

# Evolutionary dynamics of populations with genotype-phenotype map

Esther Ibáñez Marcelo, Tomas Alarcon Cor

Biomat 2013: Mathematics of Planet Earth

CENTRE DE RECERCA MATEMÀTICA

17-21 June 2013

# Complexity of the genotype-phenotype map

- There is a complex relation between genotype and phenotype.
- Map between genotype-phenotype is not one-to-one: many genotypes are compatible with the same phenotype.
- Try to relate genotype and phenotype networks.

genotype

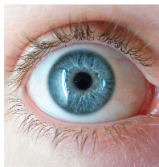
phenotype

codes for



**Phenotype**= Blue Eyes

**Phenotype**=Brown Eyes



# Complexity of the genotype-phenotype map

- There is a complex relation between genotype and phenotype.
- Map between genotype-phenotype is not one-to-one: many genotypes are compatible with the same phenotype.
- Try to relate genotype and phenotype networks.

genotype

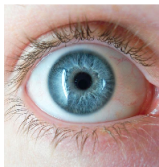
phenotype

codes for



**Phenotype**= Blue Eyes

**Phenotype**=Brown Eyes



# Complexity of the genotype-phenotype map

- There is a complex relation between genotype and phenotype.
- Map between genotype-phenotype is not one-to-one: many genotypes are compatible with the same phenotype.
- Try to relate genotype and phenotype networks.

genotype

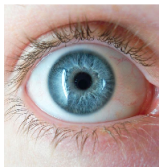
phenotype

codes for



**Phenotype**= Blue Eyes

**Phenotype**=Brown Eyes



# Complexity of the genotype-phenotype map

- There is a complex relation between genotype and phenotype.
- Map between genotype-phenotype is not one-to-one: many genotypes are compatible with the same phenotype.
- Try to relate genotype and phenotype networks.

genotype

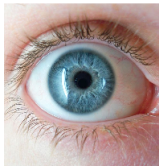
phenotype

codes for



**Phenotype**= Blue Eyes

**Phenotype**=Brown Eyes



# Problems from the evolutionary point of view

- While genes are passed on from one generation to the next, **selective pressure acts at the level of phenotypes.**
- **Models assigning fitness values** to genotypes or genes (mutant variants, different alleles, ...) have been proposed, but, in general, **is not a valid approach.**
- A more appropriate approach consists of considering models that **take into account the genotype and the phenotype**, or a **model where selection acts at the level of phenotypes.**

# Problems from the evolutionary point of view

- While genes are passed on from one generation to the next, **selective pressure acts at the level of phenotypes.**
- **Models assigning fitness values** to genotypes or genes (mutant variants, different alleles, ...) have been proposed, but, in general, **is not a valid approach.**
- A more appropriate approach consists of considering models that **take into account the genotype and the phenotype**, or a **model where selection acts at the level of phenotypes.**

# Problems from the evolutionary point of view

- While genes are passed on from one generation to the next, **selective pressure acts at the level of phenotypes.**
- **Models assigning fitness values** to genotypes or genes (mutant variants, different alleles, ...) have been proposed, but, in general, is **not a valid approach.**
- A more appropriate approach consists of considering models that **take into account the genotype and the phenotype**, or a **model where selection acts at the level of phenotypes.**



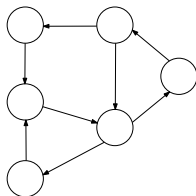
# Problems from the evolutionary point of view

- While genes are passed on from one generation to the next, **selective pressure acts at the level of phenotypes.**
- **Models assigning fitness values** to genotypes or genes (mutant variants, different alleles, ...) have been proposed, but, in general, is **not a valid approach.**
- A more appropriate approach consists of considering models that **take into account the genotype and the phenotype**, or a **model where selection acts at the level of phenotypes.**

# Model GRNs & Genotype & Phenotype

**GRNs** are modelled as graphs:

- **nodes** correspond to the **genes** ( $g_i$ )
- **regulatory interactions** between them as directed **edges**



- Each regulatory edge has an associated weight  $w_{ij} = \pm 1, 0$ .

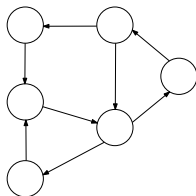
GRN and associated weights define a **genotype**.

