La matematica dell'evoluzione

Dal Fenomeno al Modello, Fossanova, 8 marzo 2018

Dario Benedetto, Sapienza, Dipartimento di Matematica

La matematica dell'evoluzione

Dal Fenomeno al Modello, Fossanova, 8 marzo 2018

Dario Benedetto, Sapienza, Dipartimento di Matematica

Mi occupo di

- meccanica dei fluidi
- teorie cinetiche
- metodi probabilistici per l'analisi automatica dei testi

Vi propongo un percorso tra **genetica** e **matematica**

- perché la genetica ha una storia interessante che si è intersecata con vari aspetti della matematica
- perché nella genetica c'é del bel materiale per attività didattico-laboratoriali in matematica per le scuole superiori
- perché mi permetterà di fare qualche piccola osservazione sulla modellistica matematica nell'insegnamento

Vi propongo un percorso tra **genetica** e **matematica**

- perché la genetica ha una storia interessante che si è intersecata con vari aspetti della matematica
- perché nella genetica c'é del bel materiale per attività didattico-laboratoriali in matematica per le scuole superiori
- perché mi permetterà di fare qualche piccola osservazione sulla modellistica matematica nell'insegnamento

Sarà un racconto inizialemnte storico, ma "di parte", con una selezione degli argomenti utili a questi scopi.

Perché io?

Perché io?

- dallo studio dei testi alla storia delle lingue
- la storia delle lingue è parallela alla storia dei popoli
- la storia dei popoli ha aspetti genetici
- intanto insegnavo matematica per biotecnologie, quindi ho dovuto cercare nuovi spunti didattici, studiando un po' di genetica
- ho iniziato a collaborare al Piano Lauree Scientifiche, sperimentando questi spunti

Laboratori per il Piano Lauree Scientifiche nelle scuole

- **PLS Matematica in moto**
- **PLS Genetica e probabilità**
- PLS Teoria delle catastrofi
- TFA su modelli matematici, probabilità, statistica

(materiale reperibile su http://brazil.mat.uniroma1.it/dario/)

con la collaborazione di Carlotta Maffei, Mirko Degli Esposti, Corrado Mascia, Cristina Ipsevich, Elena Possamai, Giuliana Massotti, Luigi Regogliosi e tanti altri...

L'insegnamento della matematica

- la formazione degli insegnanti in matematica è tradizionalmente vicina ai settori teorici (logica, algebra, geometria)
- la didattica della matematica risente di questa impostazione: rigore metodologico, approccio sintetico, calcolo, ma l'applicazione è un aspetto marginale
- le attuali indicazioni nazionali spingono invece verso la modellistica, la probabilità e la statistica

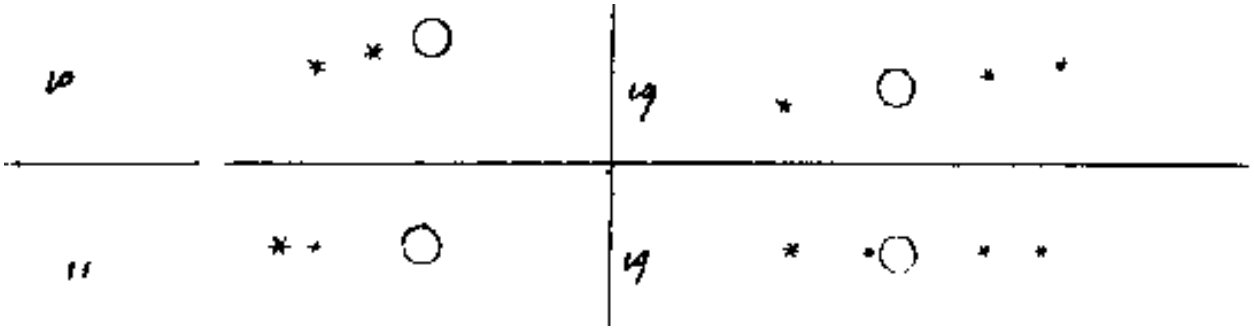
L'insegnamento della matematica

- la formazione degli insegnanti in matematica è tradizionalmente vicina ai settori teorici (logica, algebra, geometria)
- la didattica della matematica risente di questa impostazione: rigore metodologico, approccio sintetico, calcolo, ma l'applicazione è un aspetto marginale
- le attuali indicazioni nazionali spingono invece verso la modellistica, la probabilità e la statistica

Il precorso verso la modellistica ruoterà intorno all'idea didattica della Matematica come Scienza del cambiamento.

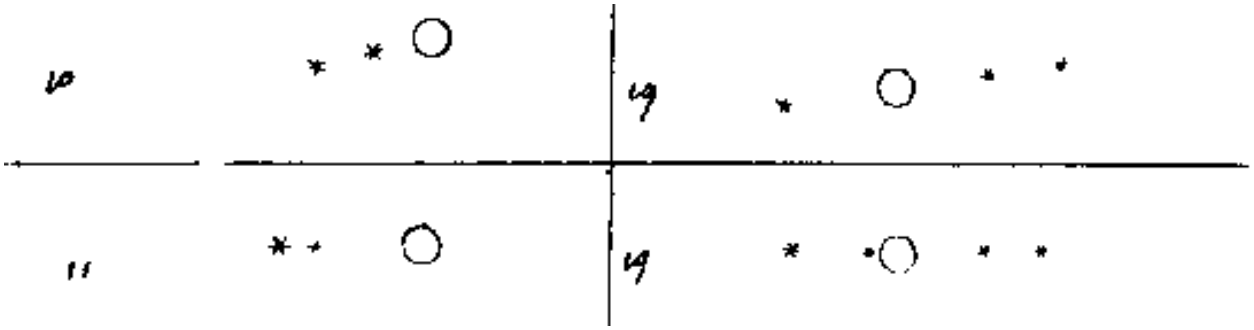
Un cambiamento in cielo

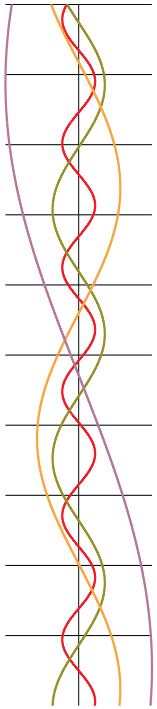
L'osservazione dei 4 satelliti di Giove



Un cambiamento in cielo

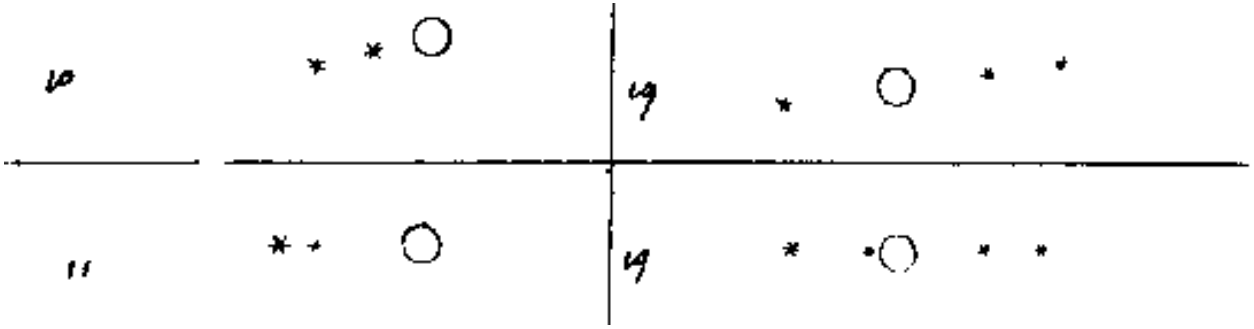
L'osservazione dei 4 satelliti di Giove

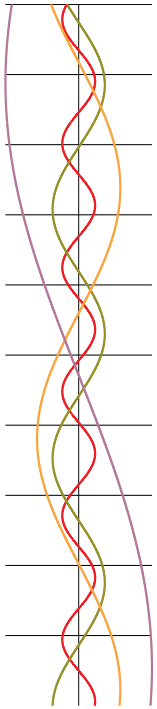




Un cambiamento in cielo

L'osservazione dei 4 satelliti di Giove





La comprensione delle leggi del moto ha portato alla nascita della Fisica moderna e dell'Analisi Matematica.

Newton

Le equazioni di Newton sono **leggi di evoluzione deterministiche**

Newton

Le equazioni di Newton sono **leggi di evoluzione deterministiche**

- **modellizzano** il fenomeno
- permettono di **prevedere il futuro** con certezza partire dal presente
- ci dicono come **modificare la sua evoluzione**

Newton

Le equazioni di Newton sono **leggi di evoluzione deterministiche**

- **modellizzano** il fenomeno
- permettono di **prevedere il futuro** con certezza partire dal presente
- ci dicono come **modificare la sua evoluzione**

Le leggi della Fisica sono i prototipi dei modelli matematici

Newton

Le equazioni di Newton sono **leggi di evoluzione deterministiche**

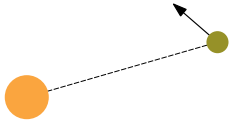
- **modellizzano** il fenomeno
- permettono di **prevedere il futuro** con certezza partire dal presente
- ci dicono come **modificare la sua evoluzione**

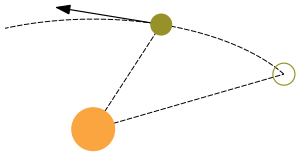
Le leggi della Fisica sono i prototipi dei modelli matematici
(ma sono più di un modello, pretendono di dirci come è fatto l'universo).

Newton

La legge del moto $F = ma$ è infatti una legge di "aggiornamento" del sistema:

presente	$F = ma$	futuro
velocità e posizione	$\longrightarrow$ accelerazione $\longrightarrow$	velocità e posizione





Il caso più semplice

In assenza di forze, la velocità è costante, cioè c'è un rapporto di proporzionalità costante tra lo spazio percorso e il tempo impiegato. Il moto è rettilineo uniforme.

Il caso più semplice

In assenza di forze, la velocità è costante, cioè c'è un rapporto di proporzionalità costante tra lo spazio percorso e il tempo impiegato. Il moto è rettilineo uniforme.

Ci sono due aspetti in questa parola "rettilineo"

- il moto descrive una retta nello spazio
- posizione e tempo descrivono nel piano cartesiano spazio$\times$tempo una retta

Il caso più semplice

La retta viene introdotta come ente geometrico, e poi se ne ricerca la sua **equazione** usando la similitudine tra triangoli.

Fissati due suoi punti (t_1, x_1) e (t_2, x_2), l'equazione è

$$\frac{x - x_1}{x_2 - x_1} = \frac{t - t_1}{t_2 - t_1}$$

Il caso più semplice

Modellisticamente, la **legge lineare** è il più semplice modello di evoluzione di un fenomeno:

la velocità di cambiamento costante

$$\text{futuro} = \text{presente} + \text{velocità} \times \text{tempo}$$

In formule $x(t) = x_0 + v(t - t_0)$.

Un'osservazione didattica

- la formula è uno strumento noto, che lo studente usa quando gli viene chiesto
- l'idea che sta dietro al punto di vista modellistico sulle leggi lineari richiede più attenzione e fatica

Un'osservazione didattica

- la formula è uno strumento noto, che lo studente usa quando gli viene chiesto
- l'idea che sta dietro al punto di vista modellistico sulle leggi lineari richiede più attenzione e fatica

Ho introdotto qui il modello lineare come il modello più semplice, per poter passare a quello successivo, che non è quello suggerito dall'algebra (cioè le parabole...).

Altri modelli di cambiamento: Malthus

Nel 1798 T. Malthus pubblica il suo *Saggio sul principio della popolazione e i suoi effetti sullo sviluppo futuro della società*, in cui formula il primo semplice **modello matematico** del cambiamento della numerosità di una popolazione.

N_0	numerosità iniziale	124 426
N_1	numerosità dopo il primo anno	129 403
N_2	numerosità dopo il secondo anno	134 579
$\vdots$		
N_t	numerosità dopo t anni	?

Altri modelli di cambiamento: Malthus

Nel 1798 T. Malthus pubblica il suo *Saggio sul principio della popolazione e i suoi effetti sullo sviluppo futuro della società*, in cui formula il primo semplice **modello matematico** del cambiamento della numerosità di una popolazione.

N_0	numerosità iniziale	124 426
N_1	numerosità dopo il primo anno	129 403
N_2	numerosità dopo il secondo anno	134 579
$\vdots$		
N_t	numerosità dopo t anni	?

La popolazione aumenta di circa 5000 individui l'anno... quindi sembra anche questo una evoluzione a velocità costante.

Malthus

Malthus, invece, scopre che é il **rapporto** tra la numerosità di un anno e quella dell'anno precedente ad essere praticamente **costante**

$$N_1/N_0 \approx N_2/N_1 \approx N_{t+1}/N_t \approx \text{ costante}$$

Perché?

Malthus

Malthus, invece, scopre che é il **rapporto** tra la numerosità di un anno e quella dell'anno precedente ad essere praticamente **costante**

$$N_1/N_0 \approx N_2/N_1 \approx N_{t+1}/N_t \approx \text{ costante}$$

Perché il numero di nati e il numero di morti in un anno è **proporzionale** alla numerosità della popolazione in quell'anno: se su 1000 000 di persone nascono 1 000 bambini, su 2000 000 ne nasceranno 2000.

Malthus

$$\text{futuro} - \text{presente} = \text{costante} \times \text{presente}$$

In formule:

$$N_{t+1} = N_t + aN_t = (1 + a)N_t$$

dove a è il saldo tra **tasso di mortalità e natalità**.

Anche questa è una legge deterministica, che dà come previsione

$$N_t = (1 + a)^t N_0$$

Malthus

$$\text{futuro} - \text{presente} = \text{costante} \times \text{presente}$$

In formule:

$$N_{t+1} = N_t + aN_t = (1 + a)N_t$$

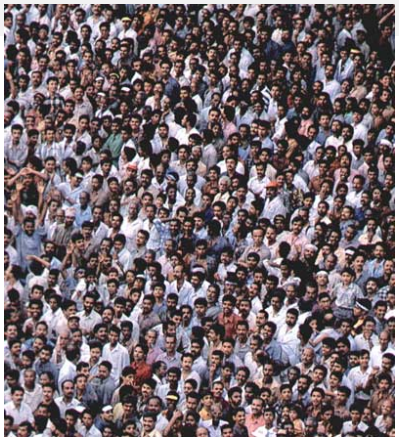
dove a è il saldo tra **tasso di mortalità e natalità**.

Anche questa è una legge deterministica, che dà come previsione

$$N_t = (1 + a)^t N_0$$

Si tratta di una **crescita esponenziale**, ed è quella dei modelli a **tasso di variazione costante**.

Malthus



Dalla legge di evoluzione Malthus prevede che

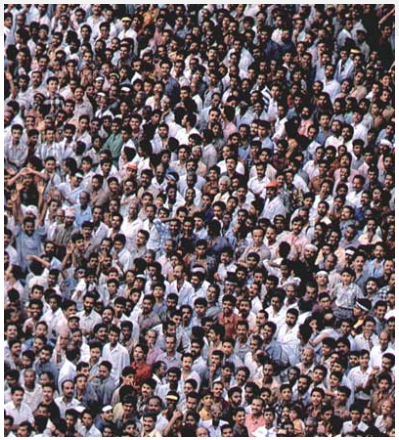
$$N_k = (1 + a)^k N_0$$

Se $a > 0$, la popolazione cresce **esponenzialmente**.

Nell'esempio la popolazione cresce del 4% all'anno, $1 + a = 1.04$

N_0	124 426
N_{10}	184 180
N_{20}	272 633
N_{50}	884 256
N_{100}	6284 129

Malthus



Dalla legge di evoluzione Malthus prevede che

$$N_k = (1 + a)^k N_0$$

Se $a > 0$, la popolazione cresce **esponenzialmente**.

Nell'esempio la popolazione cresce del 4% all'anno, $1 + a = 1.04$

N_0	124 426
N_{10}	184 180
N_{20}	272 633
N_{50}	884 256
N_{100}	6284 129

Malthus afferma che in questo caso inizieranno a scarseggiare le risorse e inizierà la **lotta per la sopravvivenza**.

Altri cambiamenti: Darwin

Le idee di Malthus, fondate su un modello matematico, hanno una grande influenza su Darwin, che nel 1859 pubblica *L'Origine delle specie*, in cui illustra e spiega l'**evoluzione delle specie**.

- Osserva somiglianze e diversità tra le specie esistenti
- Osserva fossili di specie che non esistono più e di specie che **cambiano nel tempo**

Altri cambiamenti: Darwin

Le idee di Malthus, fondate su un modello matematico, hanno una grande influenza su Darwin, che nel 1859 pubblica *L'Origine delle specie*, in cui illustra e spiega l'**evoluzione delle specie**.

- Osserva somiglianze e diversità tra le specie esistenti
- Osserva fossili di specie che non esistono più e di specie che **cambiano nel tempo**

Spiega queste osservazioni come esito della lotta per le risorse:
è la **selezione naturale**, in cui solo alcuni individui, quelli *più adattati*, riescono a riprodursi e trasmettono le loro caratteristiche ai discendenti.

Darwin

La selezione naturale non è una "guerra" in cui il più "cattivo" sopravvive, ma un qualunque fenomeno che porta individui con alcune caratteristiche **ereditabili** a riprodursi più di altri.

Darwin

La selezione naturale non è una "guerra" in cui il più "cattivo" sopravvive, ma un qualunque fenomeno che porta individui con alcune caratteristiche **ereditabili** a riprodursi più di altri.

Ma esistono altre forme di selezione, per esempio la **selezione sessuale**, che è indipendente dall'adattamento all'ambiente.

Darwin

La selezione naturale non è una "guerra" in cui il più "cattivo" sopravvive, ma un qualunque fenomeno che porta individui con alcune caratteristiche **ereditabili** a riprodursi più di altri.

In una delle **grandi estinzioni di massa**, non è sopravvissuto il dinosauro più cattivo ma il mammifero più schivo. Come dice S.J. Gould, l'evoluzione delle specie è **biografia**,

Dove interviene la Matematica?

Come si misura la variabilità?

- 1859 C. Darwin *L'origine delle specie*
- 1865
- 1877
- 1900
- 1918
- 1937
- 1942
- 1953

Dove interviene la Matematica?

Come si misura la variabilità?

- 1859 C. Darwin *L'origine delle specie*
- 1865
- 1877 F. Galton "inventa" la retta di regressione studiando l'ereditarietà dell'altezza
- 1900
- 1918
- 1937
- 1942
- 1953

Dove interviene la Matematica?

La **biometria** fa nascere la **statistica moderna** sotto la spinta dell'eugenetica e del darwinismo sociale.

- 1859 C. Darwin *L'origine delle specie*
- 1865
- 1877 F. Galton "inventa" la retta di regressione studiando l'ereditarietà dell'altezza
- 1900
- 1918 T. Fisher *The Correlation Between Relatives on the Supposition of Mendelian Inheritance* ;

in questo lavoro introduce per la prima volta il termine **varianza**

- 1937
- 1942
- 1953

Dove interviene la Matematica?

Le basi biologiche dell'ereditarietà

- 1859 C. Darwin
- 1865 G. Mendel *Verhandlungen des naturforschenden Vereins Brnn*
- 1877 F. Galton
- 1900
- 1902
- 1915
- 1918 T. Fisher
- 1937
- 1942
- 1953

Dove interviene la Matematica?

qui molta logica e probabilità elementare

- 1859 C. Darwin
- 1865 G. Mendel *Verhandlungen des naturforschenden Vereins Brnn*
- 1877 F. Galton
- 1900 riscoperta delle leggi di Mendel, indipendentemente da 4 scienziati
- 1902
- 1915
- 1918 T. Fisher
- 1937
- 1942
- 1953

Dove interviene la Matematica?

qui molta logica e probabilità elementare

- 1859 C. Darwin
- 1865 G. Mendel *Verhandlungen des naturforschenden Vereins Brnn*
- 1877 F. Galton
- 1900 riscoperta delle leggi di Mendel, indipendentemente da 4 scienziati
- 1902 Sutton-Boveri: teoria cromosomica
- 1915
- 1918 T. Fisher
- 1937
- 1942
- 1953

Dove interviene la Matematica?

qui molta logica e probabilità elementare

- 1859 C. Darwin
- 1865 G. Mendel *Verhandlungen des naturforschenden Vereins Brnn*
- 1877 F. Galton
- 1900 riscoperta delle leggi di Mendel, indipendentemente da 4 scienziati
- 1902 Sutton-Boveri: teoria cromosomica
- 1915 T.H. Morgan et. al. *The Mechanism of Mendelian Inheritance*
- 1918 T. Fisher
- 1937
- 1942
- 1953: R.Franklin, J.Watson, F.Crick: la struttura del DNA

Dove interviene la Matematica?

Come si concilia il mondo **discreto** della teoria mendeliana con le osservazioni dei biometristi?

- 1859 C. Darwin
- 1865 G. Mendel *Verhandlungen des naturforschenden Vereins Brnn*
- 1877 F. Galton
- 1900 riscoperta delle leggi di Mendel, indipendentemente da 4 scienziati
- 1902 Sutton-Boveri: teoria cromosomica
- 1915 T.H. Morgan et. al. *The Mechanism of Mendelian Inheritance*
- 1918 T. Fisher
- 1937
- 1942
- 1953: R.Franklin, J.Watson, F.Crick: la struttura del DNA

Dove interviene la Matematica?

È stato fondamentale lo studio matematico della **genetica di popolazione**

- 1859 C. Darwin
- 1865 G. Mendel *Verhandlungen des naturforschenden Vereins Brnn*
- 1877 F. Galton
- 1900 riscoperta delle leggi di Mendel, indipendentemente da 4 scienziati
- 1902 Sutton-Boveri: teoria cromosomica
- 1915 T.H. Morgan et. al. *The Mechanism of Mendelian Inheritance*
- 1918 T. Fisher
- 1937 T. Dobzhansky *Genetics and the Origin of Species*: teoria cromosomica e matematica della genetica delle popolazioni
- 1942: J. Huxley *Evolution: The Modern Synthesis*
- 1953: R.Franklin, J.Watson, F.Crick: la struttura del DNA

La sintesi moderna dell'evoluzionismo (1936–1947)

- Ereditarietà mendeliana
- Genetica di popolazione
- Selezione naturale

La sintesi moderna dell'evoluzionismo (1936–1947)

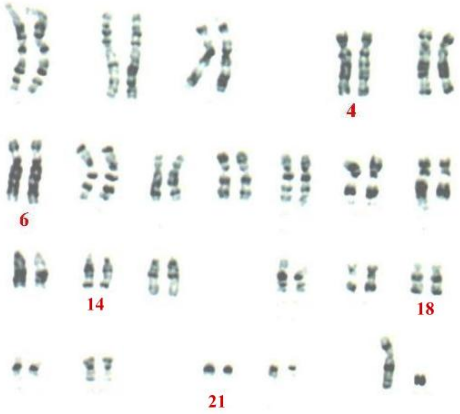
- Ereditarietà mendeliana (modelli probabilistici)
- Genetica di popolazione (modelli deterministici e probabilistici)
- Selezione naturale

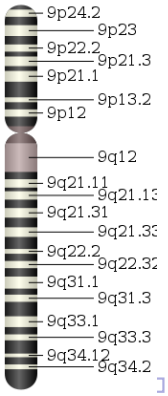
Geni e caratteri ereditari

L'informazione ereditaria è contenuta nei cromosomi, sotto forma **geni** (tratti di DNA in cui sono è scritta in codice la struttura delle proteine).

Nelle cellule di ogni individuo i cromosomi sono presenti in **coppie**:

- i cromosomi sessuali sono di due tipi diversi X e Y, e distinguono le femmine, che hanno X e X, dai maschi, che hanno X e Y
- gli altri, che codificano le stesse informazioni, ma con possibili varianti.





Le varianti dei geni

Le coppie di cromosomi non sessuali portano l'informazione genetica per le stesse caratteristiche, e questa informazione può essere disponibile con delle **varianti**, che si chimano **alleli**.

Nella coppia di cromosomi 9 c'è il gene che codifica per la presenza di particolari proteine sui globuli rossi che determinano il **gruppo sanguigno** di tipo A, B, AB, 0.

Le varianti dei geni

Le coppie di cromosomi non sessuali portano l'informazione genetica per le stesse caratteristiche, e questa informazione può essere disponibile con delle **varianti**, che si chimano **alleli**.

Nella coppia di cromosomi 9 c'è il gene che codifica per la presenza di particolari proteine sui globuli rossi che determinano il **gruppo sanguigno** di tipo A, B, AB, 0.
Questo gene ha tre varianti:

- l'allele A codifica per la presenza della proteina A,
- l'allele B codifica per la presenza della proteina B,
- l'allele 0 non codifica per nessuna proteina.

Le varianti dei geni

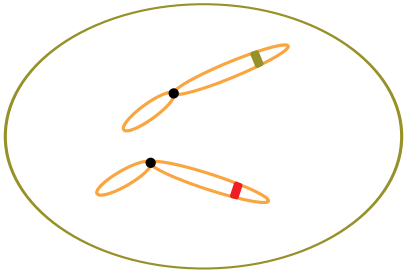
Ogni individuo ha una **coppia di alleli** (**genotipo**), e ad ogni coppia corrisponde un gruppo sanguigno (**fenotipo**)

genotipo	gruppo	genotipo	gruppo
AA	A	BB	B
$A0$	A	$B0$	B
AB	AB	00	0

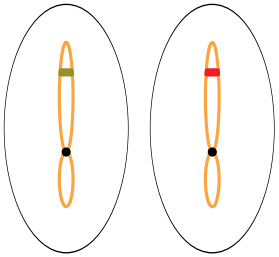
Gli alleli A e B sono codominanti, l'allele 0 è recessivo.

La trasmissione dei geni

Le cellule germinali si ottengono per meiosi e hanno un solo cromosoma per ogni coppia omologa.



