

Modelling **pathways crosstalk**
in gene/protein networks as
closed walks & cycles on graphs

introduction

basic concepts

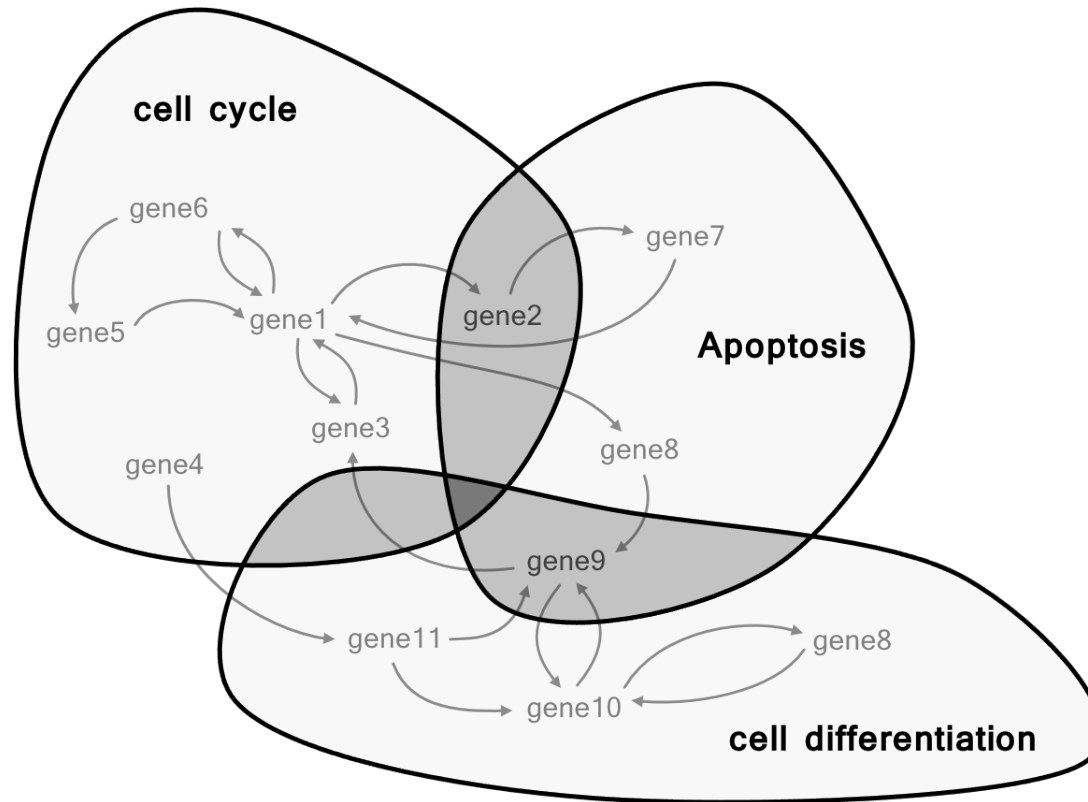
research goals

methods

results

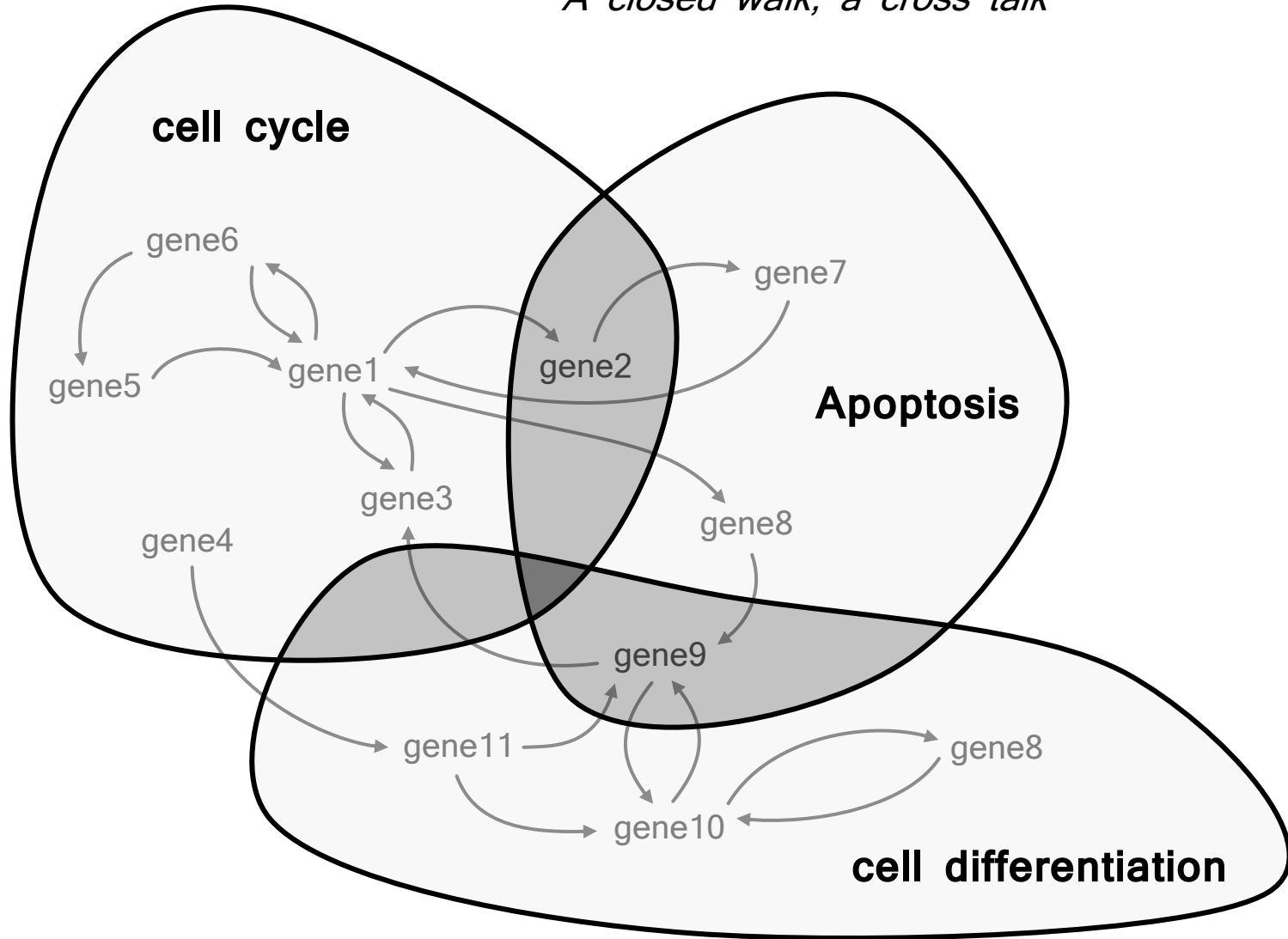
perspectives

Modelling **pathways crosstalk** in gene/protein networks as closed walks & cycles on graphs



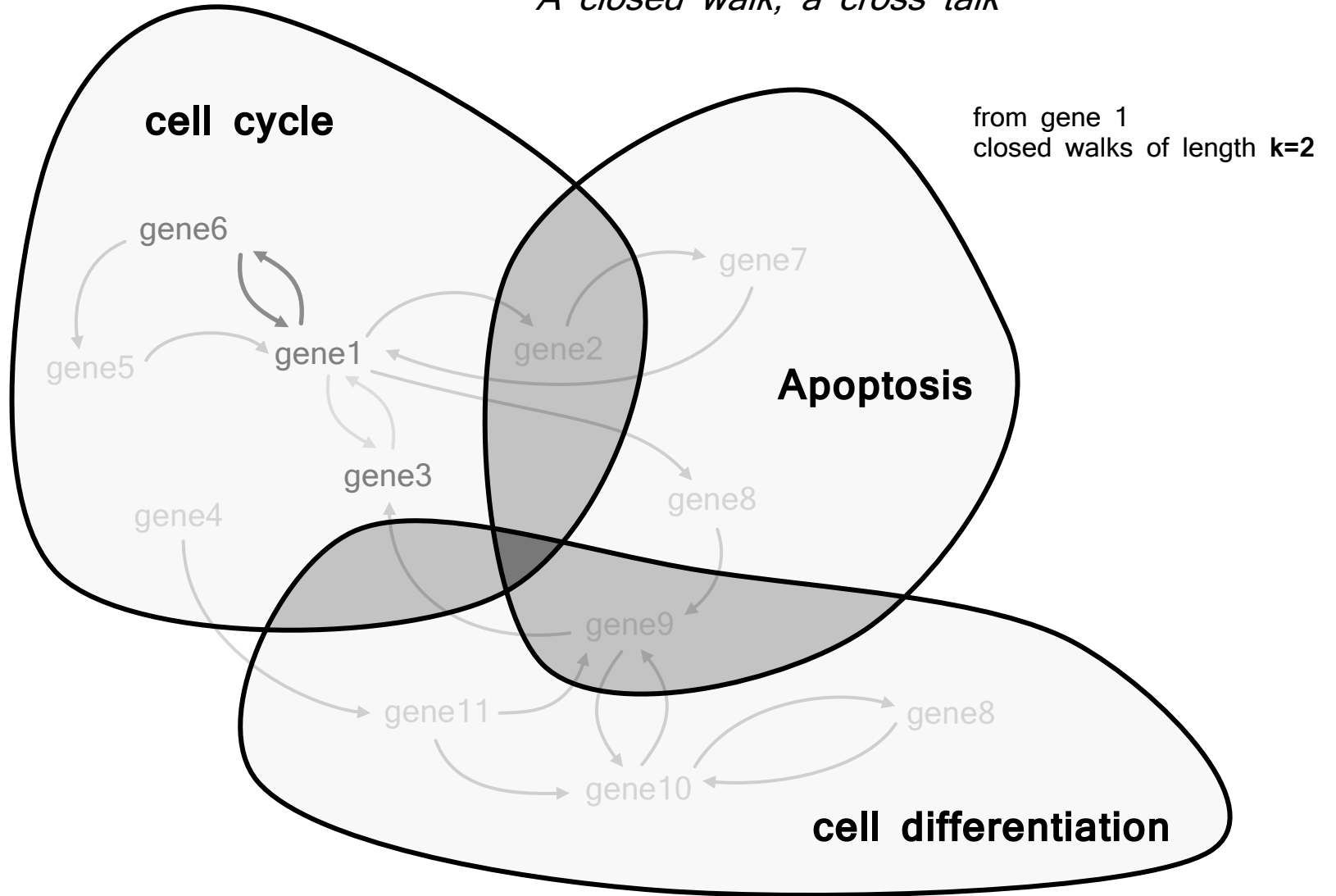
introduction
basic concepts
research goals
methods
results
perspectives

*A graph, a directed graph, nodes and edges.
A closed walk, a cross talk*



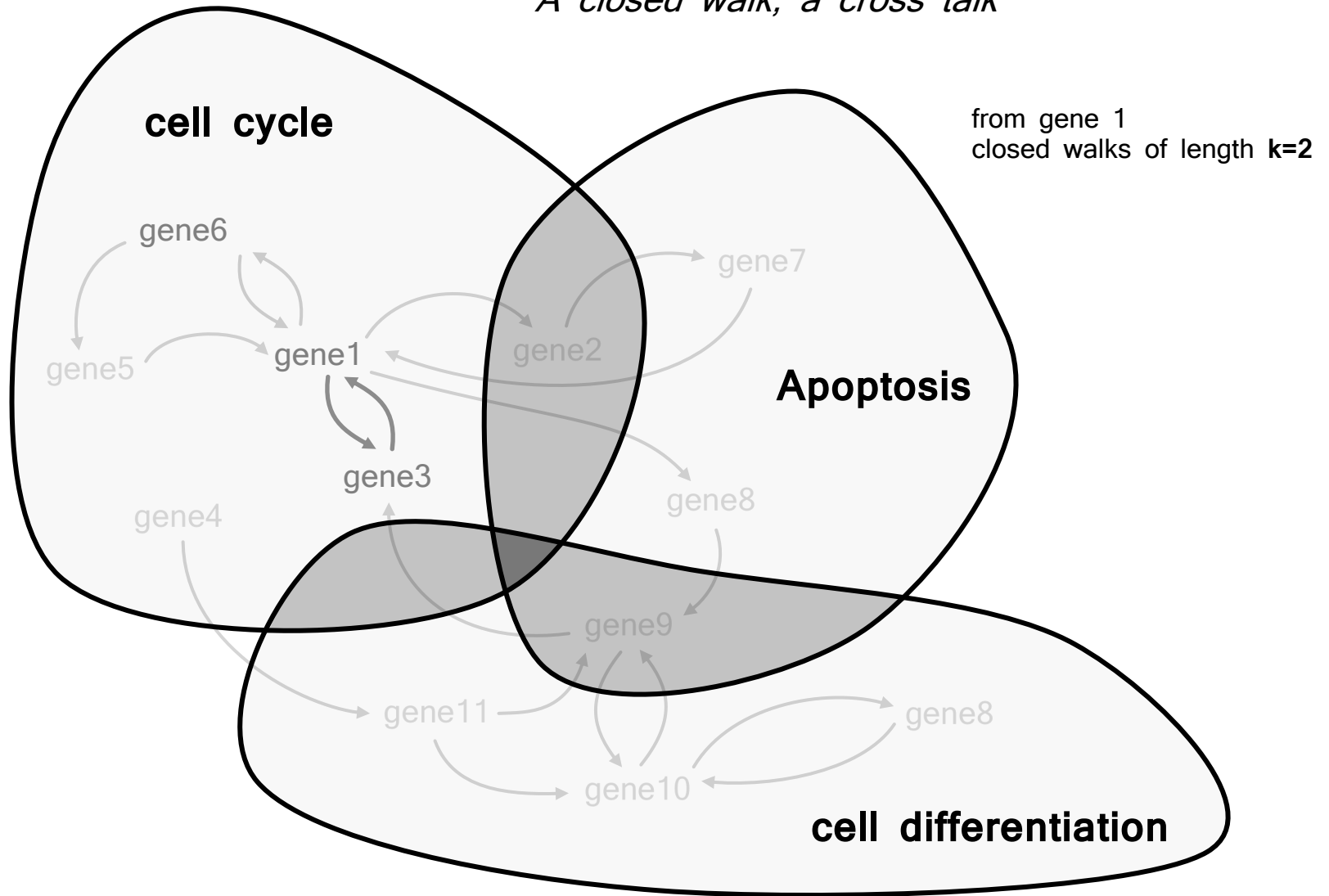
introduction
basic concepts
research goals
methods
results
perspectives

