Database	Details	X	
BIND - Biomolecular Interaction Network Database	Details	Free	PSI-MI
BioGRID - General Repository for Interaction Datasets	Details	📖	PSI-MI
BRITE - Biomolecular Relations in Information Transmission and Expression	Details	Free	
CA1Neuron - Pathways of the hippocampal CA1 neuron	Details	Free	
Cancer Cell Map - The Cancer Cell Map	Details	Free	BioPAX
CellCircuits - CellCircuits	Details	Free	
CPDB - ConsensusPathDB	Details	📖	BioPAX


Software
Database
Scientist
Efficient Communication
BioPAX

And more..

48 Pathway Visualization Tools
12 widely used literature mining tools
Enrichment, causality, reachability analysis algorithms..

Impossible to compare, combine, re-use

Solution?

- A common representation
- And services..
  - Aggregation
  - Integration
  - Comparison
  - Querying
  - Time-stamping
  - Provenance

BioPAX

- **Bio**logical **Pa**thways **Ex**change
- BioPAX is in OWL (Ontology Web Language).
- Community process:
  - www.biopax.org
  - biopax-discuss@biopax.org
  - www.biopaxwiki.org
- Released in levels
- Proposals created by editors
- Tested by stakeholders.

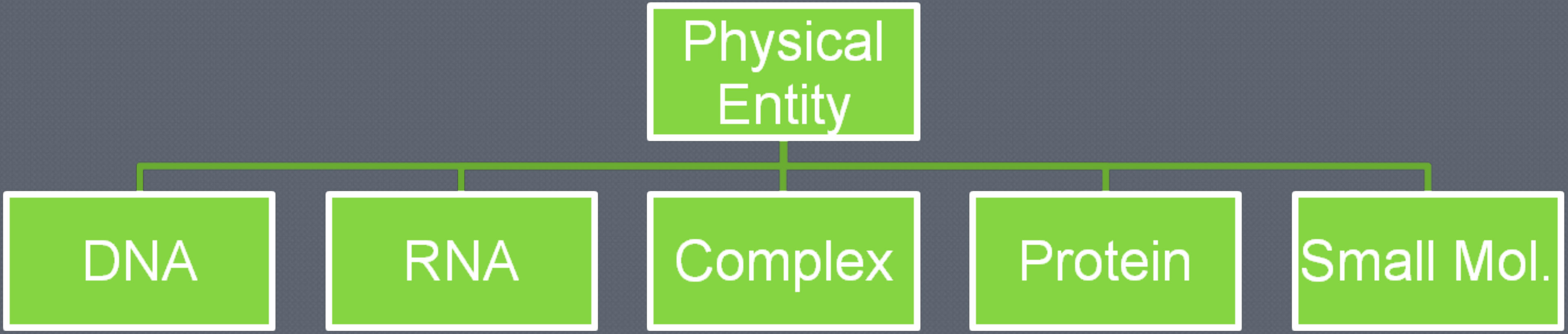
BioPAX Level 2

- Stable
- Covers:
  - Molecular Interactions
  - Metabolic Networks
  - Signaling (Partially)
- Major Data Providers:
  - BioCyc,
  - Reactome,
  - NCI/Nature PID,
  - Cancer Cell Map,
  - INOH

BioPAX Level 3

- RC4- After 3 years of development
- Covers:
  - Molecular Interactions
  - Metabolic Networks
  - Signaling Networks
  - Gene Regulation
  - Genetic Interactions
- Mojor providers testing:
  - Reactome,
  - NCI/Nature PID,
  - INOH
  - Panther
  -

Entities



Physical Entity

DNA

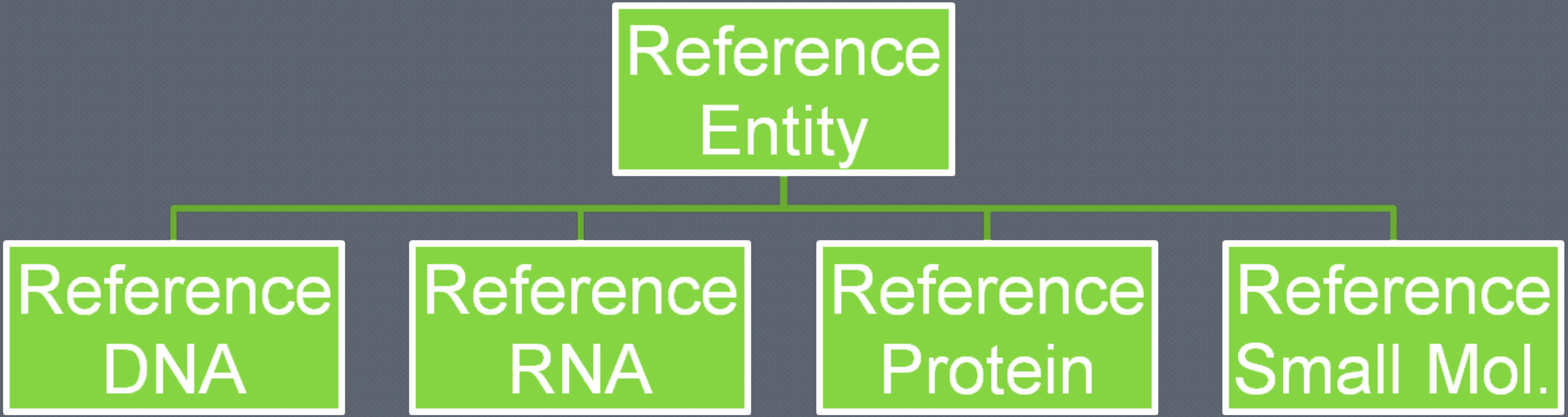
RNA

Complex

Protein

Small Mol.