$\Longrightarrow$

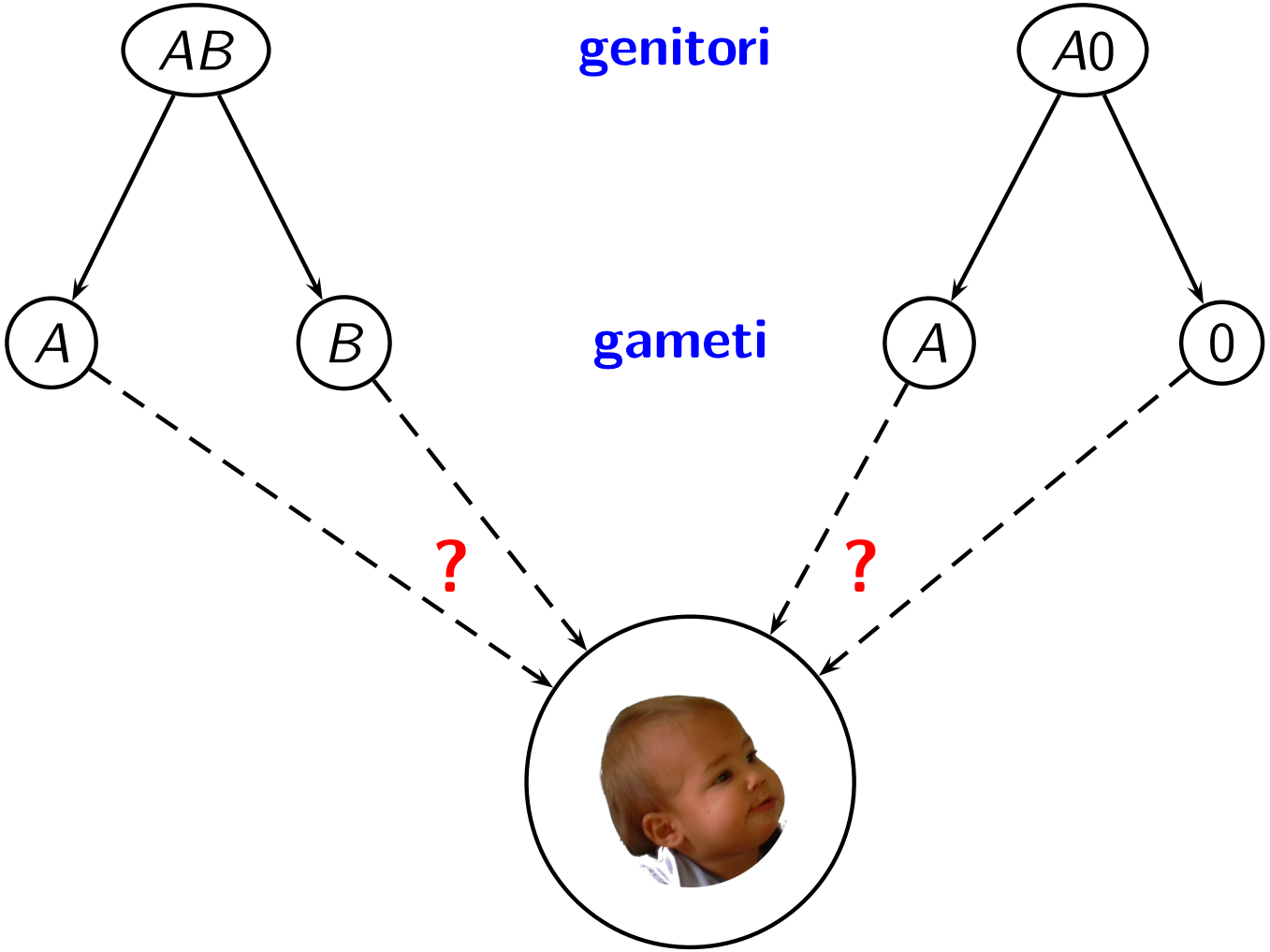


La trasmissione dei geni

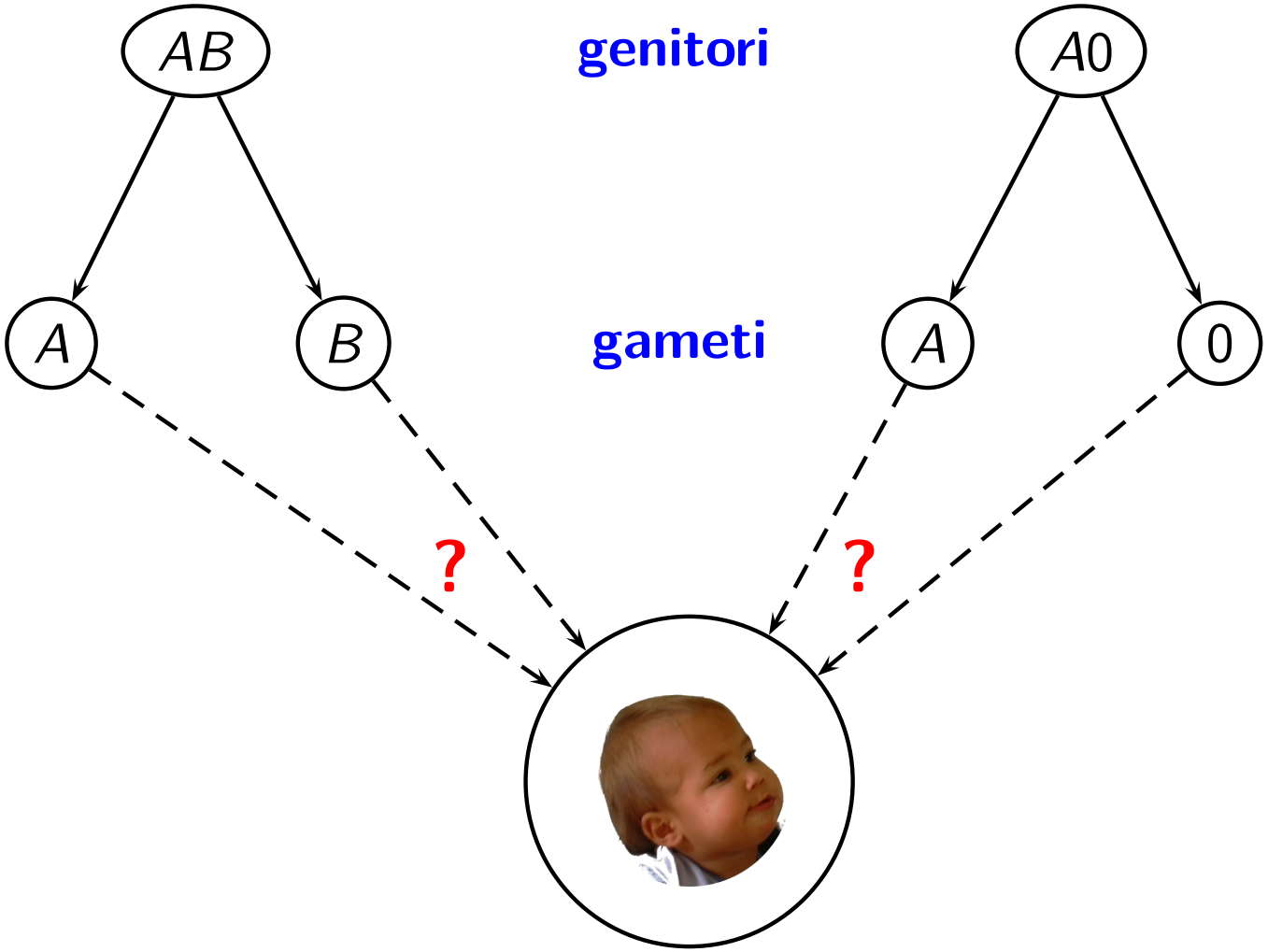
AB **genitori** $A0$

La trasmissione dei geni

La trasmissione dei geni



La trasmissione dei geni



4 possibilità: AA, $A0$, AB, $B0$, con le loro probabilità.

Genetica elementare e probabilità elementare

La genetica di un individuo è frutto del caso, e segue le regole della **probabilità elementare**

Genetica elementare e probabilità elementare

La genetica di un individuo è frutto del caso, e segue le regole della

probabilità elementare

Per questo gli esempi della genetica sono un possibile aiuto per l'insegnamento della probabilità, alternativo o complementare a quello dei giochi di dadi e carte.

Genetica elementare e probabilità elementare

Per esempio, usando il cosiddetto **approccio classico**, che recita

"probabilità = casi possibili / casi favorevoli",

si vede che i 4 genotipi della slide precedente

$$\text{genotipi:} \quad AA \quad A0 \quad AB \quad B0$$

hanno tutti probabilità 1/4.

Genetica elementare e probabilità elementare

Per esempio, usando il cosiddetto **approccio classico**, che recita

"probabilità = casi possibili / casi favorevoli",

si vede che i 4 genotipi della slide precedente

$$\text{genotipi:} \quad AA \quad A0 \quad AB \quad B0$$

hanno tutti probabilità $1/4$.

D'altra parte i corrispondenti fenotipi possibili sono solo 3

$$\text{fenotipi:} \quad A \quad B \quad AB$$

ma il fenotipo A corrisponde a AA e $A0$, e ha probabilità $1/2 = 1/4 + 1/4$.

Si usa la **somma delle probabilità degli eventi incompatibili**

Genetica elementare e probabilità elementare

Nel gioco degli alleli, ha un ruolo cruciale l'**indipendenza tra eventi**, che è alla base delle leggi di Mendel.

È così importante in genetica che esiste un modo standard di rappresentazione della regola del prodotto delle probabilità per eventi indipendenti, i **quadrati di Punnet**.

Genetica elementare e probabilità elementare

Per esempio, da due genitori "eterozigoti" $A0 \times A0$

genitore 1 / genitore 2	A (1/2)	0 (1/2)
A (1/2)	AA (1/4)	A0 (1/4)
0 (1/2)	0A (1/4)	00 (1/4)

La probabilità dei due **omoziogoti** AA e 00 è $1/4$ ciascuno.

La probabilità dell'**eterozigote** $A0$ è $1/2$.

La probabilità del fenotipo A è $1/4 + 1/2 = 3/4$.

Genetica elementare e probabilità elementare

Giocando sul fatto che il genotipo si osserva solo con test genetici, mentre il fenotipo è più visibile, si può introdurre, con un esempio realistico, anche la **probabilità condizionata.**

Genetica elementare e probabilità elementare

Giocando sul fatto che il genotipo si osserva solo con test genetici, mentre il fenotipo è più visibile, si può introdurre, con un esempio realistico, anche la **probabilità condizionata.**

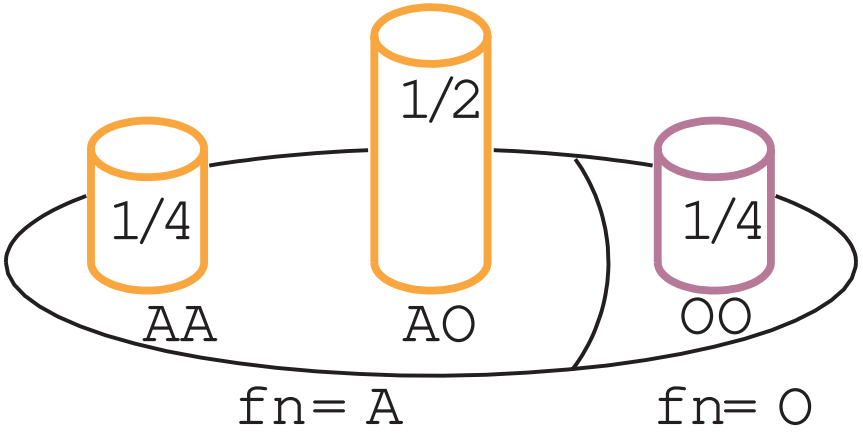
Tradizionalmente si usano esempi del tutto artificiali come

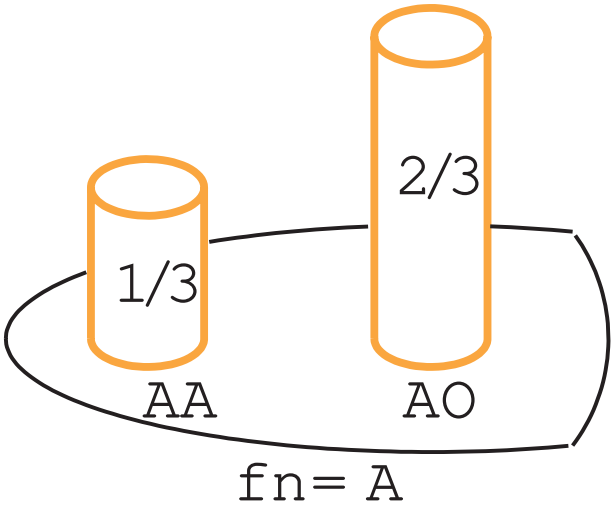
"qual è la probabilità che esce 2 a un lancio di un dado se so che è uscito un numero pari?"

Genetica elementare e probabilità elementare

Un esempio genetico (da E. Batschelet, "Introduzione alla matematica per biologi"): un figlio di genitori "eterozigoti" $A0 \times A0$ è di gruppo sanguigno A. Con quale probabilità è di genotipo AA?

Genetica elementare e probabilità elementare





Geni recessivi

Nel 1908, il genetista R. Punnet si chiedeva insieme ad un suo amico di cricket, il matematico G.H. Hardy, come mai gli alleli dominanti non finissero per essere gli unici presenti nelle popolazioni.

Geni recessivi

Nel 1908, il genetista R. Punnet si chiedeva insieme ad un suo amico di cricket, il matematico G.H. Hardy, come mai gli alleli dominanti non finissero per essere gli unici presenti nelle popolazioni.

Hardy era un "matematico puro", che si occupava di teoria dei numeri. Ciò nonostante risolse il problema di Punnet, formulando la **legge dell'equilibrio genetico di Hardy-Weinberg**.

Frequenze di genotipi e frequenze di alleli

Una popolazione di 100 individui

A0	00	00	A0	AB	00	00	00	00	B0
00	A0	AB	B0	BB	00	00	A0	00	AB
A0	A0	00	B0	AA	00	A0	00	00	B0
A0	A0	B0	A0	00	AA	00	00	B0	00
A0	AA	A0	AB	A0	B0	00	00	A0	00
A0	A0	00	00	00	B0	00	AB	00	B0
00	A0	A0	00	AA	00	B0	00	00	00
B0	00	00	A0	00	B0	A0	00	00	00
B0	00	00	B0	AB	AA	A0	A0	A0	A0
A0	A0	B0	A0	AA	A0	A0	00	A0	A0

Posso contare la frequenza relativa dei genotipi:

"AB" sul totale $= 6/100$, che è anche la probabilità di scegliere a caso un individuo di genotipo AB

Frequenze di genotipi e frequenze di alleli

Una popolazione di 100 individui

A0	00	00	A0	AB	00	00	00	00	B0
00	A0	AB	B0	BB	00	00	A0	00	AB
A0	A0	00	B0	AA	00	A0	00	00	B0
A0	A0	B0	A0	00	AA	00	00	B0	00
A0	AA	A0	AB	A0	B0	00	00	A0	00
A0	A0	00	00	00	B0	00	AB	00	B0
00	A0	A0	00	AA	00	B0	00	00	00
B0	00	00	A0	00	B0	A0	00	00	00
B0	00	00	B0	AB	AA	A0	A0	A0	A0
A0	A0	B0	A0	AA	A0	A0	00	A0	A0

Oppure posso contare la frequenza relativa degli alleli:

"B" sul totale $= 23/200 = 2.3/100$, che è la probabilità di scegliere un allele B a caso

La legge di Hardy-Weinberg

Come sono distribuiti i genotipi dei figli?

A0	00	00	A0	AB	00	00	00	00	B0
00	A0	AB	B0	BB	00	00	A0	00	AB
A0	A0	00	B0	AA	00	A0	00	00	B0
A0	A**0**	B0	A0	00	AA	00	00	B0	00
A0	AA	A0	AB	A0	B0	00	00	A0	00
A0	A0	00	00	00	B0	00	AB	00	B0
00	A0	A0	00	AA	00	B0	00	00	00
B0	00	00	A0	00	B0	A0	00	00	00
B0	00	00	**B**0	AB	AA	A0	A0	A0	A0
A0	A0	B0	A0	AA	A0	A0	00	A0	A0

Figlio B0

La legge di Hardy-Weinberg

Come sono distribuiti i genotipi dei figli?

A0	00	00	A0	AB	00	00	00	00	B0
00	A0	AB	B0	BB	00	00	A0	00	AB
A0	A0	00	B0	AA	00	**A**0	00	00	B0
A0	A0	B0	A0	00	AA	00	00	B0	00
A0	AA	A0	AB	A0	B0	00	00	A0	00
A0	A0	00	00	00	B0	00	AB	00	B0
00	A0	A0	00	AA	00	B0	00	00	00
B0	00	00	A0	00	B0	A0	00	00	00
B0	00	00	B0	**A**B	AA	A0	A0	A0	A0
A0	A0	B0	A0	AA	A0	A0	00	A0	A0

Figlio AA

La legge di Hardy-Weinberg

Come sono distribuiti i genotipi dei figli?

A0	00	00	A0	AB	00	00	00	00	B0
00	A0	AB	B0	BB	00	00	A0	00	AB
A0	A0	00	B0	AA	00	**A**0	00	00	B0
A0	A0	B0	A0	00	AA	00	00	B0	00
A0	AA	A0	AB	A0	B0	00	00	A0	00
A0	A0	00	00	00	B0	00	AB	00	B0
00	A0	A0	00	AA	00	B0	00	00	00
B0	00	00	A0	00	B0	A0	00	00	00
B0	00	00	B0	**A**B	AA	A0	A0	A0	A0
A0	A0	B0	A0	AA	A0	A0	00	A0	A0

È come scegliere a caso due alleli!

Ma allora la frequenza degli alleli **non cambia** con le generazioni!

La legge di Hardy-Weinberg

È un modello matematico, vero sotto ipotesi semplificative

- distribuzione degli alleli identica tra maschi e femmine
- popolazione abbastanza grande da poter considerare casuale la relazione di genitorialità
- il gene deve essere **neutro per la selezione**
- è il modello di base per l'analisi fuori dall'equilibrio, cioè quando c'è selezione

La legge di Hardy-Weinberg

È un modello matematico, vero sotto ipotesi semplificative

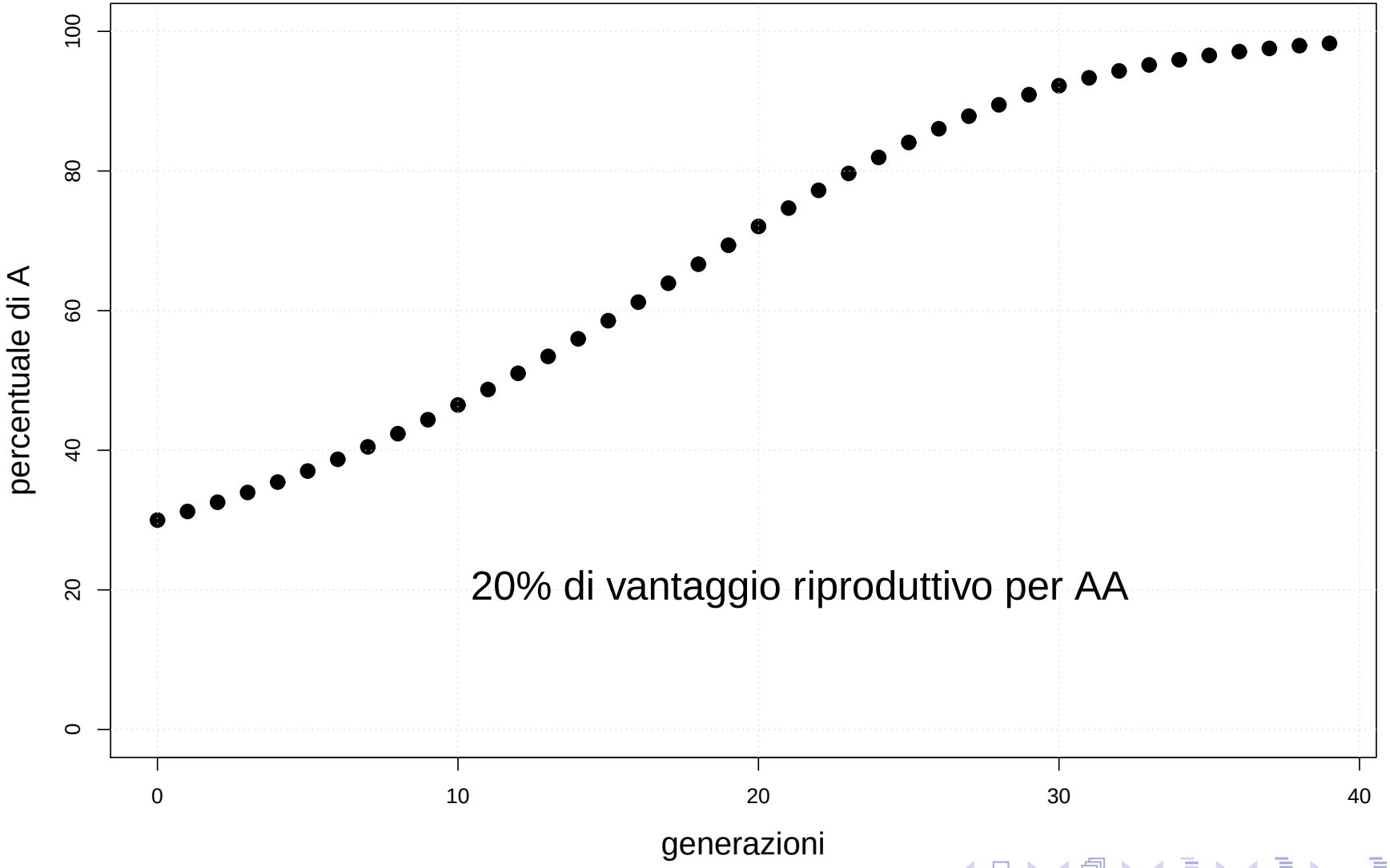
- distribuzione degli alleli identica tra maschi e femmine
- popolazione abbastanza grande da poter considerare casuale la relazione di genitorialità
- il gene deve essere **neutro per la selezione**
- è il modello di base per l'analisi fuori dall'equilibrio, cioè quando c'è selezione

Molti laboratori didattici di tipo modellistico introducono i **sistemi dinamici discreti** (leggi di aggiornamento a tempi discreti). È lo strumento con cui si può analizzare l'effetto della selezione.

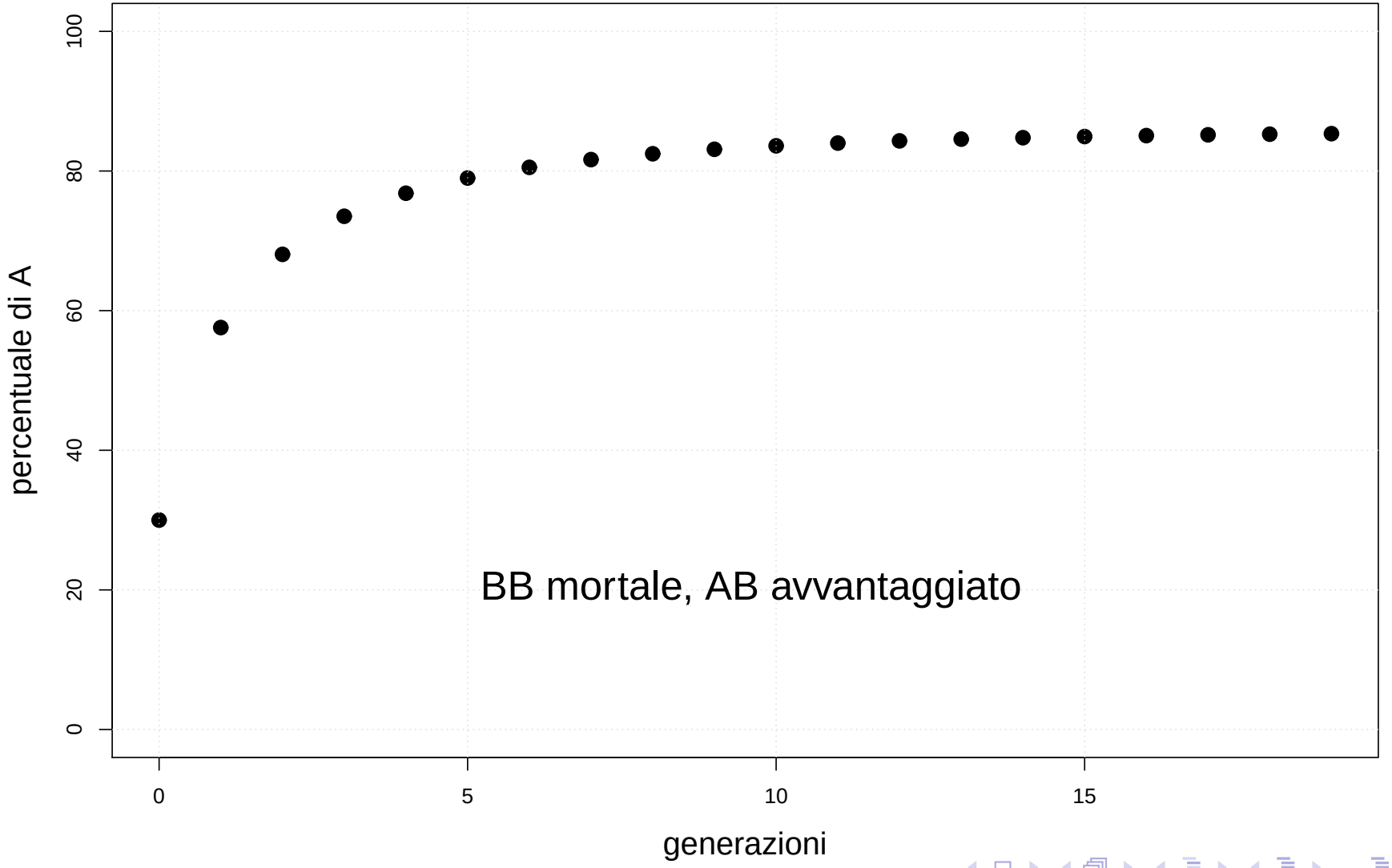
Un esempio di selezione

Per semplificare, consideriamo il caso di soli due alleli, A e B, e ipotizziamo che gli individui AA abbiano un maggior successo riproduttivo. Nella generazione dei figli la frequenza di alleli A sarà più alta, e questo fenomeno si riprodurrà ad ogni generazione.

Un esempio di selezione



Un esempio più interessante



Dove siamo arrivati.

La fonte della variabilità genetica è dovuta alle mutazioni.

Su questa variabilità agisce la selezione, che porta all'estinzione le varianti alleliche sfavorite.

Dove siamo arrivati.

La fonte della variabilità genetica è dovuta alle mutazioni.
Su questa variabilità agisce la selezione, che porta all'estinzione le varianti alleliche sfavorite.

Ma perché le popolazioni non sono tutte uguali?
Come nascono nuove specie?

Dove siamo arrivati.

La fonte della variabilità genetica è dovuta alle mutazioni.
Su questa variabilità agisce la selezione, che porta all'estinzione le varianti alleliche sfavorite.

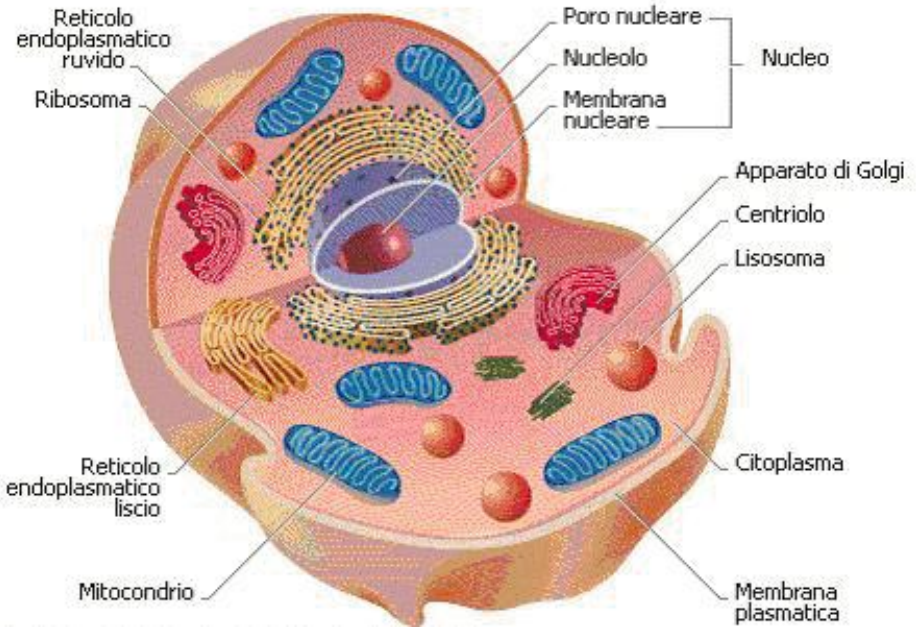
Ma perché le popolazioni non sono tutte uguali?
Come nascono nuove specie?

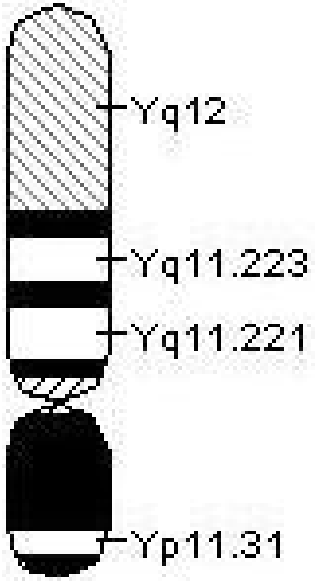
C'è un fenomeno cruciale, la **deriva genetica**.

La deriva genetica

Semplifichiamo ancora: esiste del DNA che è presente in una sola copia, e si eredita da un solo genitore, dunque niente complicazioni sentimentali...

- i geni del cromosoma Y, ereditato, così come è, dal padre
- i geni nel DNA dei mitocondri delle cellule, che provengono dalla duplicazione dei mitocondri dell'ovuolo materno





La deriva genetica

Anche questi geni sono soggetti a mutazione, dunque non sono gli stessi nella popolazione.
Vediamo che succede, per il **solo effetto del caso** a una popolazione con più alleli, con una simulazione che si piò fare con un dado, per una popolazione di 6 individui.

Una simulazione genetica

Ogni gruppo di lavoro si divide in due gruppi, A e B, che hanno due diverse varianti di una caratteristica genetica.

Ogni gruppo ha inizialmente nella popolazione 3 individui: nella colonna della generazione 0, quella iniziale, segnate in ordine 3 volte A e 3 volte B.

Per ogni generazione, ogni individuo ha un certo numero di discendenti, ma il totale non eccede 6 (le risorse vitali sono molto scarse...).