*A graph, a directed graph, nodes and edges.
A closed walk, a cross talk*



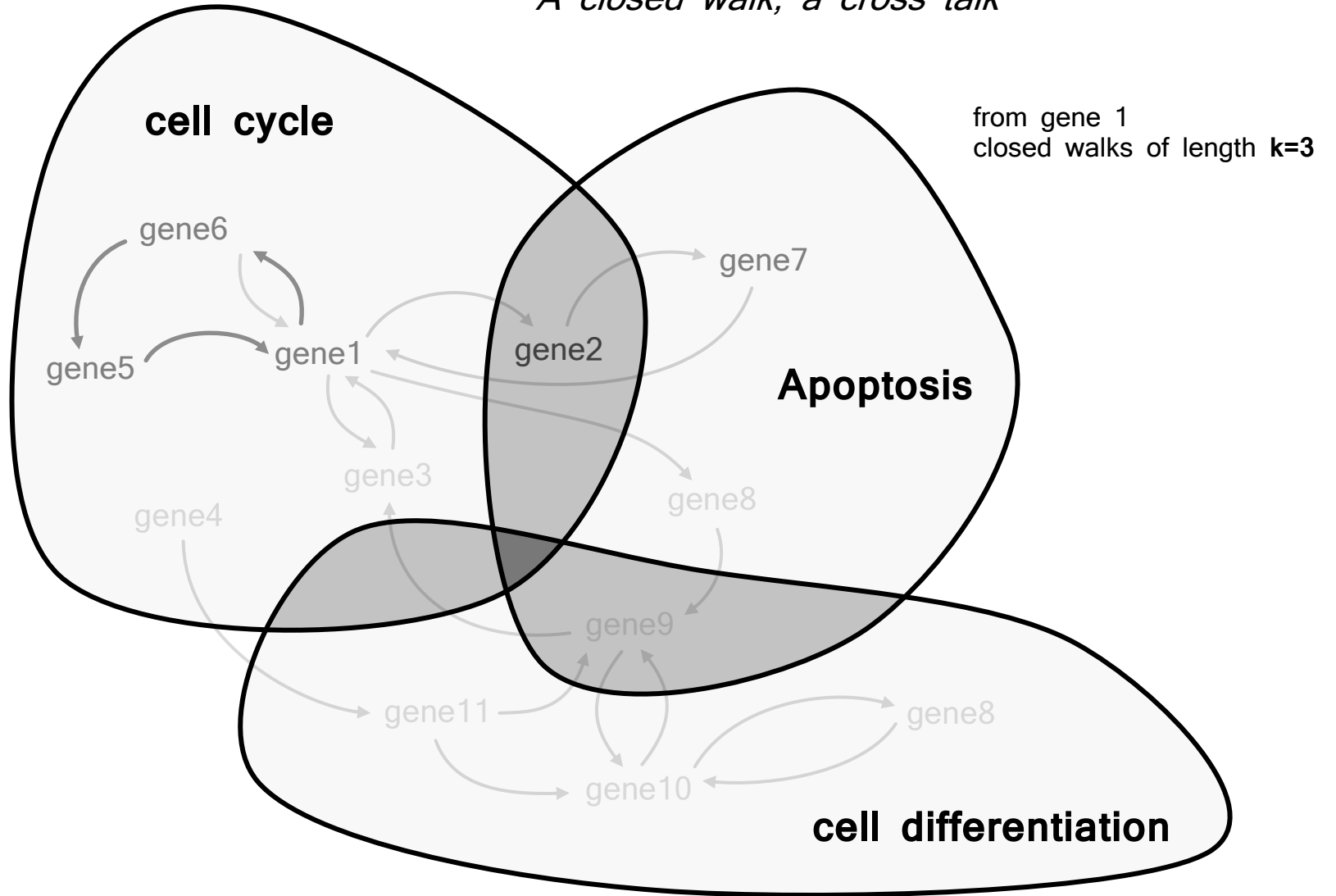
introduction
basic concepts
research goals
methods
results
perspectives

*A graph, a directed graph, nodes and edges.
A closed walk, a cross talk*



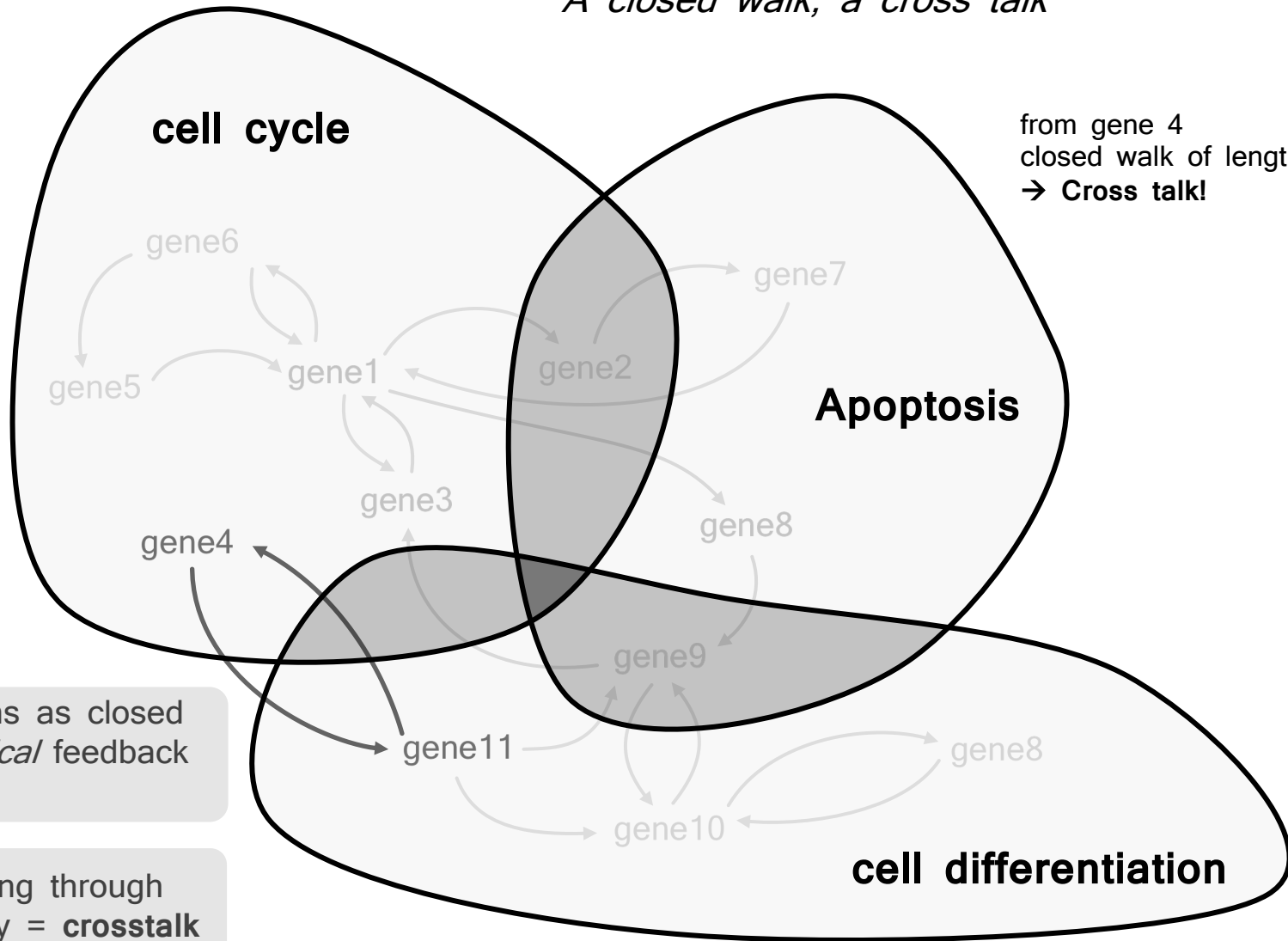
introduction
basic concepts
research goals
methods
results
perspectives

*A graph, a directed graph, nodes and edges.
A closed walk, a cross talk*



introduction
basic concepts
research goals
methods
results
perspectives

*A graph, a directed graph, nodes and edges.
A closed walk, a cross talk*



from gene 4
closed walk of length $k=2$
→ **Cross talk!**

gene interactions as closed walks \approx *biological* feedback loops

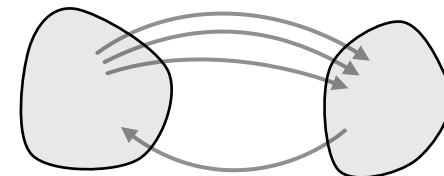
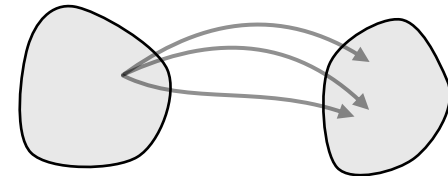
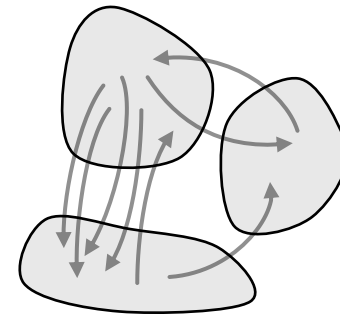
closed walk going through another pathway = **crosstalk**

introduction
basic concepts
research goals
methods
results
perspectives

Are some functional pathways more prone to crosstalk than the others? If so, which ones?

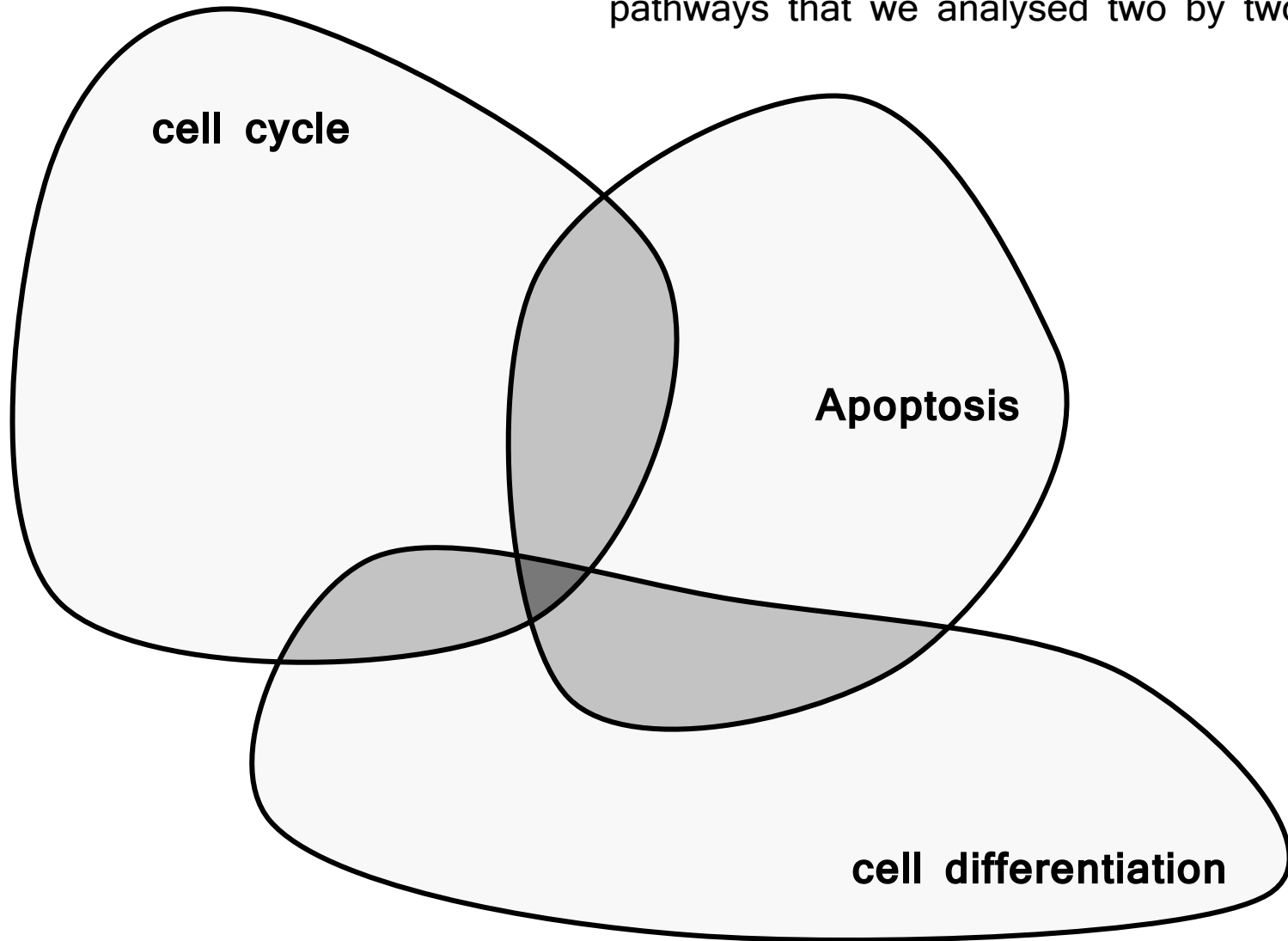
What are the genes that are apparent “entry points” for the crosstalk with other pathways?

Is crosstalk a symmetric process, i.e. do both pathways “need each other” the same, or is one rather “exploiting” the other for input stimuli?



introduction
basic concepts
research goals
methods
results
perspectives

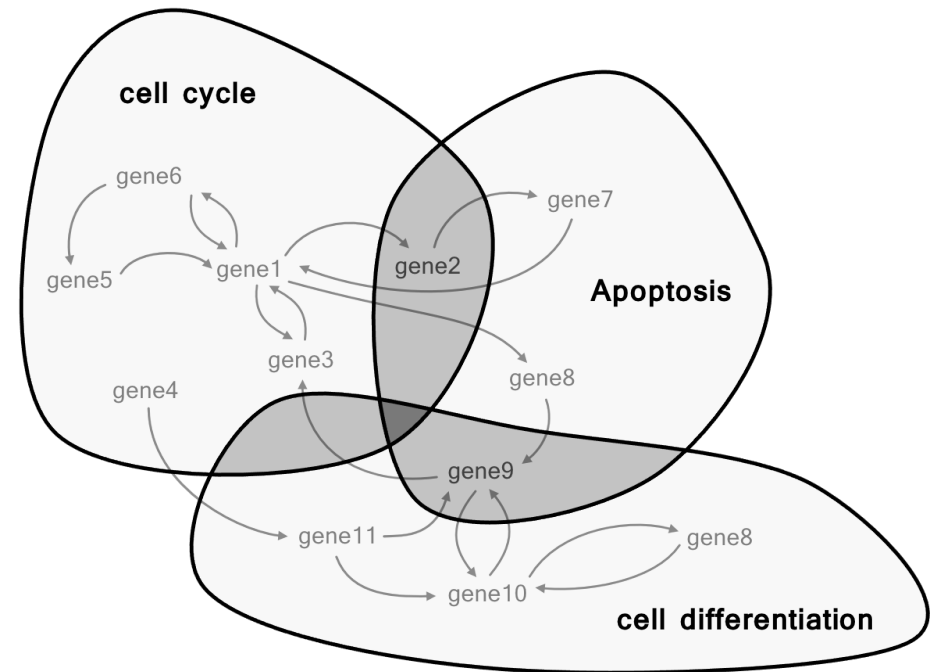
We restricted our analysis to three functional pathways that we analysed two by two



introduction
basic concepts
research goals
methods
results
perspectives

Modelling **pathways crosstalk** in gene/protein networks as closed walks & cycles on graphs