- Each node has associated a state variable  $g_i(t)$   
 $\mathbf{g}(t) = (g_i(t)) = \pm 1, i = 1, \dots, N$   
( $N$  = number of genes)
- The final state of the dynamics  $\mathbf{g}(\infty)$  define the **phenotype**.

# Model GRNs & Genotype & Phenotype

**GRNs** are modelled as graphs:

- **nodes** correspond to the **genes** ( $g_i$ )
- **regulatory interactions** between them as directed edges



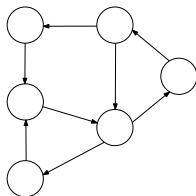
- Each regulatory edge has an associated weight  $w_{ij} = \pm 1, 0$ .

- Each node has associated a state variable  $g_i(t) = \pm 1, i = 1, \dots, N$  ( $N$  = number of genes)
- The final state of the dynamics  $g(\infty)$  define the **phenotype**.

# Model GRNs & Genotype & Phenotype

**GRNs** are modelled as graphs:

- **nodes** correspond to the **genes** ( $g_i$ )
- **regulatory interactions** between them as directed **edges**



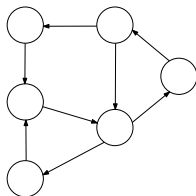
- Each regulatory edge has an associated weight  $w_{ij} = \pm 1, 0$ .

- Each node has associated a state variable  $g(t) = (g_i(t)) = \pm 1, i = 1, \dots, N$  ( $N$  = number of genes)
- The final state of the dynamics  $g(\infty)$  define the **phenotype**.

# Model GRNs & Genotype & Phenotype

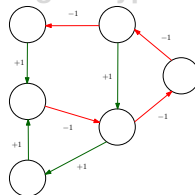
**GRNs** are modelled as graphs:

- **nodes** correspond to the **genes** ( $g_i$ )
- **regulatory interactions** between them as directed **edges**



- Each regulatory edge has an associated weight  $w_{ij} = \pm 1, 0$ .

GRN and associated weights define a **genotype**.



- Each node has associated a state variable

$$\mathbf{g}(t) = (g_i(t)) = \pm 1, i = 1, \dots, N$$

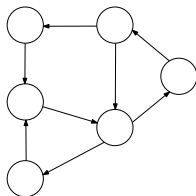
( $N$  = number of genes)

- The final state of the dynamics  $\mathbf{g}(\infty)$  define the **phenotype**.

# Model GRNs & Genotype & Phenotype

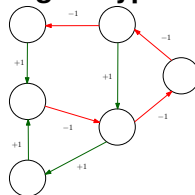
**GRNs** are modelled as graphs:

- **nodes** correspond to the **genes** ( $g_i$ )
- **regulatory interactions** between them as directed **edges**



- Each regulatory edge has an associated weight  $w_{ij} = \pm 1, 0$ .

GRN and associated weights define a **genotype**.



- Each node has associated a state variable

$$\mathbf{g}(t) = (g_i(t)) = \pm 1, i = 1, \dots, N$$

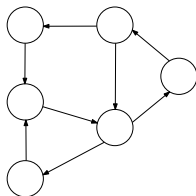
( $N$  = number of genes)

- The final state of the dynamics  $\mathbf{g}(\infty)$  define the **phenotype**.

# Model GRNs & Genotype & Phenotype

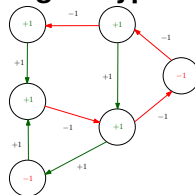
**GRNs** are modelled as graphs:

- **nodes** correspond to the **genes** ( $g_i$ )
- **regulatory interactions** between them as directed **edges**



- Each regulatory edge has an associated weight  $w_{ij} = \pm 1, 0$ .

GRN and associated weights define a **genotype**.

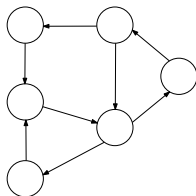


- Each node has associated a state variable  
 $\mathbf{g}(t) = (g_i(t)) = \pm 1, i = 1, \dots, N$   
( $N$  = number of genes)
- The final state of the dynamics  $\mathbf{g}(\infty)$  define the **phenotype**.

# Model GRNs & Genotype & Phenotype

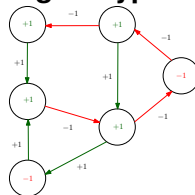
**GRNs** are modelled as graphs:

- **nodes** correspond to the **genes** ( $g_i$ )
- **regulatory interactions** between them as directed **edges**



- Each regulatory edge has an associated weight  $w_{ij} = \pm 1, 0$ .