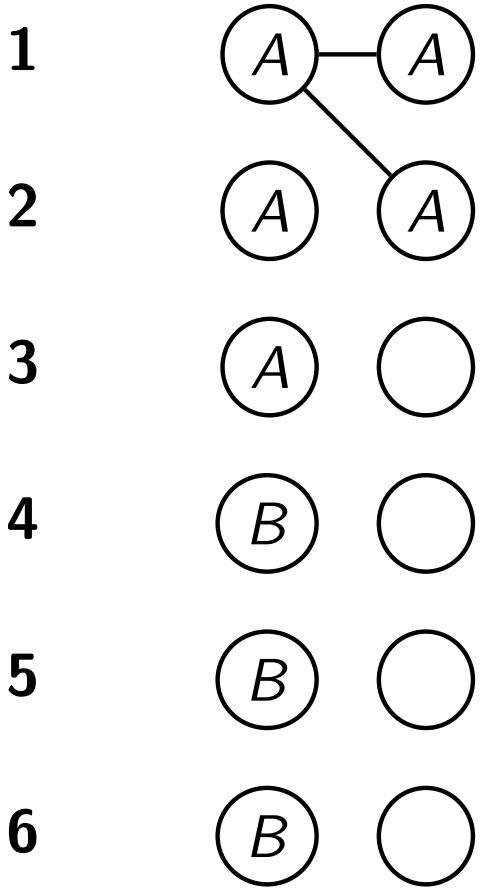
Simuleremo il **numero di figli** di ciascuno con il lancio di 6 dadi.

Una simulazione genetica

1	A	
2	A	
3	A	
4	B	
5	B	
6	B	

pos. **1,2,3,4,5,6**
Esce **1,1,2,4,4,6**

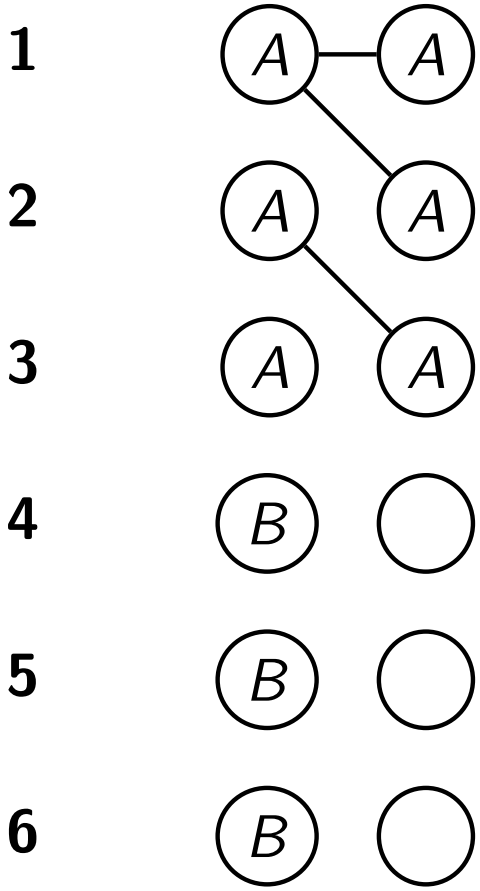
Una simulazione genetica



pos. **1,2,3,4,5,6**

Esce **1,1,2,4,4,6**

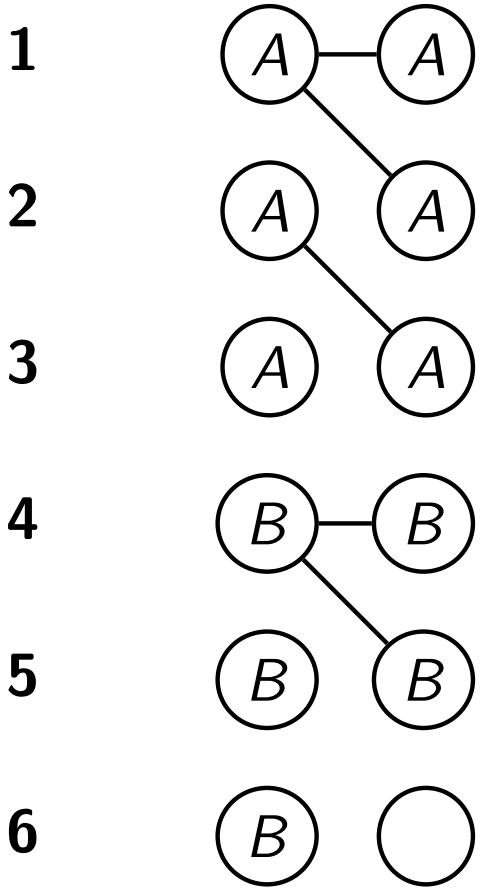
Una simulazione genetica



pos. **1,2,3,4,5,6**

Esce **1,1,2,4,4,6**

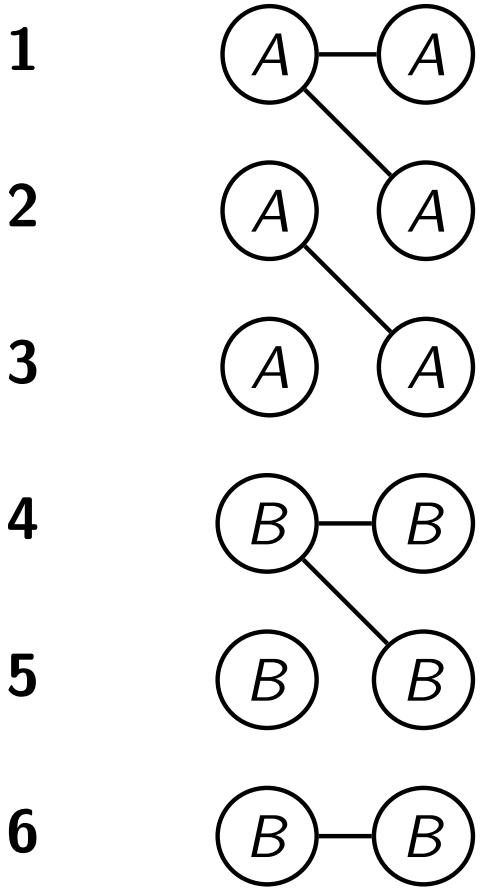
Una simulazione genetica



pos. **1,2,3,4,5,6**

Esce **1,1,2,4,4,6**

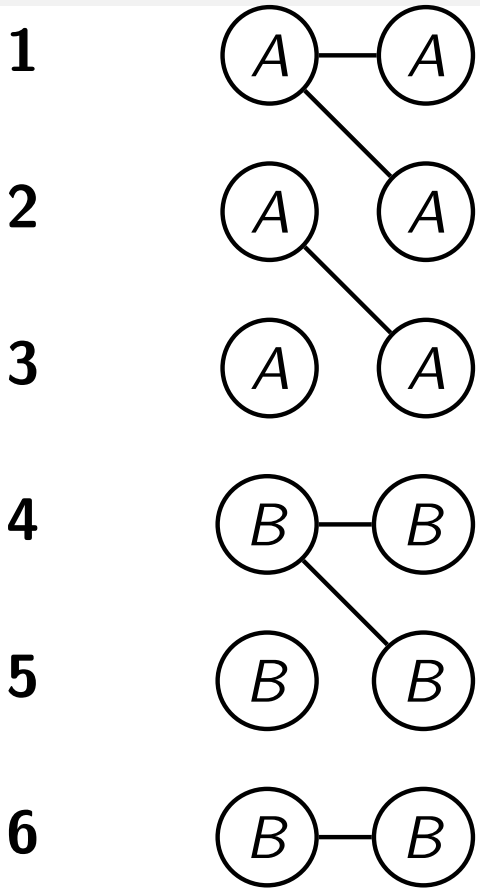
Una simulazione genetica



pos. **1,2,3,4,5,6**

Esce **1,1,2,4,4,6**

Una simulazione genetica



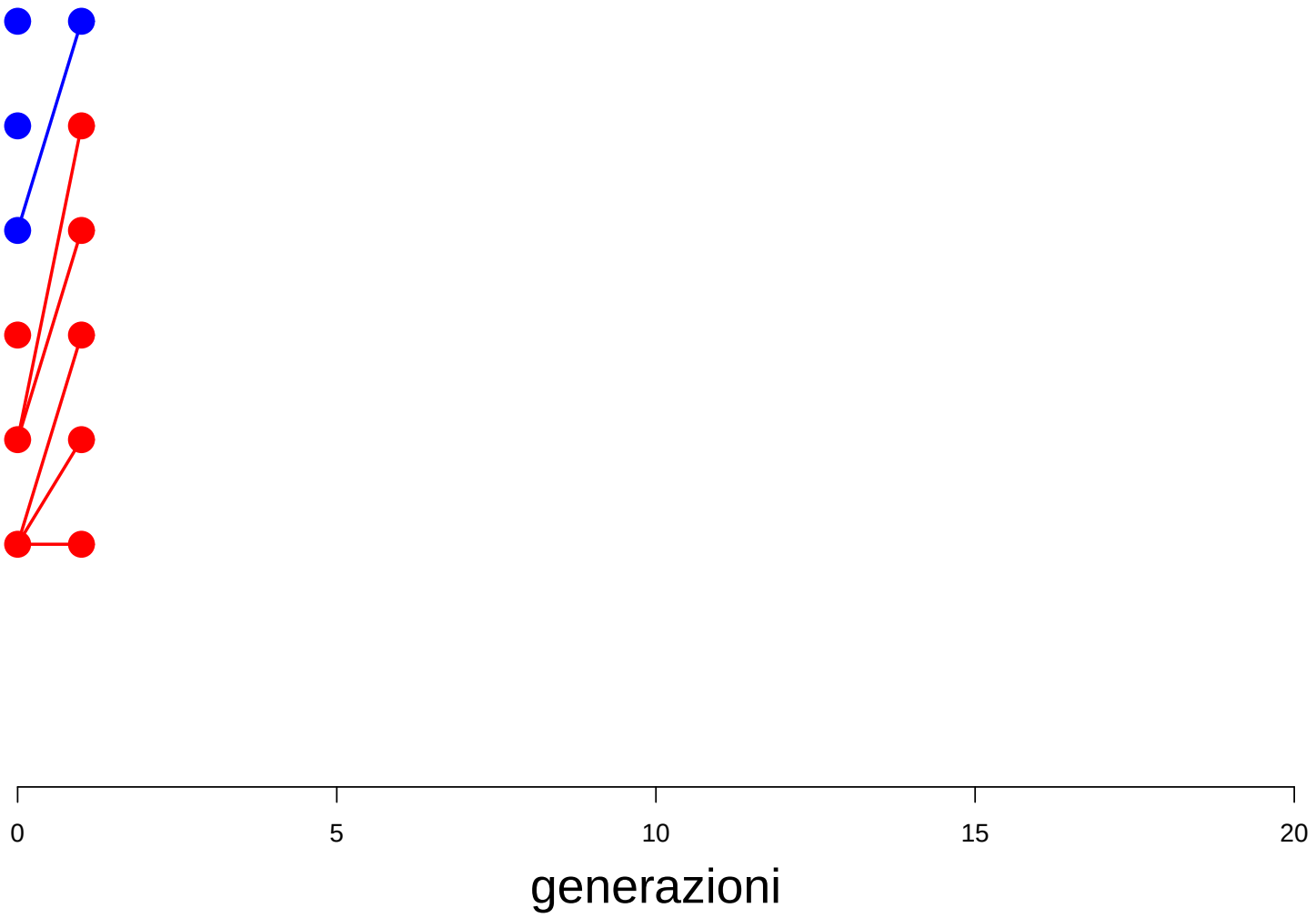
pos. **1,2,3,4,5,6**

Esce **1,1,2,4,4,6**

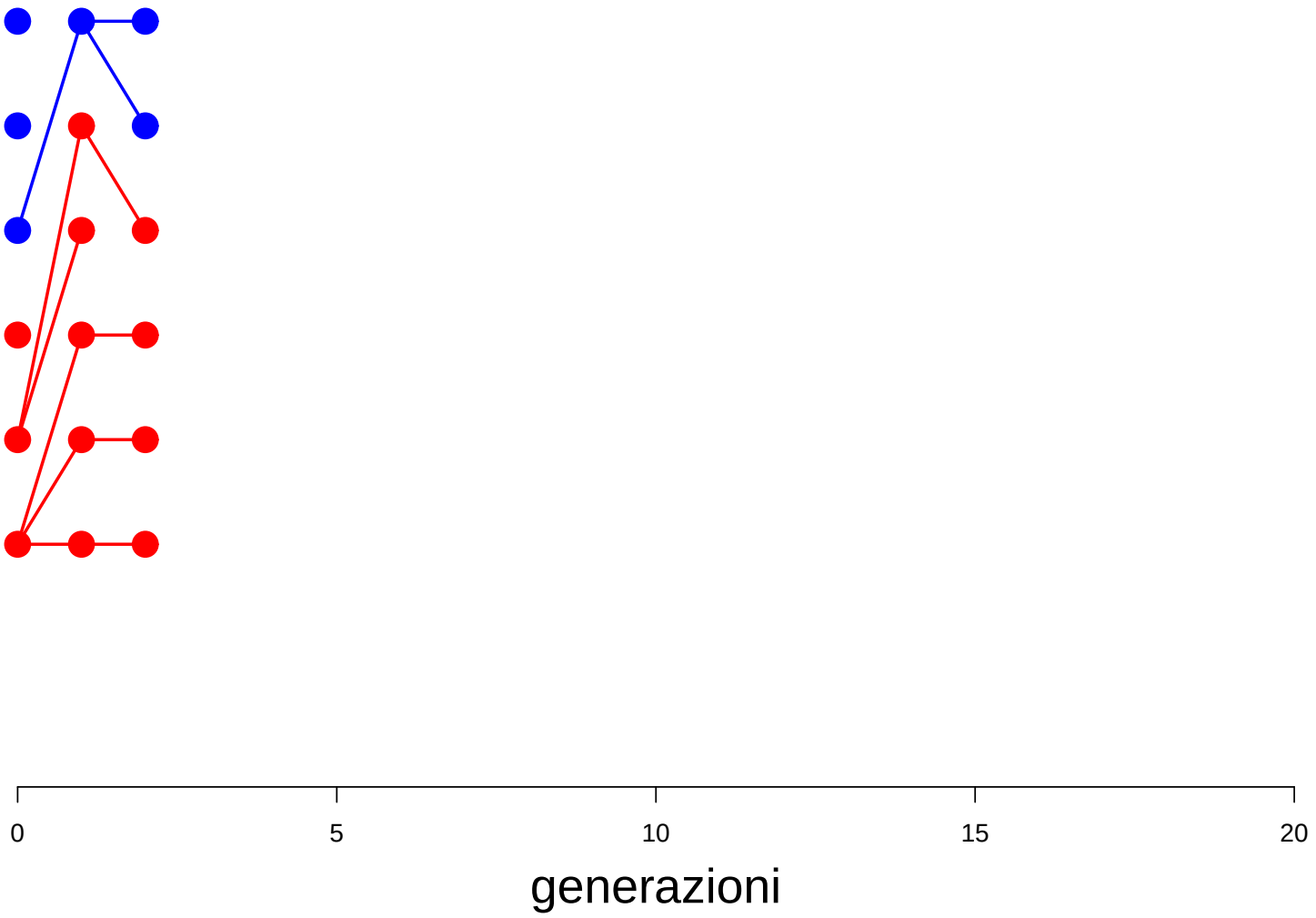
Inizialmente, la frequenza di A è del 50%, quella di B è del 50%.

Cosa accadrà al passare delle generazioni?

