Mathématiques du hasard et de l'évolution

S. Méléard, Ecole Polytechnique, Centre de Mathématiques Appliquées

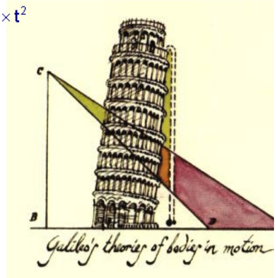
CEMRACS - CIRM 23 juillet 2013

Avant Darwin

1604, Galilée comprend la loi mathématique décrivant la chute d'un corps.

$$\vec{d} = \frac{1}{2} \times 9{,}8 \frac{m}{s^2} \times t^2$$

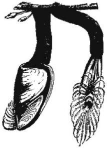
$$\vec{d} = 4{,}9 \frac{m}{s^2} t^2$$

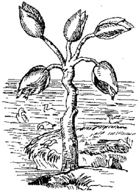


En Biologie, les connaissances sont encore proches de l’âge de pierre...

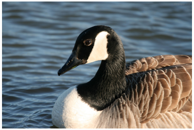
Portraict de l'Arbre qui porte des fueilles, lesquelles tombées sur terre se tournent en oyseaux volants, & celles qui tombent dans les eaux se muent en poissons.

Tiré de Duret, L'Histoire admirable des Plantes, Paris, 1605.





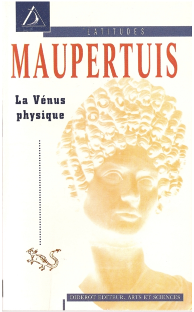
The Barnacle Tree (from Gerard's "Herball")



Une première vision évolutive

Pierre-Louis Moreau de Maupertuis (1698-1759).

Mathématicien, astronome, géographe, naturaliste. Précurseur de la génétique moderne.



Chapitre III

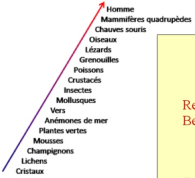
PRODUCTIONS DE NOUVELLES ESPECES

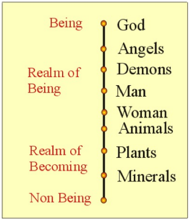
“La nature contient le fonds de toutes ces variétés : mais le hasard ou l’art les mettent en oeuvre... Nous voyons paraître des races de chiens, de pigeons, de serins qui n’étaient point auparavant dans la nature. Ce n’ont été d’abord que des individus fortuits ; l’art et les générations répétées en ont fait des espèces.”

Apparition de la variabilité individuelle

Lamarck (1744-1829) "Philosophie zoologique" (1809)

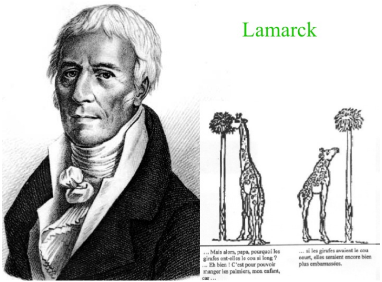
Avant Lamarck: une image linéaire de la grande Chaîne des Etres!

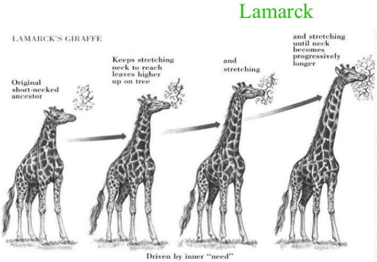




L'image linéaire de la grande Chaîne des Etres est remplacée par un arbre buissonnant, ancêtre des processus de branchement et du coalescent.

Lamarck croit à la transmission aux descendants des caractères acquis qui évoluent au cours de la vie.



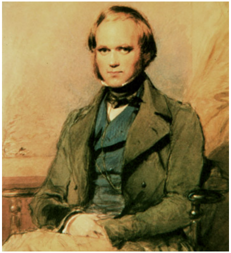


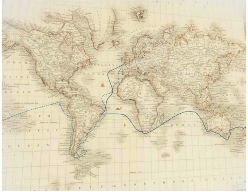
Lamarck a longtemps pensé qu'il existait des « espèces constantes » mais il écrit, en 1802,

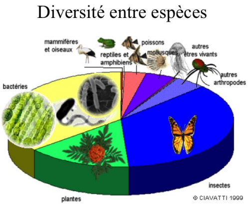
« Maintenant, je suis convaincu que j'étois dans l'erreur à cet égard et qu'il n'y a dans la nature que des individus »

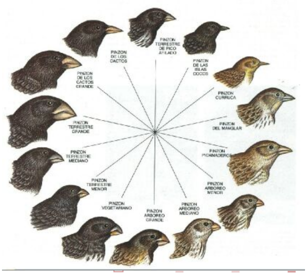
DARWIN (1809-1882)