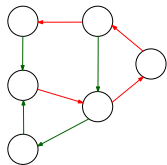
GRN and associated weights define a **genotype**.



- Each node has associated a state variable  $\mathbf{g}(t) = (g_i(t)) = \pm 1, i = 1, \dots, N$  ( $N$  = number of genes)
- The final state of the dynamics  $\mathbf{g}(\infty)$  define the **phenotype**.



## GRN



Genotypes

+  $g(0)$   
initial condition

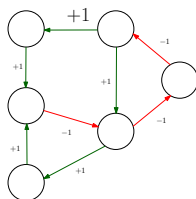
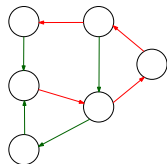
+  
DYNAMICS

PHENOTYPE

$g(\infty)$   
(final state)

# Genotypes

GRN

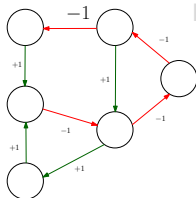


+  $g(0)$   
initial condition

PHENOTYPE

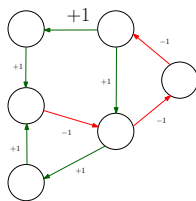
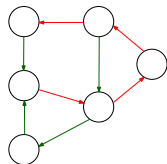
$g(\infty)$   
(final state)

+  
DYNAMICS



# Genotypes

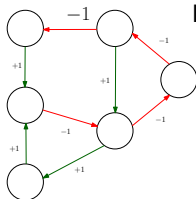
GRN



1 mutation

+  $g(0)$   
initial condition

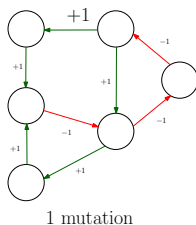
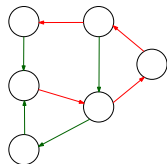
+  
DYNAMICS



PHENOTYPE  
 $g(\infty)$   
(final state)

# Genotypes

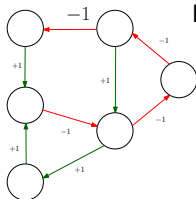
GRN



+  $g(0)$   
initial condition

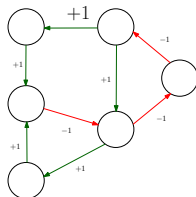
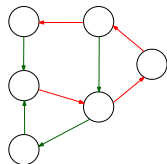
+  
DYNAMICS

PHENOTYPE  
 $g(\infty)$   
(final state)

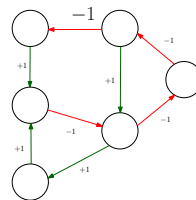


## Genotypes

### GRN



1 mutation

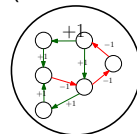


+  $g(0)$   
initial condition

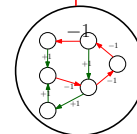
+  
DYNAMICS

## PHENOTYPE

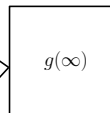
$g(\infty)$   
(final state)



← 1 mutation



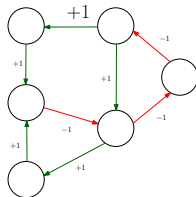
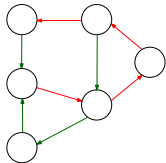
Genotypes



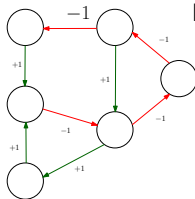
Phenotype

## Genotypes

### GRN



1 mutation

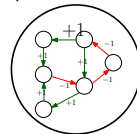


+  $g(0)$   
initial condition

+  
DYNAMICS

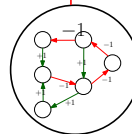
## PHENOTYPE

$g(\infty)$   
(final state)



Phenotype 1

$g(\infty)$



Phenotype 2

$g(\infty)$

Genotypes

# Dynamics of gene regulatory networks

## Dynamics

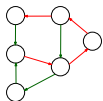
1 
$$I_i(t) = \sum_{j=1}^N w_{ij} g_j(t), \quad N = \#genes$$

2

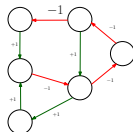
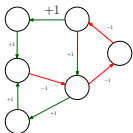
$$g_i(t+1) = \begin{cases} 1 & \text{if } I_i(t) \geq 0 \\ -1 & \text{if } I_i(t) < 0 \end{cases}$$

- The final state of the dynamics  $\mathbf{g}(\infty)$  define the **phenotype**.
- Not always it has a final state, we can obtain **cycles** of different lengths.
- **Viability conditions:** Discard or accept some final states depending on the length of the final state cycle.