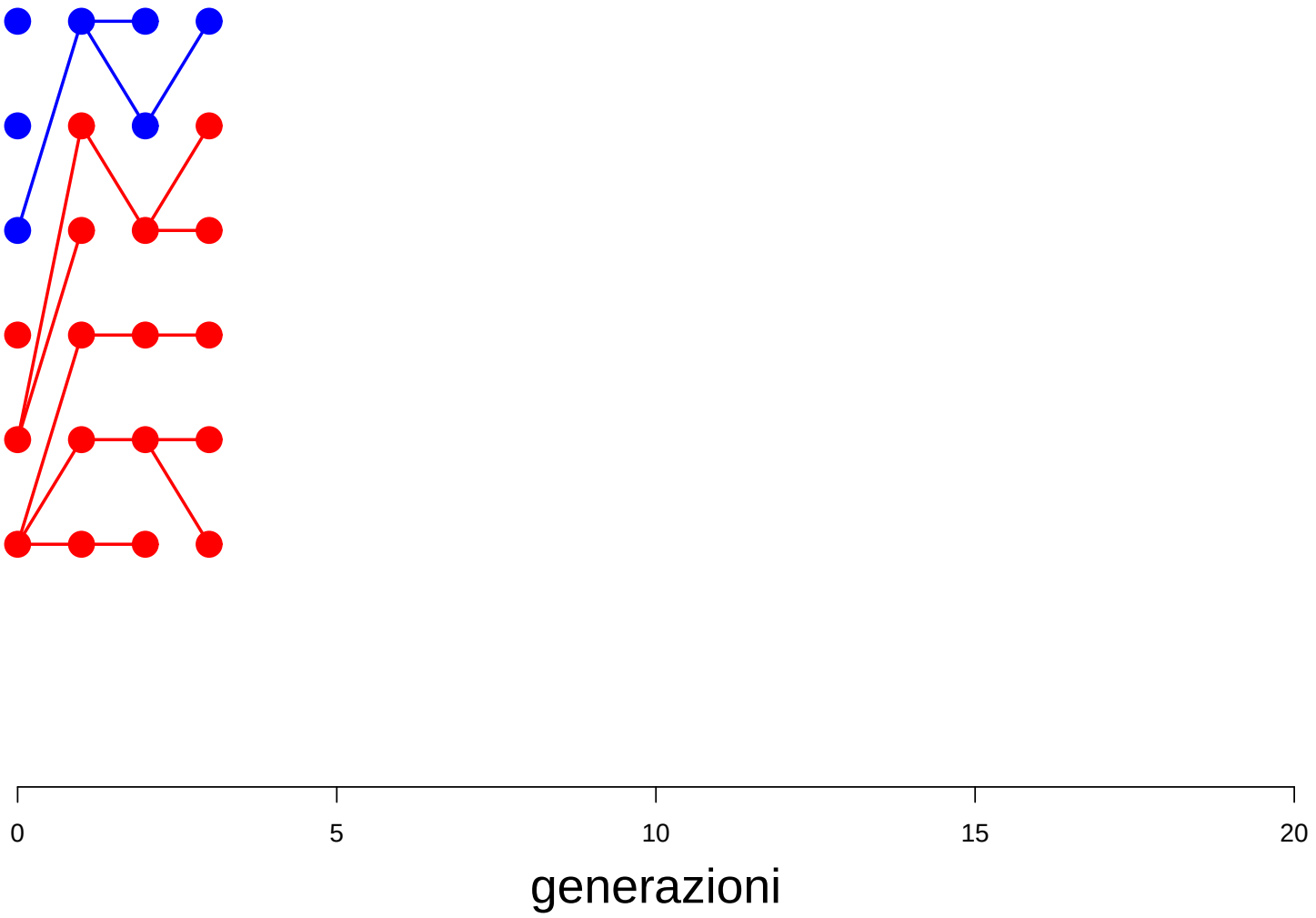
Una simulazione genetica



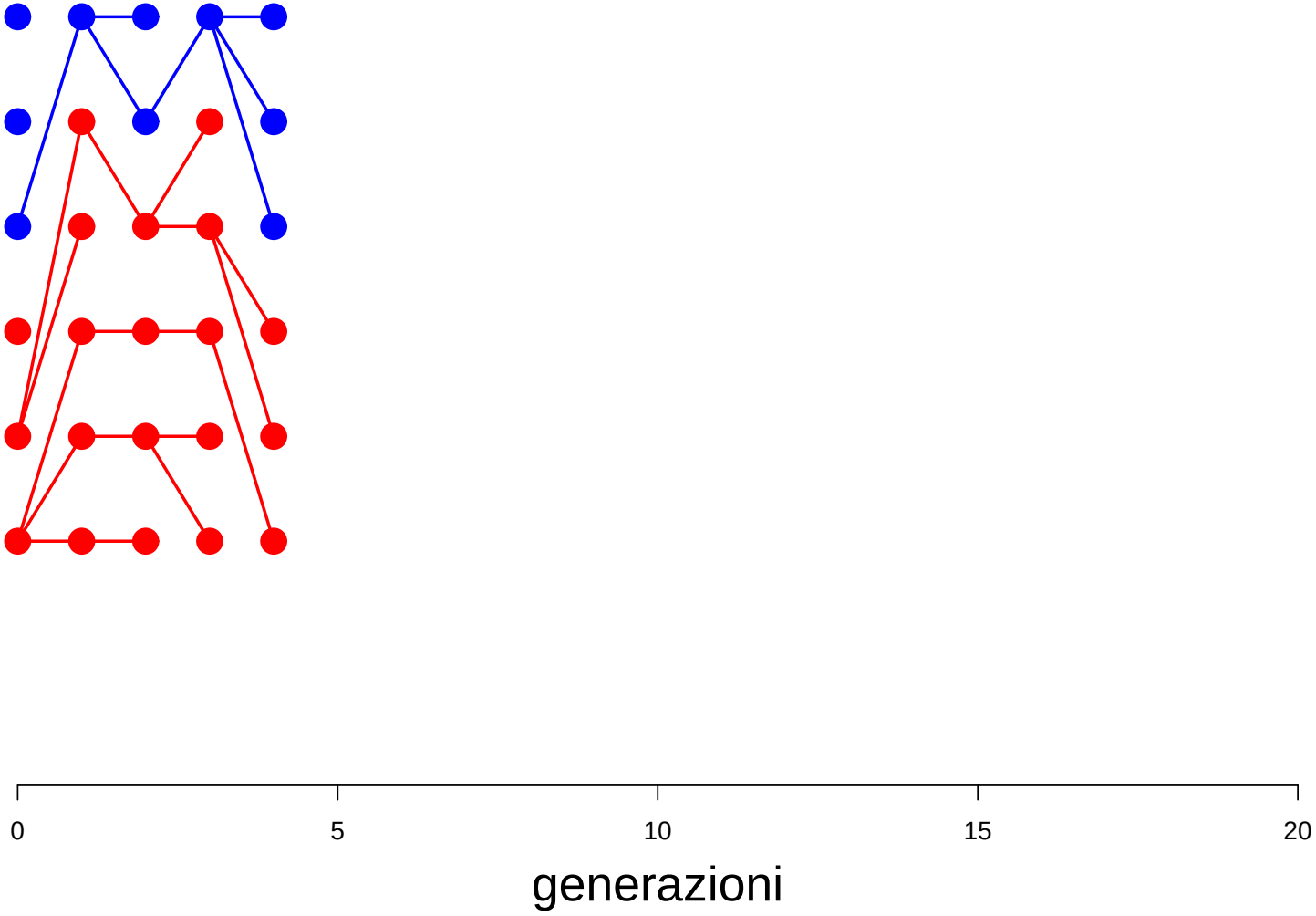
Una simulazione genetica



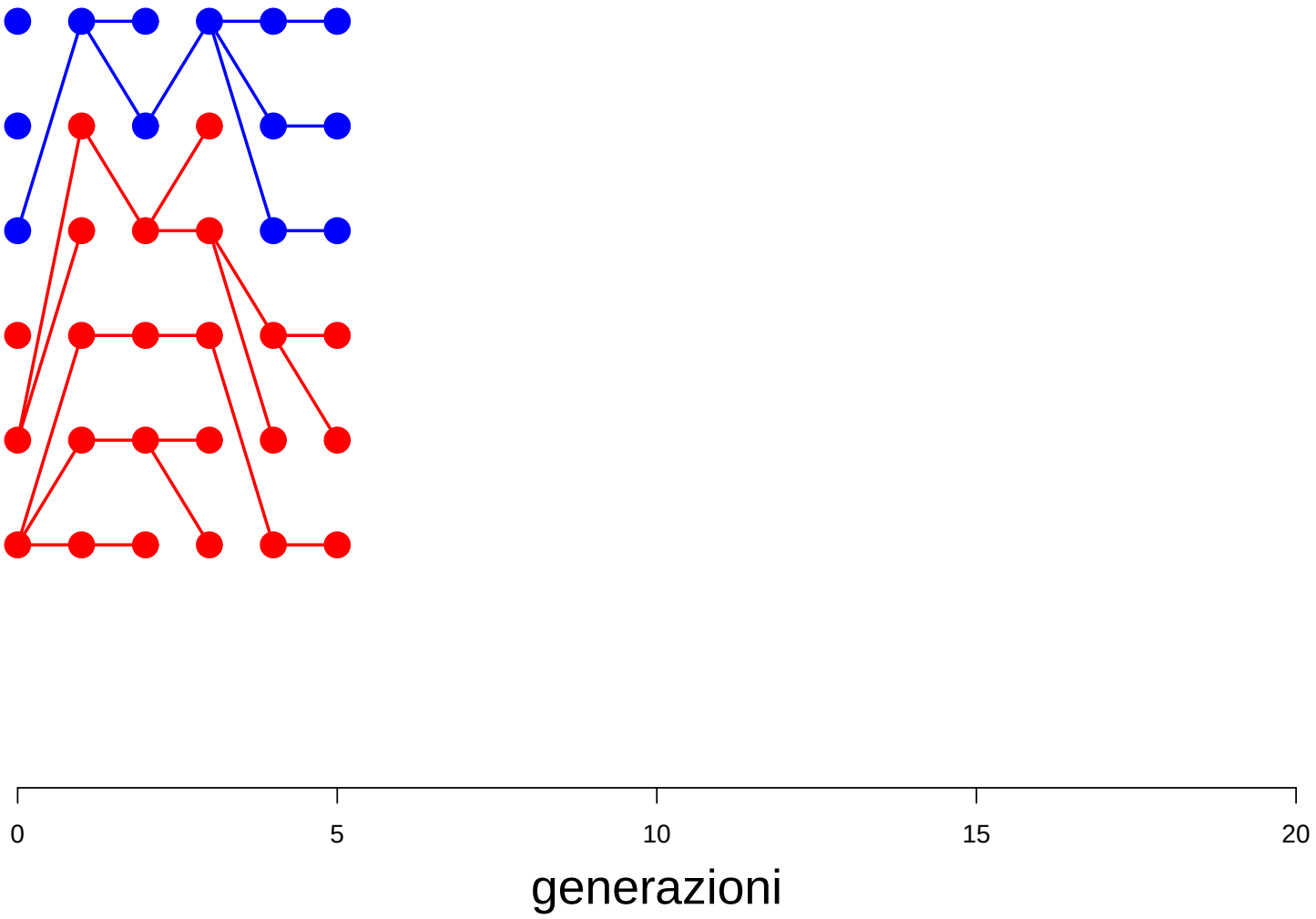
Una simulazione genetica



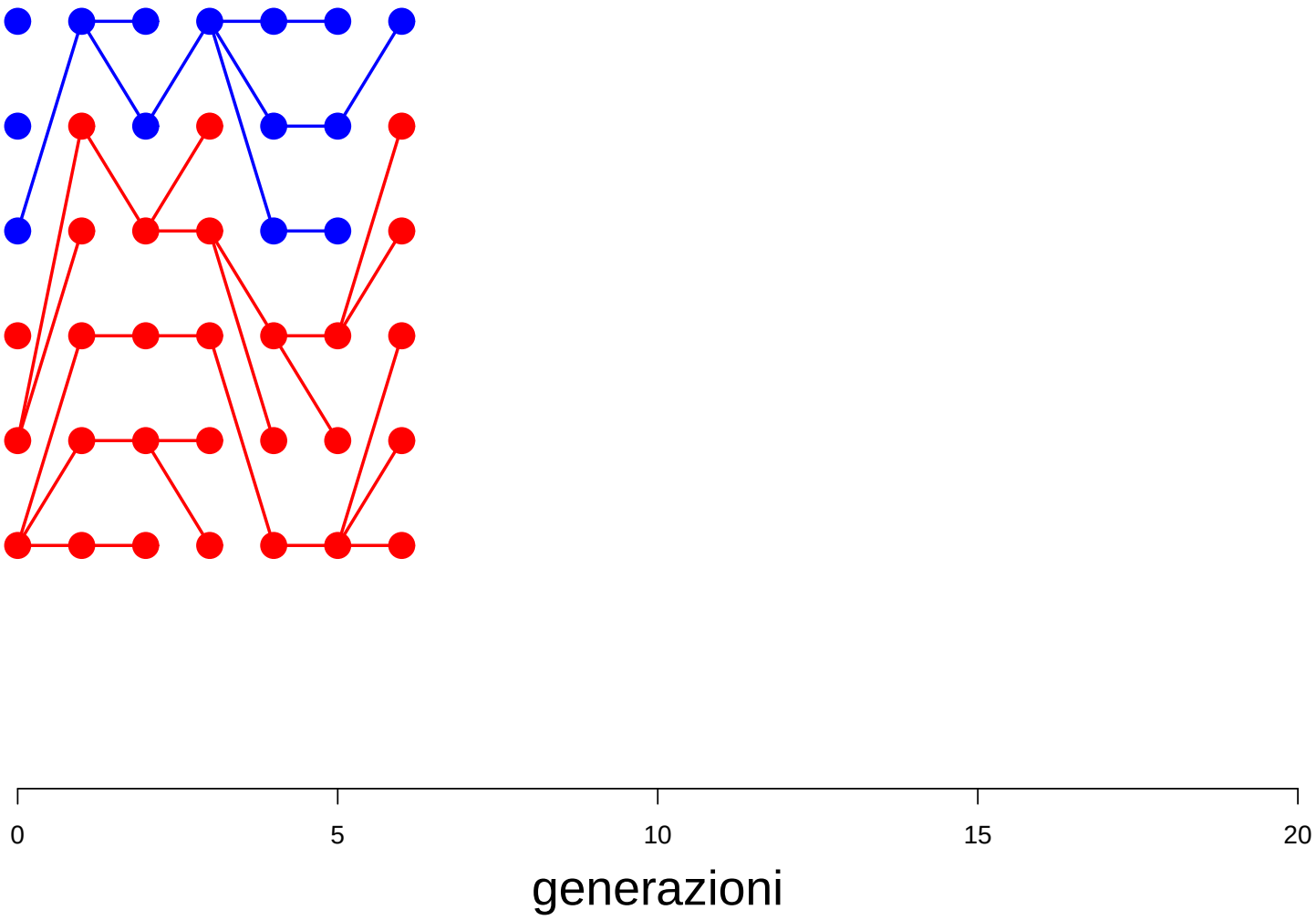
Una simulazione genetica



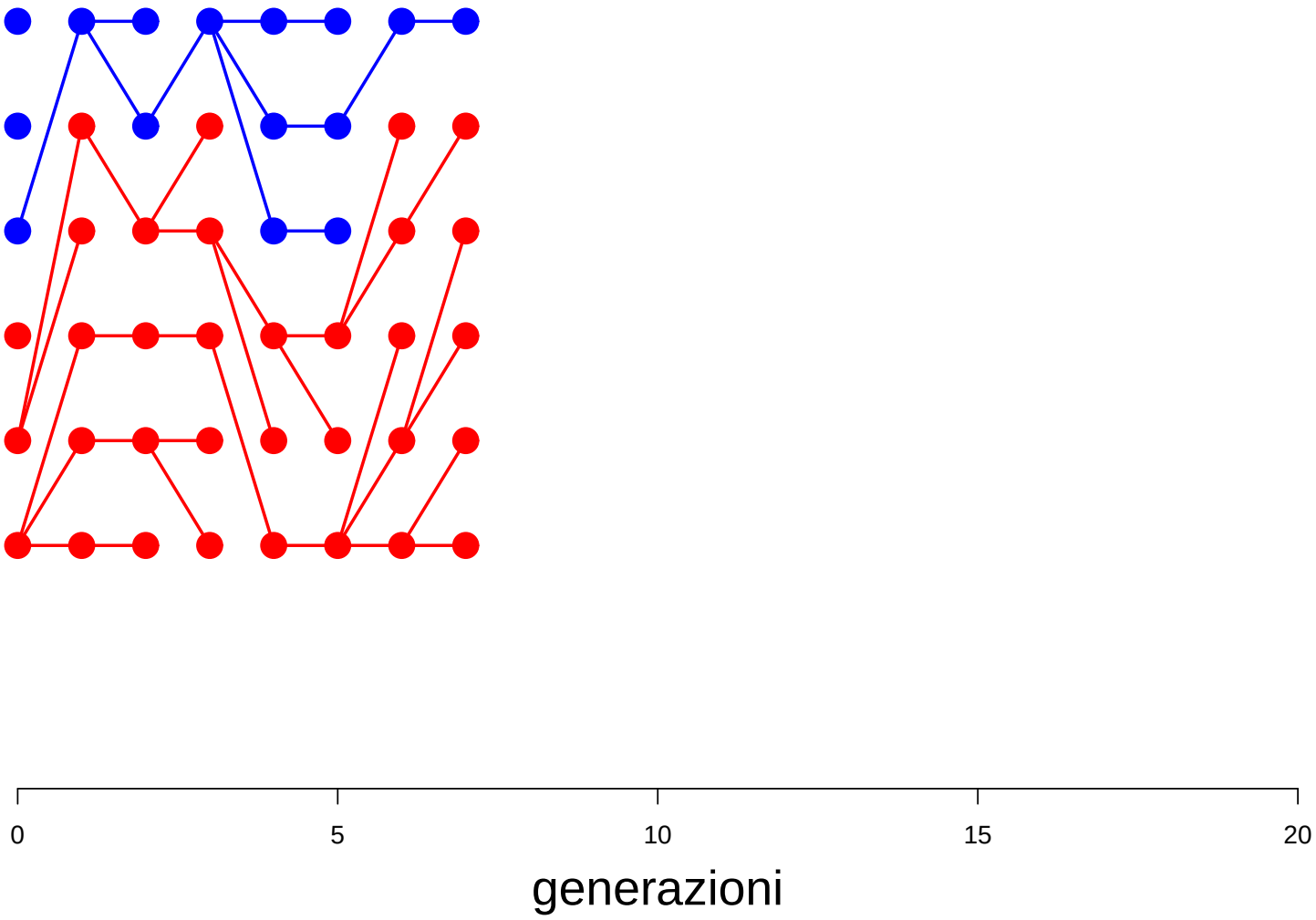
Una simulazione genetica



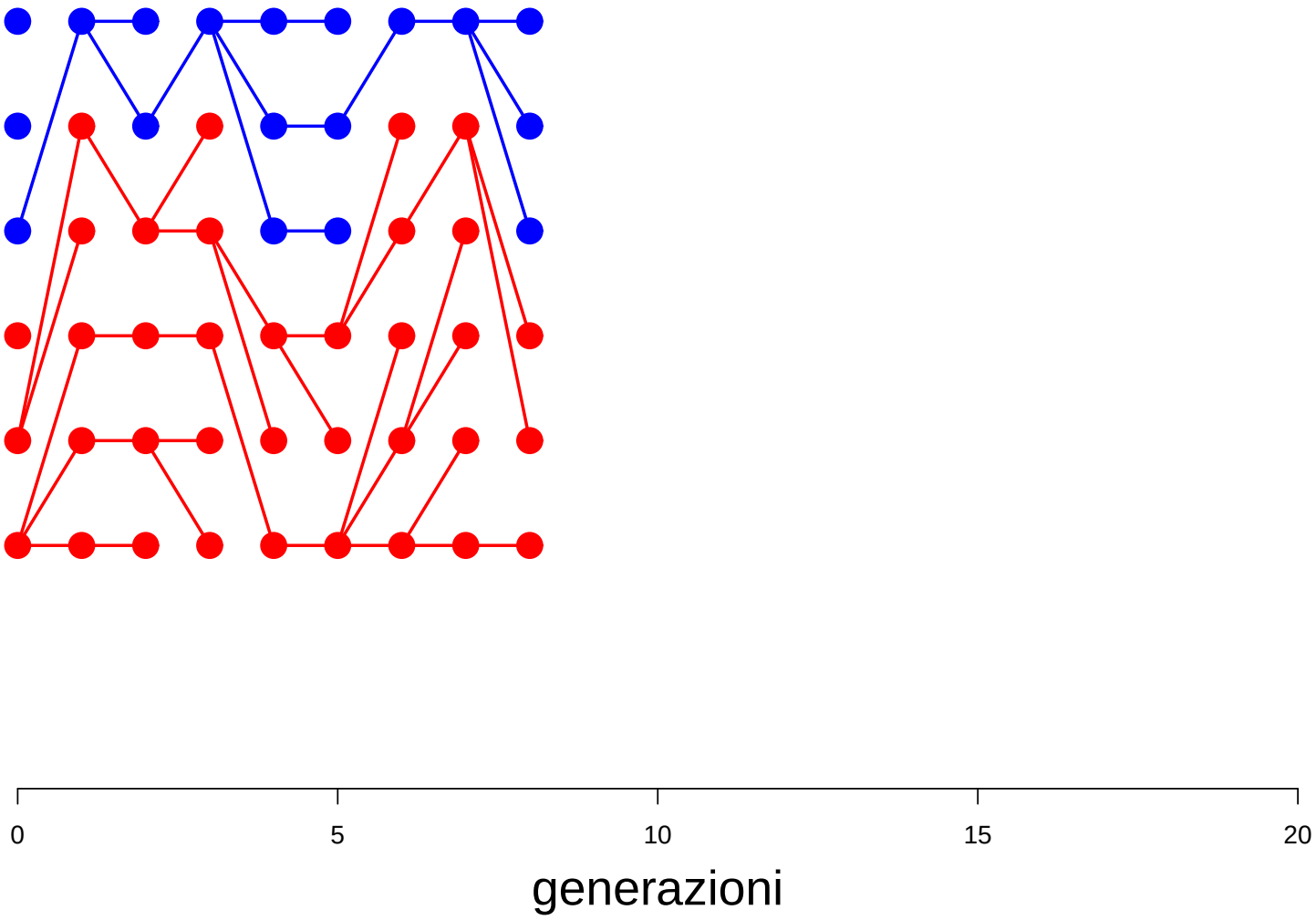
Una simulazione genetica



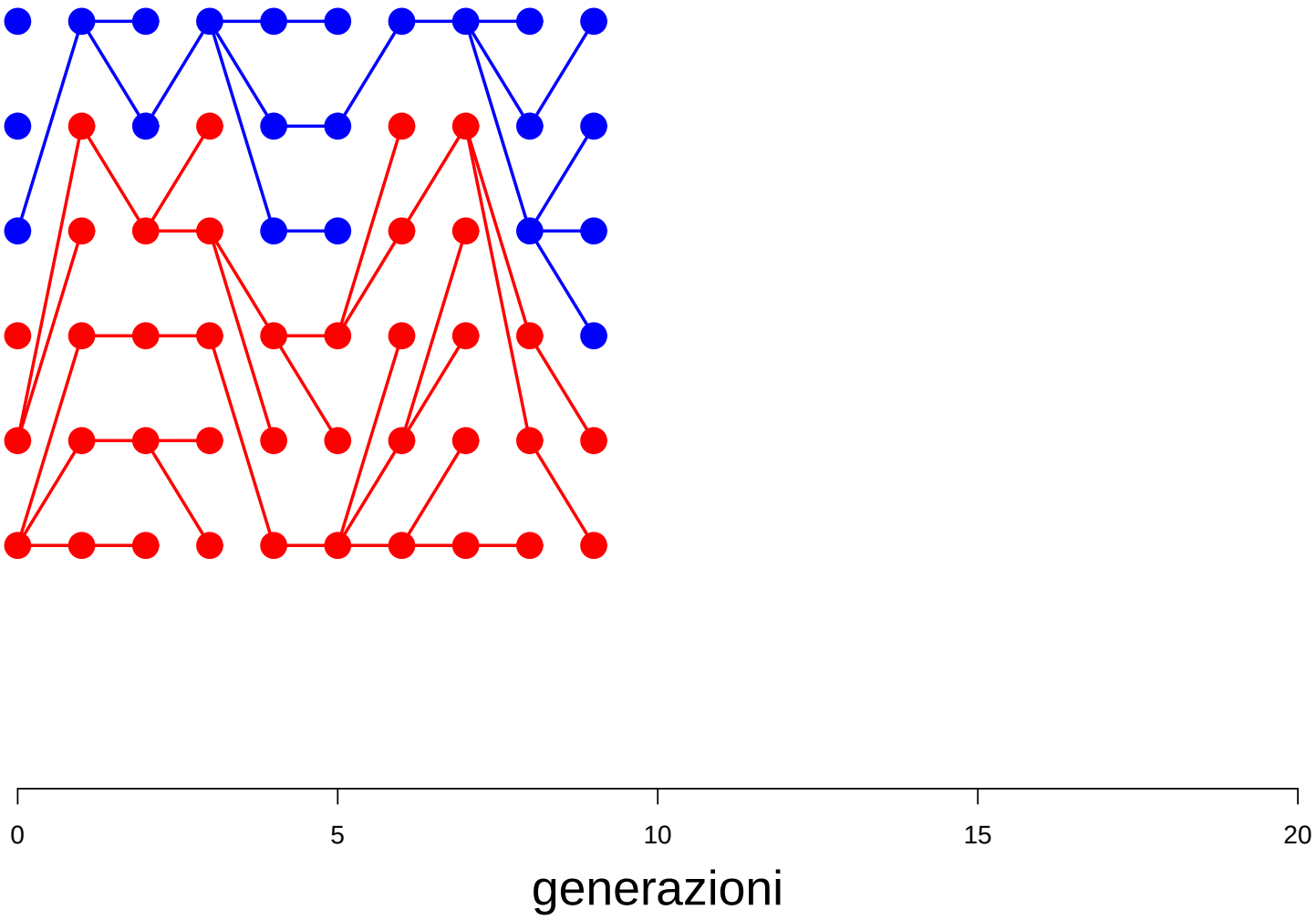
Una simulazione genetica



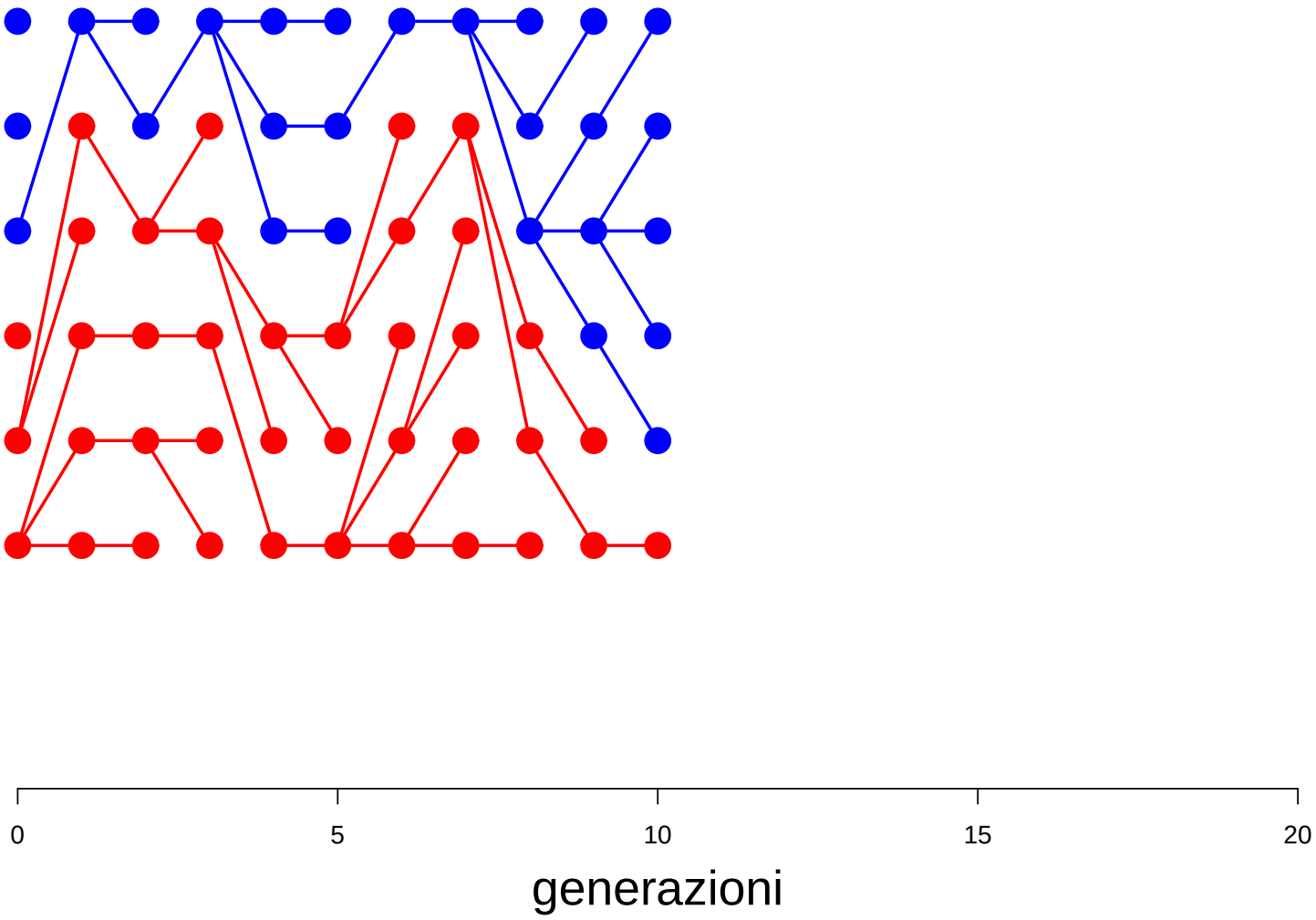
Una simulazione genetica



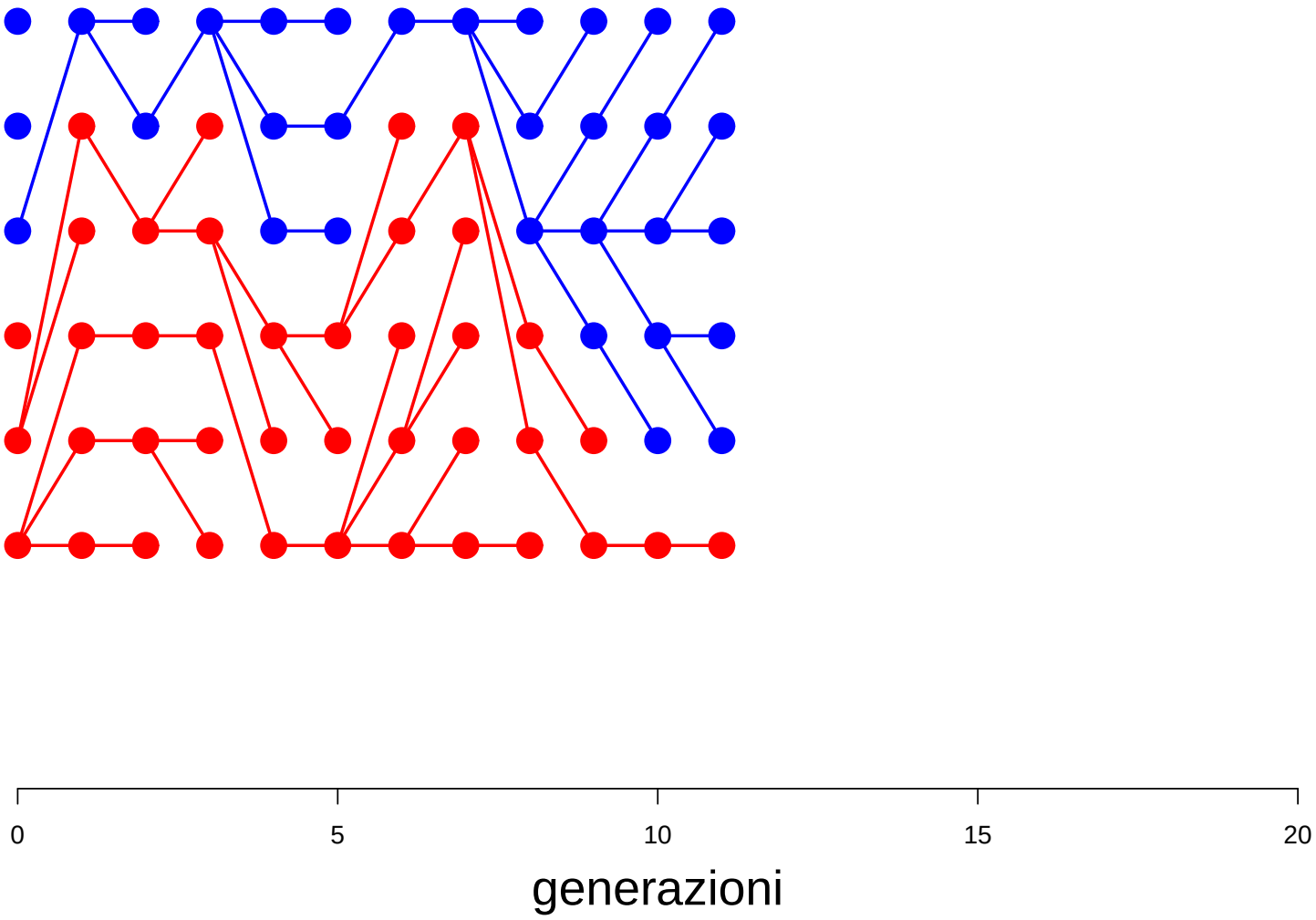
Una simulazione genetica



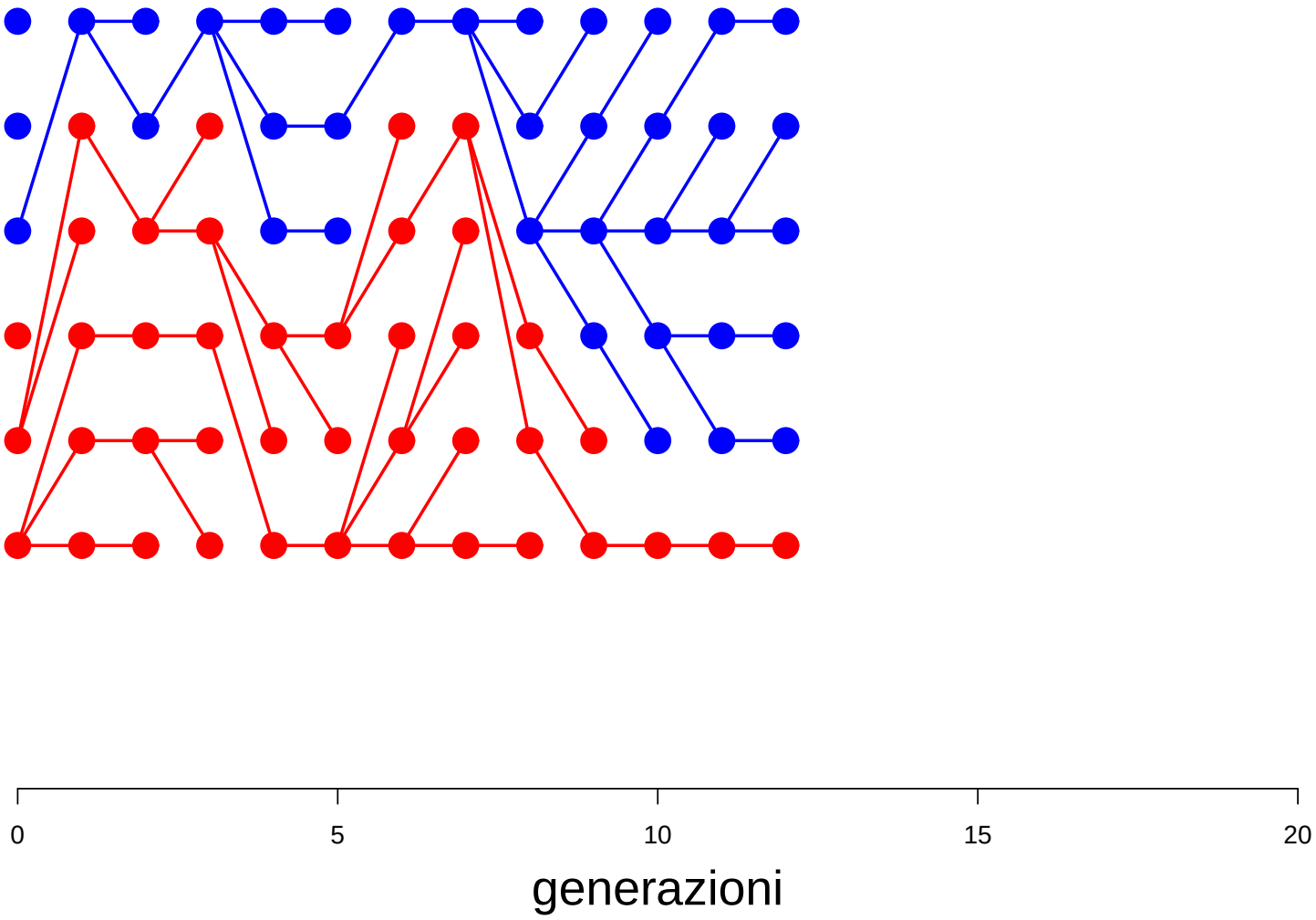
Una simulazione genetica



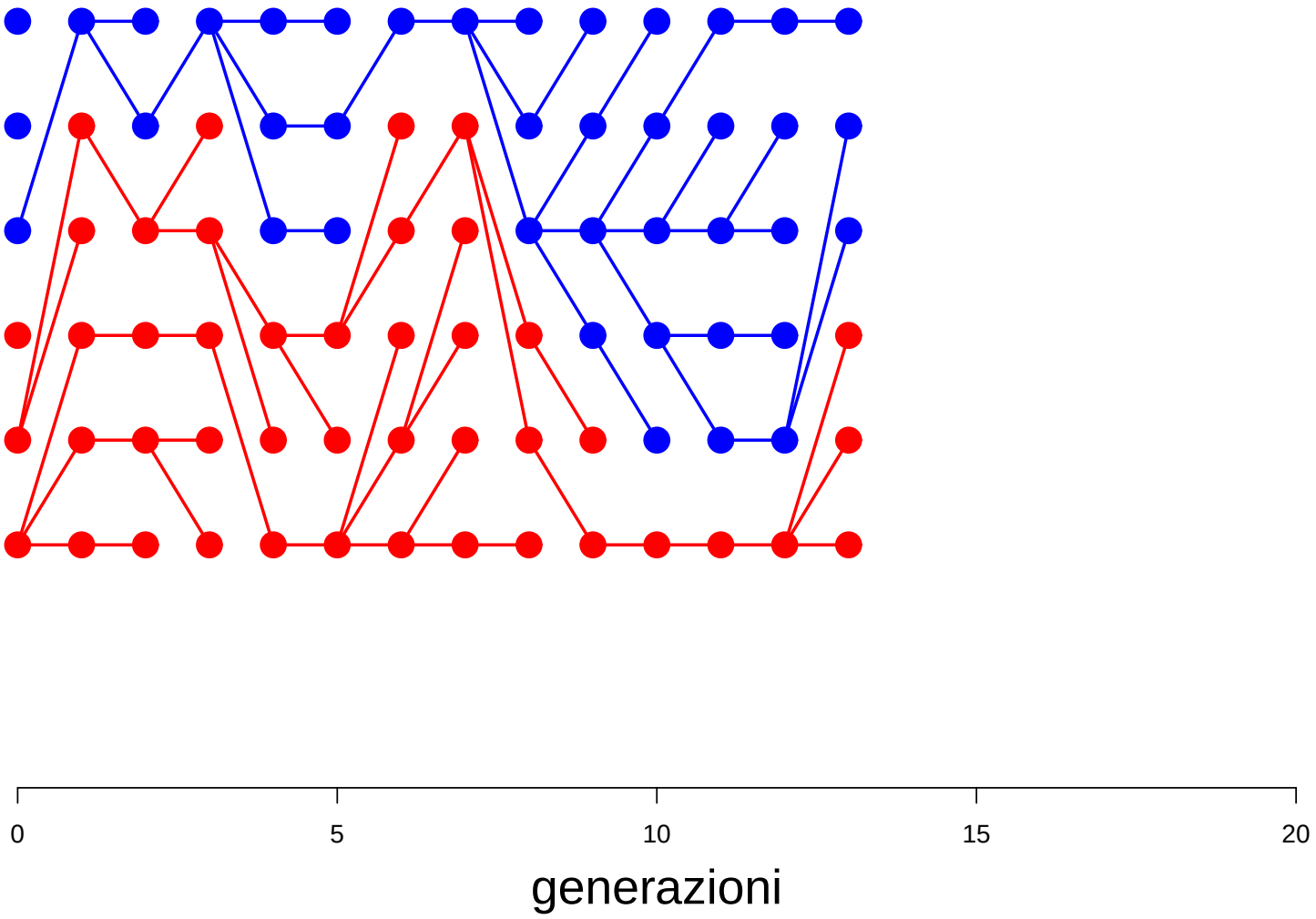
Una simulazione genetica



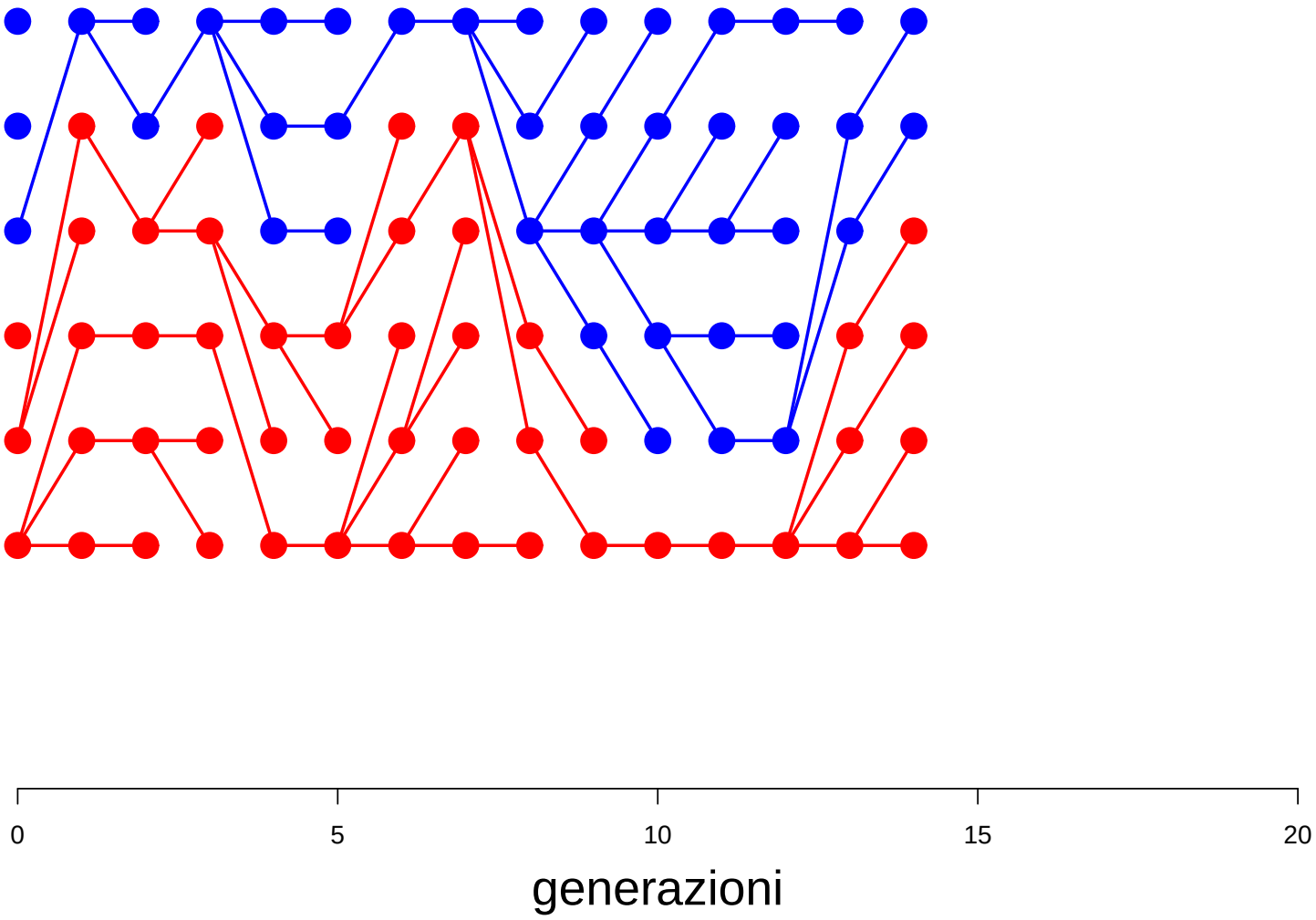
Una simulazione genetica



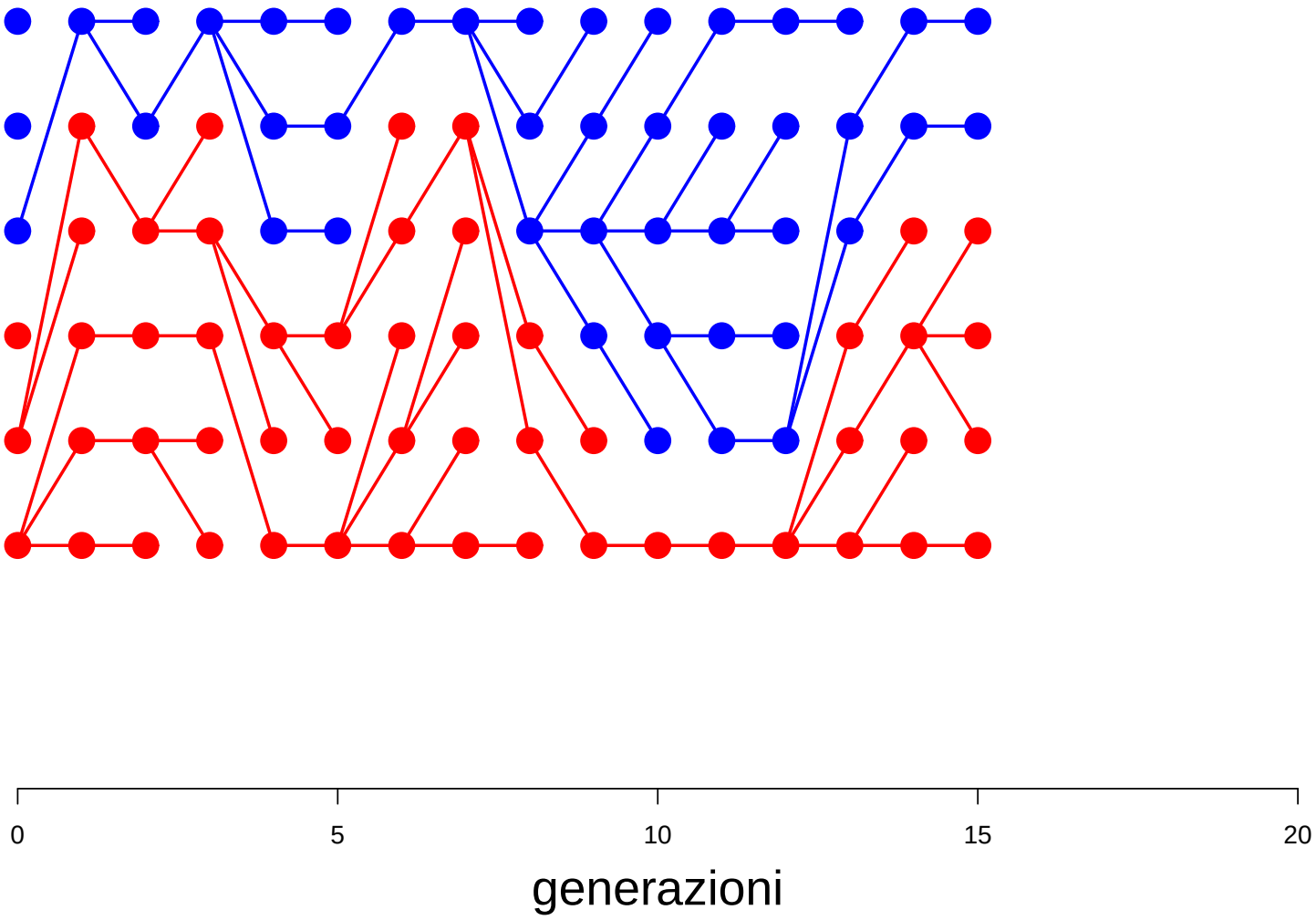
Una simulazione genetica



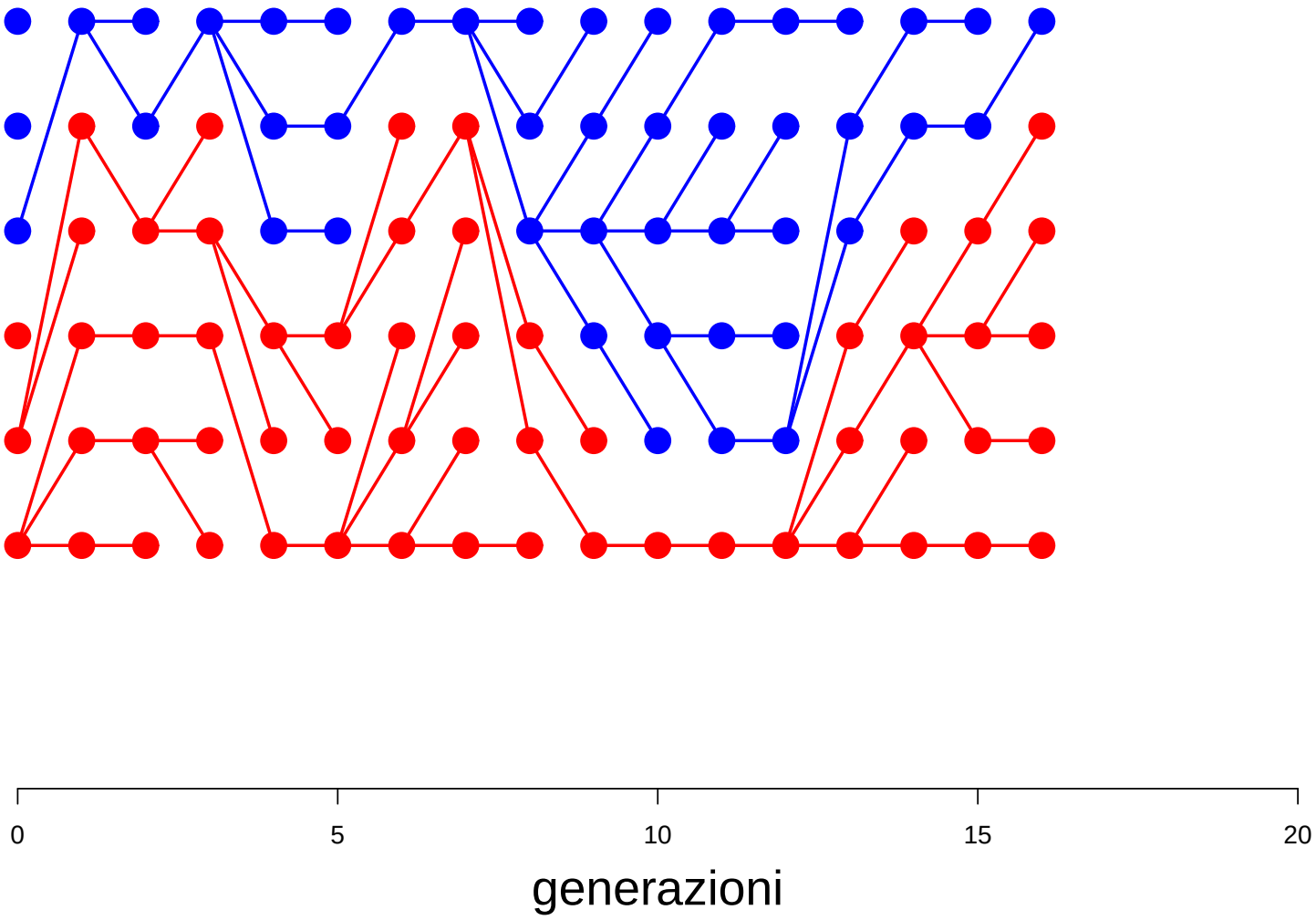
Una simulazione genetica



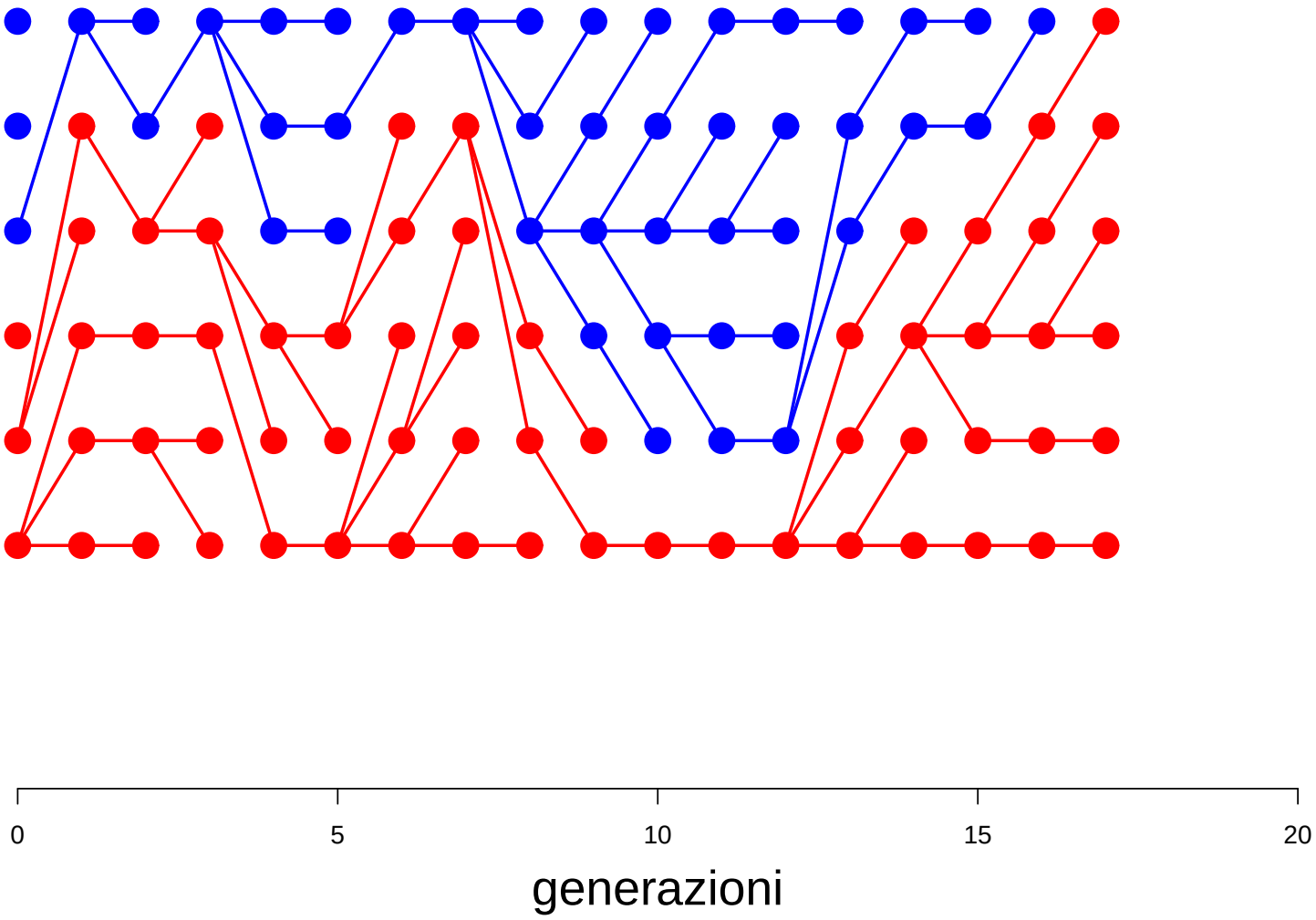
Una simulazione genetica