Source	Target	Sign
Gene1	Gene2	1
Gene1	Gene3	1
Gene1	Gene8	1
Gene1	Gene6	1
Gene6	Gene1	1
Gene5	Gene1	1
Gene3	Gene1	1



introduction
basic concepts
research goals
methods
results
perspectives

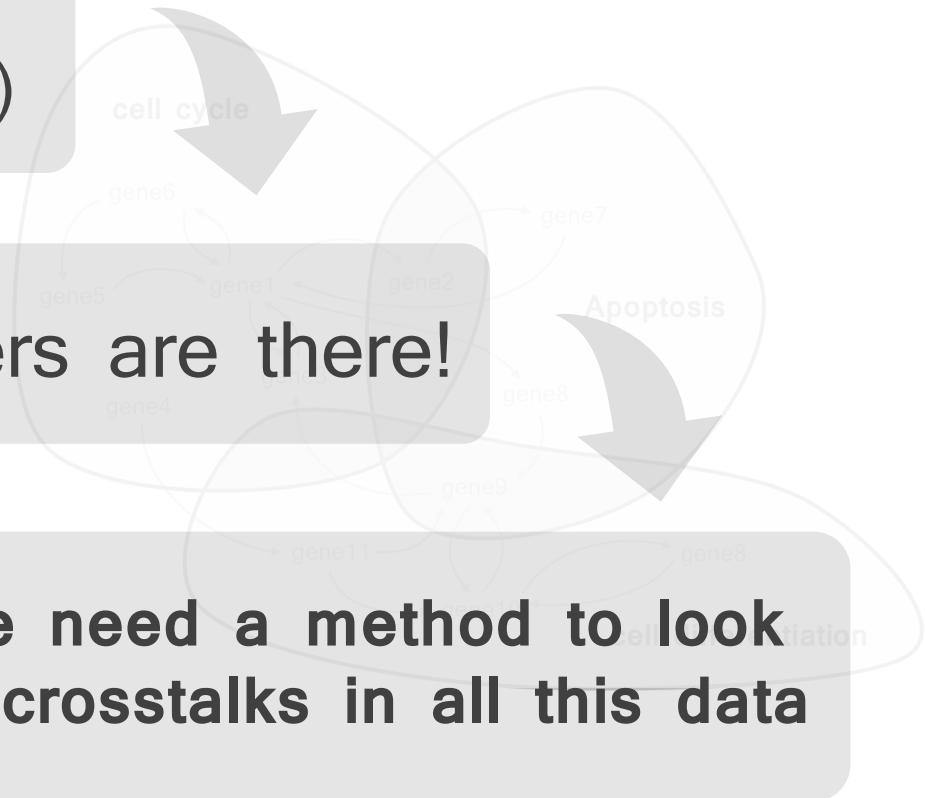
Modelling **pathways crosstalk** in gene/protein networks as closed walks & cycles on graphs

the datasets are huge!
(edgelist: 118'355 entries)

Gene1	Gene2	1
Gene1	Gene3	1
Gene1	Gene8	1
Gene1	Gene6	1
Gene6	Gene1	1
Gene5	Gene1	1
Gene3	Gene1	1

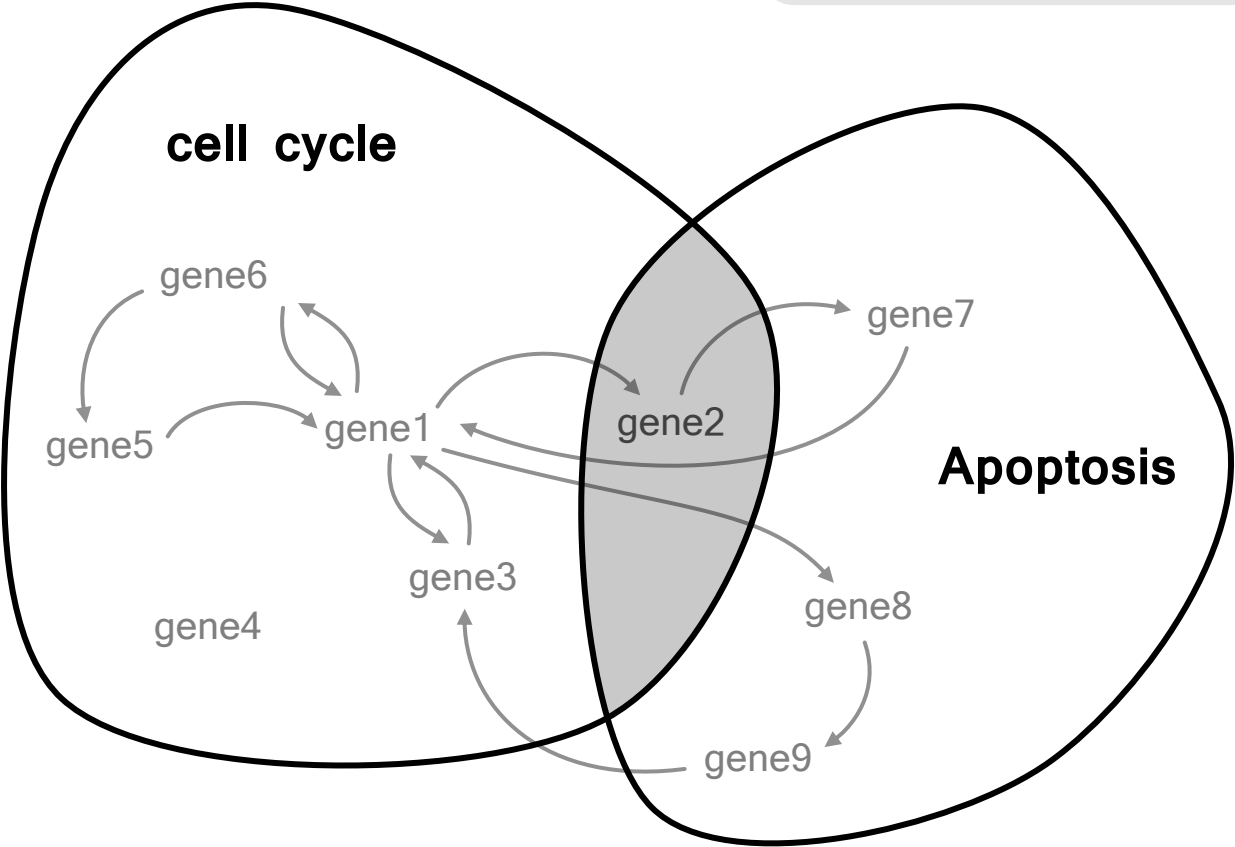
But the answers are there!

We need a method to look
for crosstalks in all this data



- introduction
- basic concepts
- research goals
- methods**
- results
- perspectives

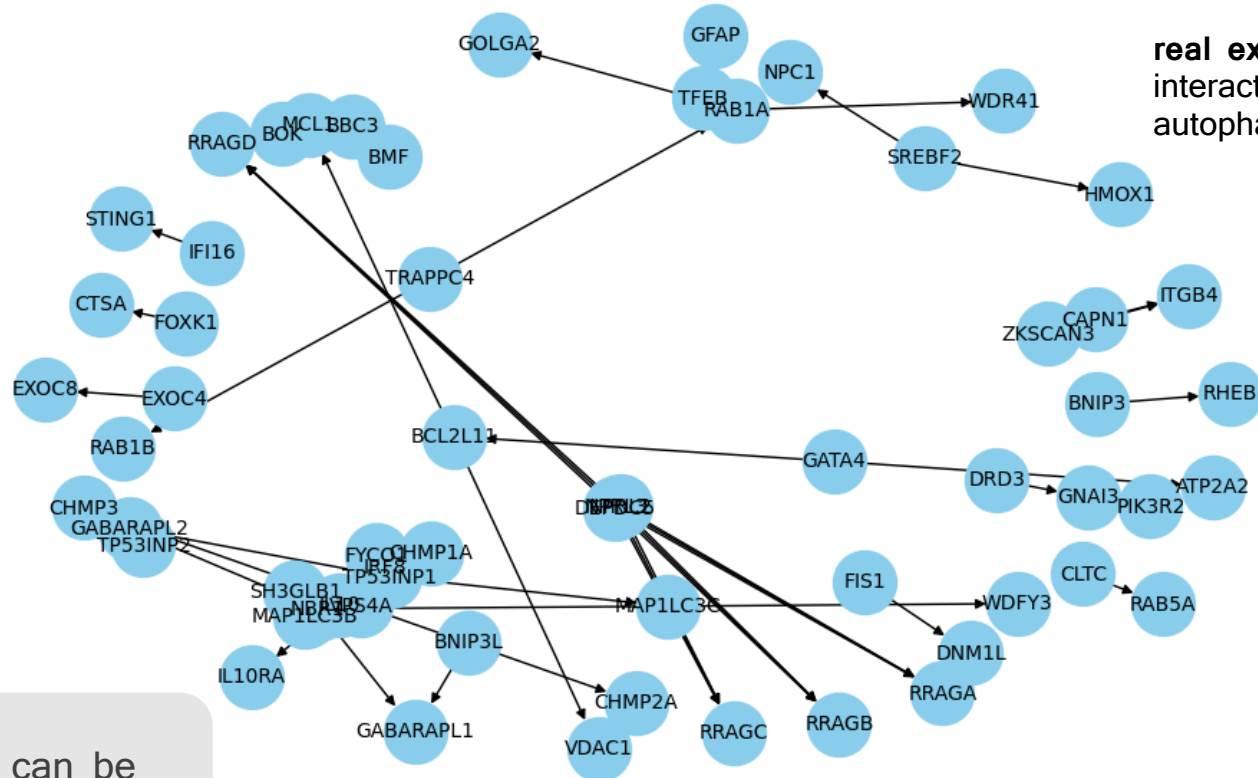
gene interactions can be represented as nodes and edges on a directed graph



introduction
basic concepts
research goals
methods
results
perspectives

*A graph, a directed graph, nodes and edges.
A closed walk, a cycle*

Directed Graph of Autophagy pathway

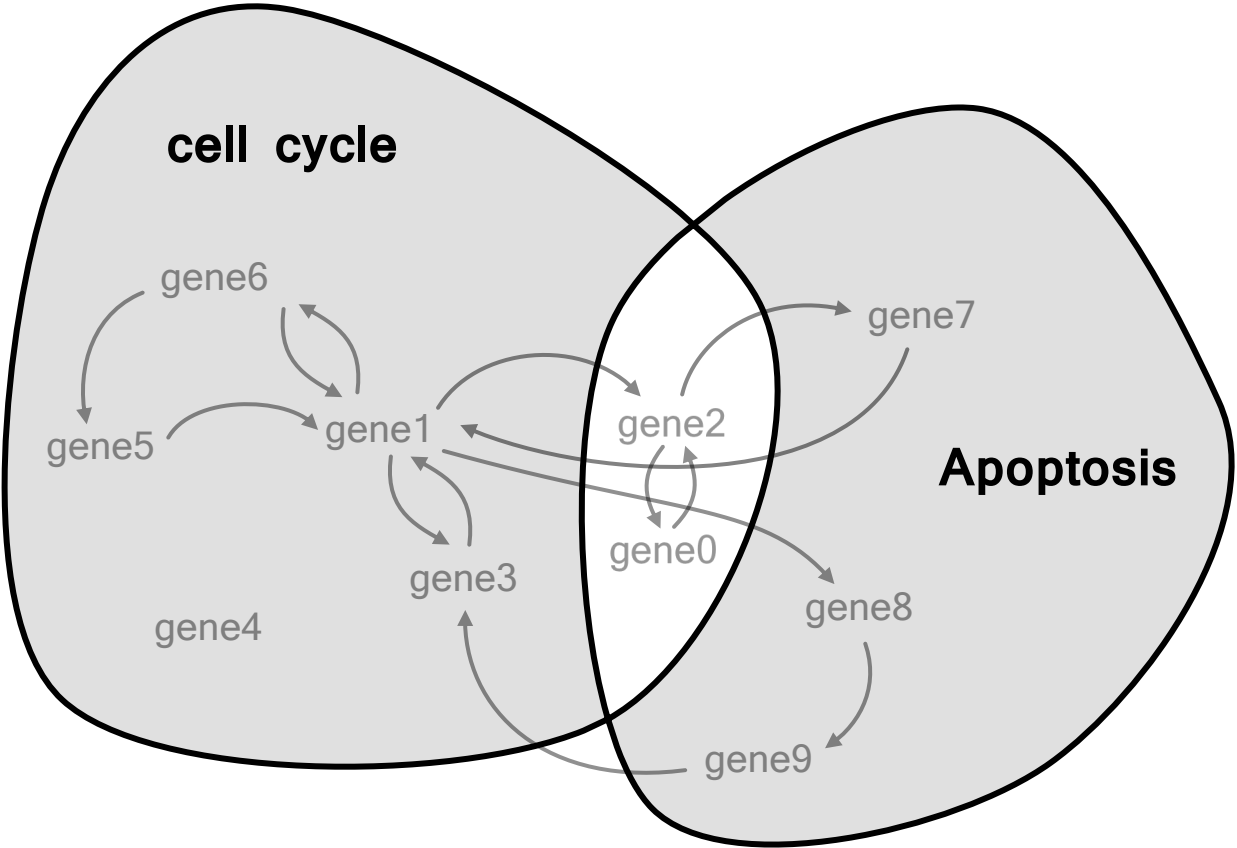


real example of a graph:
interacting genes inside the autophagy pathway

gene interactions can be represented as nodes and edges on a directed graph

introduction
basic concepts
research goals
methods
results
perspectives

we analyse gene pairs to find interactions
we exclude the intersection



introduction
basic concepts
research goals
methods
results
perspectives

we analyse gene pairs to find interactions
we exclude the intersection

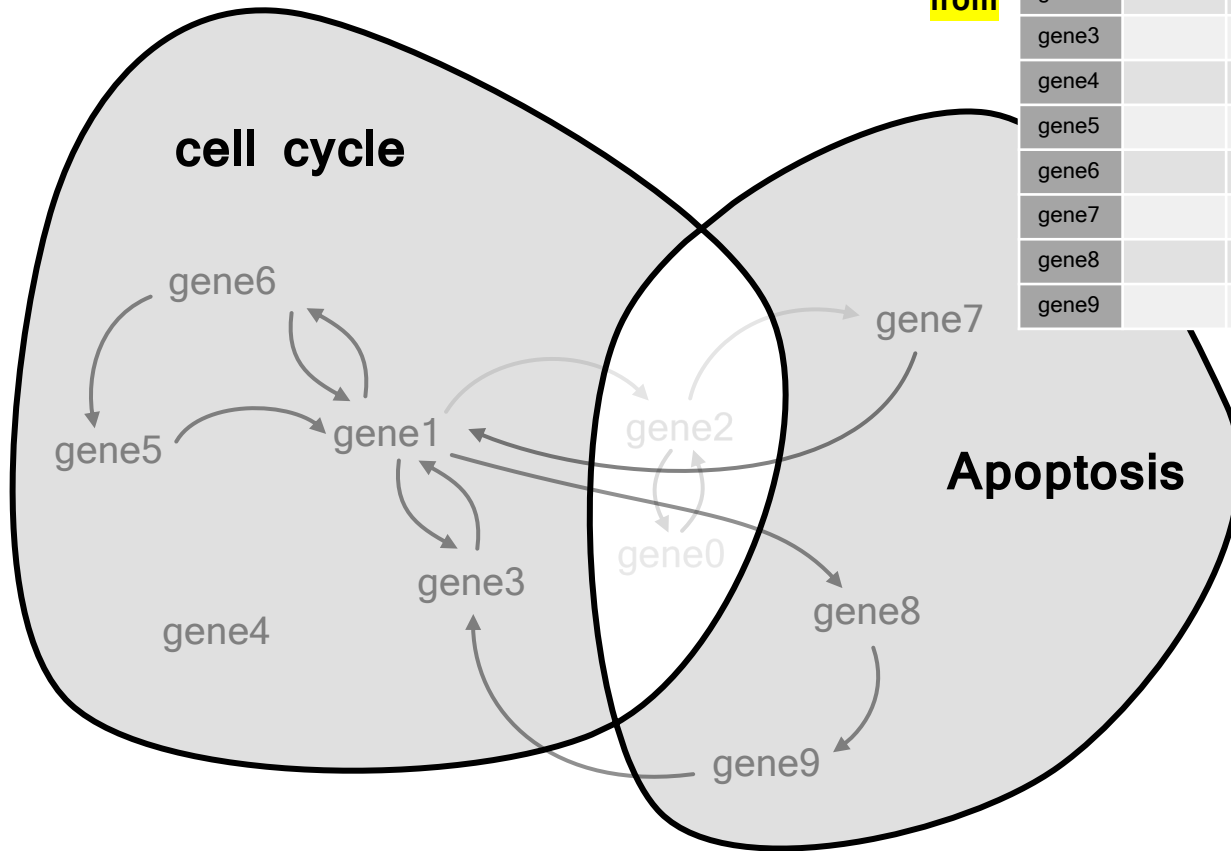


from the graph we can build an **adjacency matrix**

to

from

	gene1	gene3	gene4	gene5	gene6	gene7	gene8	gene9
gene1								
gene3								
gene4								
gene5								
gene6								
gene7								
gene8								
gene9								