GRN



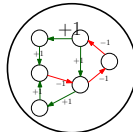
Genotypes



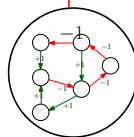
+  $g(0)$   
initial condition

+  
DYNAMICS

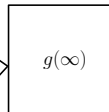
$g(\infty)$   
(final state)  
BIPARTITE NETWORK



← 1 mutation



Genotypes



Phenotype



# Population dynamics

- **Population dynamics:** Multi-type [Wright-Fisher model](#). Start with a **fix population**, chose one random for proliferation, remove another randomly.
- **Evolutionary dynamics:** Each individual picked up for proliferation is assigned a **probability of mutation**: change one sign in the genotype matrix.
- **Selective pressure:** If a mutation induces a **non-viable**<sup>1</sup> phenotype, discard and go through another Wright-Fisher step.

---

<sup>1</sup>We are going to work with viabilities 5, 7, 10. It means that we are accepting, as a final state, cycles until length 5, 7, or 10.

# Representation of the genotype-phenotype space

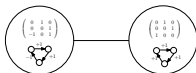
## Pseudo-bipartite graph

- Nodes:

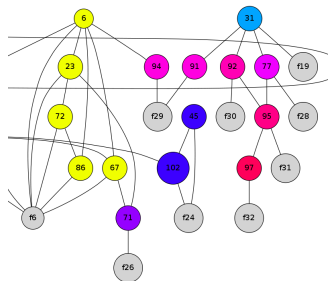
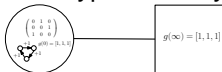
- ▶ Genotypes
- ▶ Phenotypes (grey circles)

- Edges:

- ▶ Genotype-Genotype (if only 1 mutation between them)



- ▶ Genotype-Phenotype



# Robustness and evolvability

- **Robustness:** measure the resilience to mutations. Retain the same phenotype upon mutations in genotype.
- **Robustness** is equivalent to **clustering**<sup>2</sup>.
- **Evolvability:** measure the ability for adaptation. Also called innovation.
- **Evolvability** is equivalent to **percolation**<sup>3</sup>.
- **Neutral network:** join all genotypes which the same phenotype. Usually genotypes are connected (an edge if only 1 mutation)

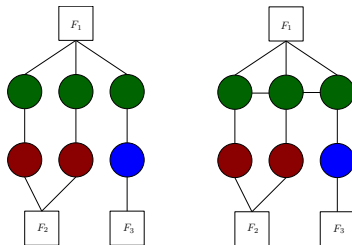
---

<sup>2</sup>**Clustering:**  $c_u = \frac{2T(u)}{\deg(u)(\deg(u)-1)}$ , where  $T(u)$  is the number of triangles through node  $u$  and  $\deg(u)$  is the degree of  $u$ .

<sup>3</sup>**Percolation:** point when giant component appears. 

# Robustness and evolvability

- **Robustness**: measure the resilience to mutations. Retain the same phenotype upon mutations in genotype.
- **Robustness** is equivalent to **clustering**<sup>2</sup>.
- **Evolvability**: measure the ability for adaptation. Also called innovation.
- **Evolvability** is equivalent to **percolation**<sup>3</sup>.
- **Neutral network**: join all genotypes which the same phenotype. Usually genotypes are connected (an edge if only 1 mutation)



<sup>2</sup>**Clustering**:  $c_u = \frac{2T(u)}{\deg(u)(\deg(u)-1)}$ , where  $T(u)$  is the number of triangles through node  $u$  and  $\deg(u)$  is the degree of  $u$ .