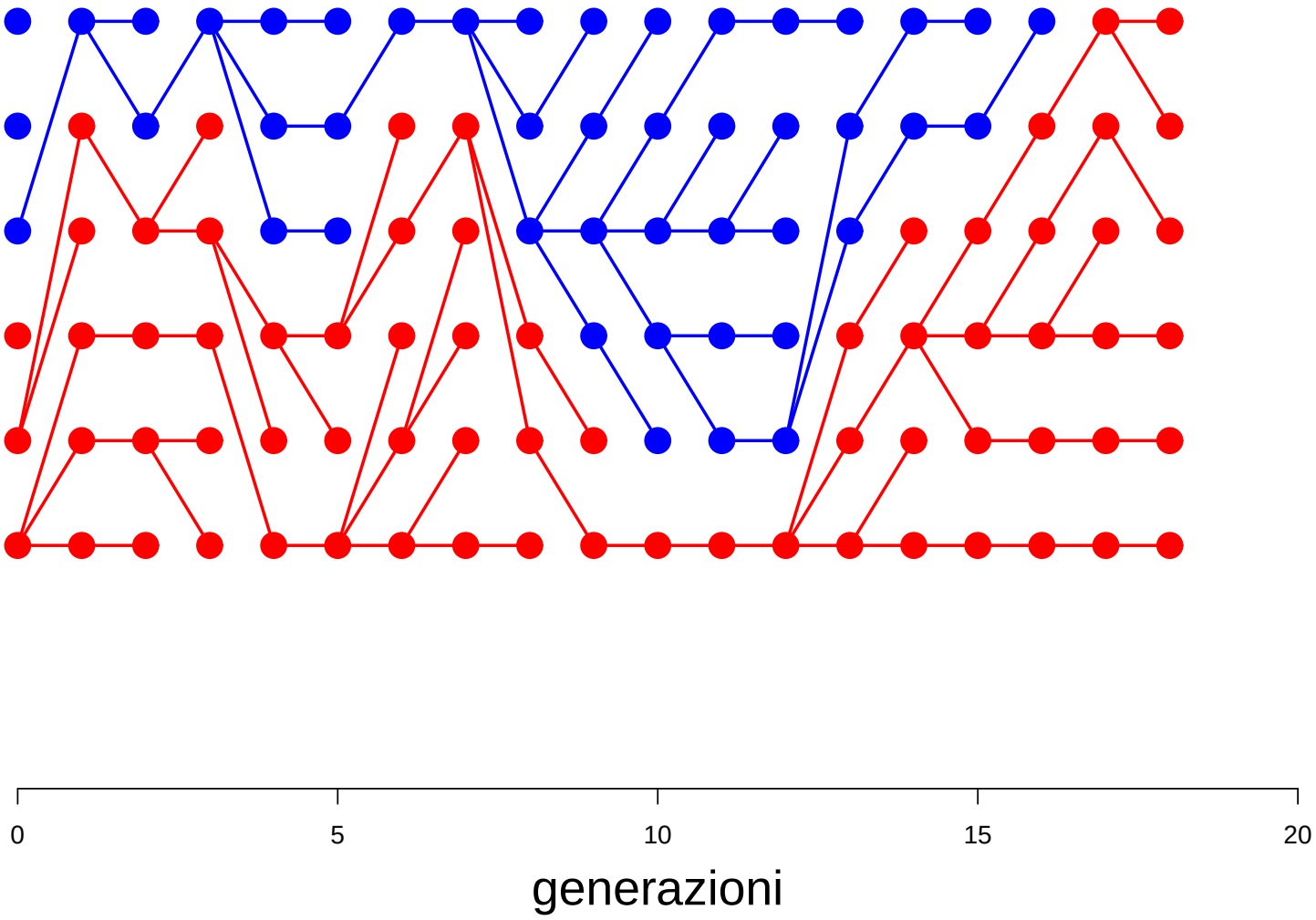
Una simulazione genetica



Una simulazione genetica



Una simulazione genetica



Una simulazione genetica

Tutti gli attuali individui, nell'ultima colonna, sono figli dei soli individui rossi nella prima colonna.

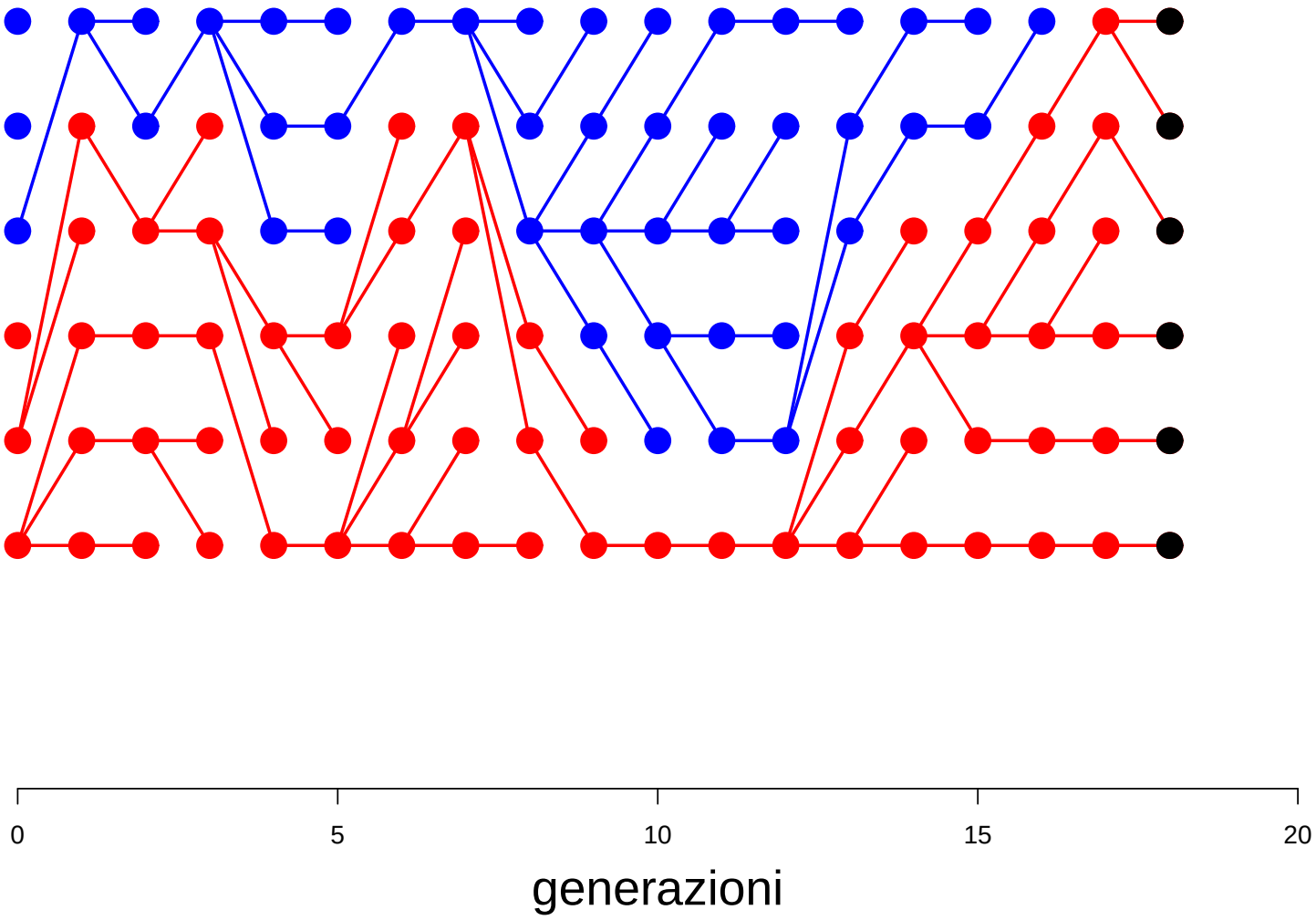
In termini di probabilistici, il modello descritto è una catena di Markov a 7 stati (il numero di "rossi" n che può variare da 0 a 6), con due stati assorbenti, $n = 0$ e $n = 6$, da cui il sistema non può uscire.

A parte i valori delle probabilità di transizione, questo modello è concettualmente identico alla passeggiata aleatoria con stati assorbenti, di cui è un esempio il ben noto problema dell'ubriaco, della casa e il fosso.

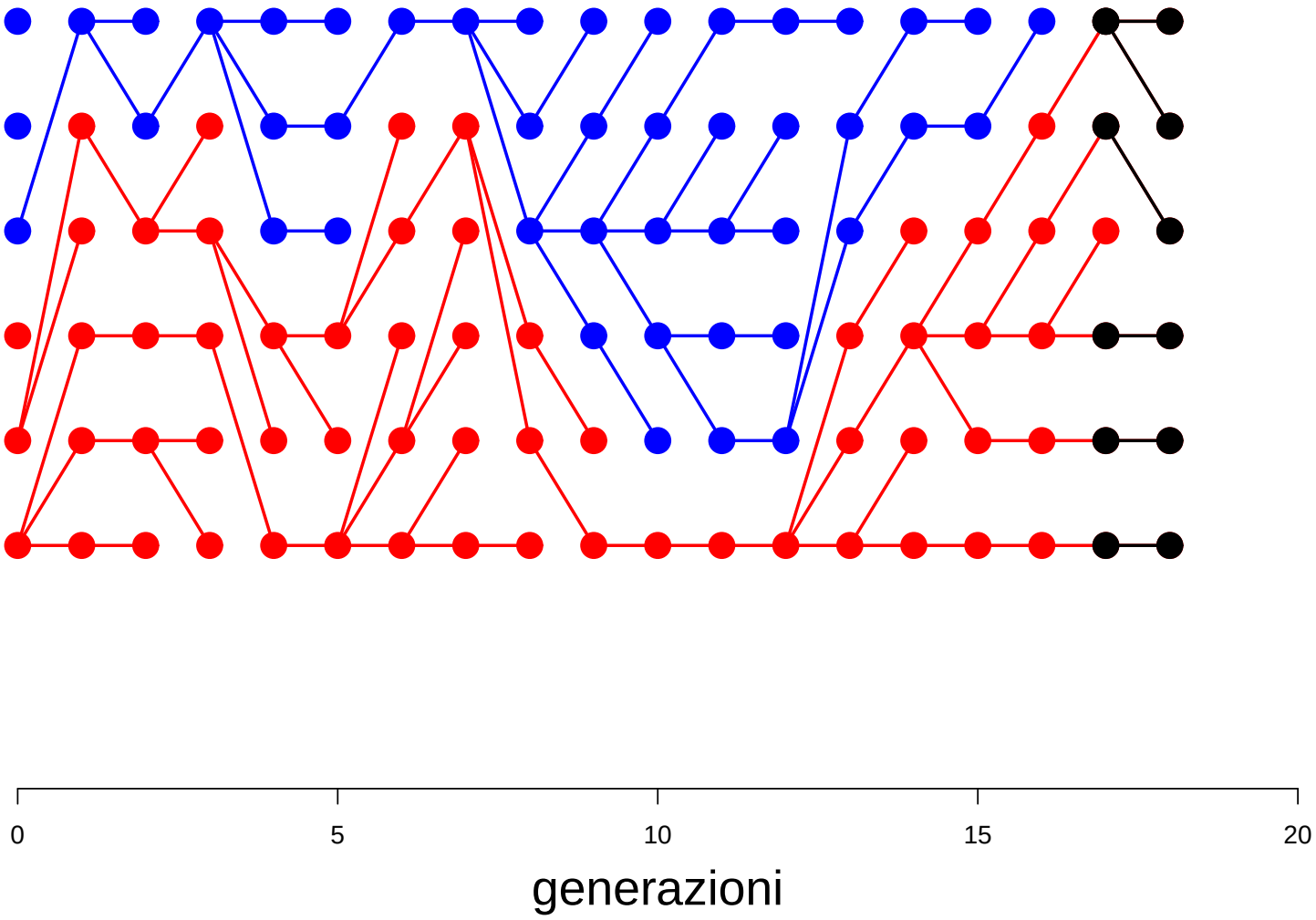
Una simulazione genetica

Si può fare un'altra importante osservazione su questo modello, ricostruendo la linea genealogica dell'ultima generazione, cioè osservando, a partire dalla fine, chi sono gli antenati degli individui dell'ultima generazione.