Entities


Reference Entity

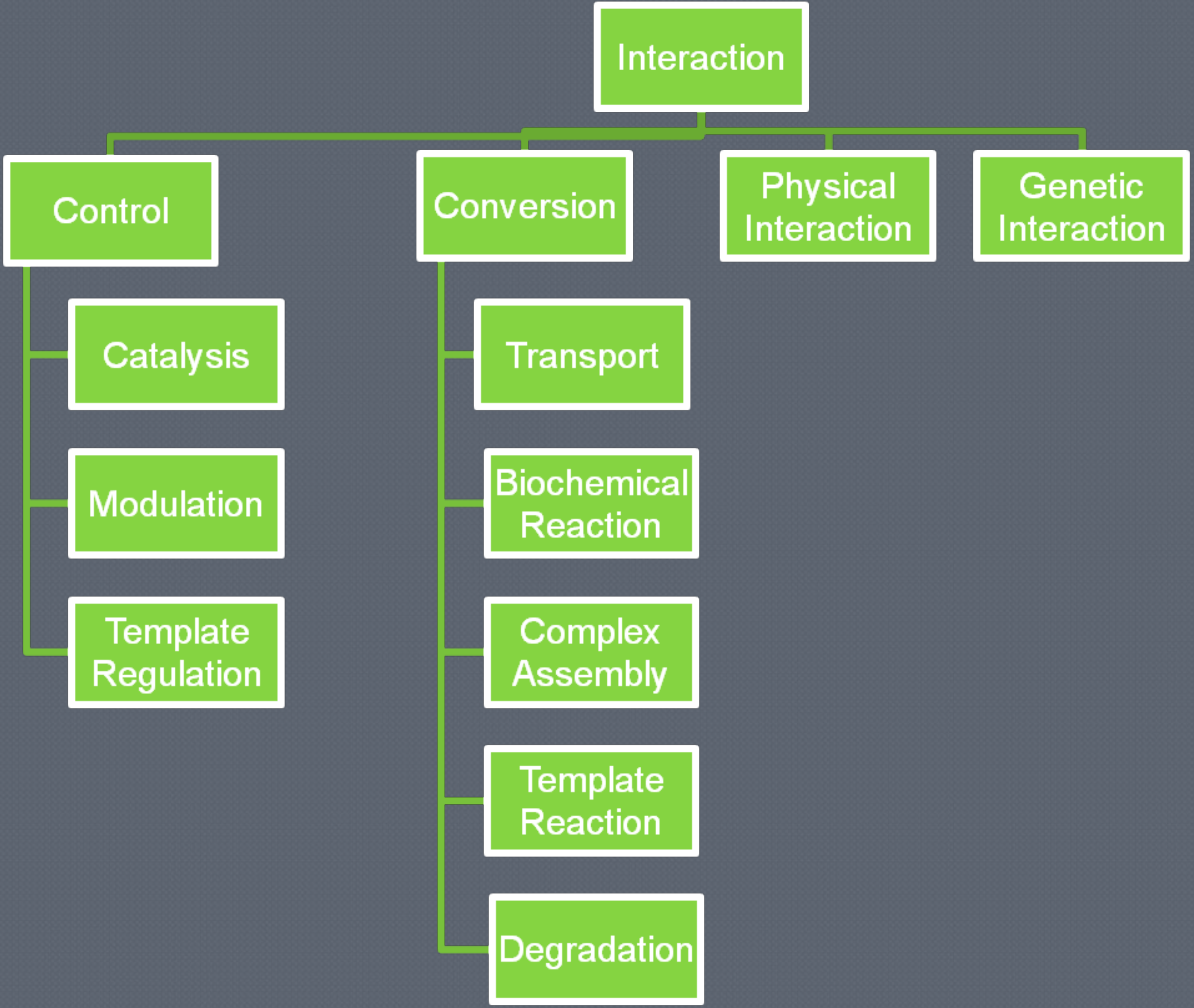
Reference DNA

Reference RNA

Reference Protein

Reference Small Mol.