introduction
basic concepts
research goals
methods
results
perspectives

we analyse gene pairs to find interactions
we exclude the intersection

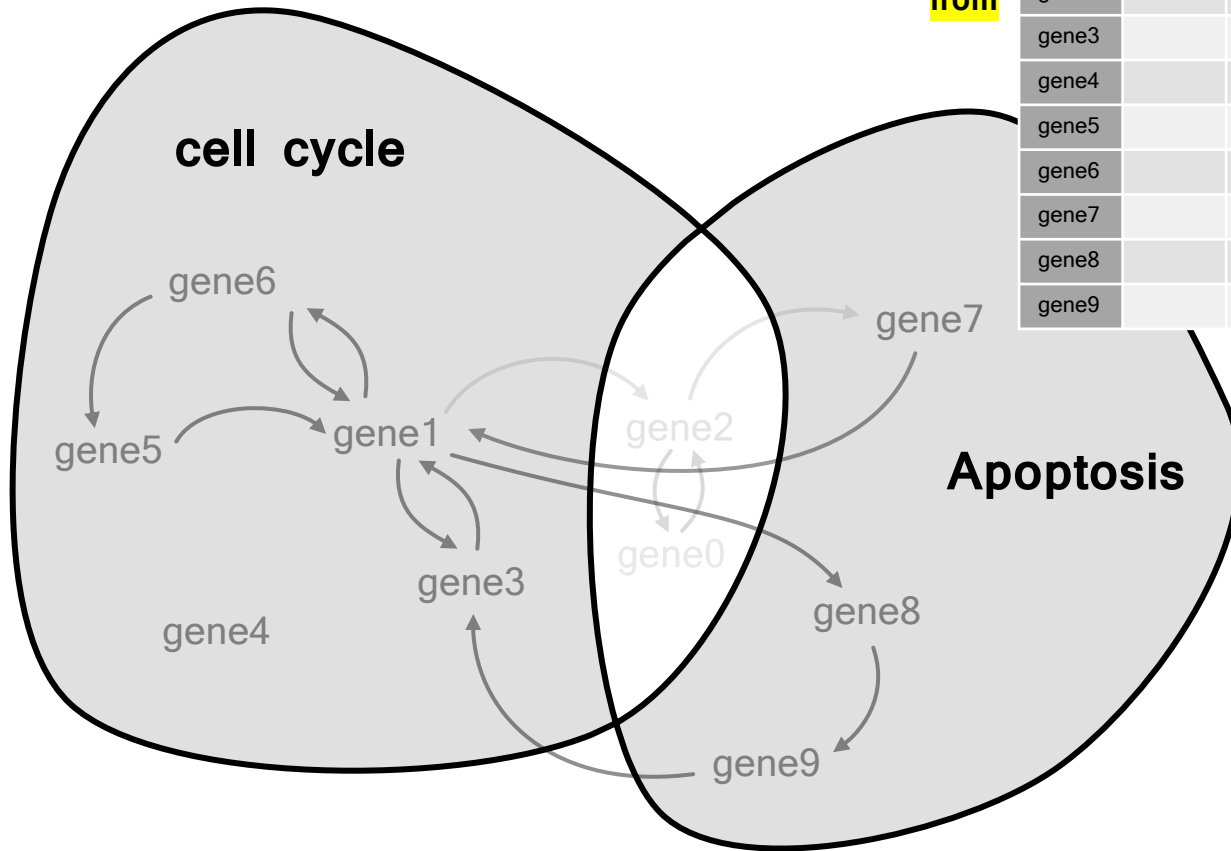


from the graph we can build an **adjacency matrix**

to

from

	gene1	gene3	gene4	gene5	gene6	gene7	gene8	gene9
gene1	0	1	0	0	1	0	1	0
gene3								
gene4								
gene5								
gene6								
gene7								
gene8								
gene9								



introduction
basic concepts
research goals
methods
results
perspectives

we analyse gene pairs to find interactions
we exclude the intersection

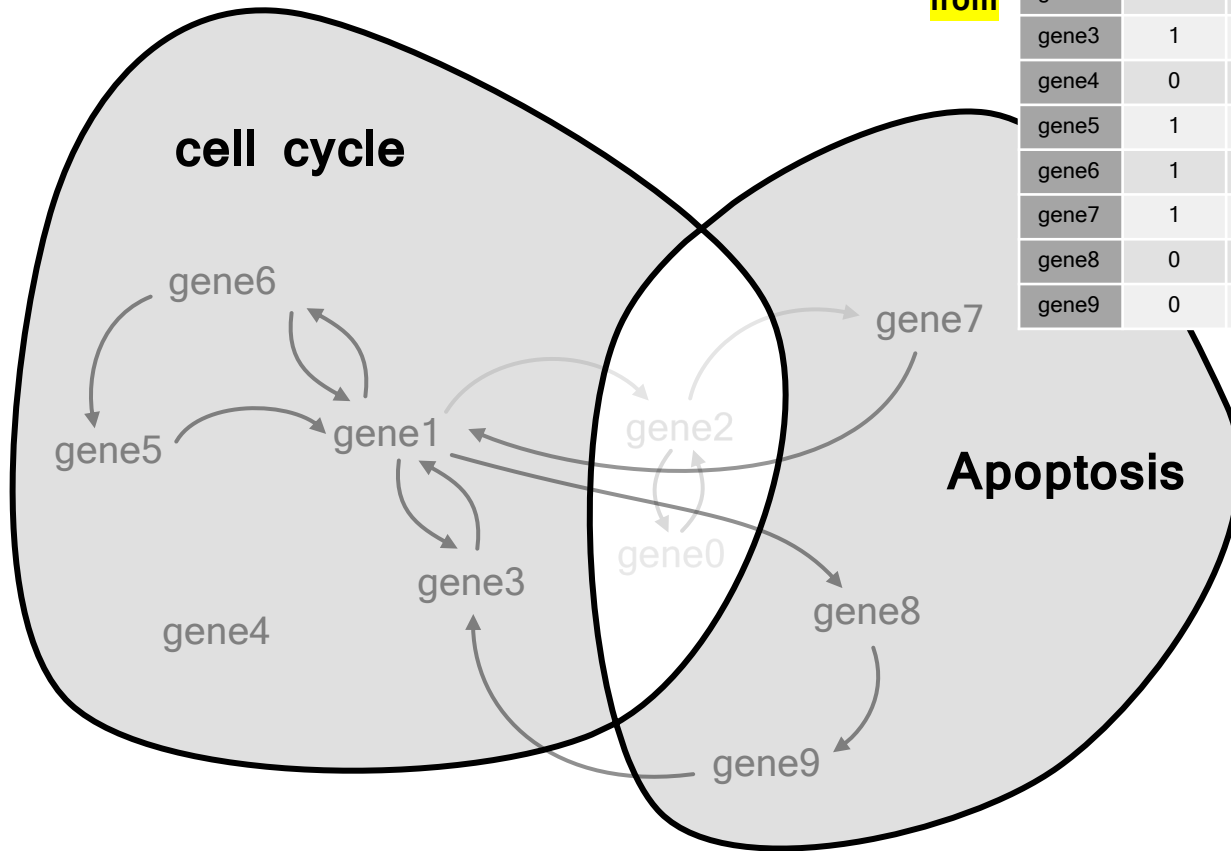


from the graph we can build an **adjacency matrix**

to

from

	gene1	gene3	gene4	gene5	gene6	gene7	gene8	gene9
gene1	0	1	0	0	1	0	1	0
gene3	1	0	0	0	0	0	0	0
gene4	0	0	0	0	0	0	0	0
gene5	1	0	0	0	0	0	0	0
gene6	1	0	0	1	0	0	0	0
gene7	1	0	0	0	0	0	0	0
gene8	0	0	0	0	0	0	0	1
gene9	0	1	0	0	0	0	0	0



introduction
basic concepts
research goals
methods
results
perspectives

we analyse gene pairs to find interactions
we exclude the intersection

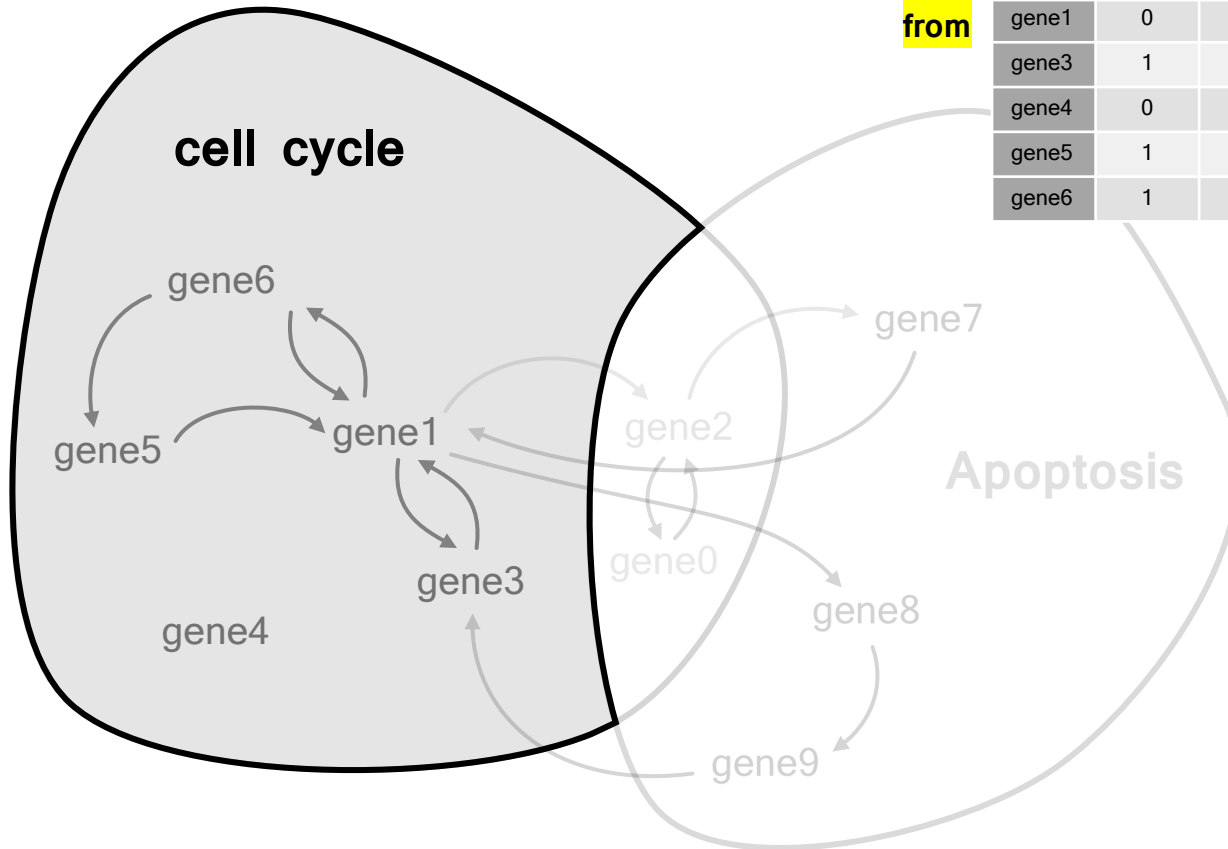


from the graph we can build an **adjacency matrix**

to

from

	gene1	gene3	gene4	gene5	gene6
gene1	0	1	0	0	1
gene3	1	0	0	0	0
gene4	0	0	0	0	0
gene5	1	0	0	0	0
gene6	1	0	0	1	0



introduction
basic concepts
research goals
methods
results
perspectives

we analyse gene pairs to find interactions
we exclude the intersection

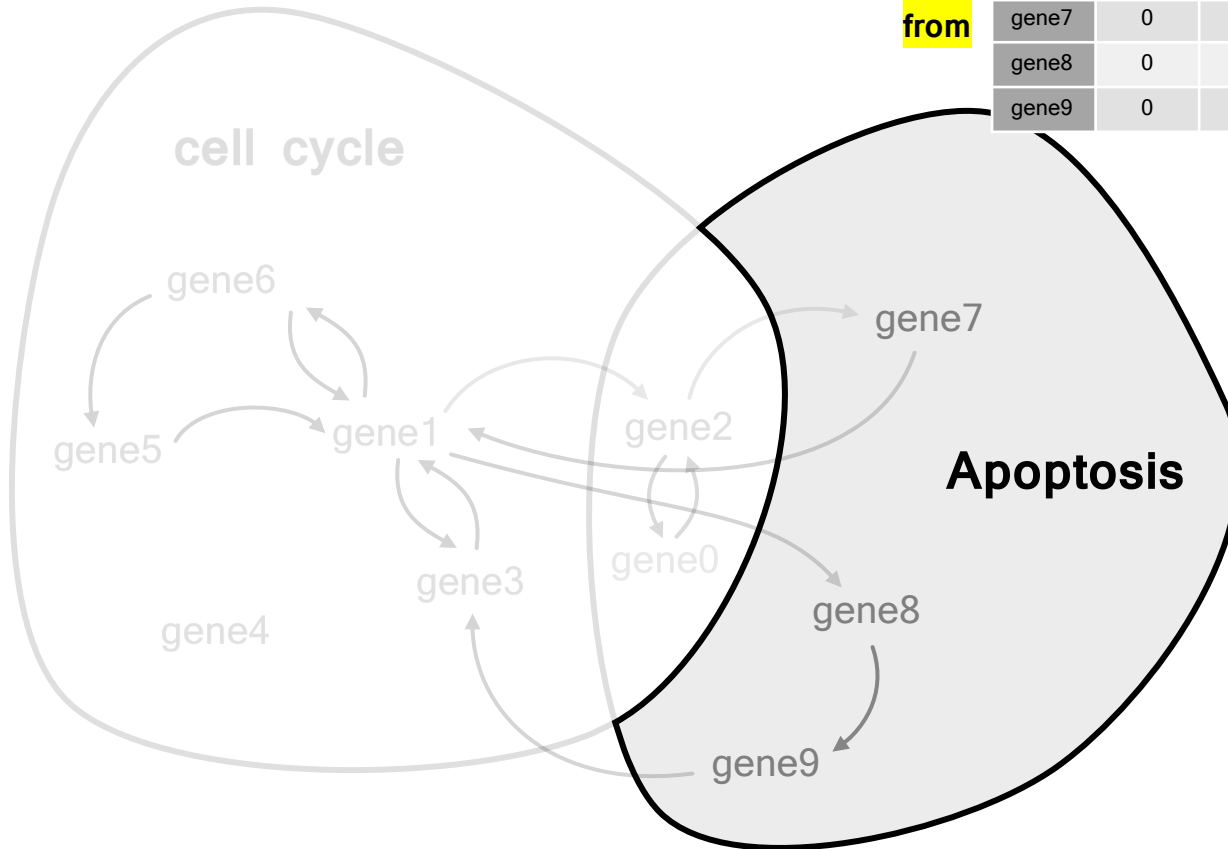


from the graph we can build an **adjacency matrix**

to

from

	gene7	gene8	gene9
gene7	0	0	0
gene8	0	0	1
gene9	0	0	0

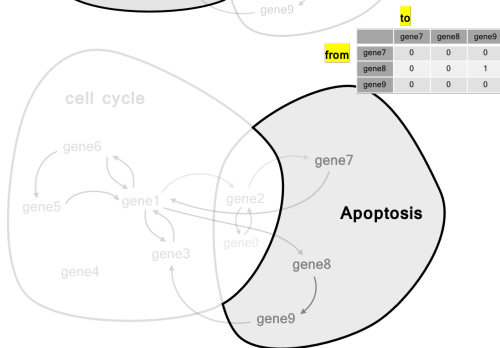
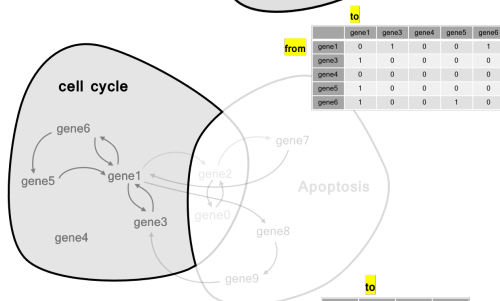
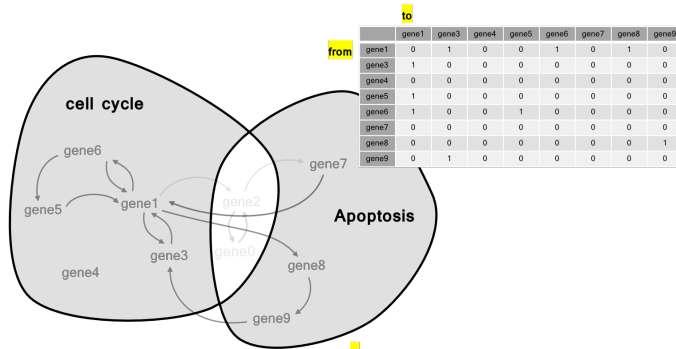