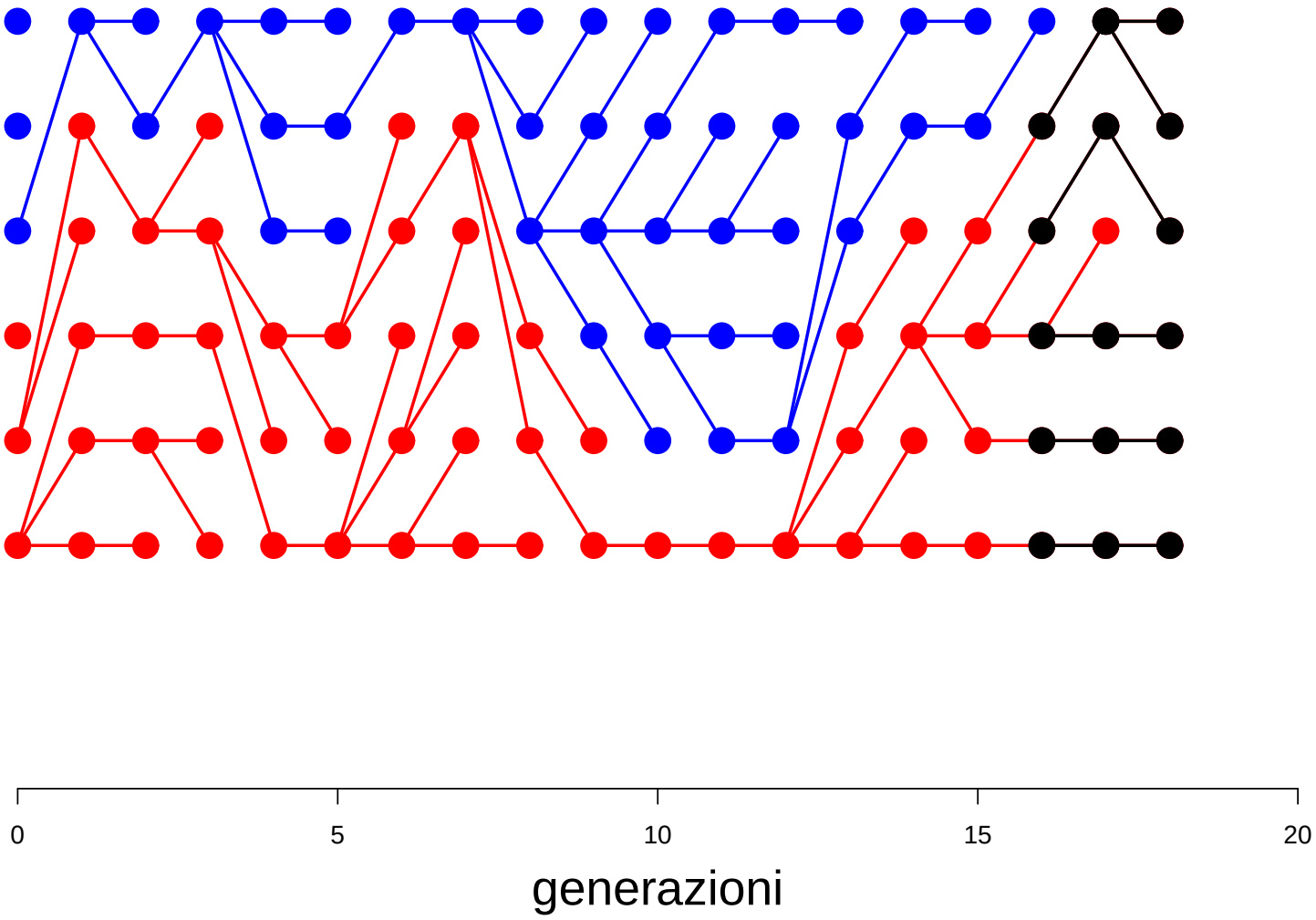
Una simulazione genetica



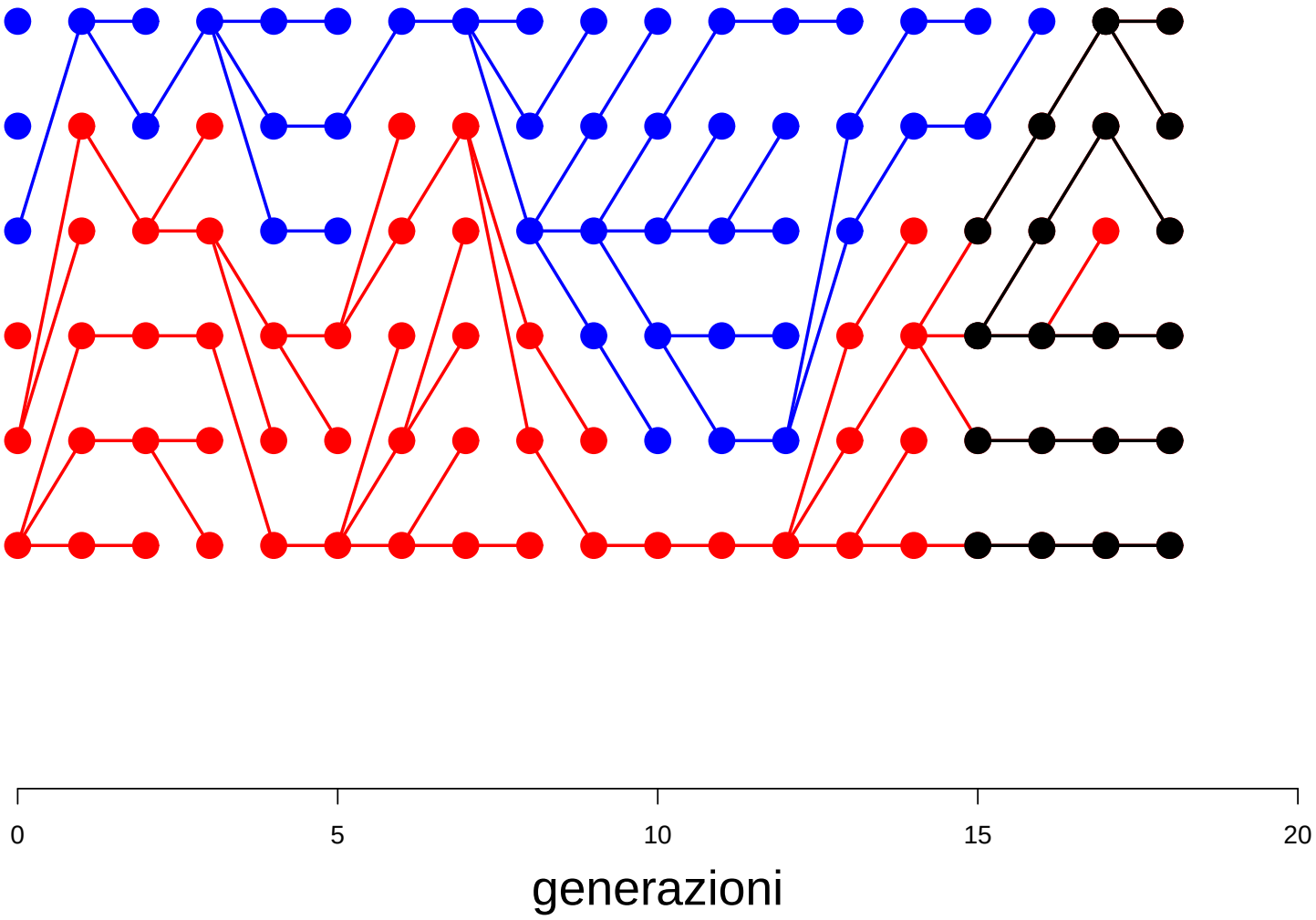
Una simulazione genetica



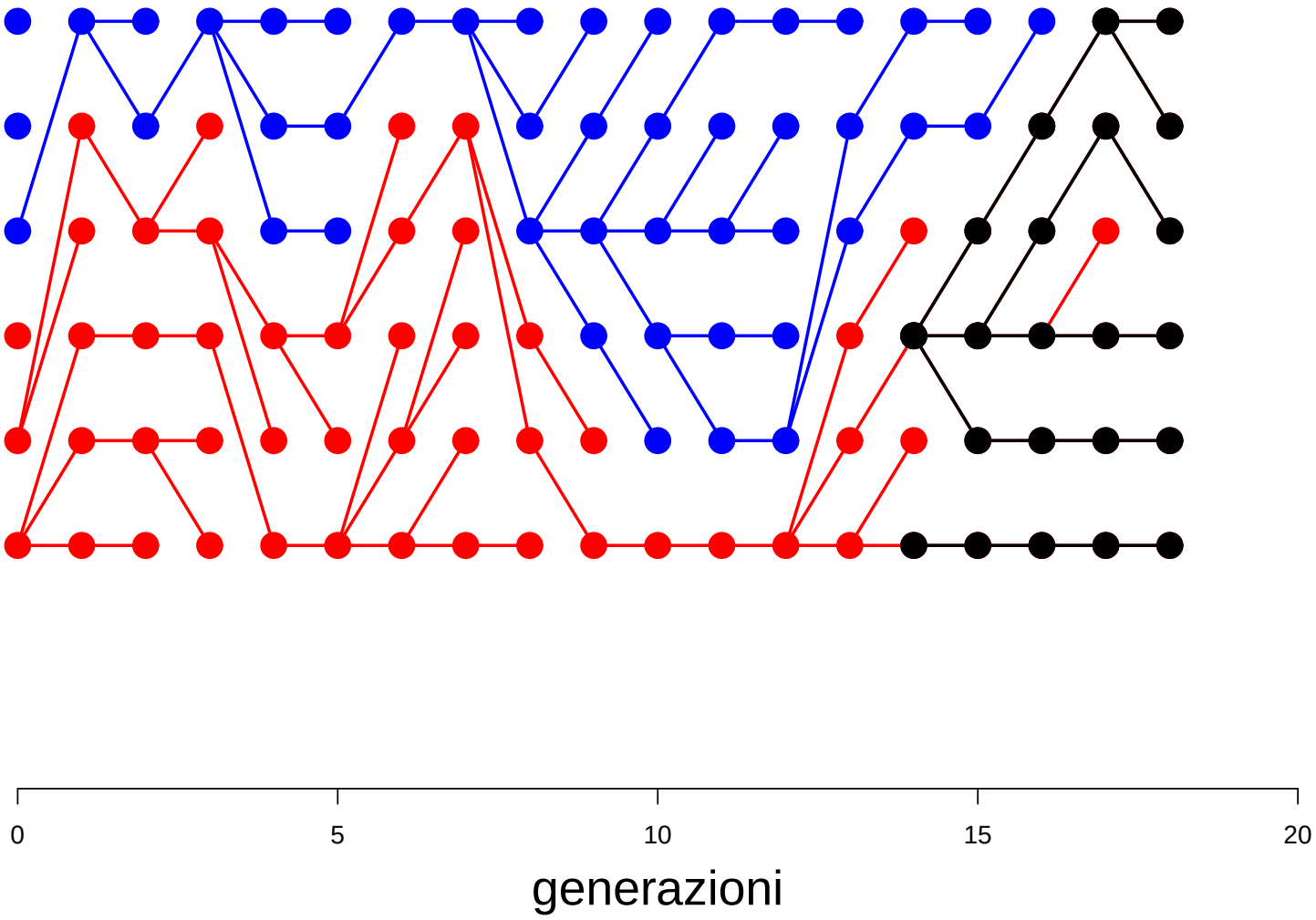
Una simulazione genetica



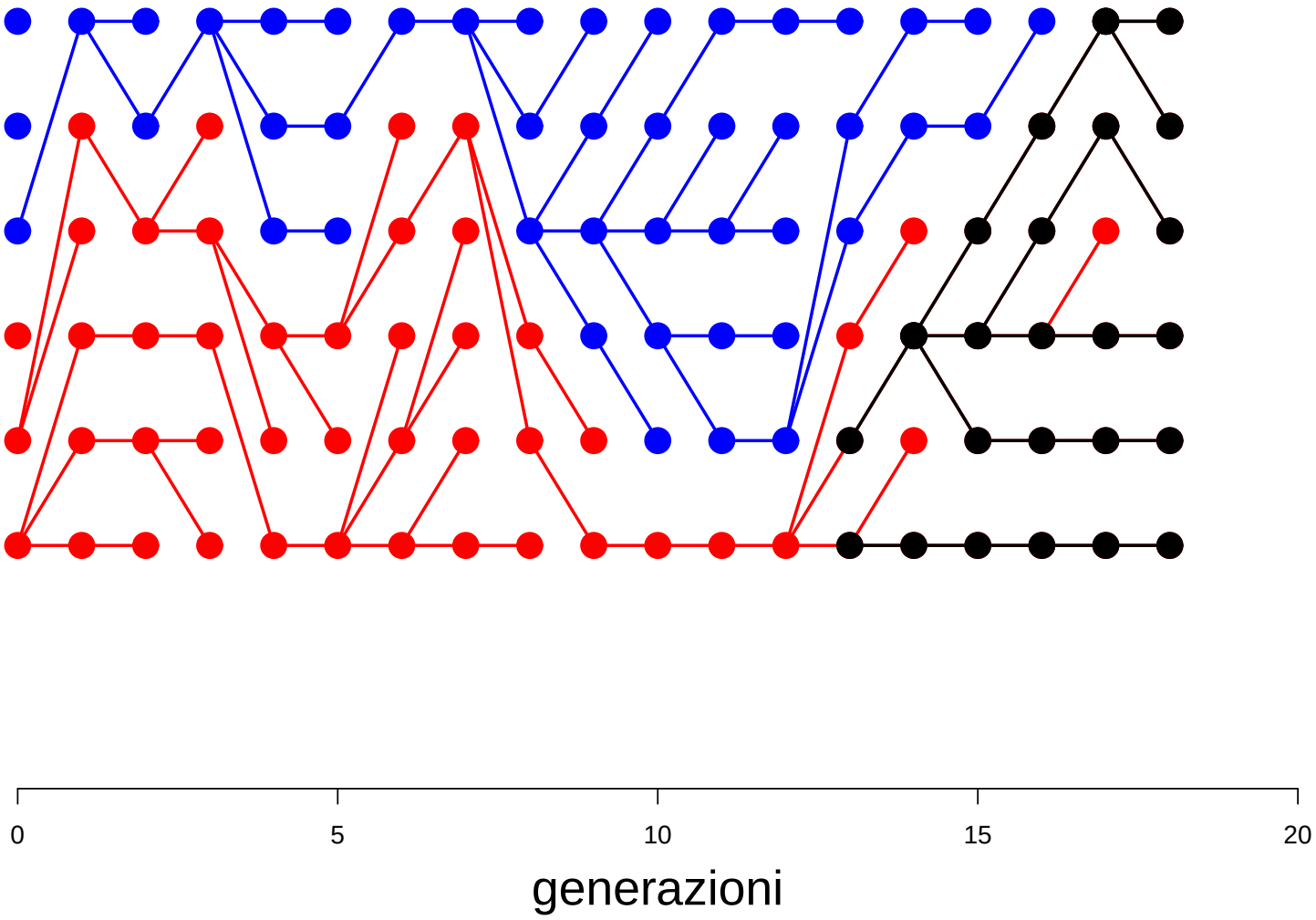
Una simulazione genetica



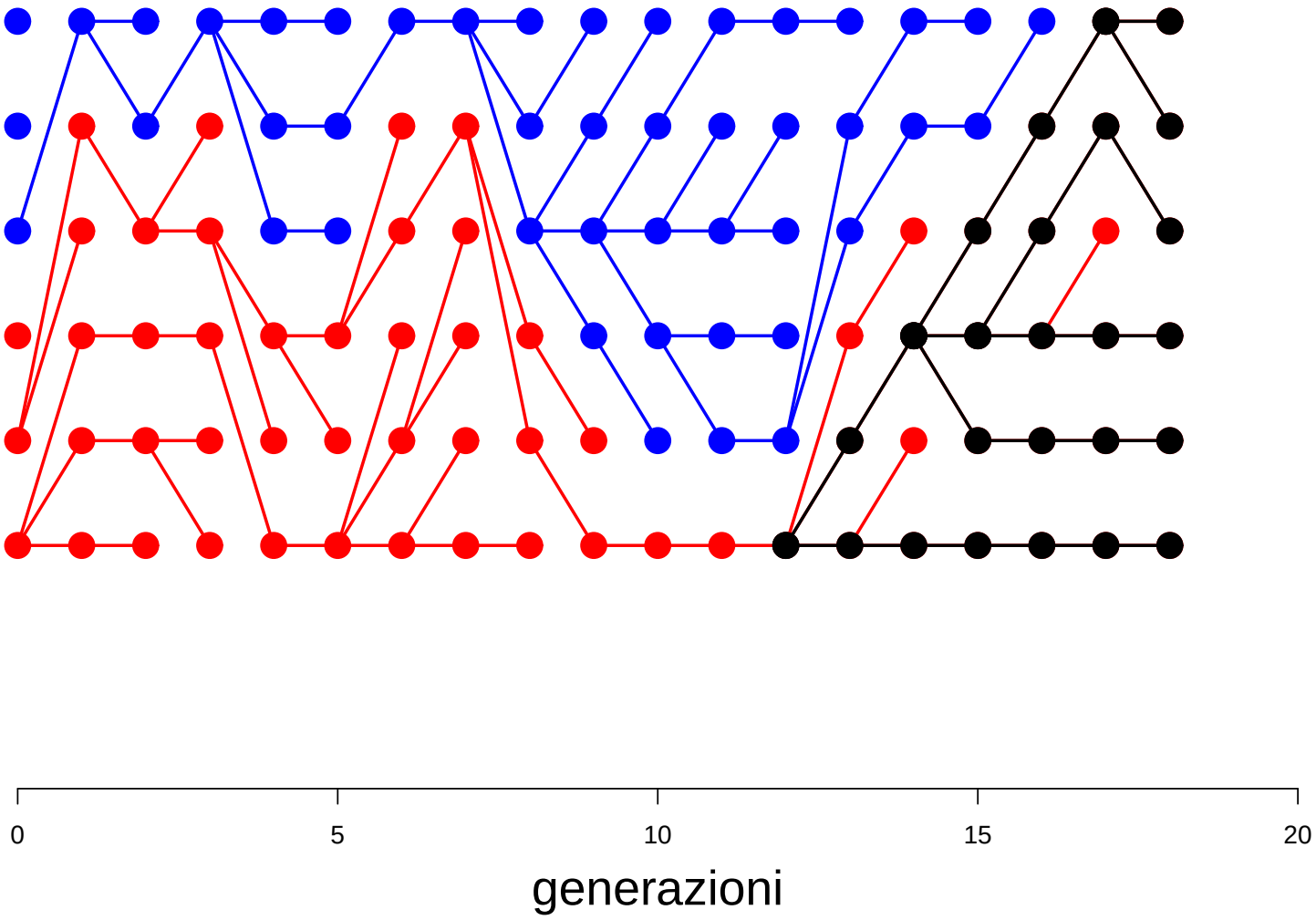
Una simulazione genetica



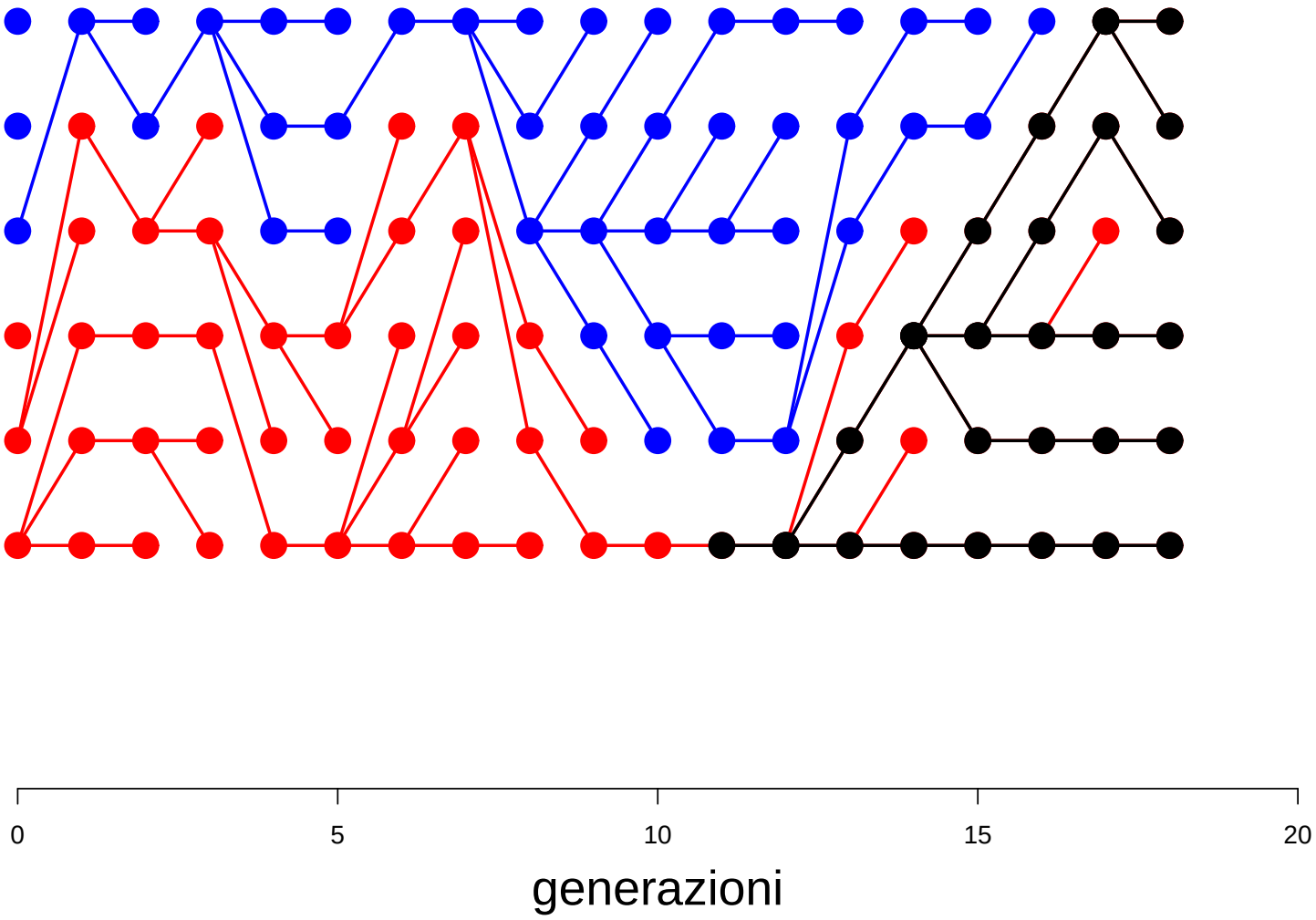
Una simulazione genetica



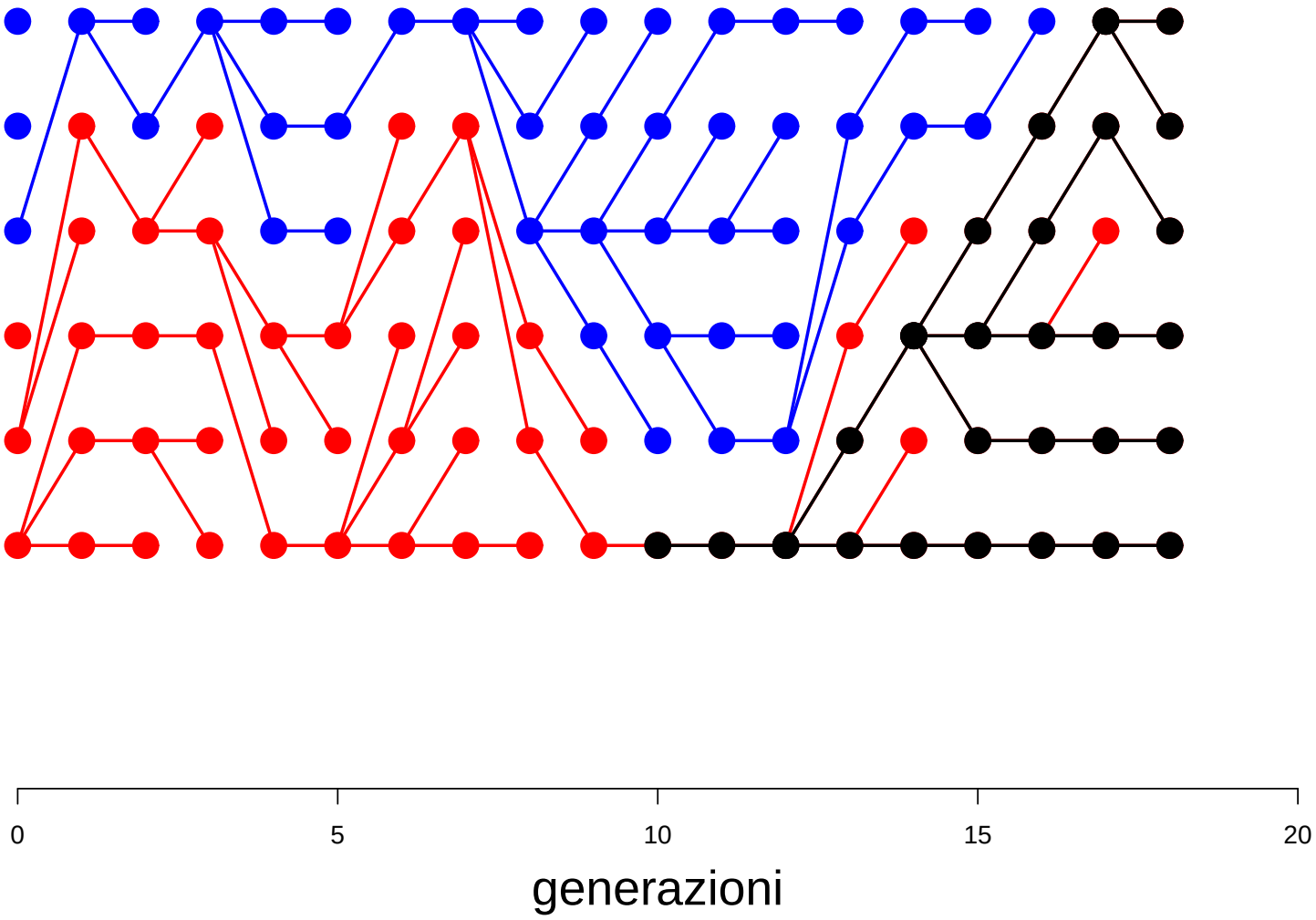
Una simulazione genetica



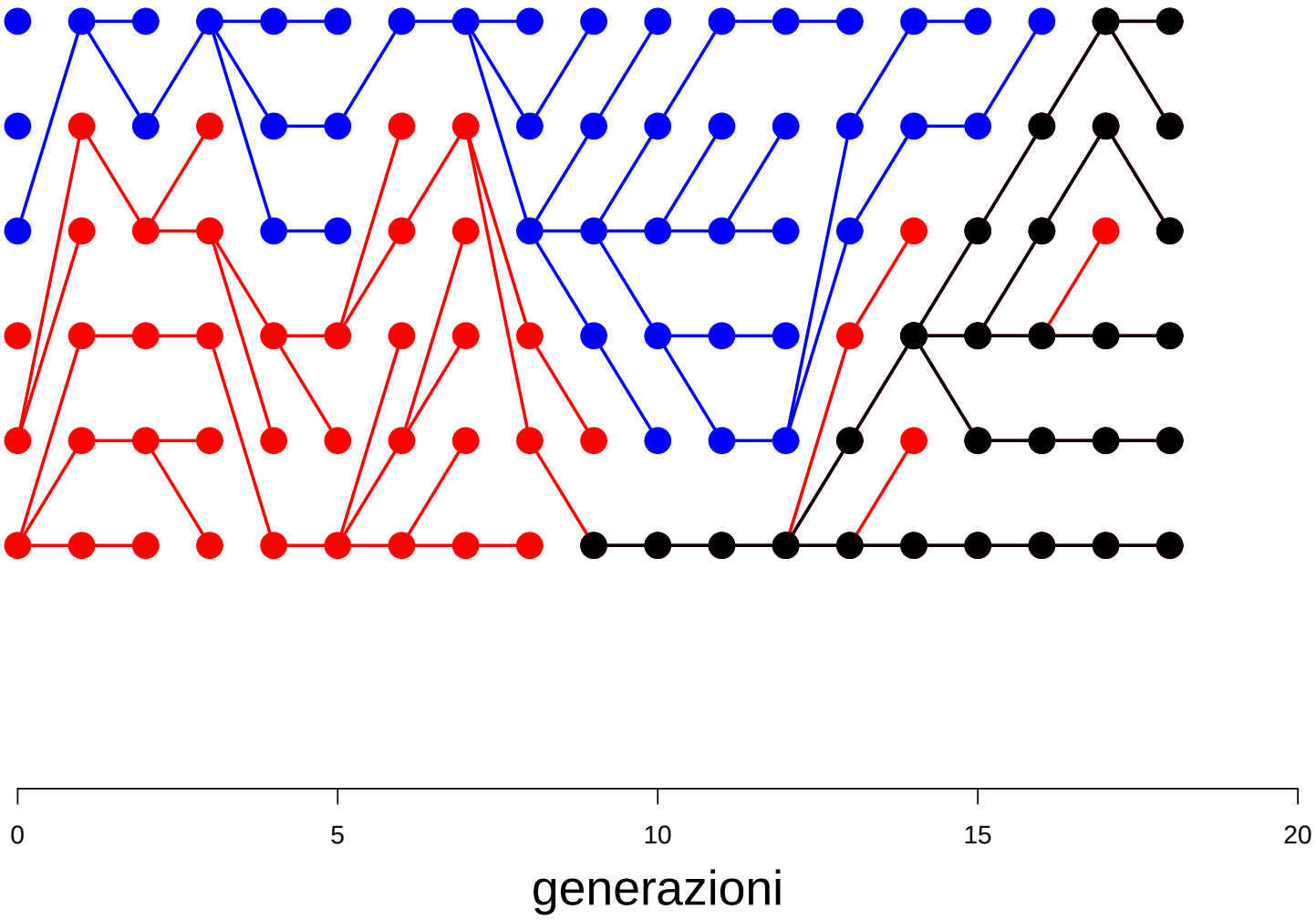
Una simulazione genetica



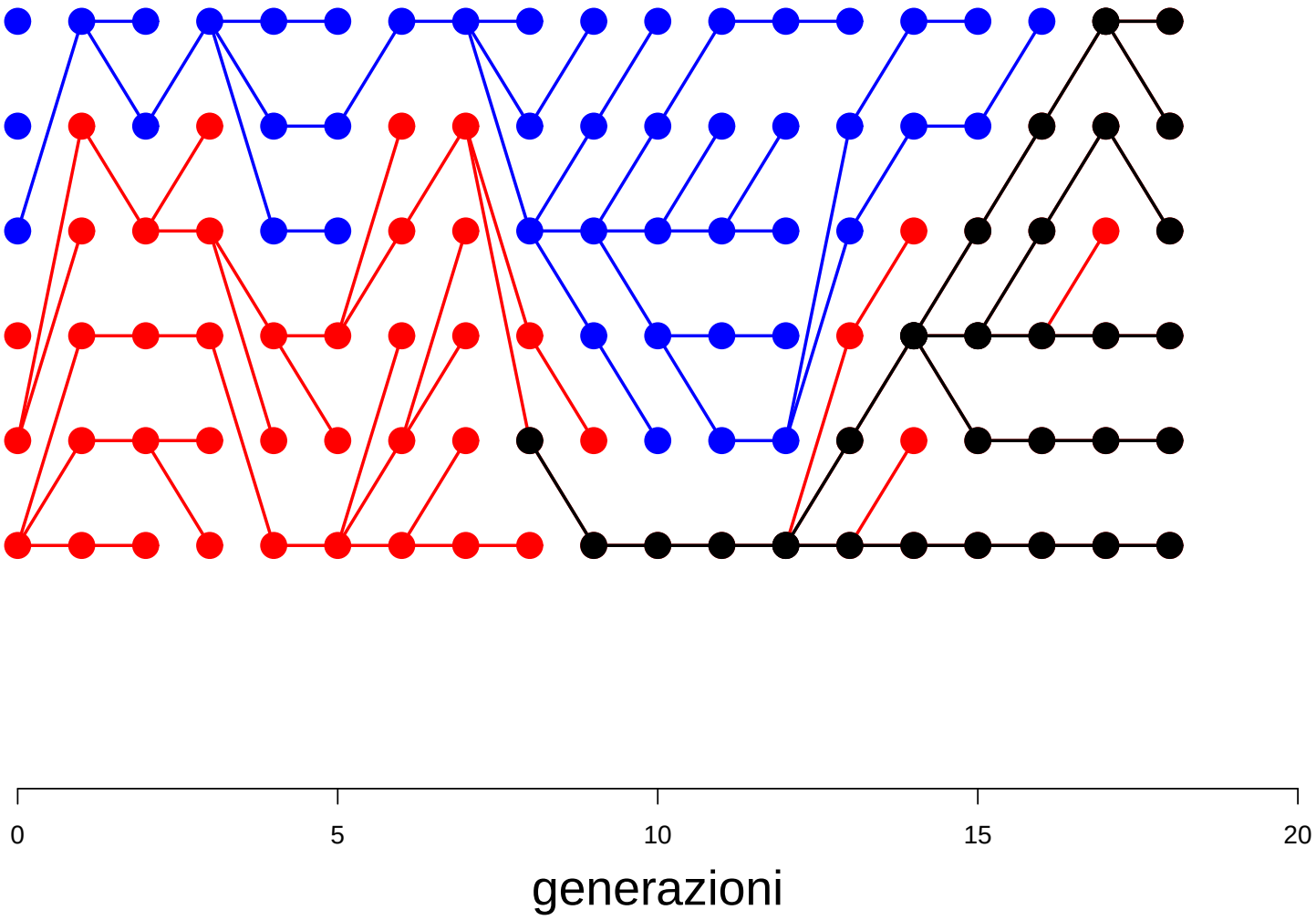
Una simulazione genetica



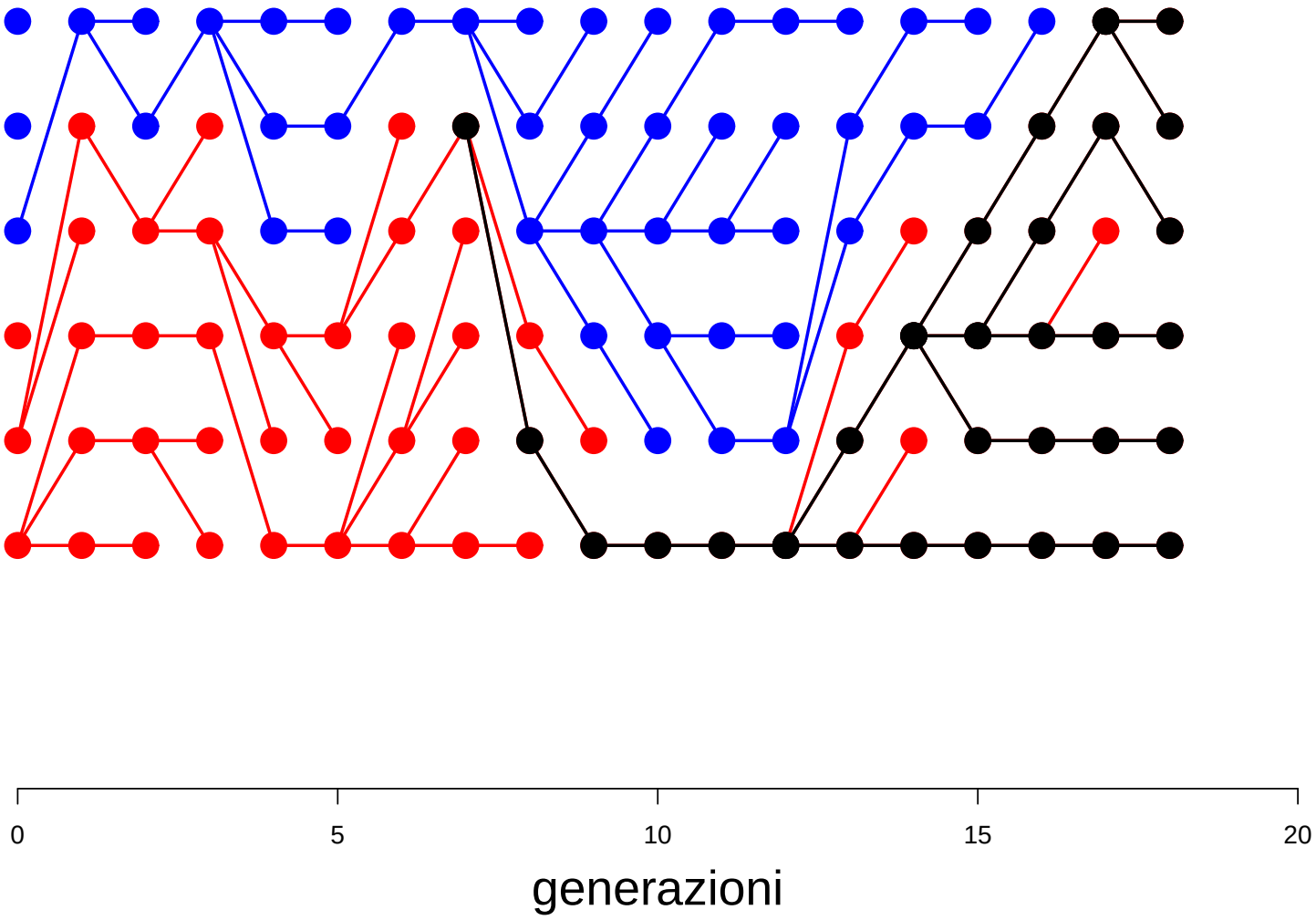
Una simulazione genetica



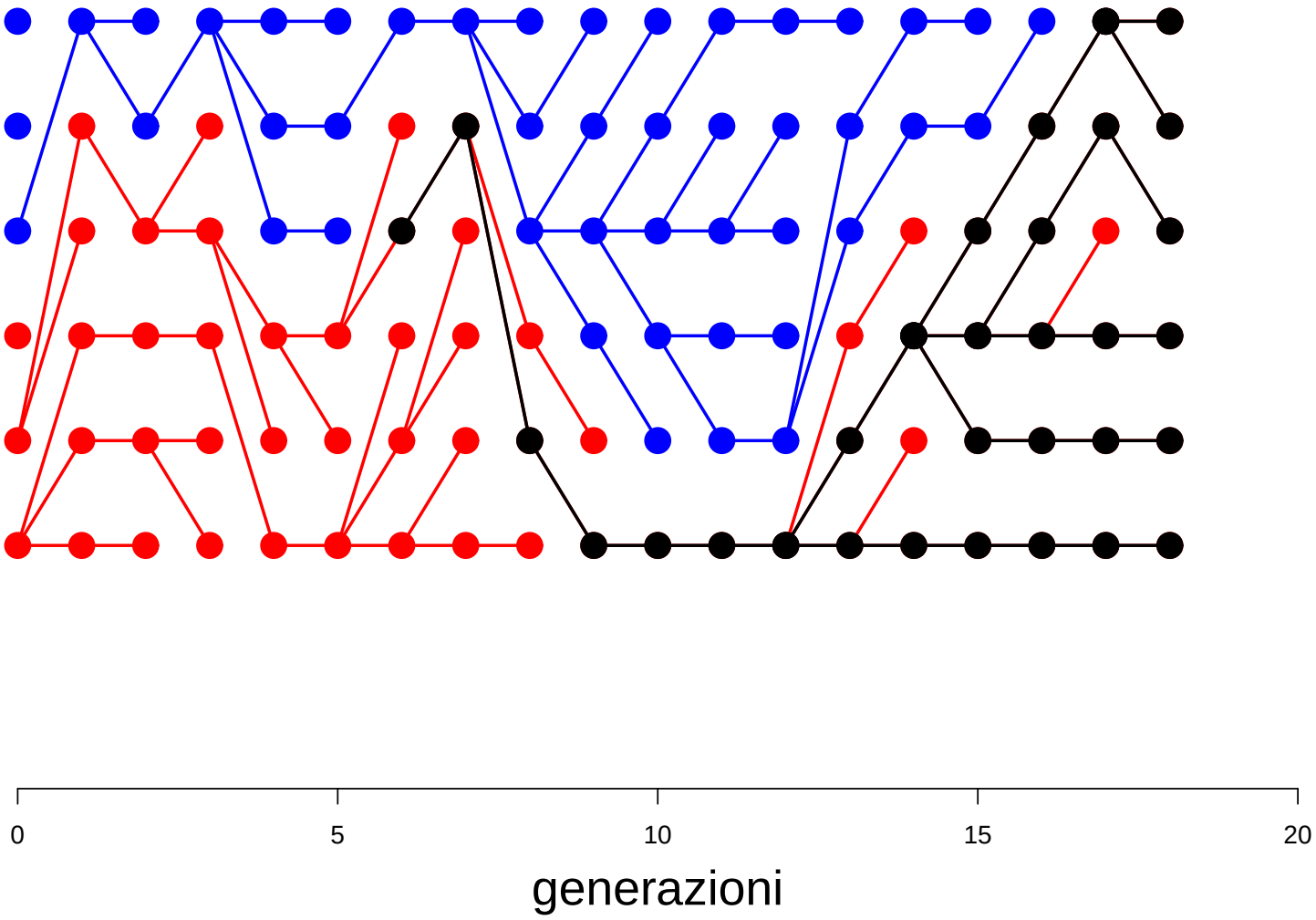
Una simulazione genetica



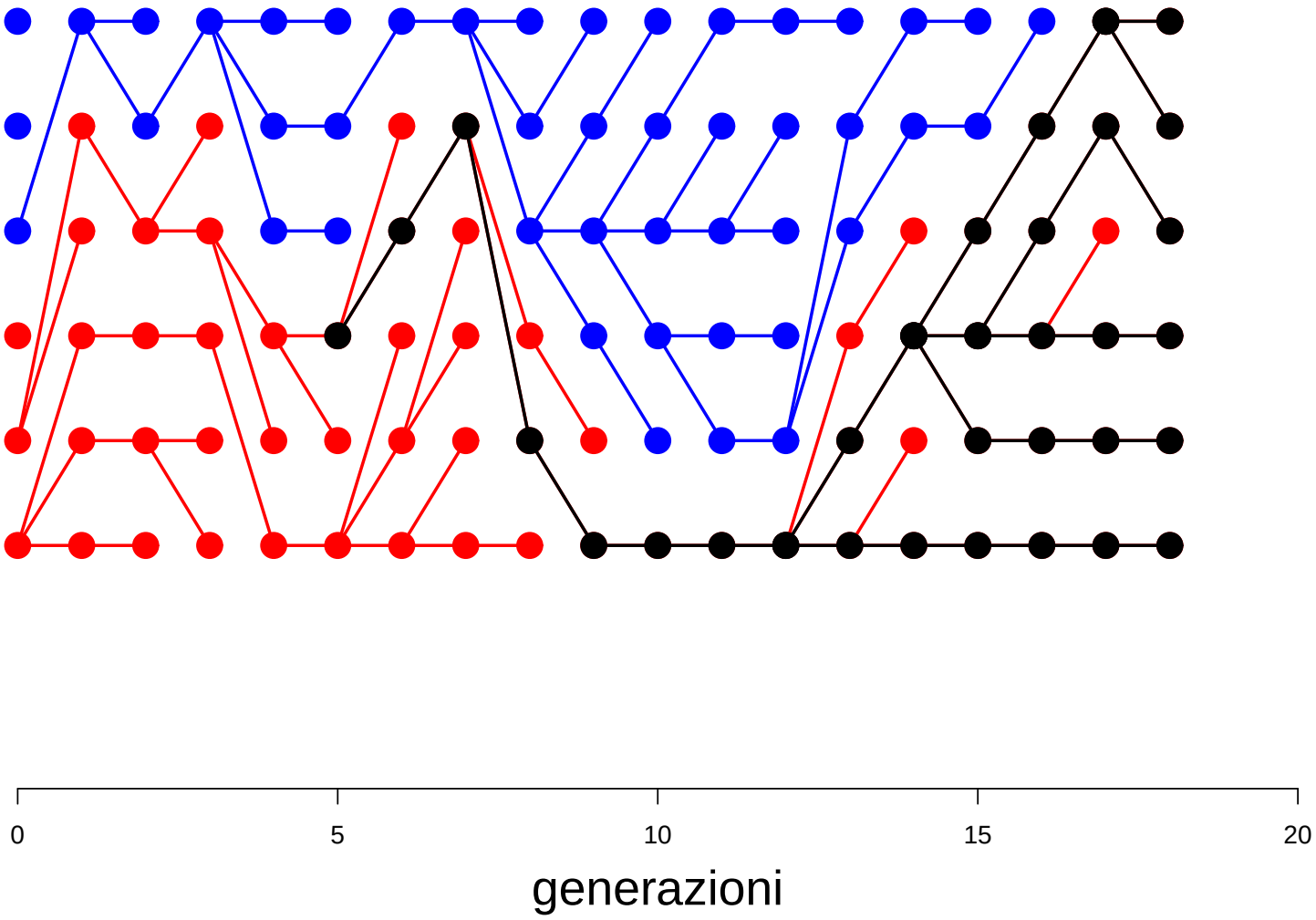
Una simulazione genetica



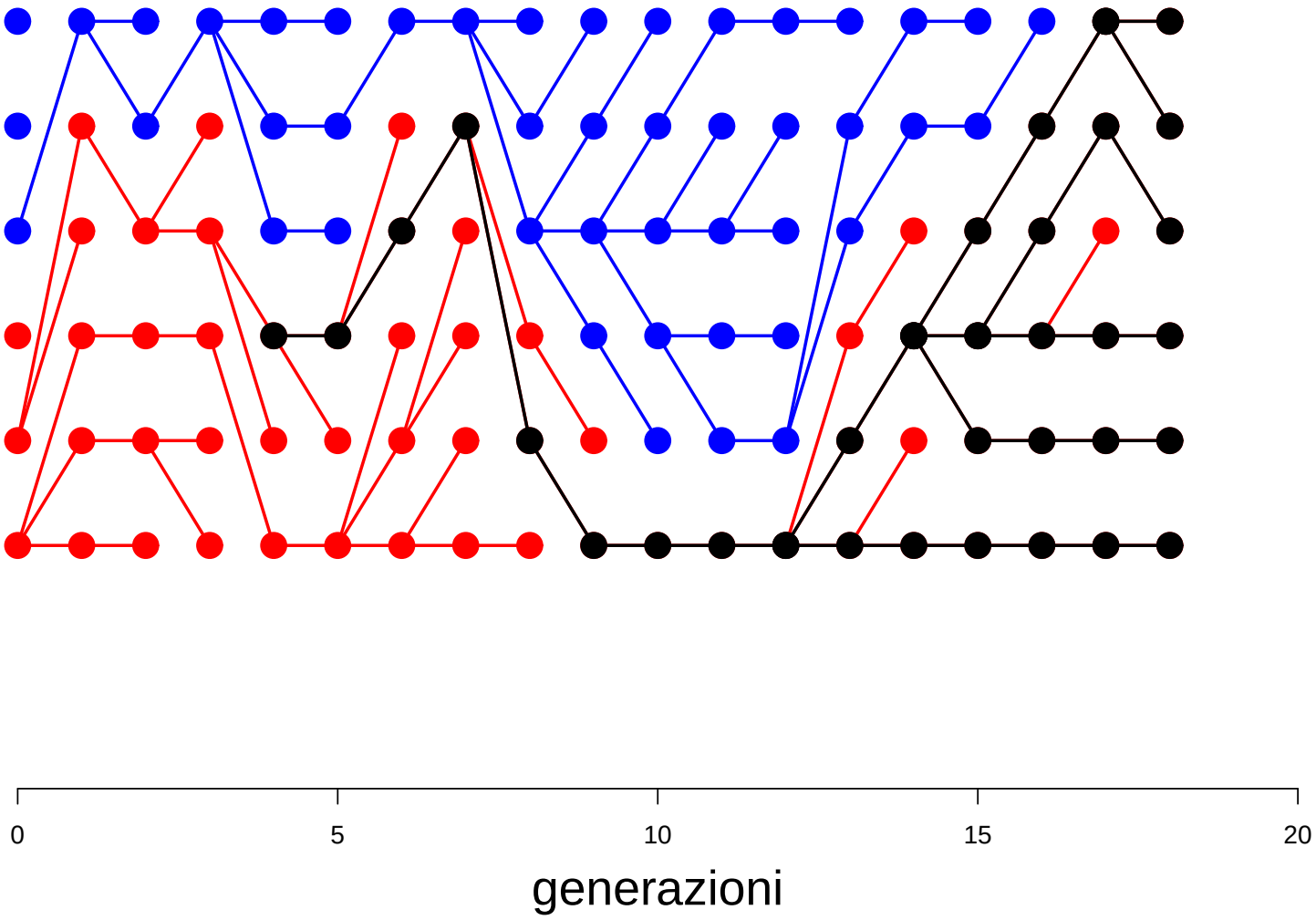
Una simulazione genetica



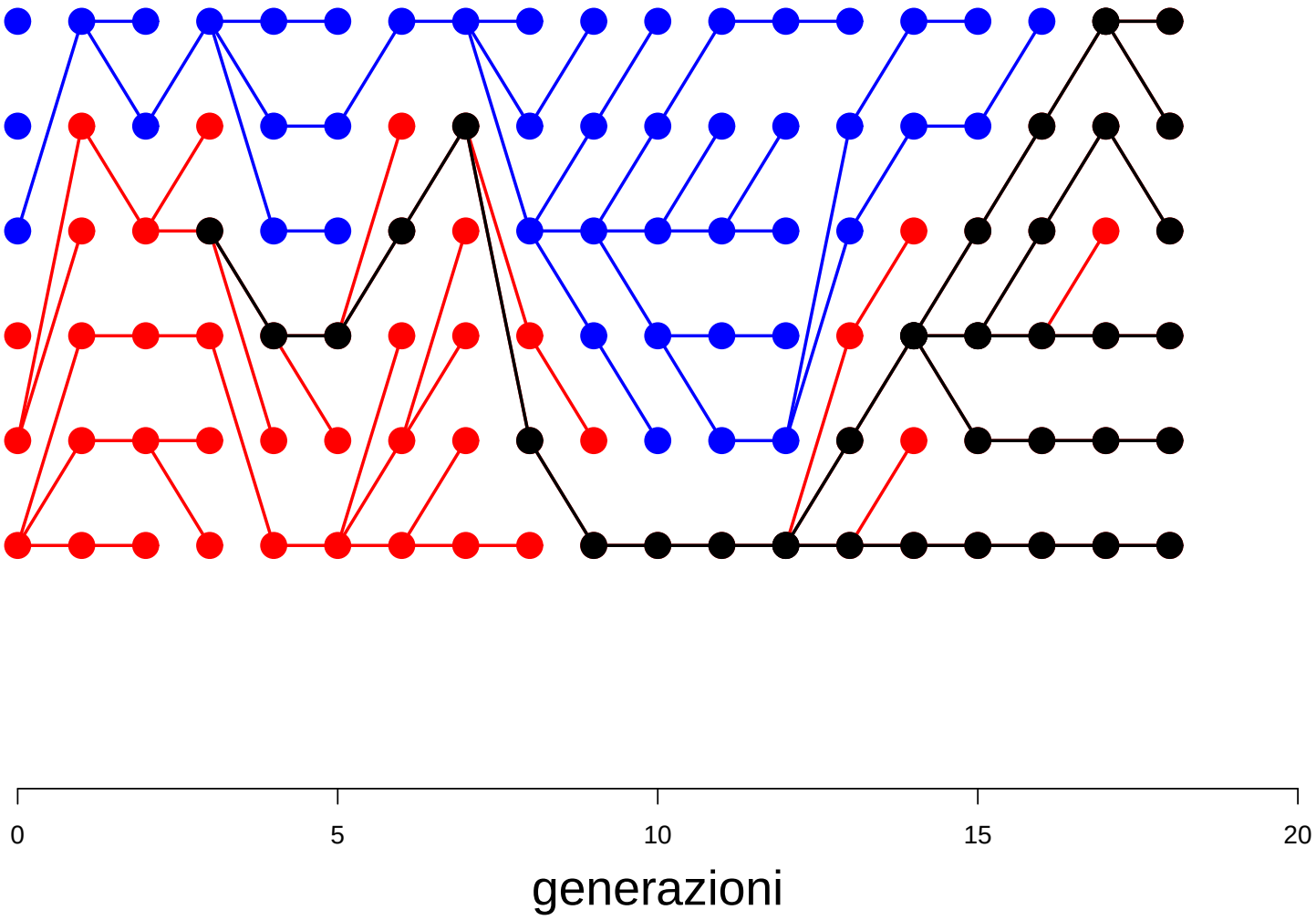
Una simulazione genetica



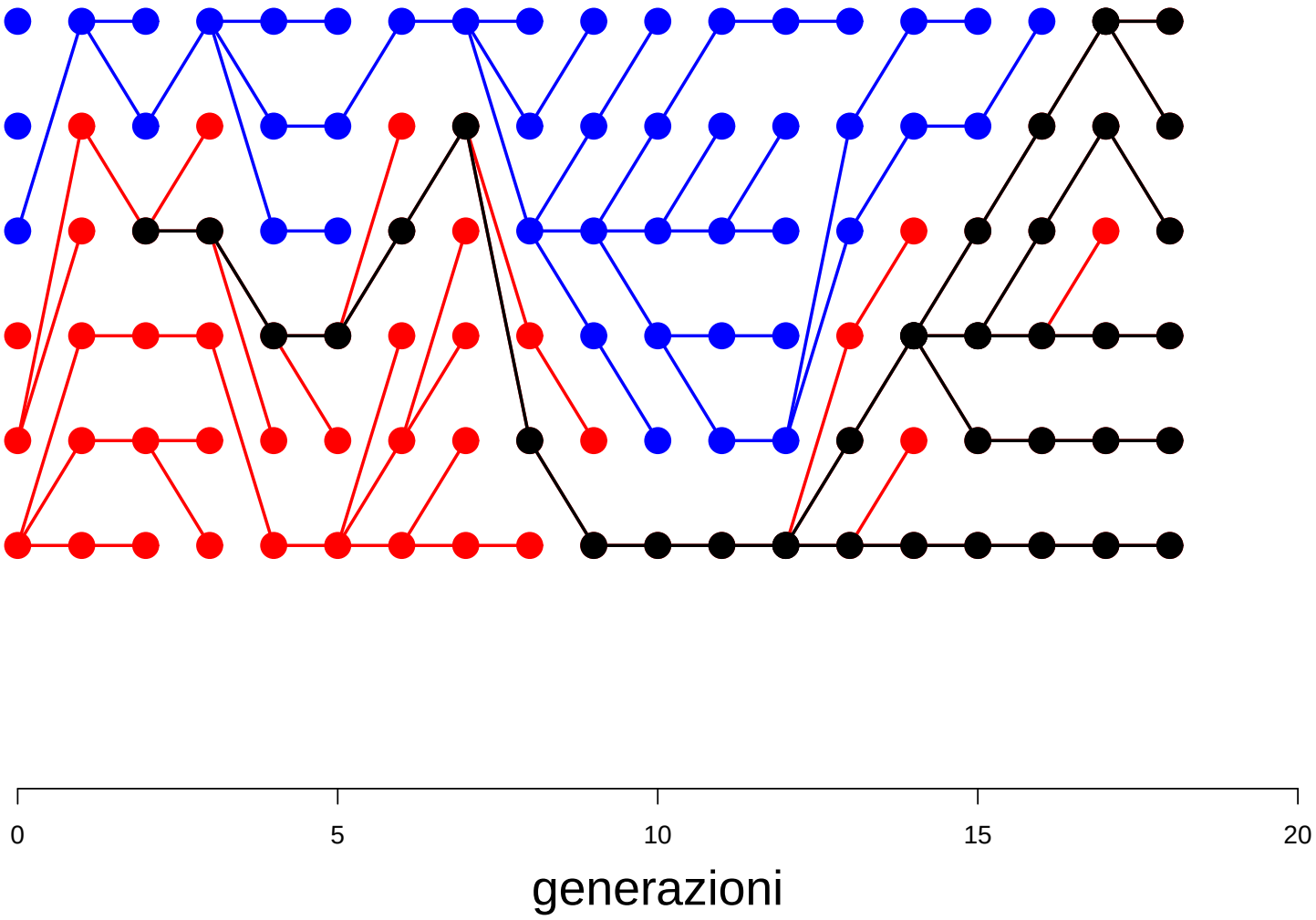
Una simulazione genetica



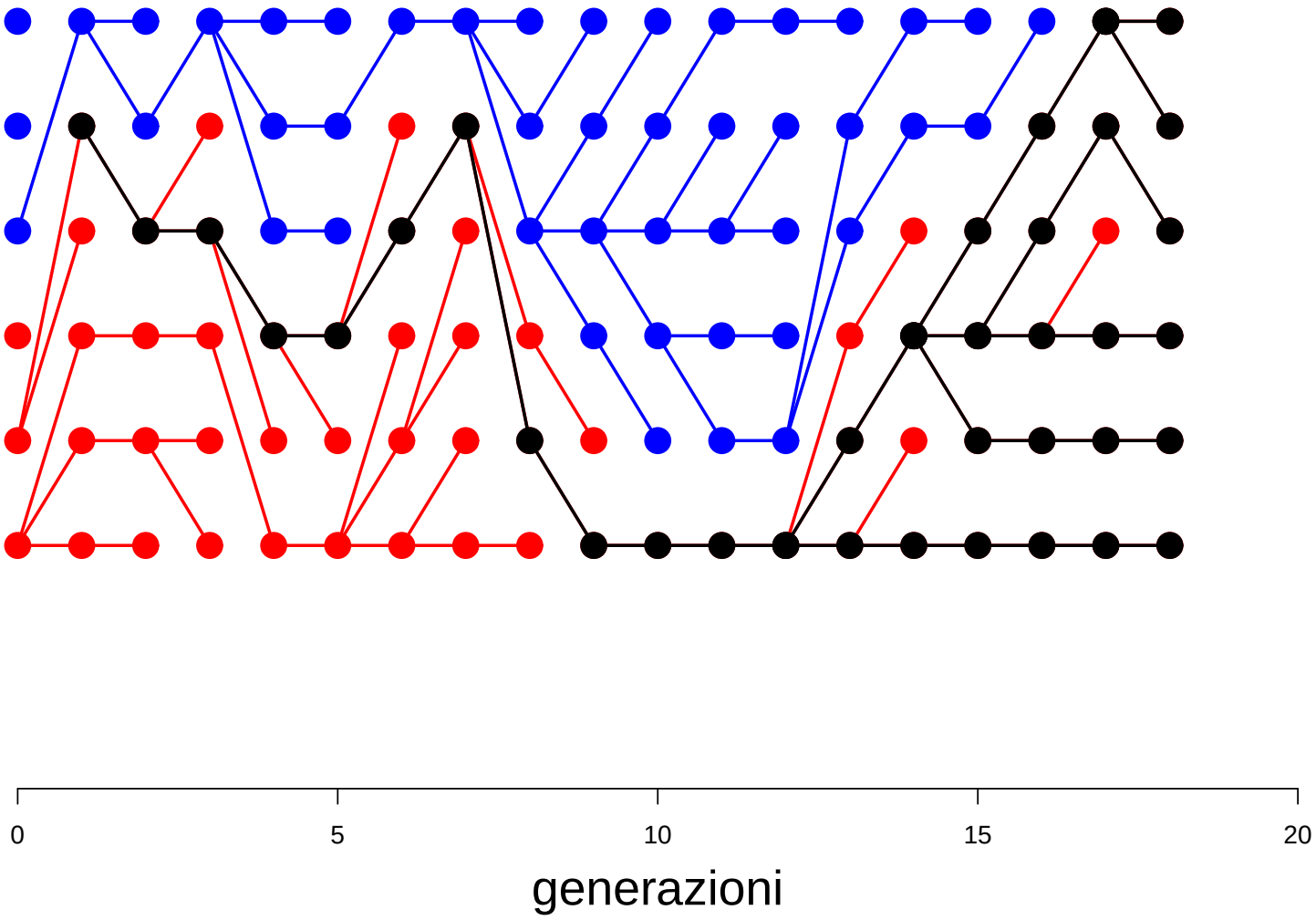
Una simulazione genetica



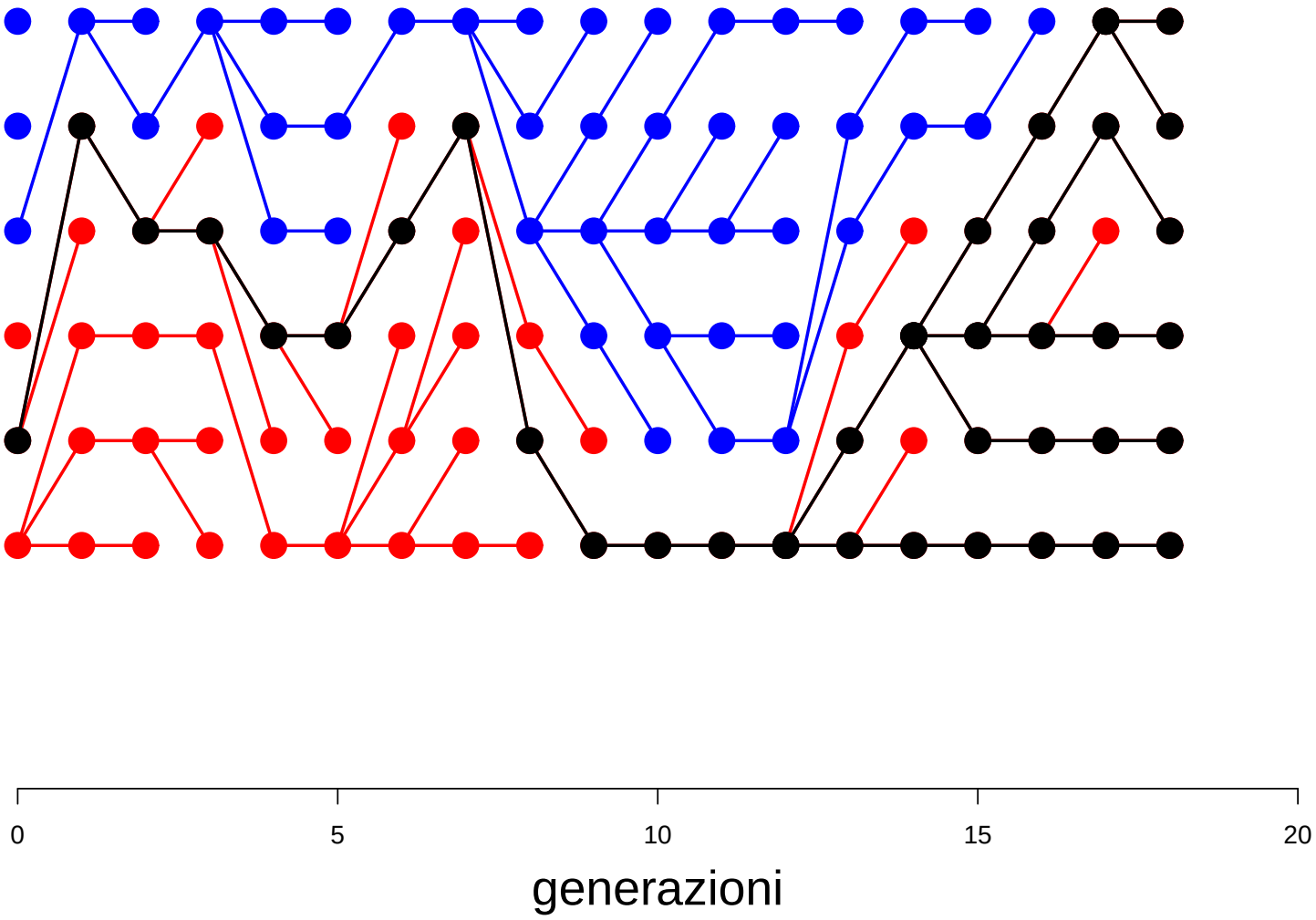
Una simulazione genetica



Una simulazione genetica



Una simulazione genetica



Adamo ed Eva

Questa simulazione ci indica che la limitatezza della popolazione umana nella storia ha delle sorprendenti implicazioni.

- È esistita una donna, **l'Eva mitrocondriale**, che è la più recente donna dalla quale tutti discendiamo
- È esistito un uomo, **l'Adamo Y-cromosomiale**, che è il più recente uomo antenato di tutti i maschi viventi

Adamo ed Eva

Non erano marito e moglie...

Adamo ed Eva

Non erano marito e moglie...

Tenendo in conto le mutazioni, osservano i DNA attuali, si può stimare l'età in cui sono vissuti:

- 100 000-170 000 anni fa: Eva mitrocondriale, ed è vissuta in Africa
- 200 000-300 000 anni fa: Adamo Y-cromosomiale, ed è vissuto in Africa

Adamo ed Eva

Non erano marito e moglie...

Tenendo in conto le mutazioni, osservano i DNA attuali, si può stimare l'età in cui sono vissuti:

- 100 000-170 000 anni fa: Eva mitrocondriale, ed è vissuta in Africa
- 200 000-300 000 anni fa: Adamo Y-cromosomiale, ed è vissuto in Africa

La prima volta che ho raccontato questa storia, la stima per l'Adamo era di 60 000 anni fa. Perchè questa stima sia cambiata è una storia interessante, che si allaccia ad un'altra questione, a cui ora accennerò.

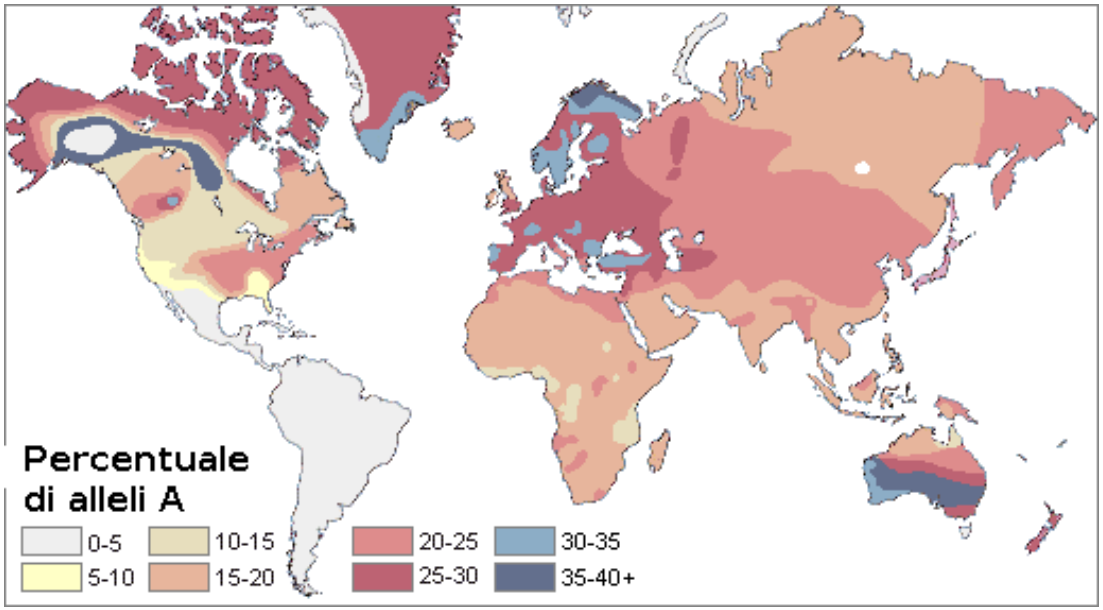
La conseguenze della deriva genetica

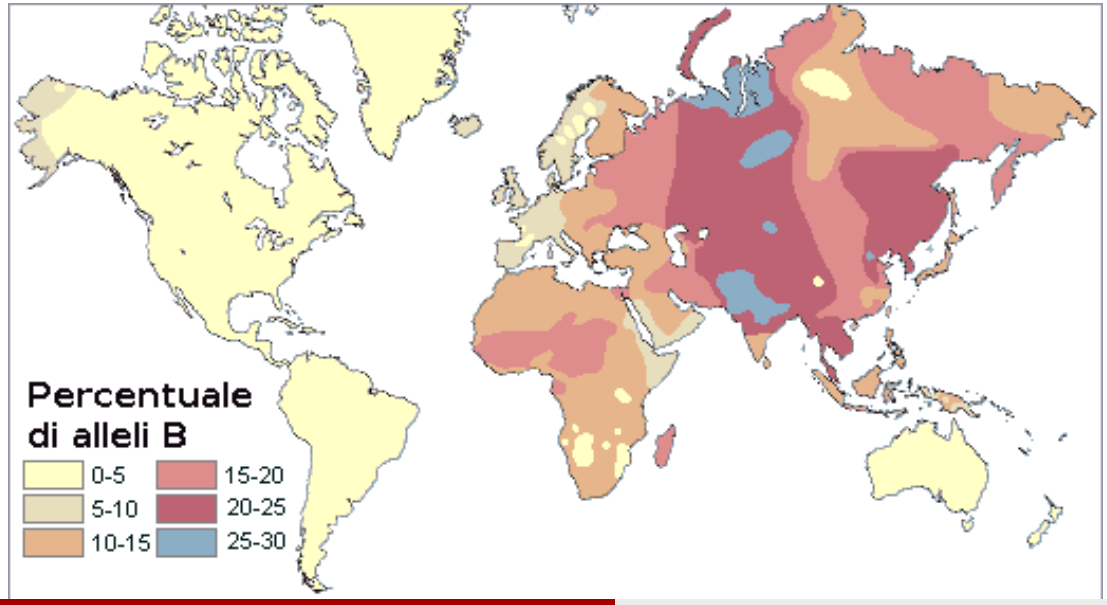
La deriva genetica precede, nel lungo periodo, l'estinzione di tutte le varianti alleliche tranne una, a meno di selezione e mutazione.