Interactions



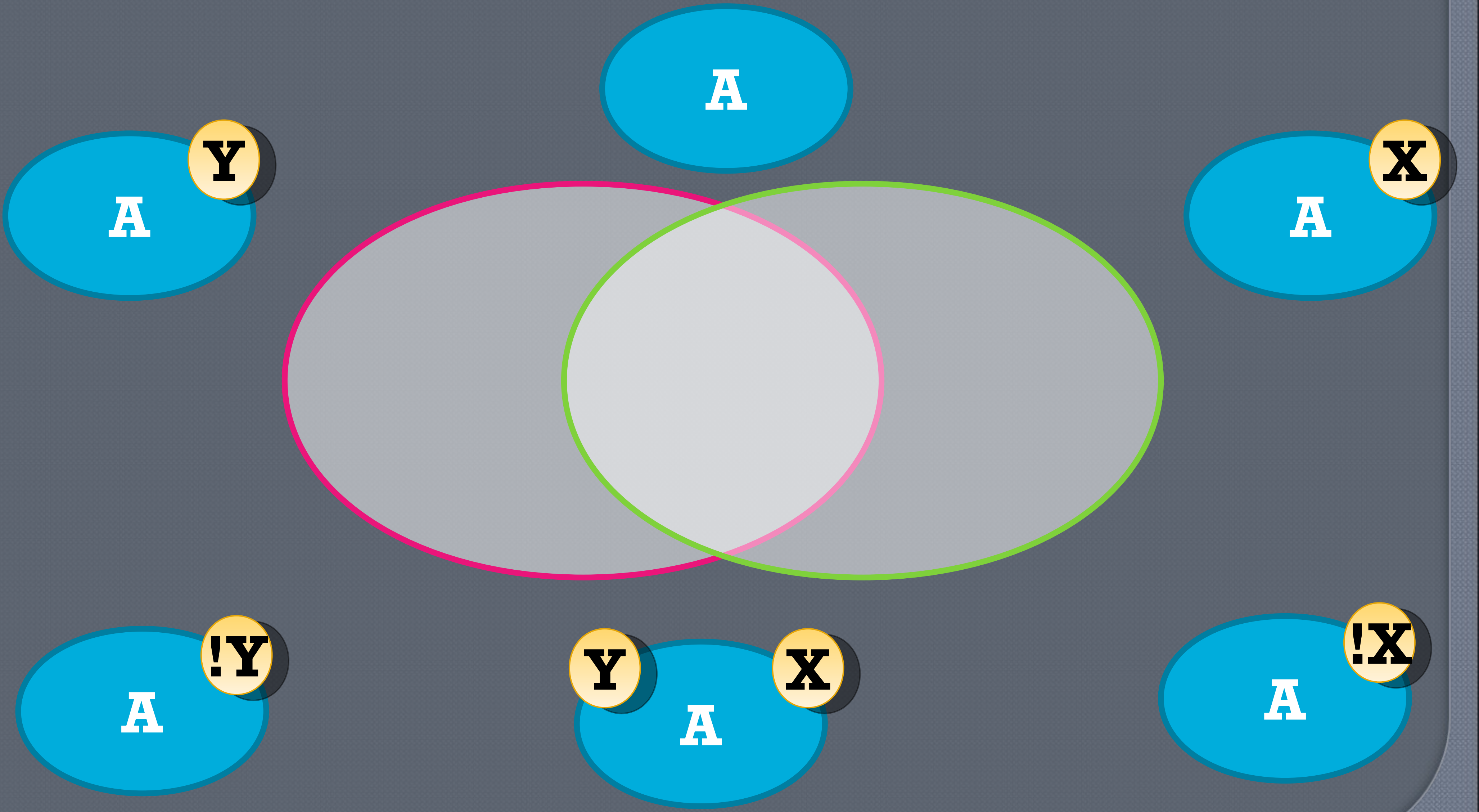
- Interaction
  - Control
    - Catalysis
    - Modulation
    - Template Regulation
  - Conversion
    - Transport
    - Biochemical Reaction
    - Complex Assembly
    - Template Reaction
    - Degradation
  - Physical Interaction
  - Genetic Interaction

Physical Entity and Reference Entity

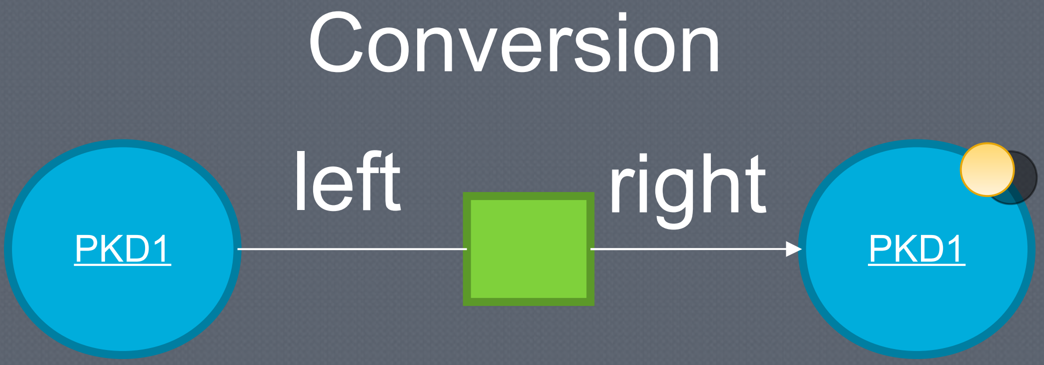
- rPKD1 is a Reference Protein

- PKD1(p) is a Protein
- PKD1(p) has reference PKD1
- P-Ser147 is a ModificationFeature
- PKD1(p) has feature P-Ser147