# Robustness and evolvability

- **Robustness:** measure the resilience to mutations. Retain the same phenotype upon mutations in genotype.
- **Robustness** is equivalent to **clustering**<sup>2</sup>.
- **Evolvability:** measure the ability for adaptation. Also called **innovation**.
- **Evolvability** is equivalent to **percolation**<sup>3</sup>.
- **Neutral network:** join all genotypes which the same phenotype. Usually genotypes are connected (an edge if only 1 mutation)

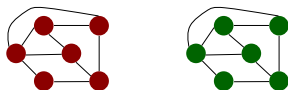
---

<sup>2</sup>**Clustering:**  $c_u = \frac{2T(u)}{\deg(u)(\deg(u)-1)}$ , where  $T(u)$  is the number of triangles through node  $u$  and  $\deg(u)$  is the degree of  $u$ .

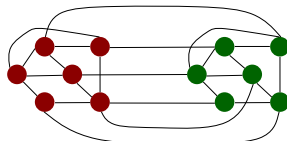
<sup>3</sup>**Percolation:** point when giant component appears.

# Robustness and evolvability

- **Robustness**: measure the resilience to mutations. Retain the same phenotype upon mutations in genotype.
- **Robustness** is equivalent to **clustering**<sup>2</sup>.
- **Evolvability**: measure the ability for adaptation. Also called innovation.
- **Evolvability** is equivalent to **percolation**<sup>3</sup>.
- **Neutral network**: join all genotypes which the same phenotype. Usually genotypes are connected (an edge if only 1 mutation)



Two robust phenotypes  
without innovation



Innovation

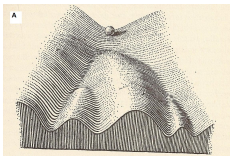
---

<sup>2</sup>**Clustering**:  $c_u = \frac{2T(u)}{\deg(u)(\deg(u)-1)}$ , where  $T(u)$  is the number of triangles through node  $u$  and  $\deg(u)$  is the degree of  $u$ .

<sup>3</sup>**Percolation**: point when giant component appears.

# Canalisation and convergence

- **Canalisation:** assimilation of a change that after some generations can survive without the mutation imposed before. Also called, *buffering of the genotype* or *genetic assimilation*.



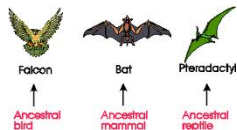
- **Convergence:** Different species of different lineages achieve the same biological traits. We relate to the fact that two or more different unrelated genotypes produce the same phenotype.
- **Convergence** is equivalent to **Hamming distance** between genotypes ( $g_1, g_2$ ).

4

<sup>4</sup>**Hamming distance:** number of mutations that we need to go from one genotype to another,  $d(g_1, g_2)$ , we only have to count the differences between the corresponding matrix associated to each genotype.

# Canalisation and convergence

- **Canalisation:** assimilation of a change that after some generations can survive without the mutation imposed before. Also called, *buffering of the genotype* or *genetic assimilation*.
- **Convergence:** Different species of different lineages achieve the same biological traits. We relate to the fact that two or more different unrelated genotypes produce the same phenotype.
- **Convergence** is equivalent to **Hamming distance** between genotypes ( $g_1, g_2$ ).





# Canalisation and convergence

- **Canalisation:** assimilation of a change that after some generations can survive without the mutation imposed before. Also called, *buffering of the genotype* or *genetic assimilation*.
- **Convergence:** Different species of different lineages achieve the same biological traits. We relate to the fact that two or more different unrelated genotypes produce the same phenotype.
- **Convergence** is equivalent to **Hamming distance** between genotypes  $(g_1, g_2)$ .

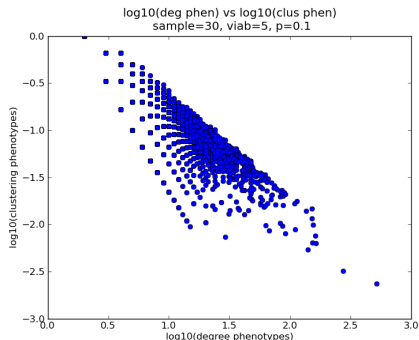
4

---

<sup>4</sup>**Hamming distance:** number of mutations that we need to go from one genotype to another,  $d(g_1, g_2)$ , we only have to count the differences between the corresponding matrix associated to each genotype.

# Some Results and Conclusions

- Strong relation between clustering and degree in the bipartite network (in phenotypes).