Popolazioni che si separano per qualche motivo, avranno, per colpa del caso, frequenze alleliche via via diverse con il passare del tempo. Nel lungo periodo, due popolazioni possono dare origine in questo modo a due specie diverse.

A partire da quanto sono diverse le frequenze di molti geni, si può stimare quanto tempo è passato, ricostruendo la storia delle popolazioni.

Le mappe degli alleli





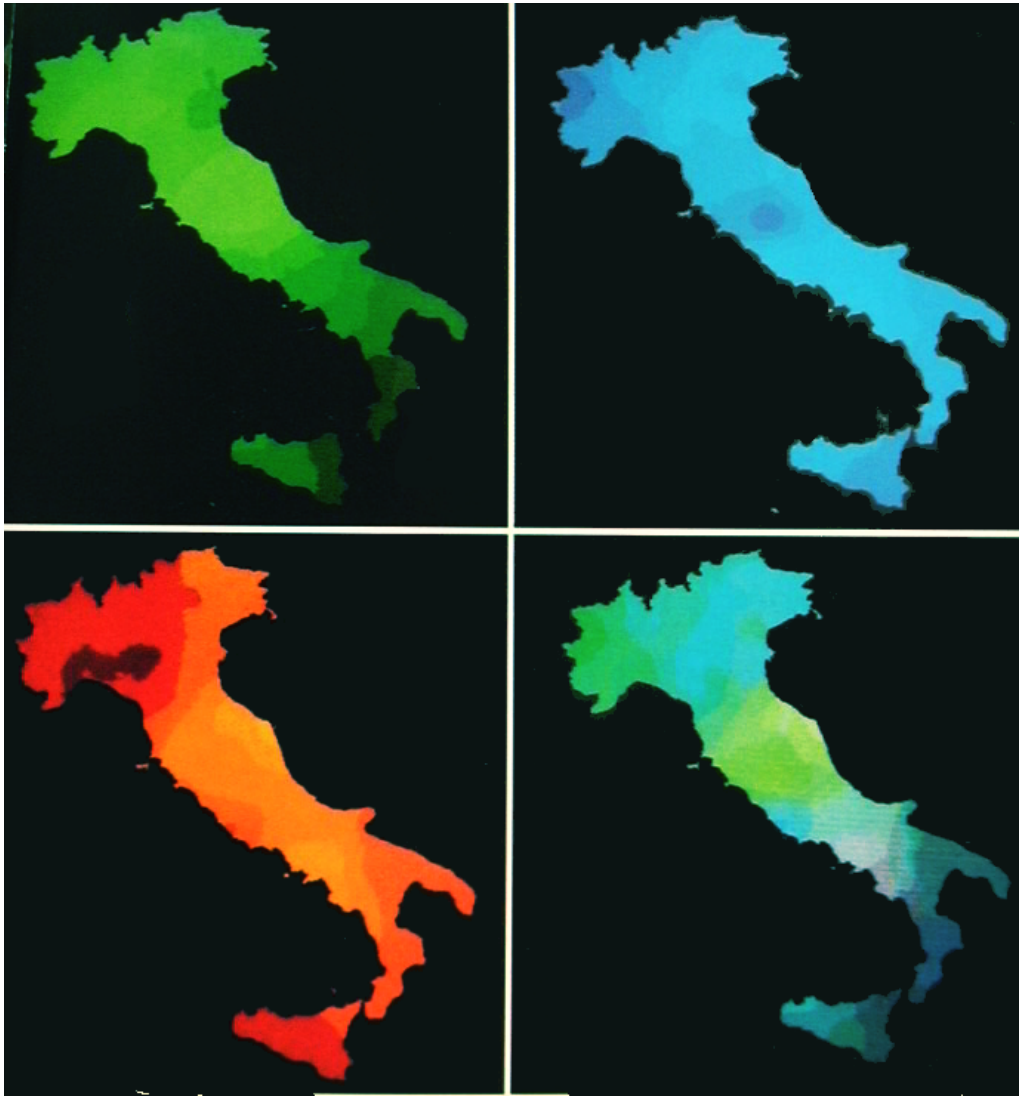
Le mappe degli alleli

Una migrazione modifica le frequenze alleliche.

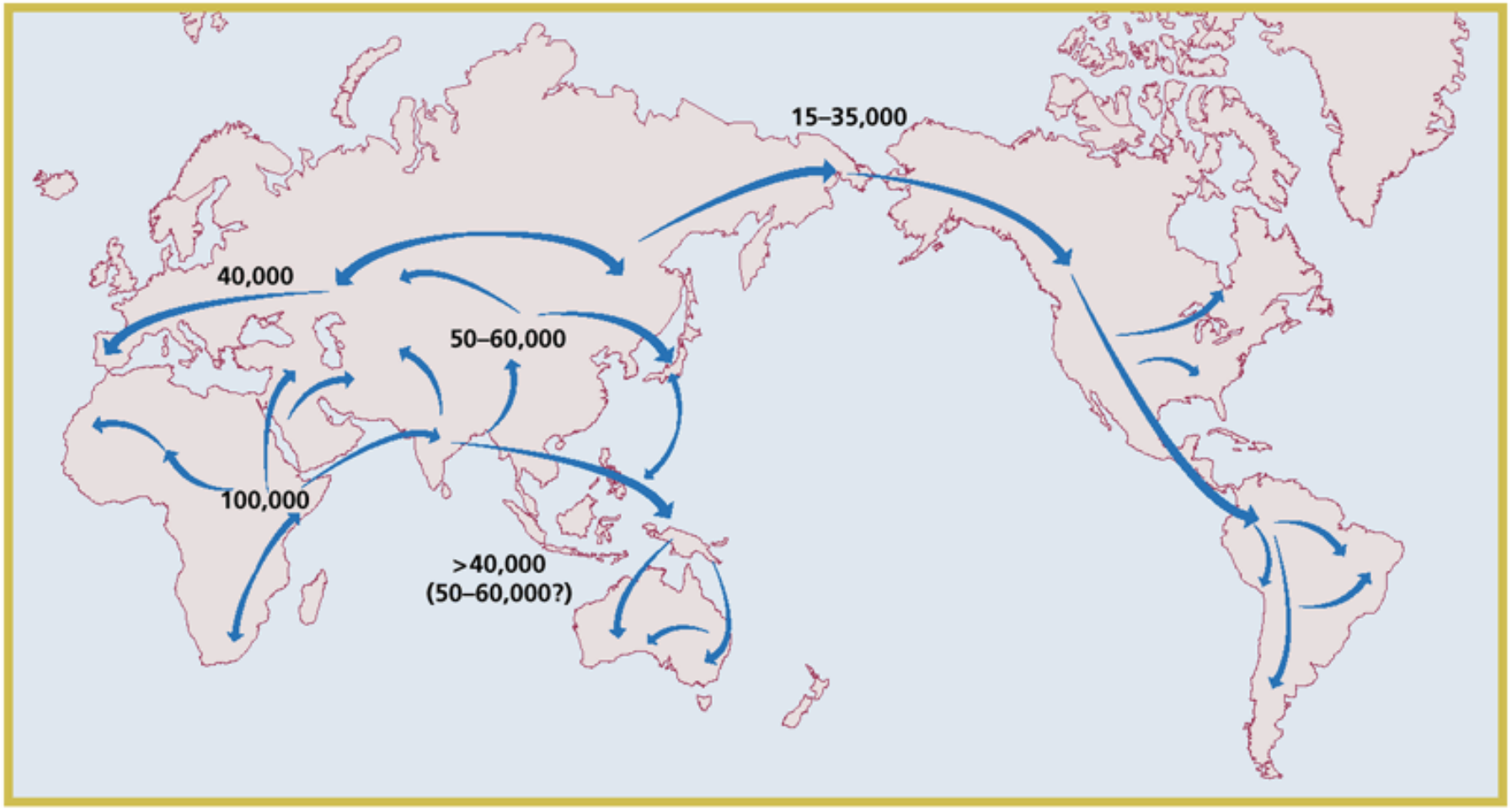
Usando le informazioni sulle frequenze degli alleli di molti geni differenti, ricostruire le migrazioni su un territorio.
Tecnichamente si ottengono (circa) dalle curve di livello delle componenti principali della varianza allelica complessiva

La storia dei popoli attraverso la genetica

L. Cavalli-Sforza **Geni, popoli e lingue** 1996



La storia dell'uomo attraverso la genetica



Altri spunti

Una storia interessante per discutere di test statistici.

- c'è un procedimento statistico per valutare l'adattamento dei dati a una teoria sottostante
- le leggi di Mendel fanno previsioni sulle frequenze di certe caratteristiche nelle piante di pisello
- ogni studente di biologia studia i test statistici che fanno vedere come i risultati degli esperimenti di Mendel siano in accordo con la sue predizioni
- Fisher disse a proposito di questi risultati "troppo belli per essere veri"
- in effetti il complesso dei risultati di Mendel è incompatibile con la statistica

Sul concetto di Razza

Nel 1972 il genetista marxista R. Lewontin pubblica l'articolo *The Apportionment of Human Diversity*
in cui mostra come la diversità umana non sia spiegabile in termini di distinzioni in razze.

La matematica per illustrare questa teoria è molto semplice, alla portata di studenti delle superiori.

Quantificare la diversità genetica

Quanta diversità c'è?

A0	00	00	A0	AB	00	00	00	00	B0
00	A0	AB	B0	BB	00	00	A0	00	AB
A0	A0	00	B0	AA	00	A0	00	00	B0
A0	A0	B0	A0	00	AA	00	00	B0	00
A0	AA	A0	AB	A0	B0	00	00	A0	00
A0	A0	00	00	00	B0	00	AB	00	B0
00	A0	A0	00	AA	00	B0	00	00	00
B0	00	00	A0	00	B0	A0	00	00	00
B0	00	00	B0	AB	AA	A0	A0	A0	A0
A0	A0	B0	A0	AA	A0	A0	00	A0	A0

Quantificare la diversità genetica

Quanta diversità c'è?

A0	00	00	A0	AB	00	00	00	00	B0
00	A0	AB	B0	BB	00	00	A0	00	AB
A0	A0	00	B0	AA	00	A0	00	00	B0
A0	A0	B0	A0	00	AA	00	00	B0	00
A0	AA	A0	AB	A0	B0	00	00	A0	00
A0	A0	00	00	00	B0	00	AB	00	B0
00	A0	A0	00	AA	00	B0	00	00	00
B0	00	00	A0	00	B0	A0	00	00	00
B0	00	00	B0	AB	AA	A0	A0	A0	A0
A0	A0	B0	A0	AA	A0	A0	00	A0	A0

Useremo la probabilità che un figlio sia eterozigote.

Quantificare la diversità genetica

$P(A)$=frequenza di A, $P(B)$=frequenza di B, $P(0)$=frequenza di 0

$$1 = (P(A) + P(B) + P(0))^2 =$$

$$P(A)^2 + P(B)^2 + P(0)^2 + 2P(A)P(0) + 2P(A)P(B) + 2P(B)P(0)$$