Semantics of Features



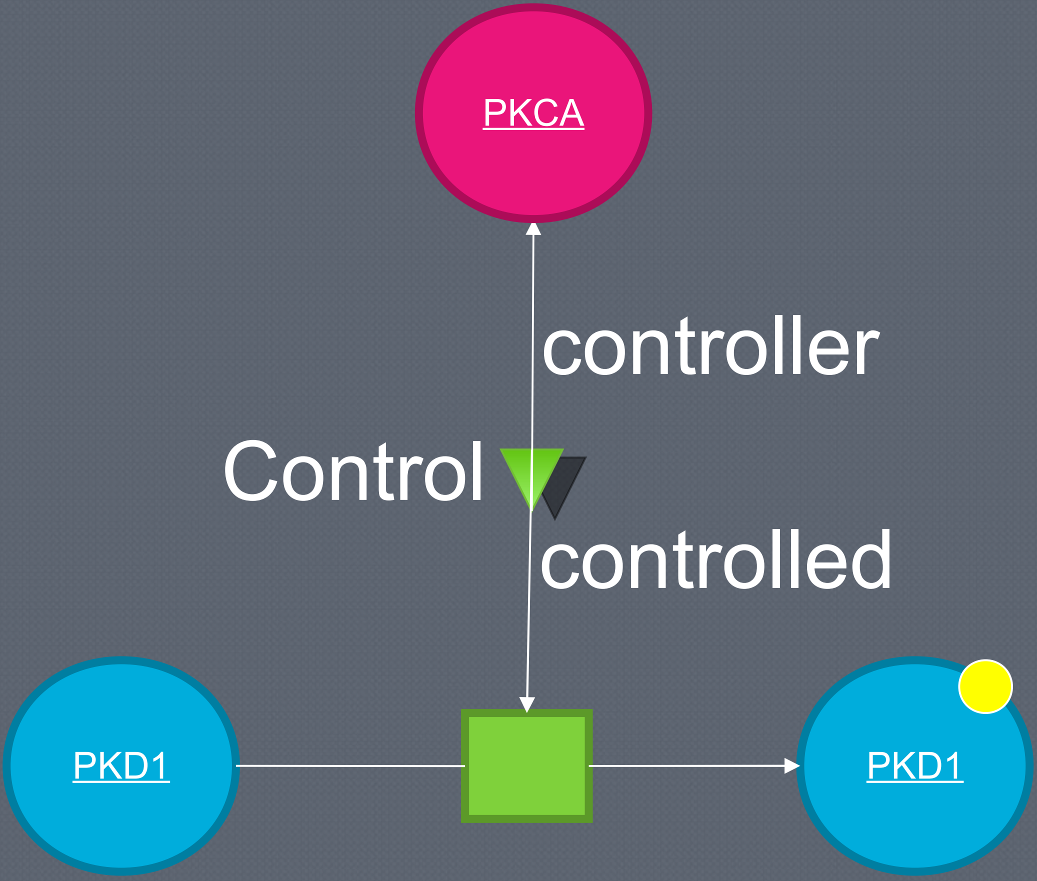
Conversion



PKD1(!p) has noFeature P-Ser147

PKD1(p) has feature P-Ser147

Control



PKCA

controller

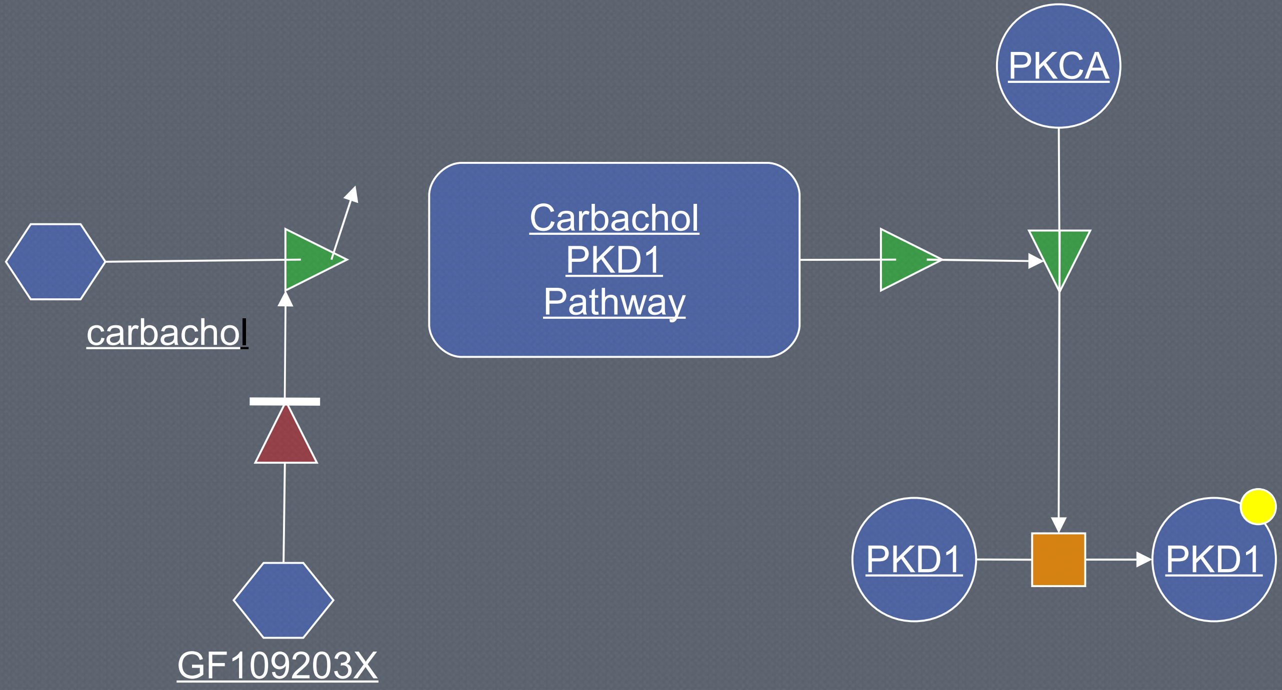
Control

controlled

PKD1

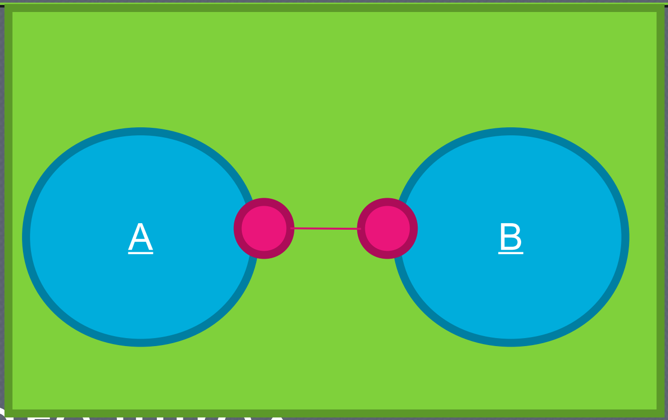
PKD1

Overview


PKCA
Carbachol
PKD1