introduction
 basic concepts
 research goals
 methods
 results
 perspectives

we analyse gene pairs to find interactions
 we exclude the intersection



from the graph we can build an **adjacency matrix**



Number of closed walks of length k starting and ending at i^{th} node:

$$N_i(k) = [A^k]_{i,i}$$

if we want the number of closed walks of length $k=2$ we take A^2

introduction
 basic concepts
 research goals
 methods
 results
 perspectives

we analyse gene pairs to find interactions
 we exclude the intersection

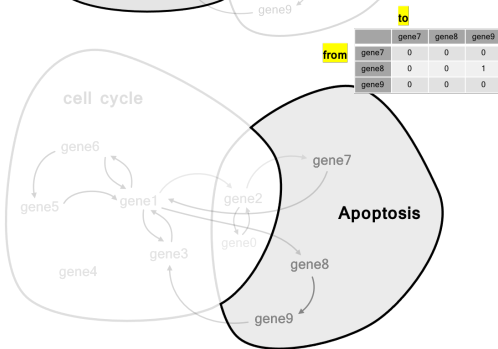
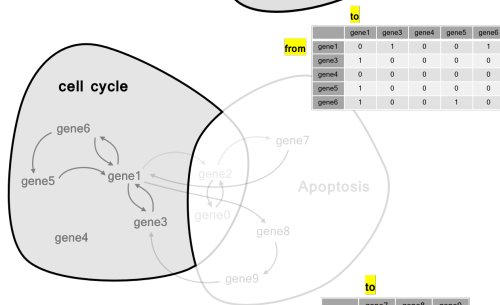
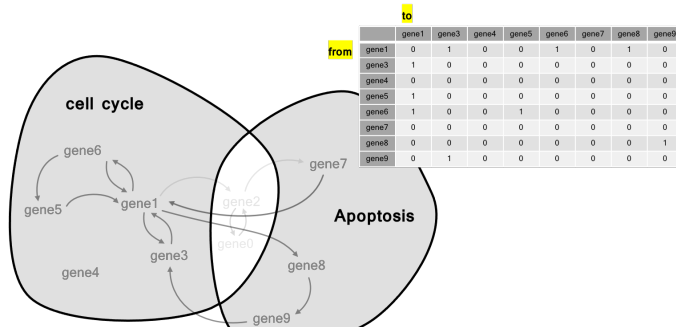


from the graph we can build an **adjacency matrix**

Number of closed walks of length k starting and ending at i^{th} node:

$$N_i(k) = [A^k]_{i,i}$$

if we want the number of closed walks of length $k=2$ we take A^2



from \ to	gene1	gene3	gene4	gene5	gene6
gene1	0	1	0	0	1
gene3	1	0	0	0	0
gene4	0	0	0	0	0
gene5	1	0	0	0	0
gene6	1	0	0	1	0

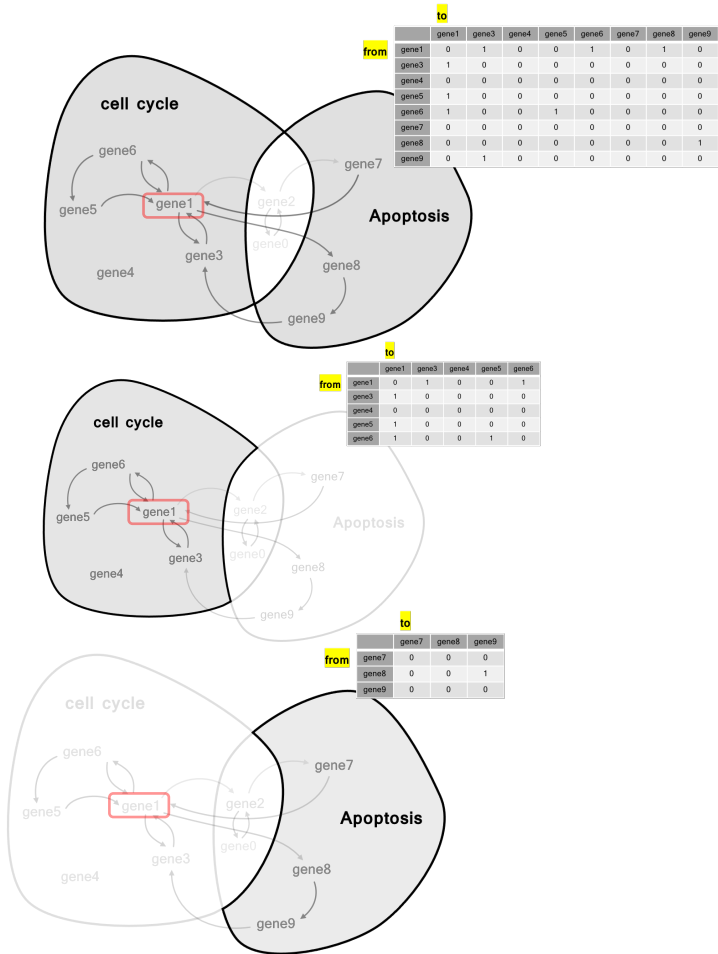
$$^2 =$$

gene1	gene3	gene4	gene5	gene6	
gene1	2	0	0	1	0
gene3	0	1	0	0	1
gene4	0	0	0	0	0
gene5	0	1	0	0	1
gene6	1	1	0	0	1

(it works!)

introduction
 basic concepts
 research goals
 methods
 results
 perspectives

but how to find the crosstalks ?
 - simple example (gene1 only)



k=2

k=3

k=4

number of closed walks of length k=2 in the total graph: 2

number of closed walks of length k=2 in cell cycle: 2

number of closed walks of length k=2 in apoptosis: 0

number of closed walks of length k=3 in the total graph: 1

number of closed walks of length k=3 in cell cycle: 1

number of closed walks of length k=3 in apoptosis: 0

number of closed walks of length k=4 in the total graph: 1

number of closed walks of length k=4 in cell cycle: 0

number of closed walks of length k=4 in apoptosis: 0

same for the rest of the genes and pathway pairs

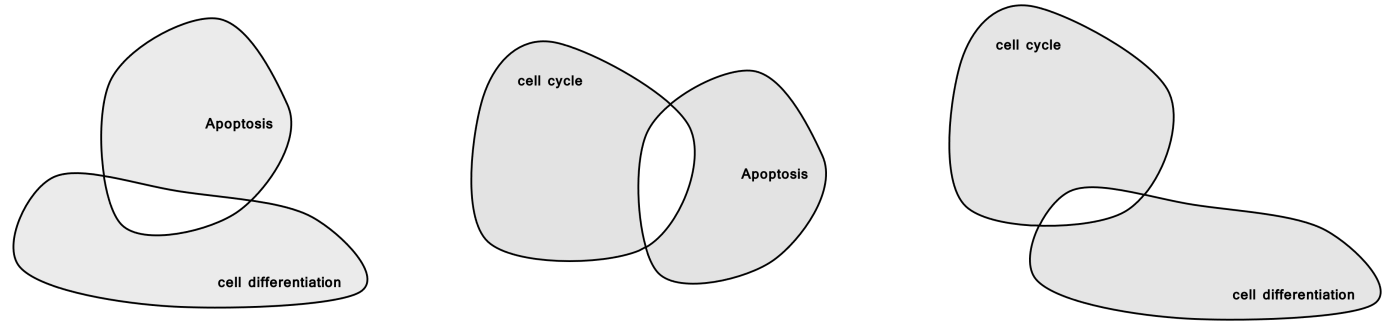
$2 - (2+0) = 0$
 crosstalks of length k=2

$1 - (1+0) = 0$
 crosstalks of length k=3

$1 - (0+0) = 1$
 crosstalks of length k=4

introduction
 basic concepts
 research goals
 methods
results
 perspectives

combinations

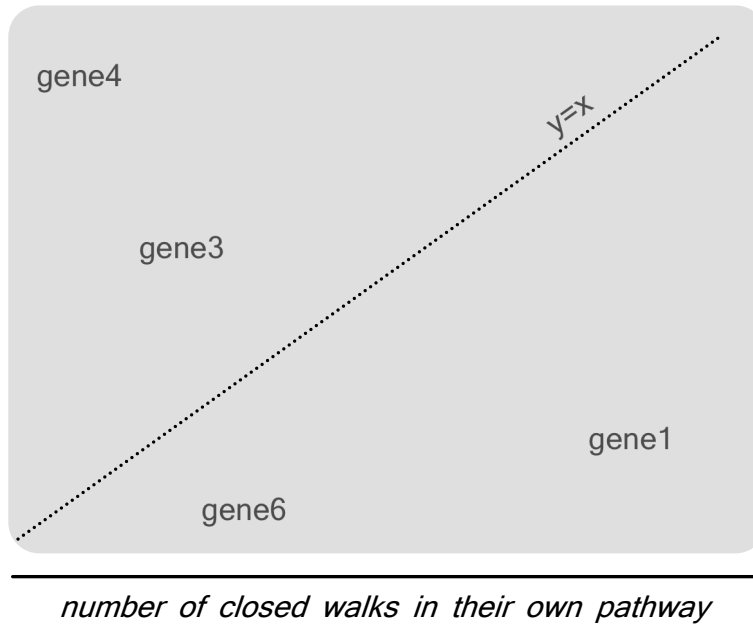


k=2

number of closed walks that cycle through the other pathway

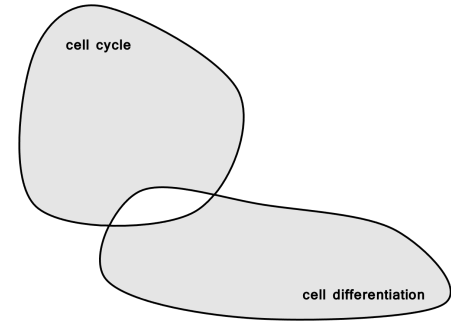
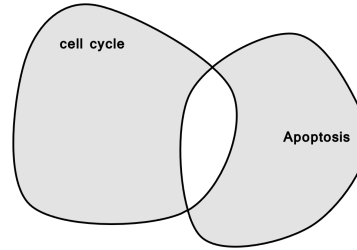
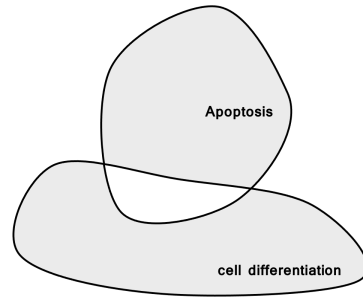
k=3

k=4

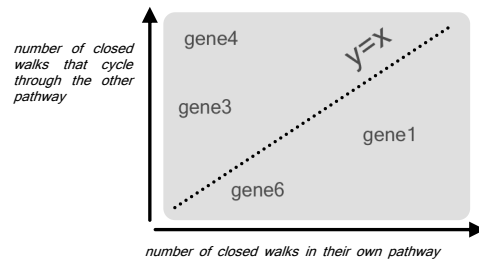


introduction
basic concepts
research goals
methods
results
perspectives

combinations



k=2

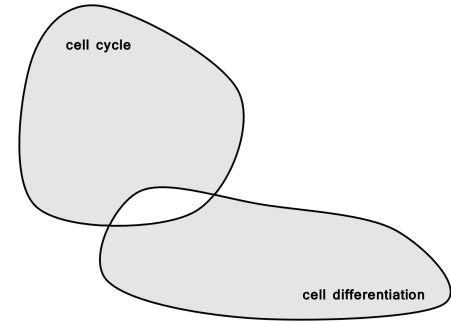
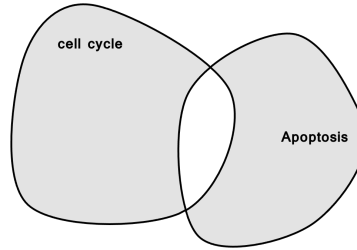
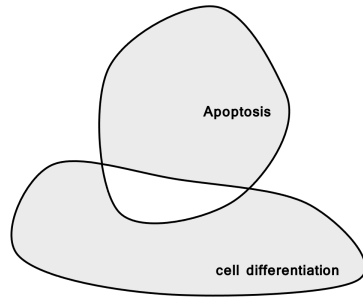


k=3

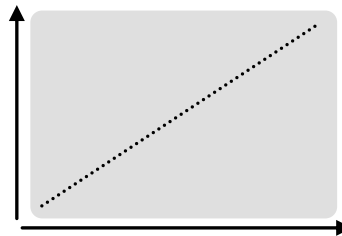
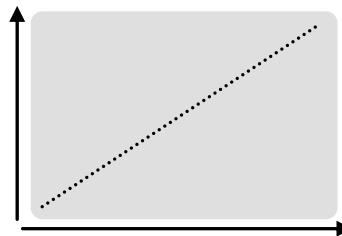
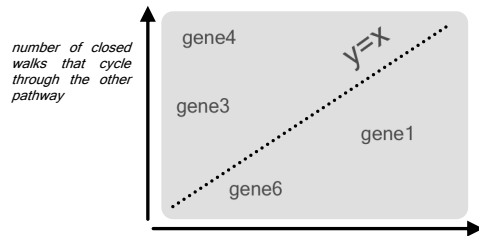
k=4

introduction
 basic concepts
 research goals
 methods
results
 perspectives

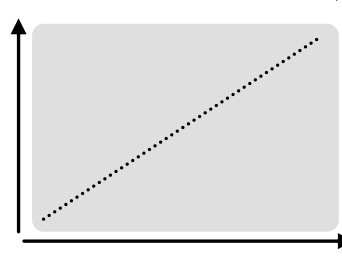
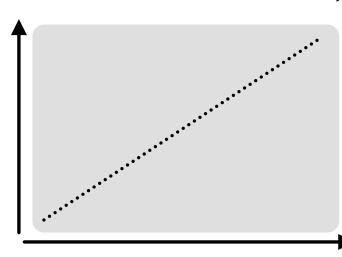
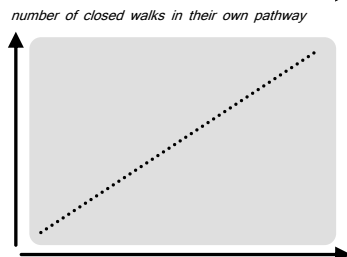
combinations