- i primi tre termini sono la probabilità dei tre genotipi omozigoti
- i doppi prodotti sono la probabilità dei tre genotipi eterozigoti; la loro somma è chiamata **eterozigosi**, ed è la probabilità di trovare un eterozigote nella popolazione

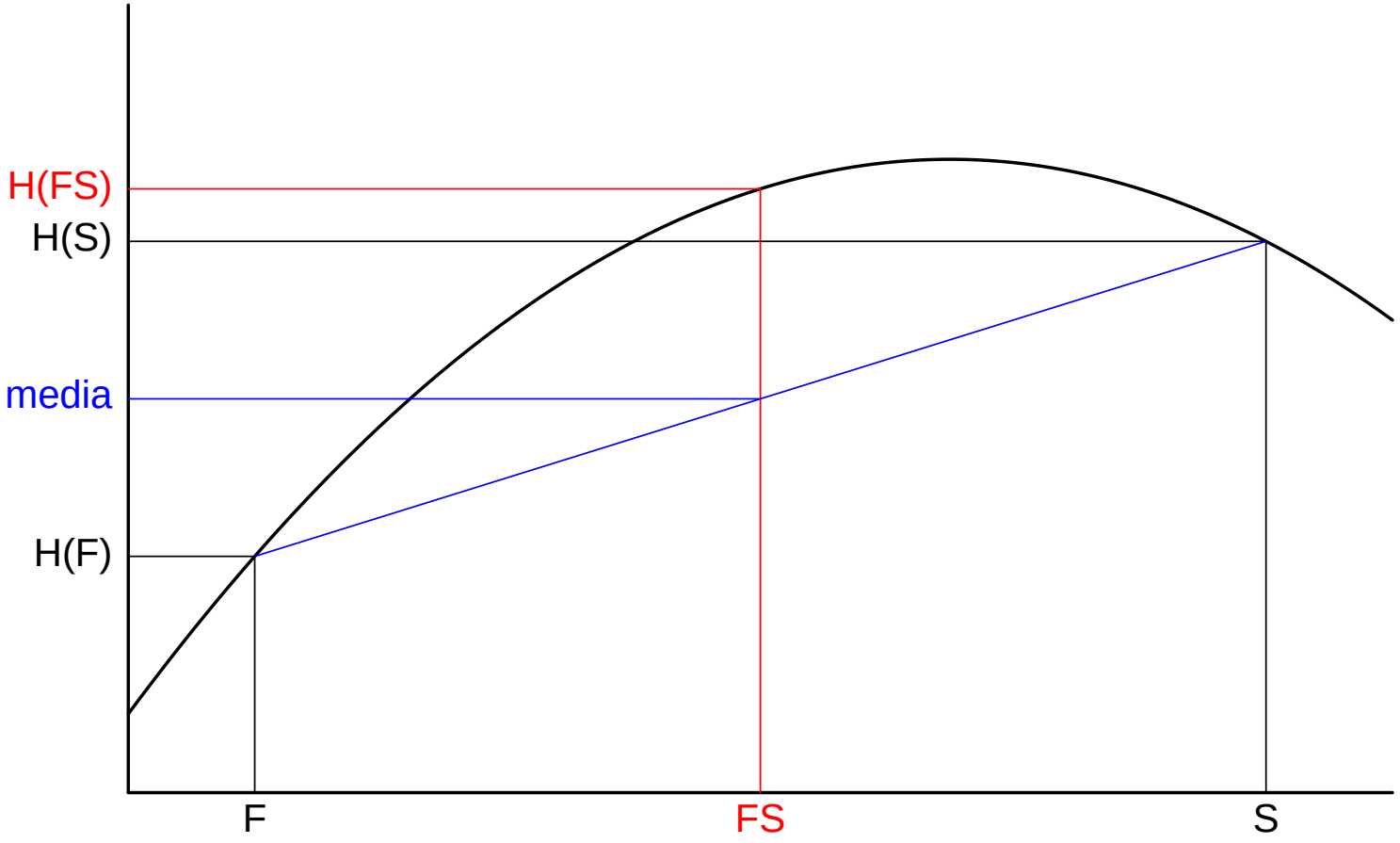
Quantificare la diversità genetica

$$H = 1 - (P(A)^2 + P(B)^2 + P(0)^2)$$

In statistica H è legato all'indice di Gini che misura appunto la diversità di una distribuzione di probabilità. Lewontin utilizzò invece l'entropia, che condivide con H le proprietà imporanti:

- H è minimo se c'è un solo allele
- H è massimo se tutti gli alleli hanno la stessa frequenza
- H aumenta se un allele si suddivide in due varianti
- H è una funzione con la concavità verso il basso

L'importanza della concavità



popolazioni

L'importanza della concavità

La richiesta di concavità equivale a chiedere che se mescoliamo due popolazioni con frequenze alleliche differenti, allora la diversità di un mix di popolazioni è è maggiore rispetto alla media delle diversità.

Lewontin usa molti alleli, e considera l'ipotesi che l'umanità sia divisa in razze, e le razze in popolazioni.

Il fallimento scientifico del concetto di razza

- $H(U)$: la diversità complessiva degli umani
- $\overline{H}_R$: la media delle diversità delle razze
- $\overline{H}_P$: la media delle diversità delle popolazioni

$$H(U) = \overline{H}_R + (H(U) - \overline{H}_R) = \overline{H}_P + (\overline{H}_R - \overline{H}_P) + (H(U) - \overline{H}_R)$$

- $\overline{H}_P$ è la diversità **dentro le popolazioni**
- $\overline{H}_R - \overline{H}_P$ è la diversità **tra le popolazioni nelle razze**
- $H(U) - \overline{H}_P$ è la diversità **tra le razze**

Il fallimento scientifico del concetto di razza

- $H(U)$: la diversità complessiva degli umani
- $\overline{H}_R$: la media delle diversità delle razze
- $\overline{H}_P$: la media delle diversità delle popolazioni

$$H(U) = \overline{H}_R + (H(U) - \overline{H}_R) = \overline{H}_P + (\overline{H}_R - \overline{H}_P) + (H(U) - \overline{H}_R)$$

- $\overline{H}_P = 85.4\%$ è la diversità **dentro le popolazioni**
- $\overline{H}_R - \overline{H}_P = 8.3\%$ è la diversità **tra le popolazioni nelle razze**
- $H(U) - \overline{H}_P = 6.3\%$ è la diversità **tra le razze**

Bibliografia

- G. Barbujani, L'invenzione delle razze, Bompiani, Milano 2007
- L.L. Cavalli-Sforza, Geni, popoli, lingue, Adelphi, Milano, 1996
- A.J.F. Griffiths, J.H. Miller, D.T. Suzuki, R. Lewontin, Genetica - Principi di Analisi Formale, Zanichelli, Bologna, 2002.
- D. Benedetto, M. Degli Esposti, C. Maffei, Matematica per le scienze della vita, CEA, Milano 2014
- D. Benedetto, C. Maffei, La matematica della diversità umana, Nuova Secondaria, gennaio 2017
- http://brazil.mat.uniroma1.it/dario/