Pathway
carbachol
GF109203X
PKD1
PKD1

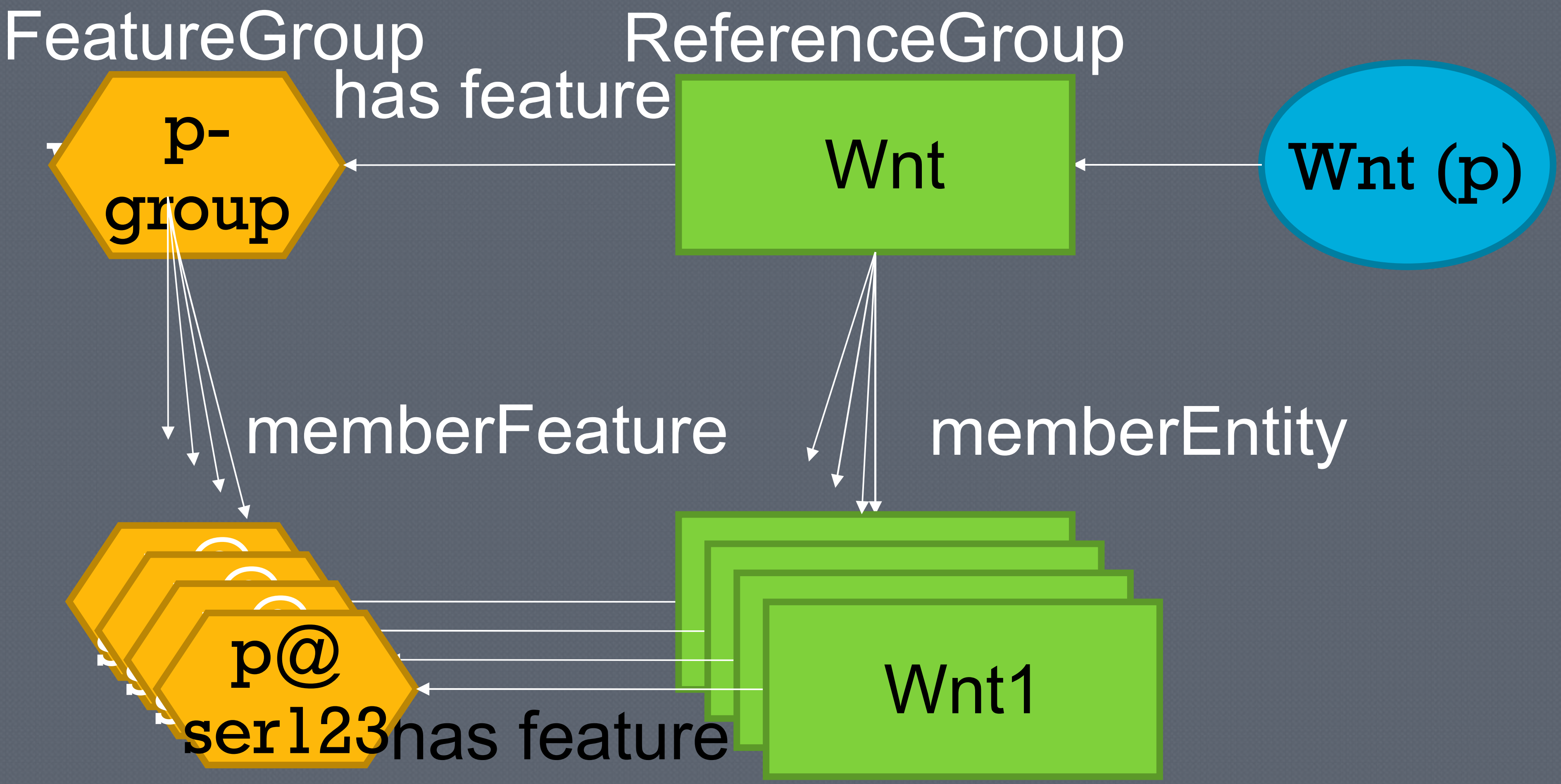
Complexes

- Binding Feature



- Also possible :
  - protein A has notFeature bindingFeaturex
  - Subcomplexes
  - Addressing the complex member

Homology Generics



FeatureGroup

ReferenceGroup

has feature

p-group

Wnt

Wnt (p)