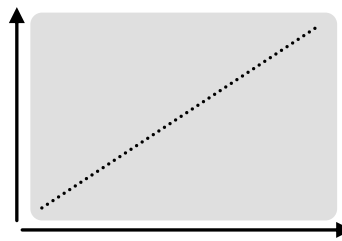
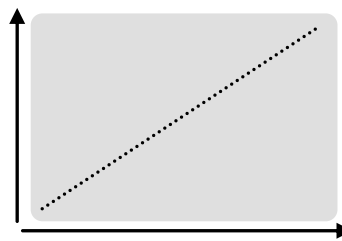
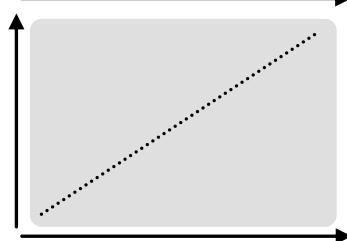
k=2



k=3

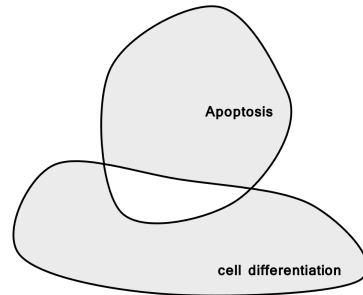


k=4



introduction
 basic concepts
 research goals
 methods
results
 perspectives

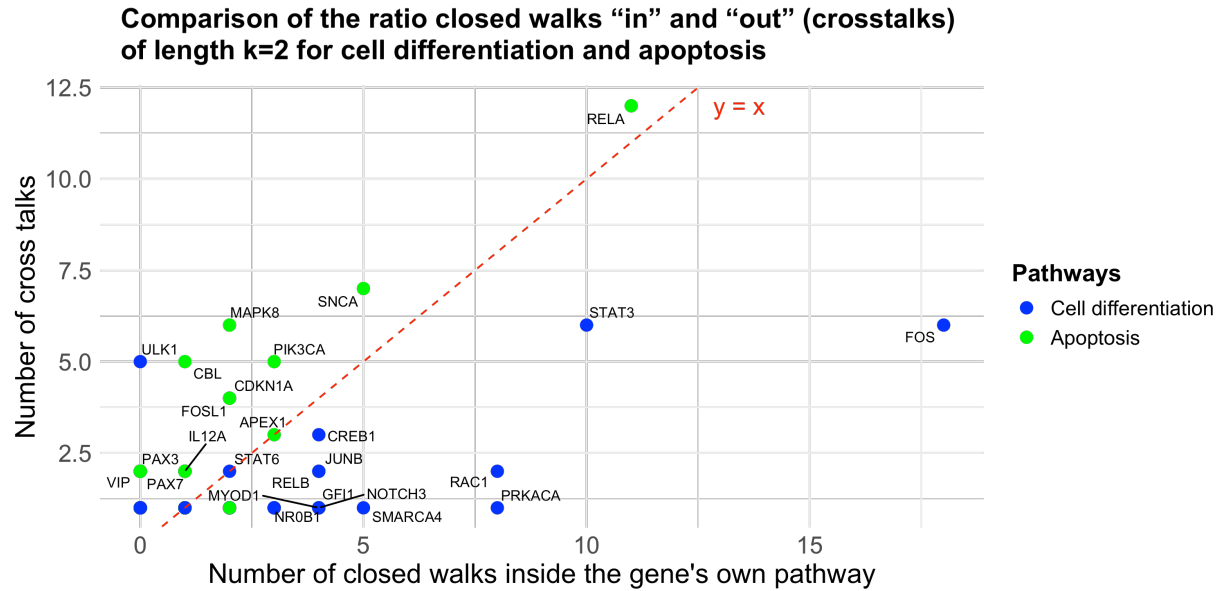
combinations



k=2

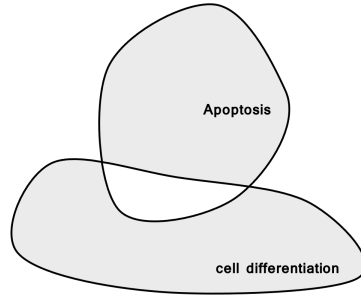
k=3

k=4

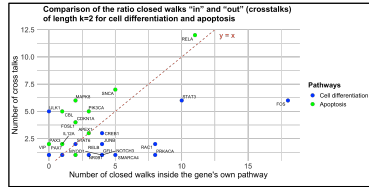


introduction
 basic concepts
 research goals
 methods
results
 perspectives

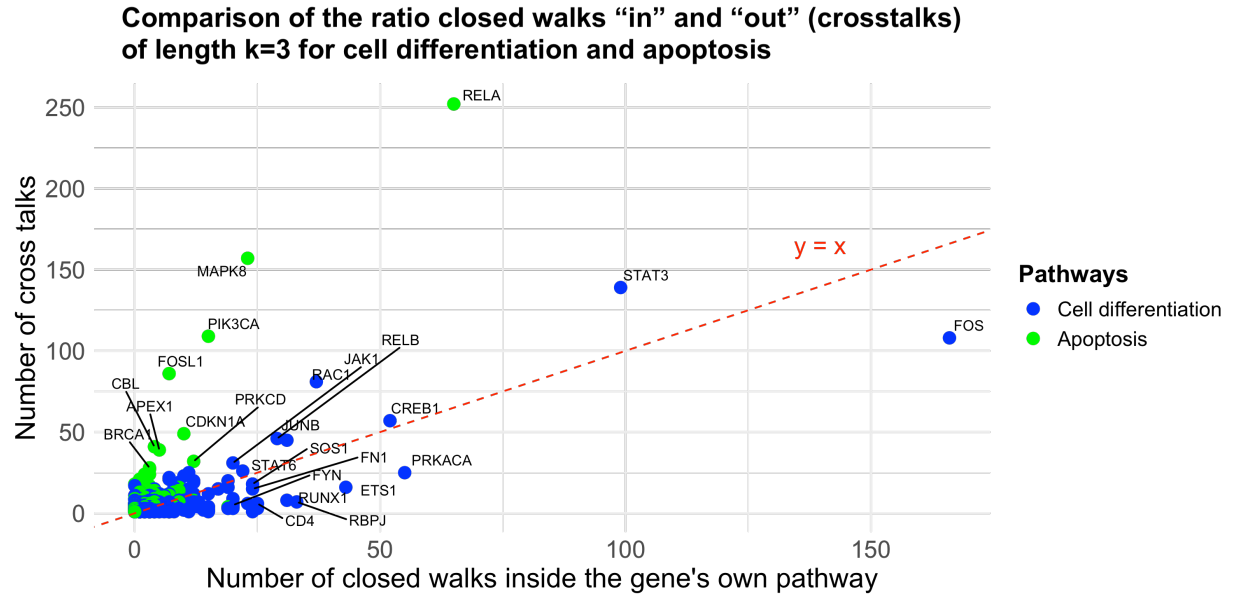
combinations



k=2



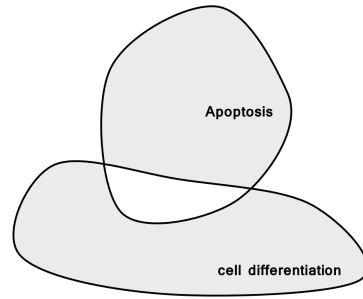
k=3



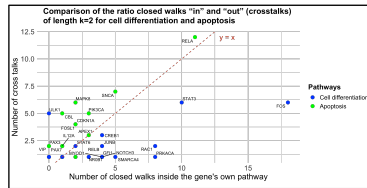
k=4

introduction
 basic concepts
 research goals
 methods
results
 perspectives

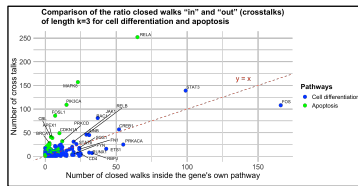
combinations



k=2

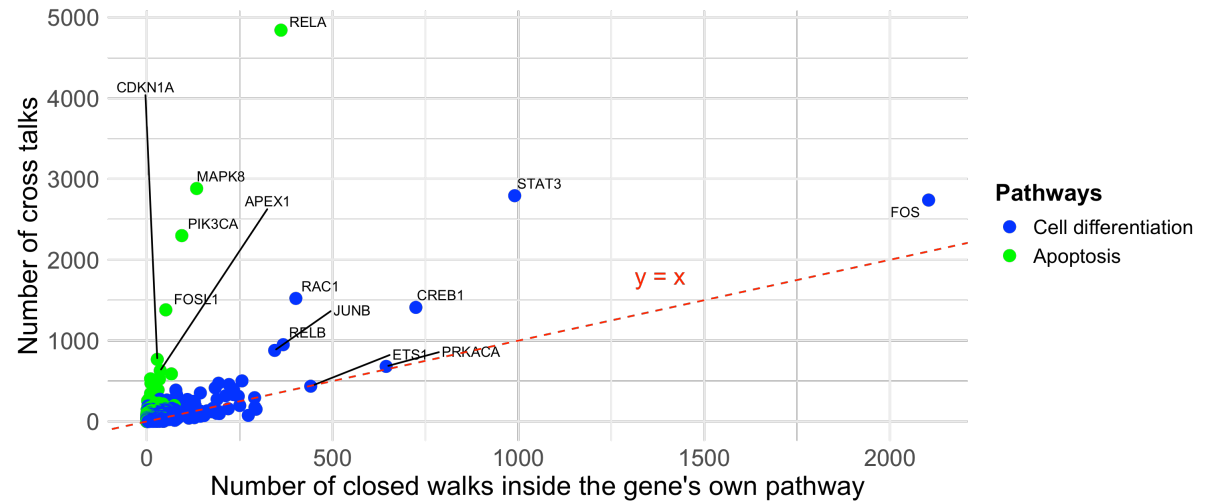


k=3



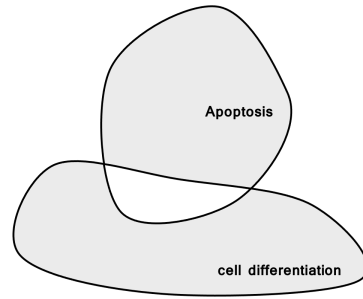
k=4

Comparison of the ratio closed walks "in" and "out" (crosstalks) of length k=4 for cell differentiation and apoptosis

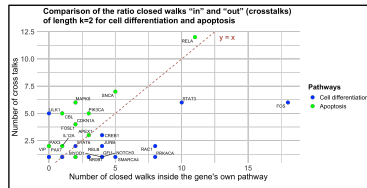


introduction
basic concepts
research goals
methods
results
perspectives

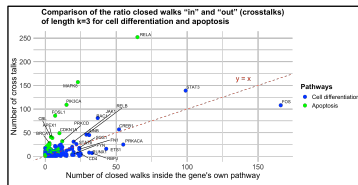
combinations



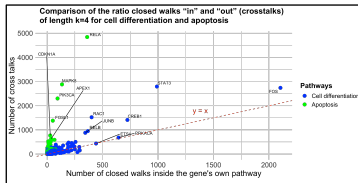
k=2



k=3

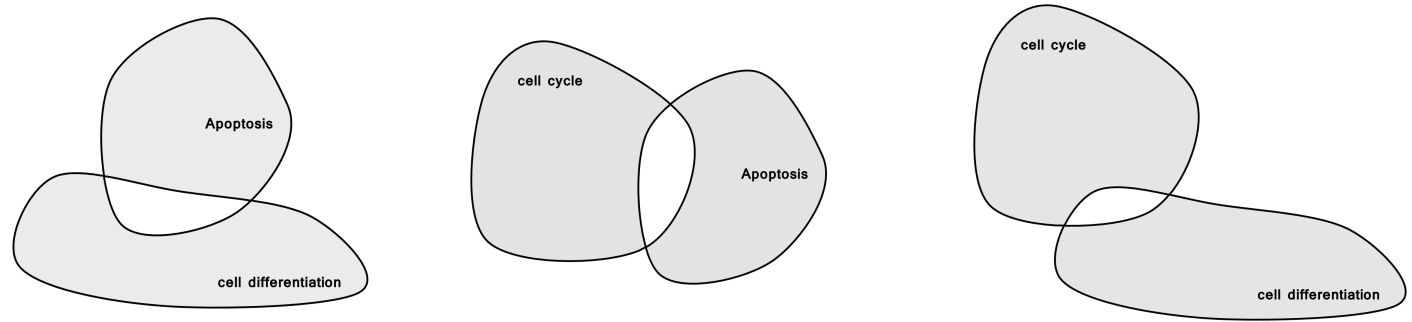


k=4



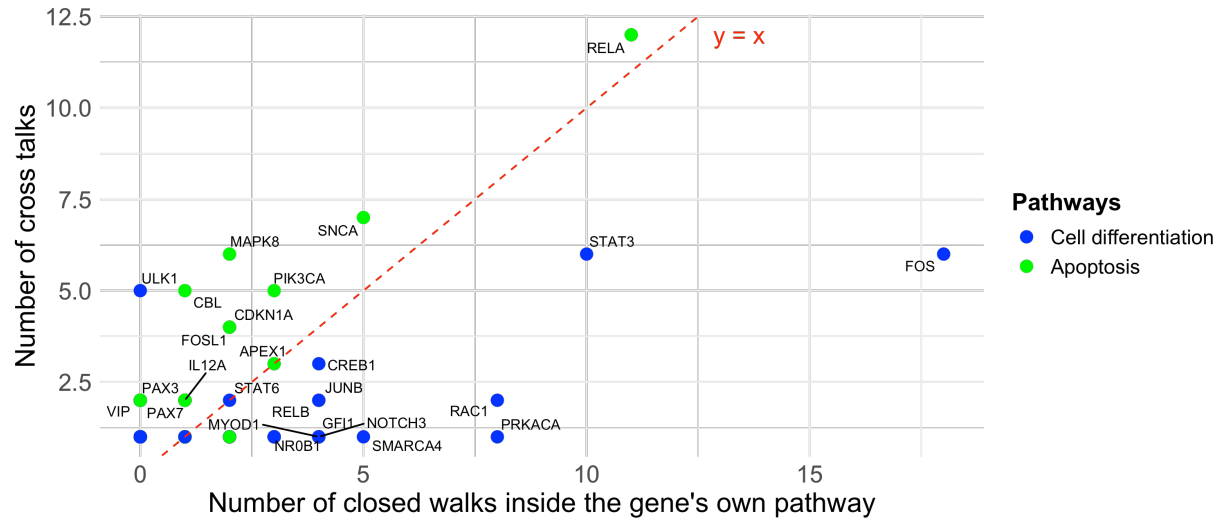
introduction
 basic concepts
 research goals
 methods
results
 perspectives

combinations



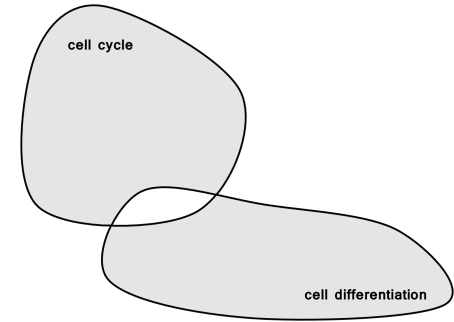
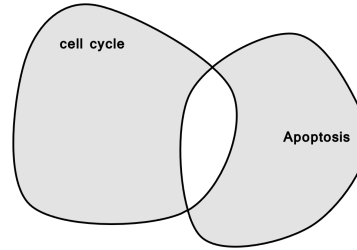
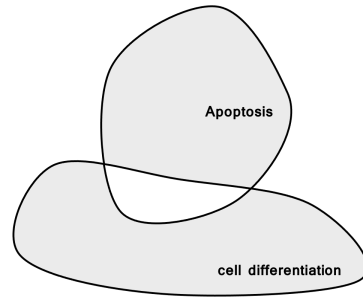
k=2