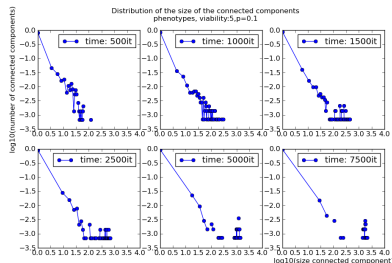
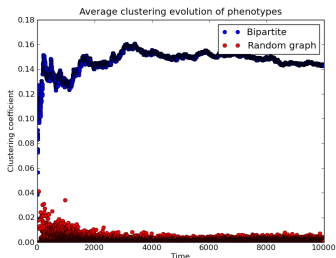
- Canalisation and evolvability coevolve.
- GRN parameters affect percolation (evolvability).
- Hamming distance measure convergence (different genotypes with a high distance can have the same phenotype). Robustness

# Some Results and Conclusions

- Strong **relation** between **clustering** and **degree** in the bipartite network (in phenotypes).
- **Canalisation** and **evolvability** coevolve.
- GRN parameters affect percolation (evolvability).
- Hamming distance measure convergence (different genotypes with a high distance can have the same phenotype). Robustness (high clustering) emerges in phenotypes associated to genotypes that exhibit high degrees of similarity.

# Some Results and Conclusions

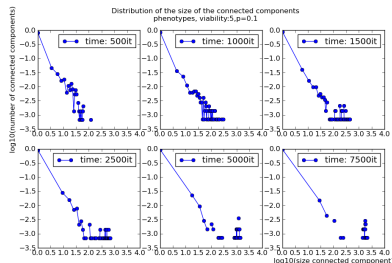
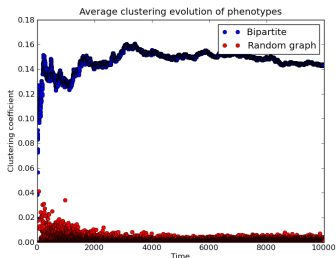
- Strong relation between clustering and degree in the bipartite network (in phenotypes).
- Canalisation and evolvability coevolve.



- GRN parameters affect percolation (evolvability).
- Hamming distance measure convergence (different genotypes with a high distance can have the same phenotype). Robustness (high clustering) emerges in phenotypes associated to genotypes

# Some Results and Conclusions

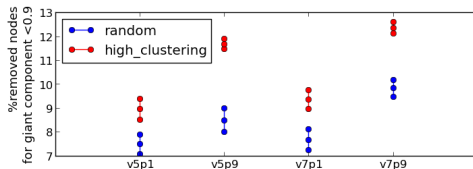
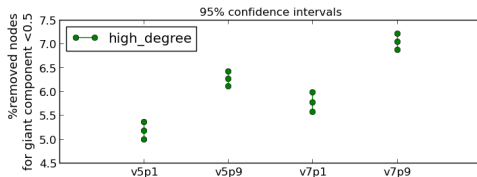
- Strong **relation** between **clustering** and **degree** in the bipartite network (in phenotypes).
- **Canalisation** and **evolvability** coevolve.



- **GRN parameters** affect percolation (**evolvability**).
- **Hamming distance** measure **convergence** (different genotypes with a high distance can have the same phenotype). Robustness (high clustering) emerges in phenotypes associated to genotypes

# Some Results and Conclusions

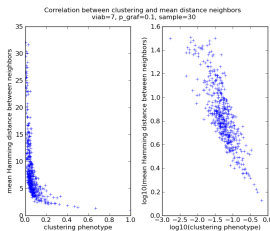
- Strong **relation** between **clustering** and **degree** in the bipartite network (in phenotypes).
- **Canalisation** and **evolvability** coevolve.
- **GRN parameters** affect percolation (**evolvability**).

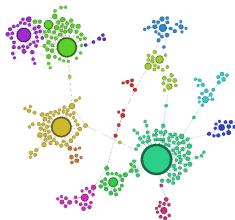
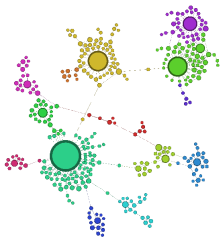


v: viabilities, p: probability Str-Watts model in GRN

# Some Results and Conclusions

- Strong **relation** between **clustering** and **degree** in the bipartite network (in phenotypes).
- **Canalisation** and **evolvability** coevolve.
- **GRN parameters** affect percolation (**evolvability**).
- **Hamming distance** measure **convergence** (different genotypes with a high distance can have the same phenotype). Robustness (high clustering) emerges in phenotypes associated to genotypes that exhibit high degrees of similarity.





Thanks for your attention