memberFeature

memberEntity

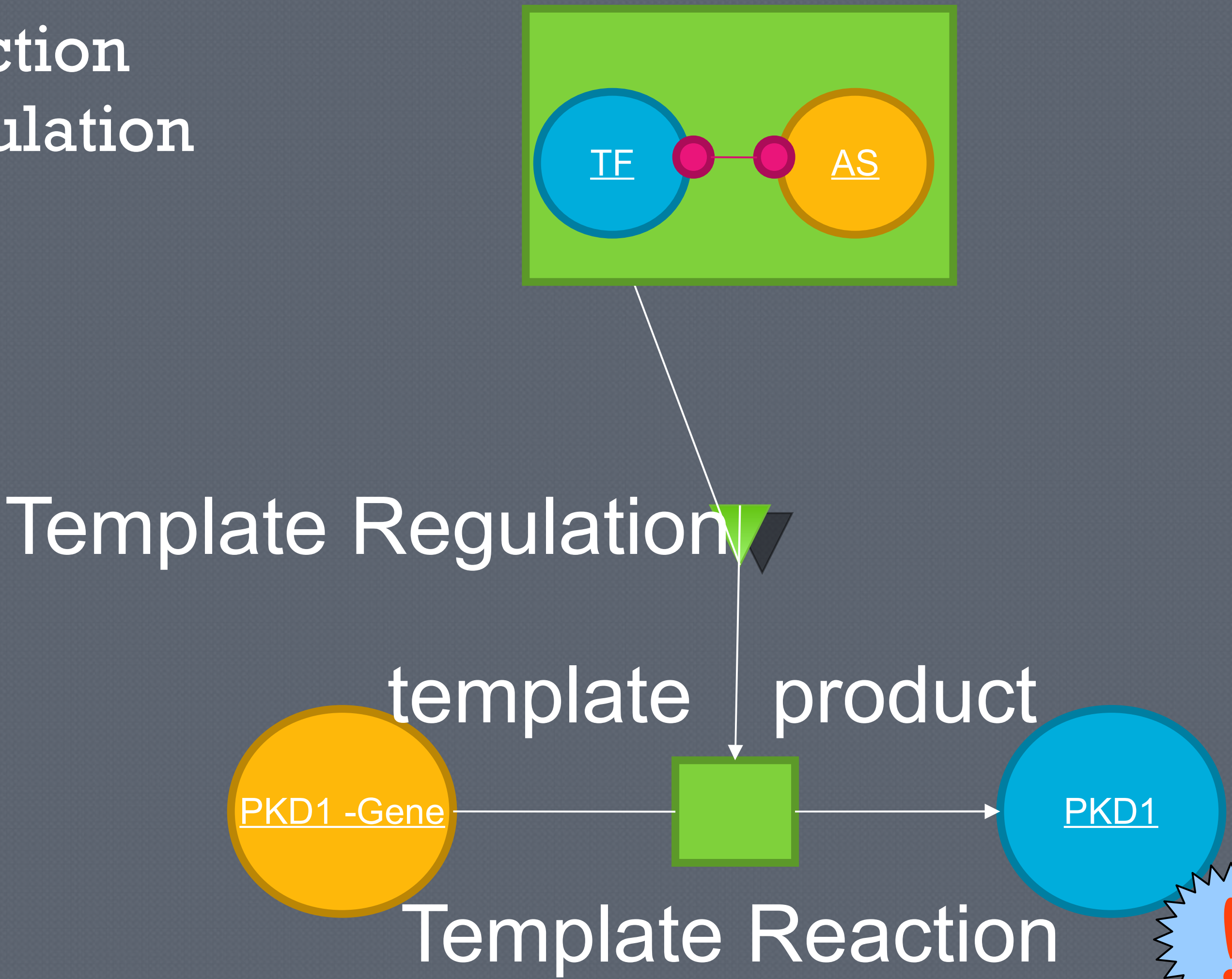
p@ ser123

has feature

Wnt1

Gene Regulation

- Two new interactions:
  - Template Reaction
  - Template Regulation



TF

AS

Template Regulation

template

product

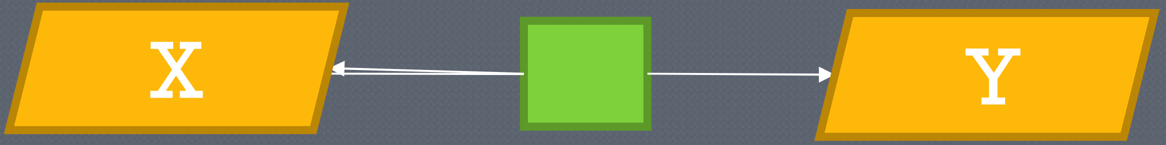
PKD1 -Gene

PKD1

Template Reaction

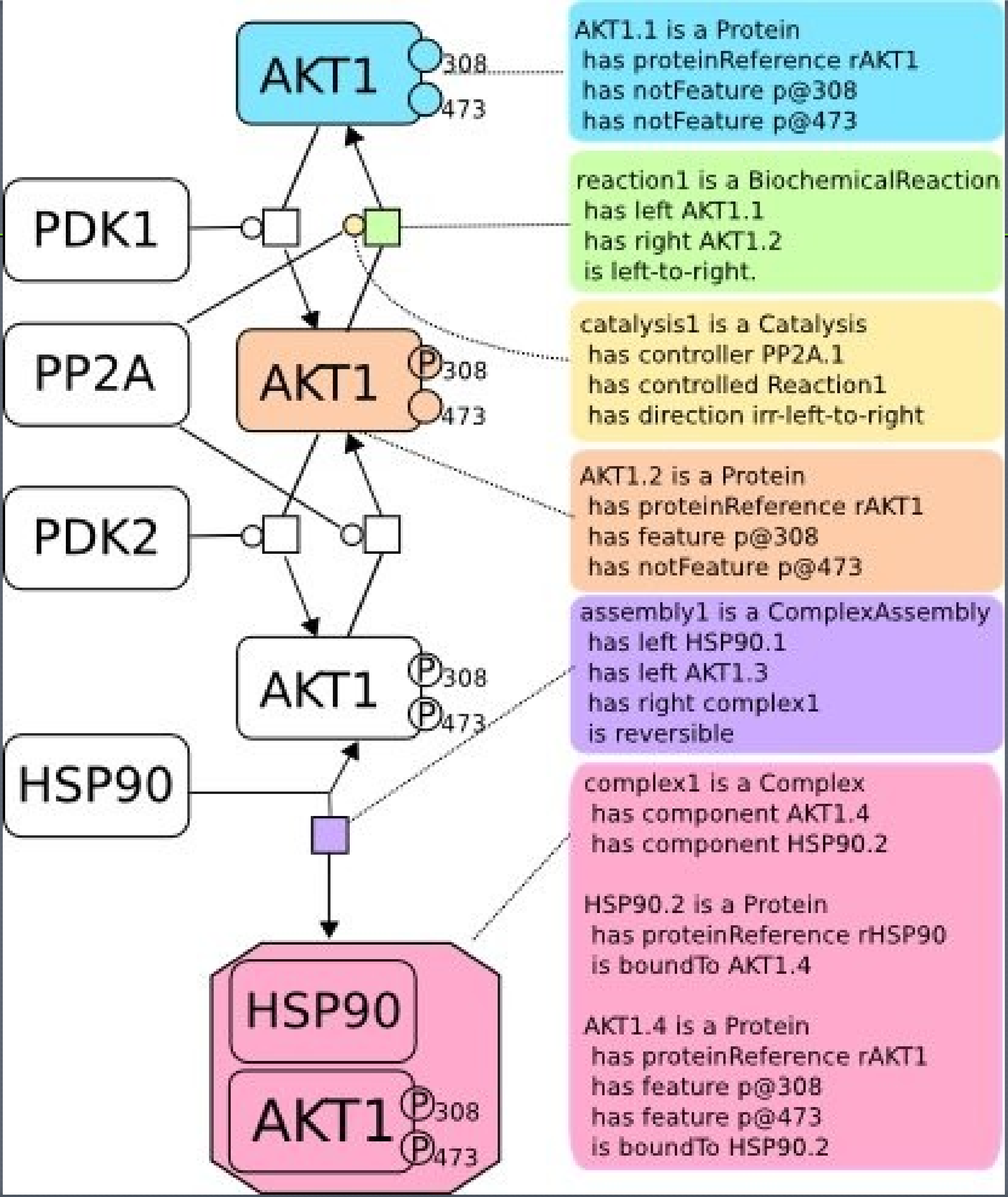
Genetic Interactions

- Effectively perturbation, phenotype n-tuples
  - X knocked out, Y knocked out, mice died



Genetic Interaction

Perturbation X knockout
Perturbation Y knockout
Phenotype Lethality



AKT1.1 is a Protein
has proteinReference rAKT1
has notFeature p@308
has notFeature p@473