Comparison of the ratio closed walks “in” and “out” (crosstalks) of length k=2 for cell differentiation and apoptosis

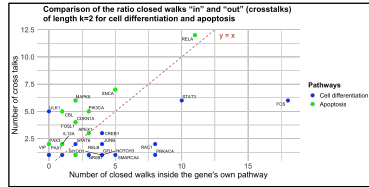


introduction
 basic concepts
 research goals
 methods
results
 perspectives

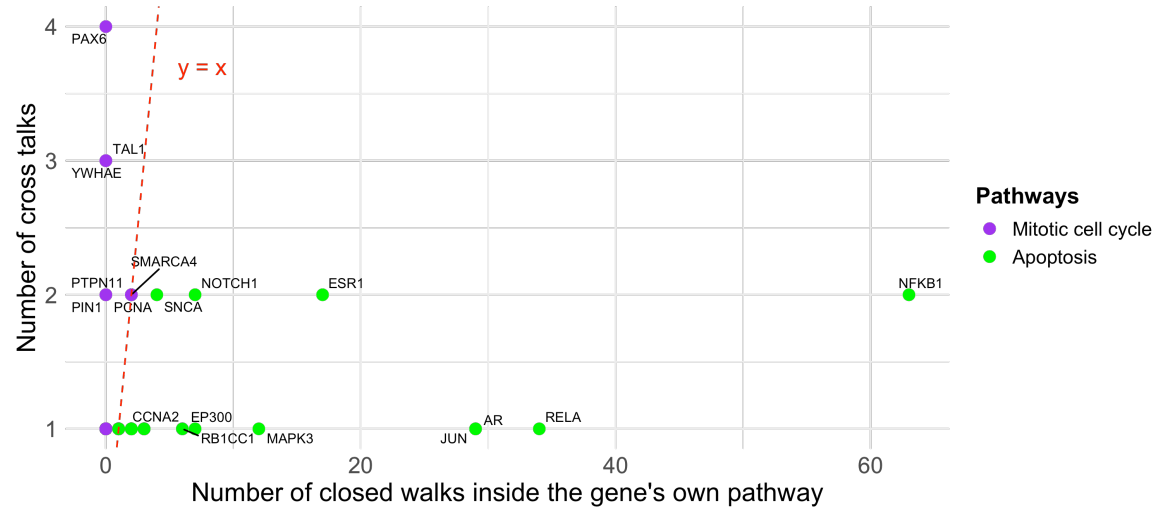
combinations



k=2

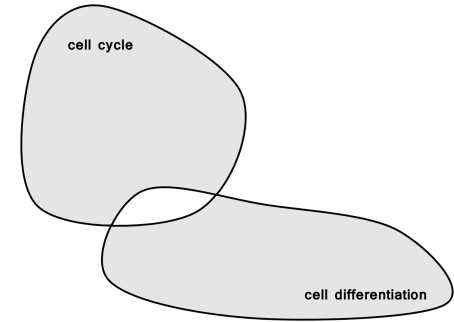
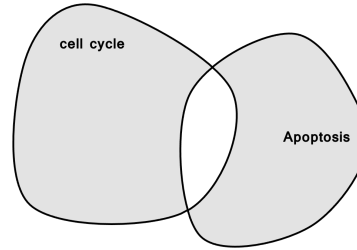
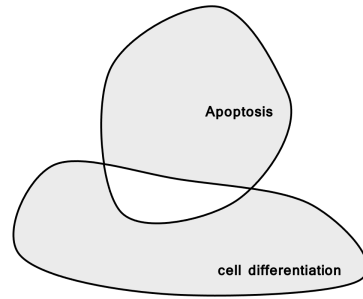


Comparison of the ratio closed walks “in” and “out” (crosstalks) of length k=2 for cell cycle and apoptosis

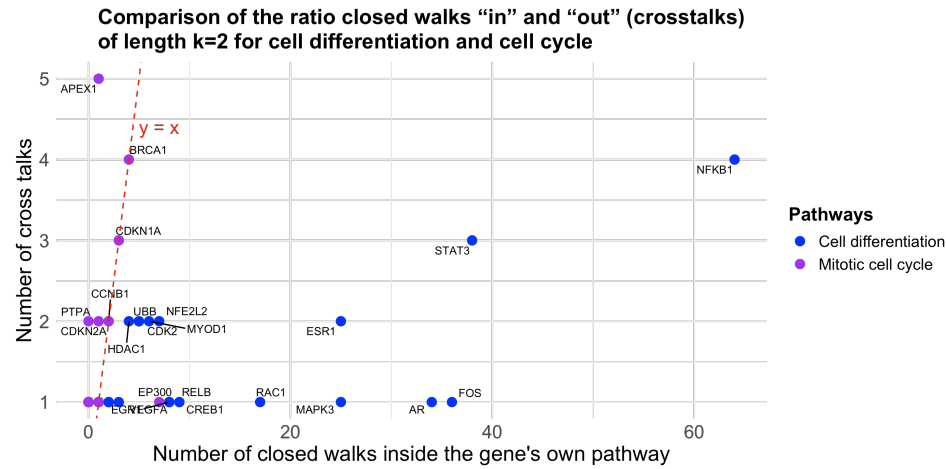
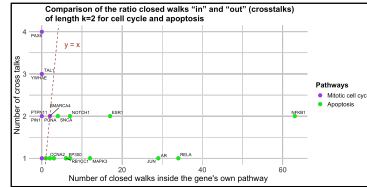
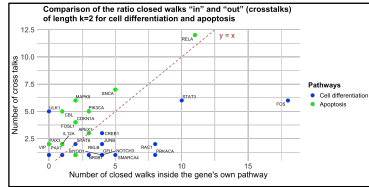


introduction
 basic concepts
 research goals
 methods
results
 perspectives

combinations

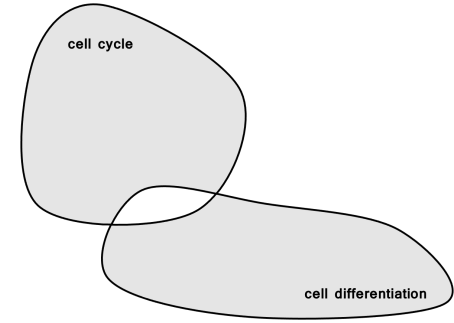
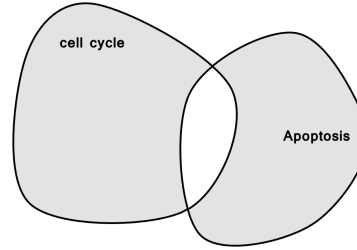
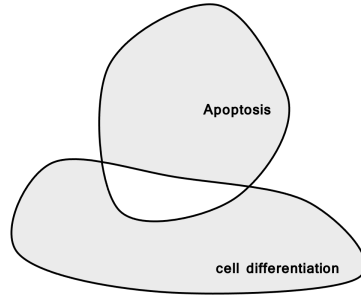


k=2

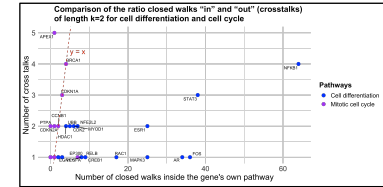
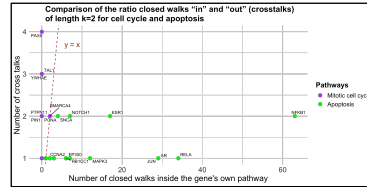
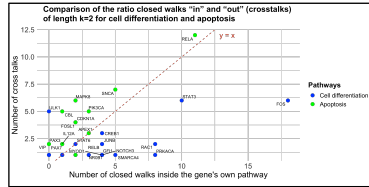


introduction
 basic concepts
 research goals
 methods
results
 perspectives

combinations



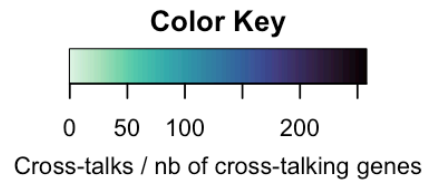
k=2



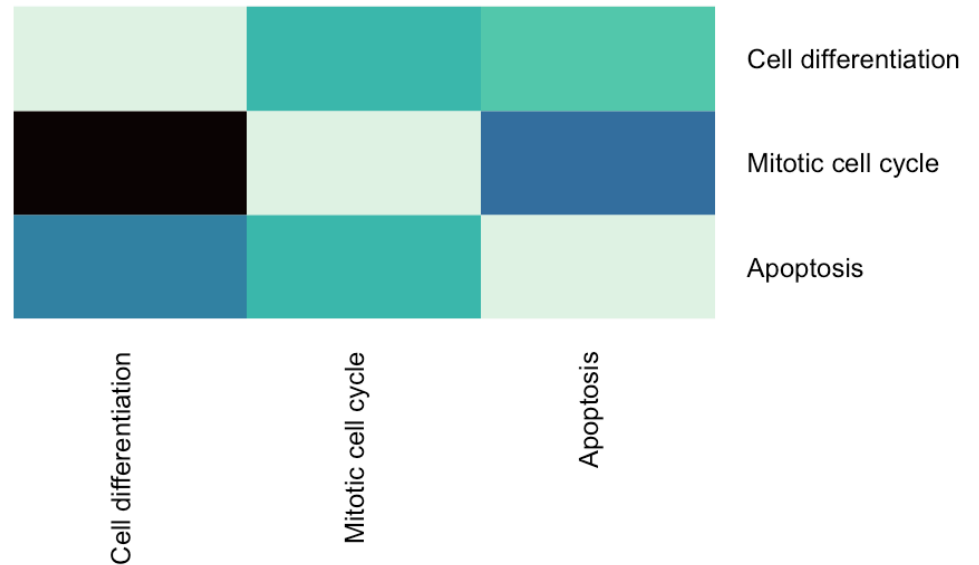
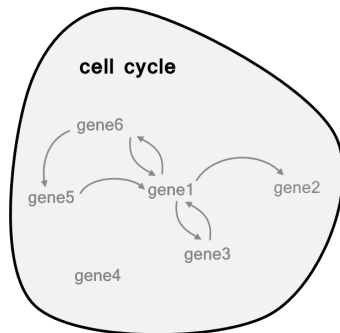
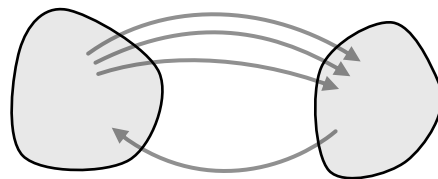
introduction
 basic concepts
 research goals
 methods
results
 perspectives

now let's talk about asymmetry

- crosstalk is an asymmetric process
- ! redundancy: the same genes that do crosstalks of length 2 also do crosstalks of length 3 and 4

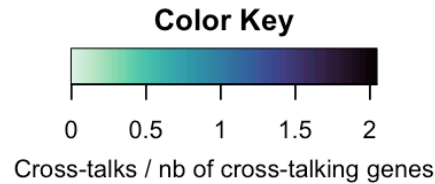


Number of Cross-Talks per gene performing Cross-Talk (of length 2, 3 and 4)

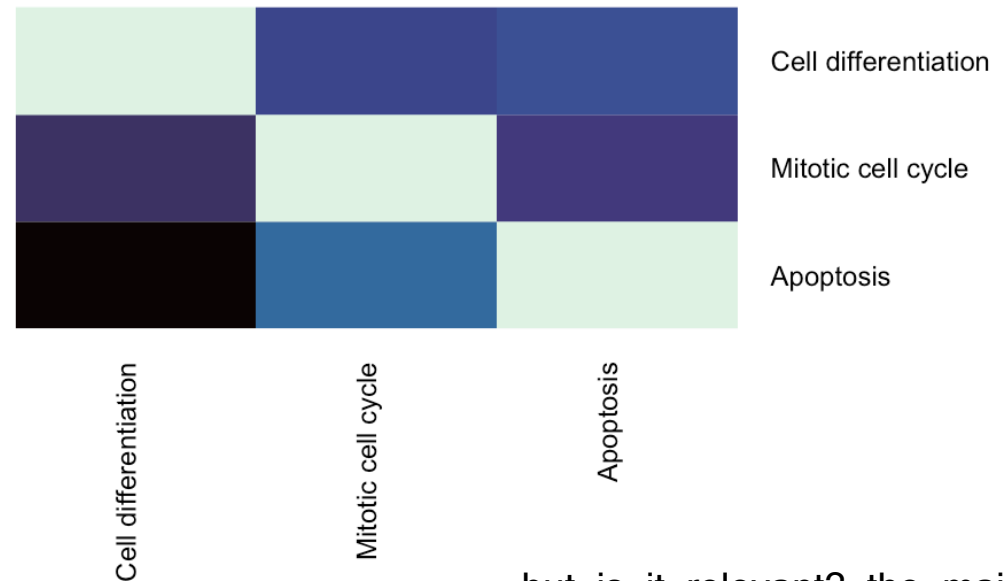


now let's talk about asymmetry

- crosstalk is an asymmetric process
- ! redundancy: the same genes that do crosstalks of length 2 also do crosstalks of length 3 and 4