Ses travaux sur l'évolution des espèces vivantes ont révolutionné la biologie. L'origine des espèces (1859)



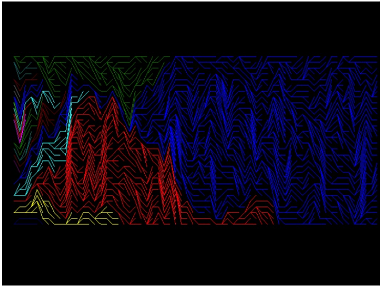






Ancêtre commun

Darwin formule l'hypothèse selon laquelle toutes les espèces vivantes ont évolué au cours du temps à partir d'un seul ou de quelques ancêtres communs grâce au processus connu sous le nom de **sélection naturelle**.



Comment expliquer la sélection naturelle aux enfants...

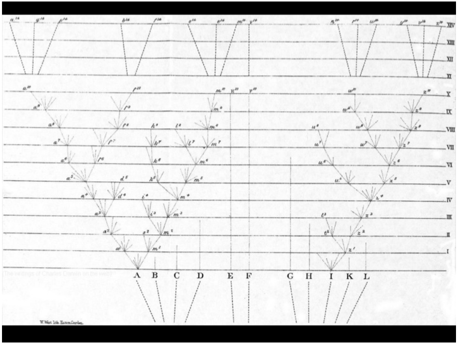

© Julius T. Csotonyi (csotonyi.com)


Darwin - précurseur d'une théorie comme l'est un mathématicien: par un raisonnement logique et objectif

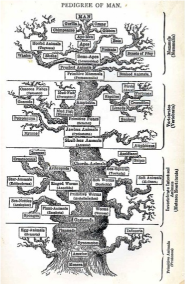
" Les années passant, j'ai profondément regretté de ne pas m'être efforcé de comprendre quelques aspects des principes essentiels des mathématiques: car les hommes dotés d'un tel entendement semblent posséder un sixième sens."

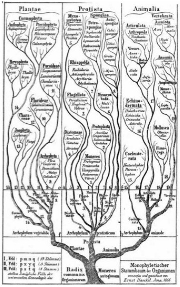
Charles Darwin, Autobiography.

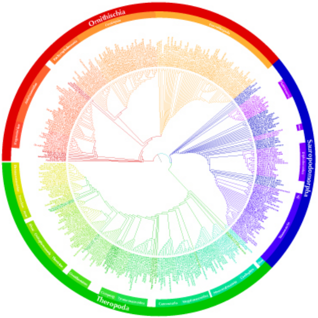
L’unique dessin de Darwin dans "l’Origine des espèces"



Origine des arbres de phylogénie



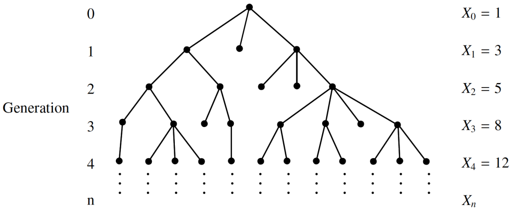




Le processus de Galton-Watson (1873)

Chaque individu d'une génération meurt et laisse place à un nombre variable d'enfants.

A chaque instant n, la variable aléatoire X_n désigne la taille de la population au temps n.



Soit p_k la probabilité pour chaque individu d'avoir k enfants: $(p_k)_k$ définit la loi de reproduction. Supposons que $k \in \{0, 1, \cdots, 5\}$.

On a

$$0 \leq p_k \leq 1 \;; \quad p_0 + p_1 + \cdots + p_5 = 1.$$

Comment évolue la taille de la population X_n quand le nombre de générations n devient très grand?

Extinction de la population à la génération n : $X_n = 0$.

Persistance de la population à la génération n : $X_n > 0$.

On introduit le polynôme

$$P(x) = p_0 + p_1 x + p_2 x^2 + \cdots + p_5 x^5.$$

C'est la fonction génératrice de la loi de reproduction.

On a

$$P(1) = 1 \ ; \ \ P(0) = p_0 = \text{ probabilité d'avoir 0 enfant.}$$

Nombre moyen d'enfants m.
Remarquons que

$$m = p_1 + 2\,p_2 + 3\,p_3 + 4\,p_4 + 5\,p_5 = P'(1).$$

Soit P_n le polynôme tel que le coefficient de x^k est la probabilité d'avoir k descendants à la génération n partant d'un seul individu à la génération 0.

On a la relation de récurrence $P_n(x) = P[P_{n-1}(x)]$.