reaction1 is a BiochemicalReaction
has left AKT1.1
has right AKT1.2
is left-to-right.

catalysis1 is a Catalysis
has controller PP2A.1
has controlled Reaction1
has direction irr-left-to-right

AKT1.2 is a Protein
has proteinReference rAKT1
has feature p@308
has notFeature p@473

assembly1 is a ComplexAssembly
has left HSP90.1
has left AKT1.3
has right complex1
is reversible

complex1 is a Complex
has component AKT1.4
has component HSP90.2

HSP90.2 is a Protein
has proteinReference rHSP90
is boundTo AKT1.4

AKT1.4 is a Protein
has proteinReference rAKT1
has feature p@308
has feature p@473
is boundTo HSP90.2

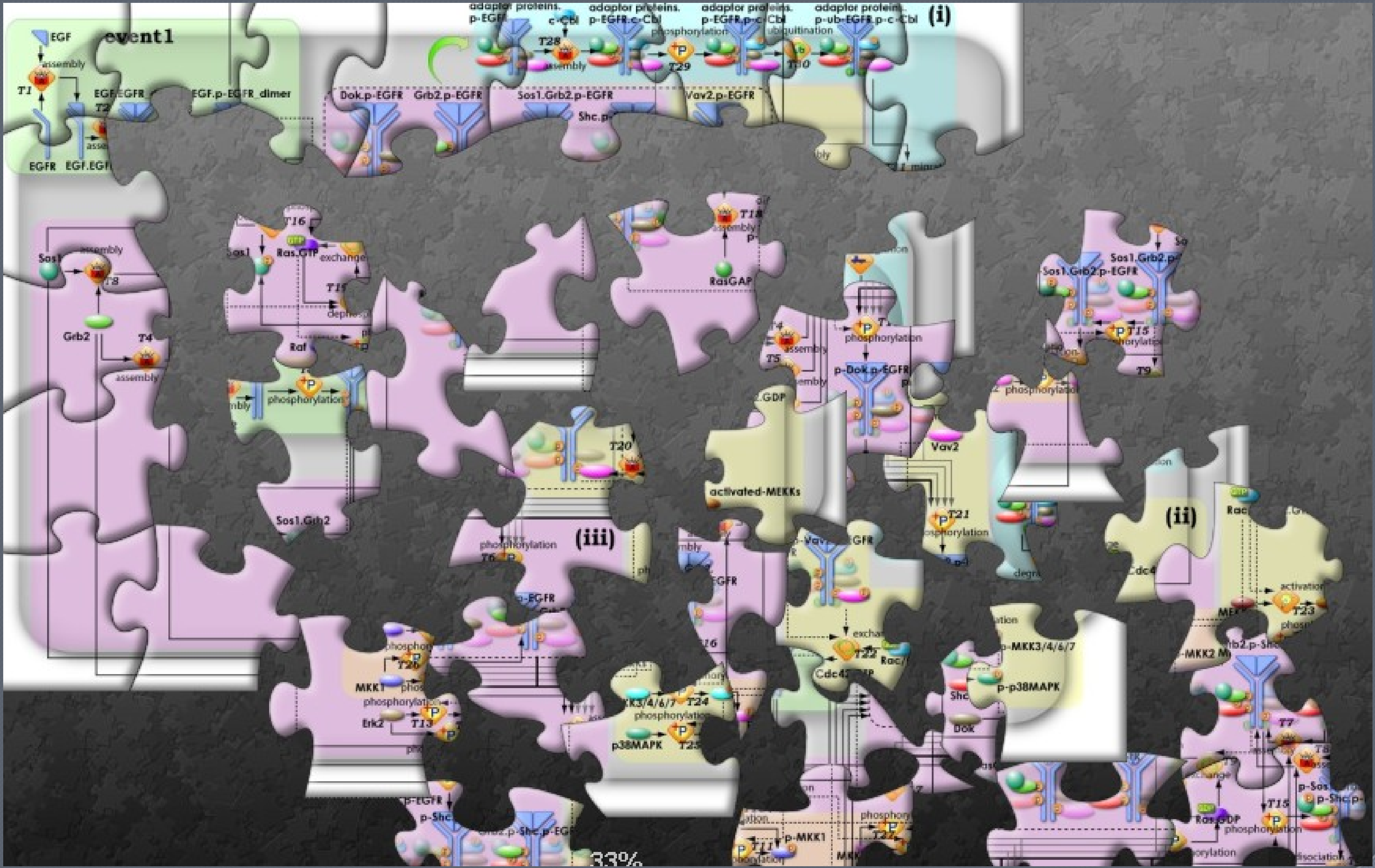
The world outside...