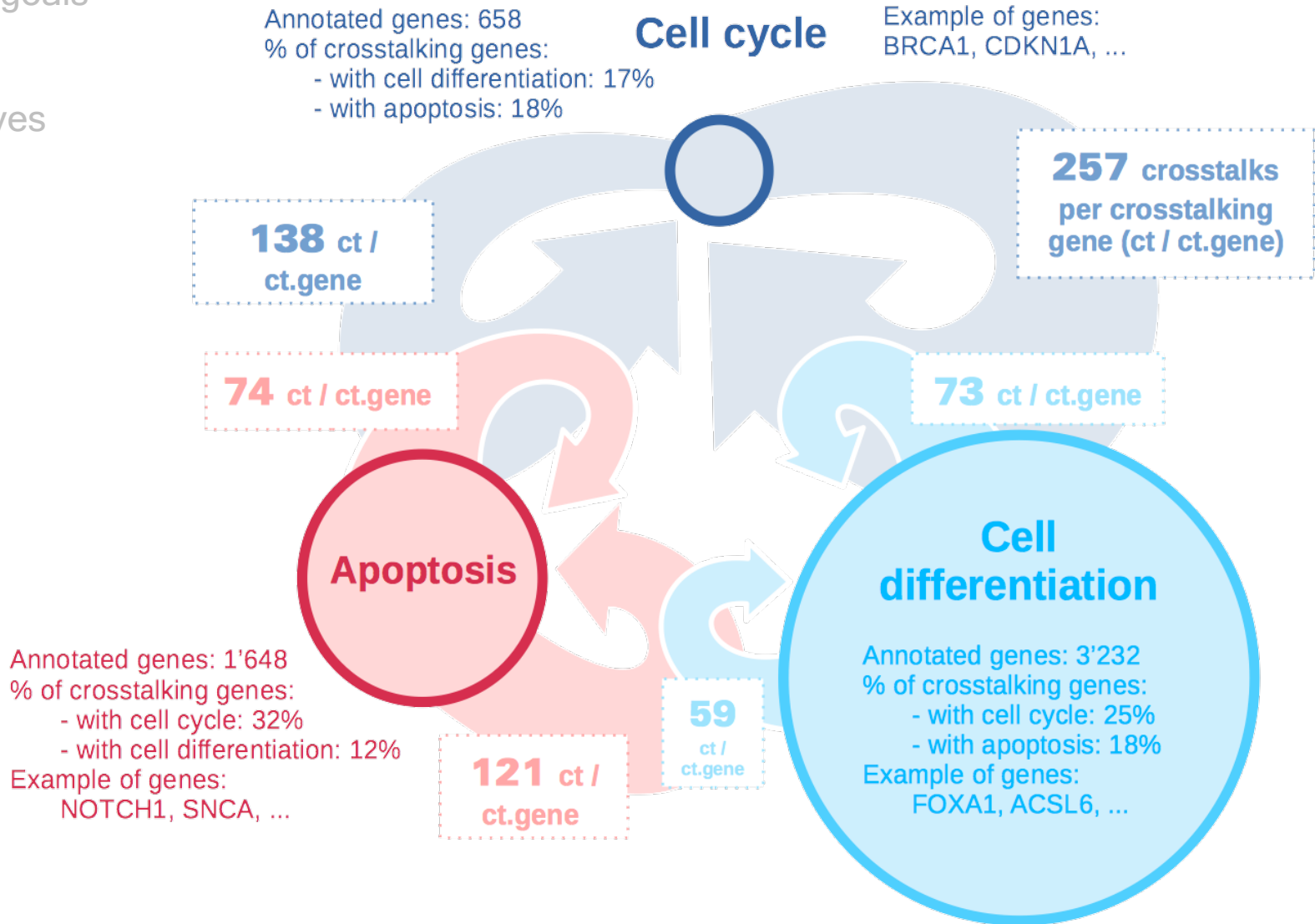


Number of Cross-Talks per gene performing Cross-Talk (of length 2)

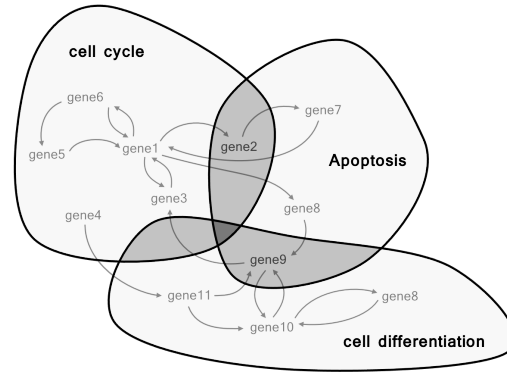


but is it relevant? the majority of crosstalking is done with $k=4$ in our analysis

introduction
basic concepts
research goals
methods
results
perspectives



introduction
basic concepts
research goals
methods
results
perspectives

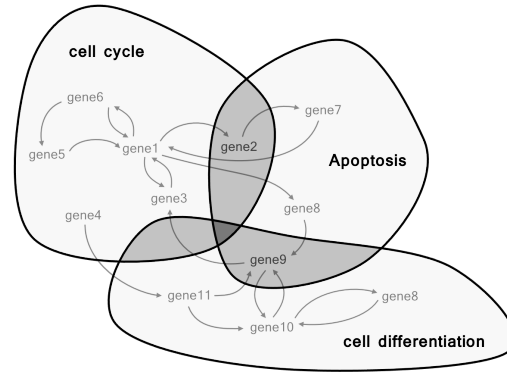


We were able to find the genes that were responsible for the most crosstalks between two pathways → entry points

Some pathways are more prone to crosstalk. There is an asymmetry between the pathways: some pathways cycle more through other pathways than the other pathways do.

Our method allows to find genes that could have been potentially mislabelled (and later reassign them)

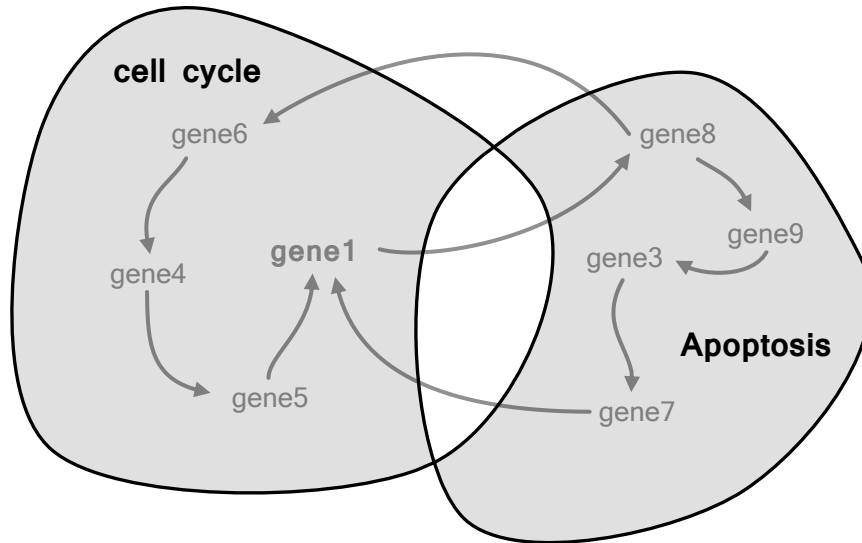
introduction
basic concepts
research goals
methods
results
perspectives



We were able to find the genes that were responsible for the most crosstalks between two pathways → entry points

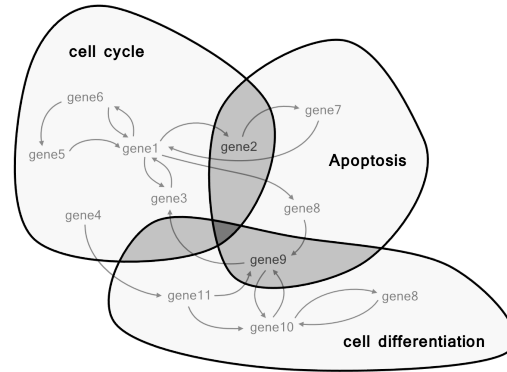
There is an asymmetry between the pathways: some pathways cycle more through other pathways than the other pathways do.

Our method allows to find genes that could have been potentially mislabelled (and later reassign them)



Find the “nature” of long crosstalks : are they “more in pathway1” or “more in pathway2”

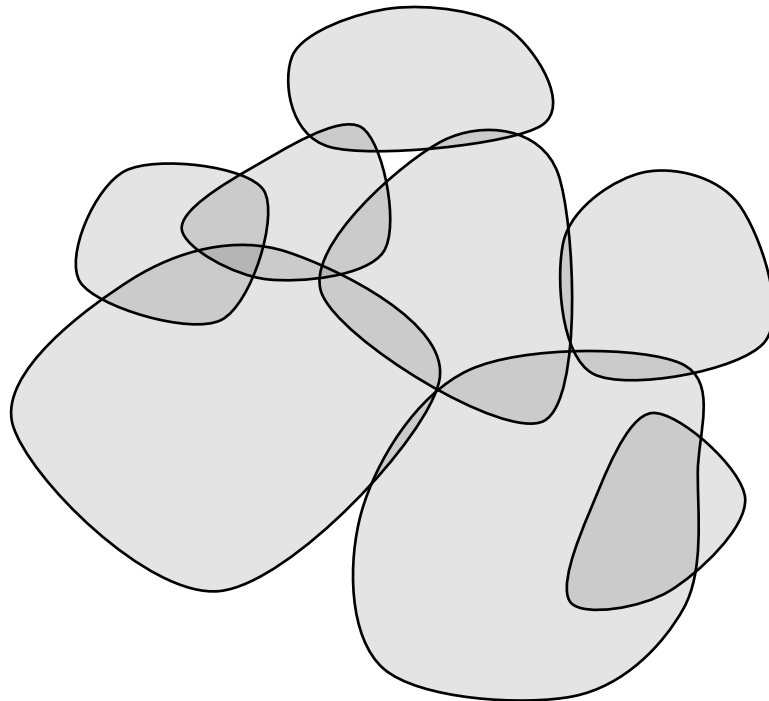
introduction
basic concepts
research goals
methods
results
perspectives



We were able to find the genes that were responsible for the most crosstalks between two pathways → entry points

There is an asymmetry between the pathways: some pathways cycle more through other pathways than the other pathways do.

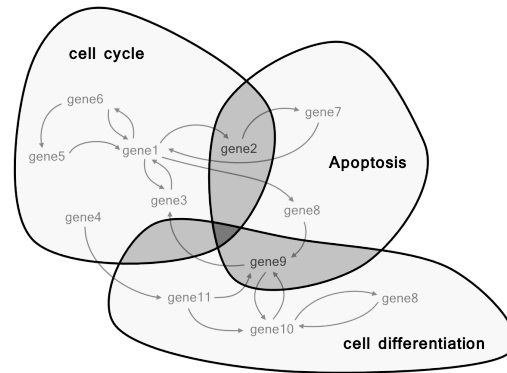
Our method allows to find genes that could have been potentially mislabelled (and later reassign them)



Find the “nature” of long crosstalks : are they “more in pathway1” or “more in pathway2”

Generalise to more pathways (intersections!)

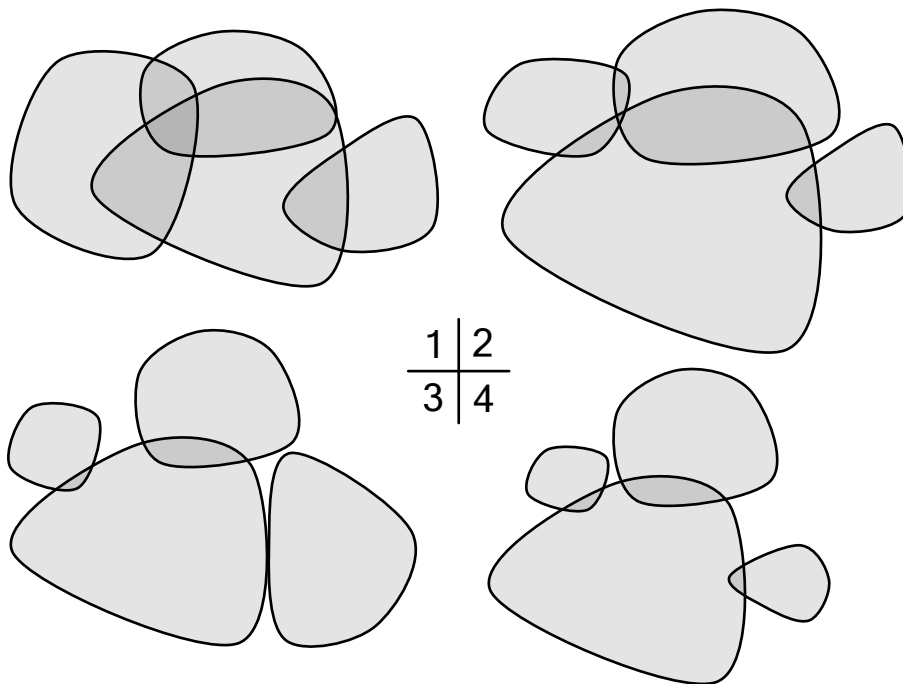
introduction
 basic concepts
 research goals
 methods
 results
 perspectives



We were able to find the genes that were responsible for the most crosstalks between two pathways → entry points

There is an asymmetry between the pathways: some pathways cycle more through other pathways than the other pathways do.

Our method allows to find genes that could have been potentially mislabelled (and later reassign them)

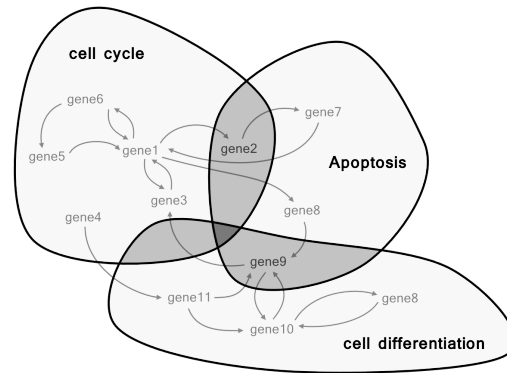


Find the “nature” of long crosstalks : are they “more in pathway1” or “more in pathway2”

Generalise to more pathways (intersections!)

Use the number of crosstalks as a potential clustering algorithm

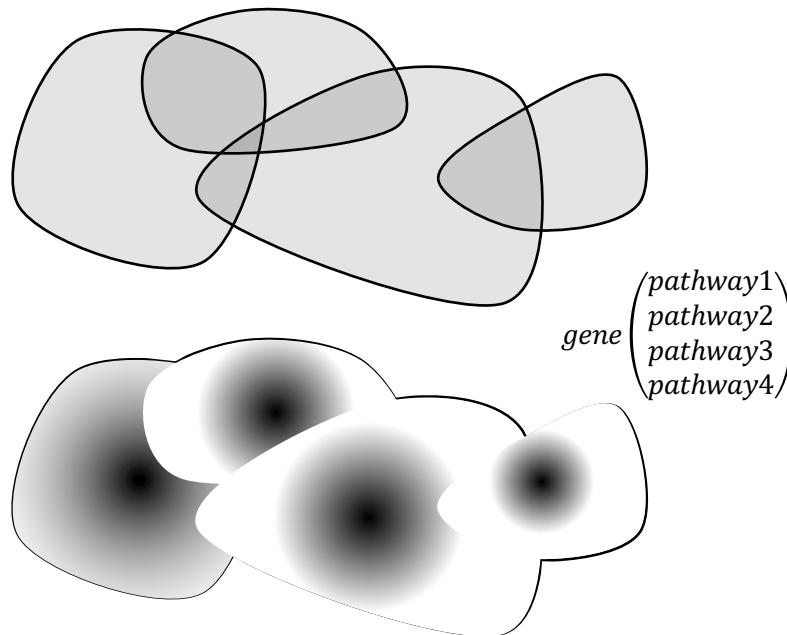
introduction
basic concepts
research goals
methods
results
perspectives



We were able to find the genes that were responsible for the most crosstalks between two pathways → entry points

There is an asymmetry between the pathways: some pathways cycle more through other pathways than the other pathways do.

Our method allows to find genes that could have been potentially mislabelled (and later reassign them)



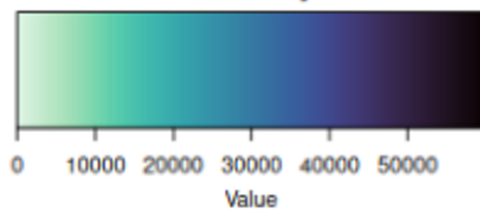
Find the “nature” of long crosstalks : are they “more in pathway1” or “more in pathway2”

Generalise to more pathways (intersections!)

Use the number of crosstalks as a potential clustering algorithm

Find a more “fluid” classification of genes

Color Key



Absolute number of crosstalks from (col) to (row)