- Extensive use of external Cvs:
  - PSI-MI for protein modifications
  - GO for cellular components
  - SO for dna regions
  -
- External linking
  - Unification
  - Relationship
  - Publication

BioPAX infrastructure

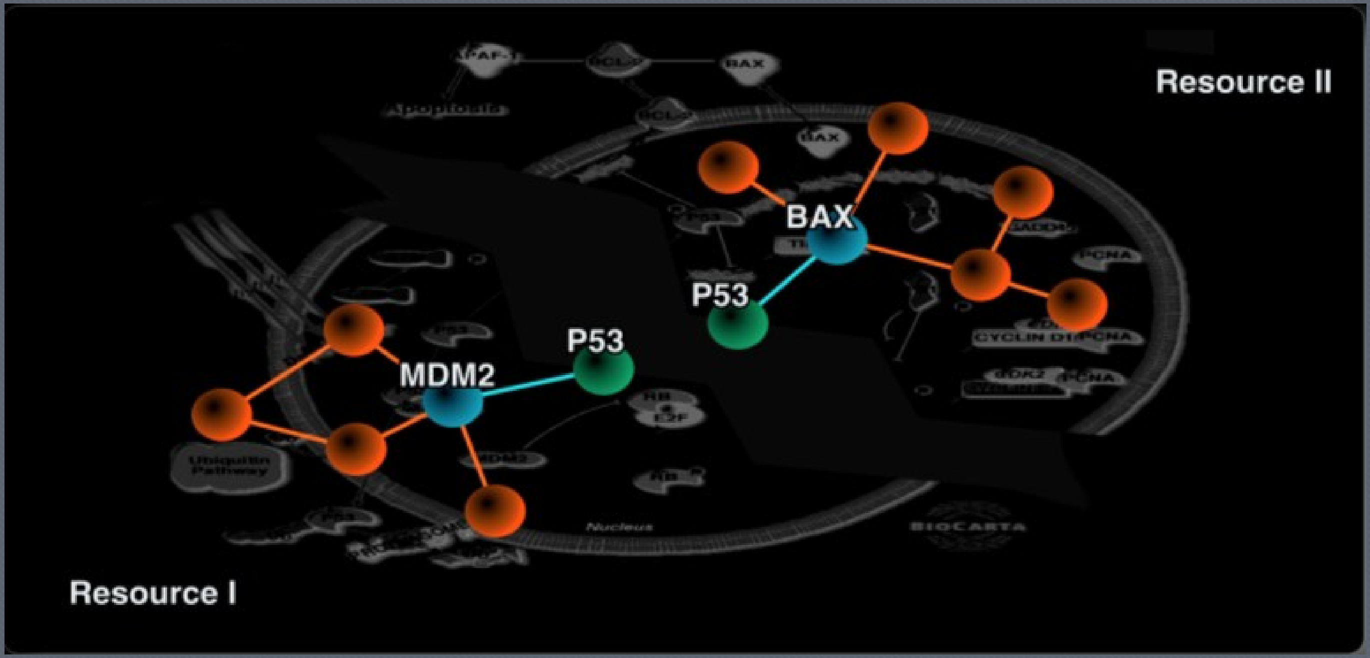
- Pathway Commons: Repository for BioPAX models: Model normalization and aggregation
- Paxtools : Java library for BioPAX manipulation and validation
- PSI-MI conversion
- Visualization : Cytoscape, PATIKA, Chisio
- caBIG

Horizontal Integration



Horizontal Integration

- Integration is difficult
  - Different notions and representations.
  - Need for concurrency.
  - Different levels of detail.
  - Incomplete/Ambiguous knowledge



A fuzzy approach

- Merge Entity References based on external IDs
- Define a similarity score between different physical entities.
- Find reactions that have similar input/output sets
- “Graph-align” networks

- Also a cross-validator!

Acknowledgements

- BioPAX community !
- Editors to the last level : Gary Bader, Ken Fukuda
- Chris Sander, Ethan Cerami, Benjamin Gross
- Özgün Babur, Arman Aksoy
-
- Support by :