Probabilité d'extinction à la génération n : $q_n = P_n(0)$. On a alors

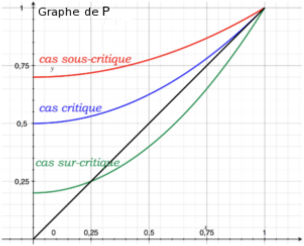
$$q_n = P(q_{n-1}).$$

C'est une suite définie par récurrence.

Probabilité que la population s'éteigne en temps long : $q = \lim_{n\to+\infty} q_n$. Elle satisfait $q = P(q)$.

Deux cas possibles:

- $m \leq 1$ (cas sous-critique ou critique): $q = 1$. **Extinction.**
- $m > 1$ (cas sur-critique): $q < 1$. **Persistance avec probabilité positive.**



Extinction des baleines noires dans l'Atlantique Nord

Caswell et al. (1999).

On peut montrer que si $m < 1$ et si la population initiale vaut K,

$$\text{Probabilité}(X_n > 0) \simeq K\, m^n.$$

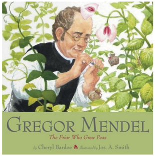
- Nombre de baleines femelles en 1994: $K = 150$.

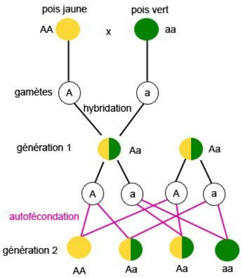
- Nombre moyen de petits: $m = 0,976$.

- On cherche n tel que Probabilité(Extinction à l'année n)=$0,99$.
- $\text{Probabilité}(X_n = 0) = 0,99 \iff \text{Probabilité}(X_n > 0) = 0,01.$
- $K\, m^n = 0,01 \iff n = 395.$
- 99% de chances de ne plus avoir de baleines en 2389.

Les petits pois de MENDEL (1822-1884) et les lois de l'hérédité

Diploïdie - Ségrégation - assortiment indépendant des caractères.





Grande population où les choix de partenaires se font au hasard:

$$\mathbb{P}(\text{allèle A}) = \mathbb{P}(\text{génotype AA}) + \frac{1}{2}\mathbb{P}(\text{génotype Aa})$$

$$\mathbb{P}(\text{AA à la génération suivante}) = \mathbb{P}(\text{ allèle A})^2$$

Loi de Hardy-Weinberg

Hardy (1877-1947). Hardy s'intéresse au problème de Mendel (Science 1908).

Il considère une grande population où les choix de partenaires se font totalement au hasard.

Soient p_n , $2q_n$ et r_n les probabilités des génotypes AA , Aa et aa à la $n$ième génération. Bien-sûr,

$$p_n + 2q_n + r_n = 1 .$$

On a vu:

$$p_1 = (p_0 + q_0)^2 \; ; \; q_1 = (p_0 + q_0)(r_0 + q_0) \; ; \; r_1 = (r_0 + q_0)^2 .$$

Remarquons que

$$q_1^2 = p_1 r_1 .$$

Mais alors, si $q_1^2 = p_1 r_1$, on obtient

$$\begin{aligned} p_2 &= (p_1 + q_1)^2 = p_1^2 + 2\,p_1 q_1 + q_1^2 \\ &= p_1^2 + 2\,p_1 q_1 + p_1 r_1 \\ &= p_1\,(p_1 + 2q_1 + r_1) = p_1 \end{aligned}$$

et de même

$$q_2 = q_1 \ ; \ r_2 = r_1 .$$

Les probabilités p_n, $2q_n$ et r_n des différents génotypes restent constantes au cours du temps.

Un gros progrès par rapport à Darwin et Galton qui n'arrivaient pas à comprendre comment la diversité pouvait perdurer.

Mais cela n'explique pas la fixation possible d'un mutant.

Comment intégrer la sélection naturelle?

Sélection naturelle et calculs de Fisher

Fisher (1890-1962) - Statisticien très célèbre. Il s'intéresse aux problèmes statistiques liés aux travaux de Galton et Mendel (1930 General theory of Natural Selection).

Hypothèse: les individus de génotype AA, Aa et aa ont des mortalités différentes avant d'atteindre l'âge de la reproduction.

Soient u, $u(1+\alpha s)$ et $u(1+\beta s)$ les probabilités de survie des différents génotypes où s est appelé l'avantage sélectif.

Si $s > 0$, le mutant a est favorable, si $s < 0$ le mutant est dit défavorable ou délétère.

Les probabilités p_n, $2q_n$ et r_n sont alors pondérées par ces probabilités de survie.

Trois limites possibles suivant les valeurs de α et β.

- Tous les individus sont de type *aa*.
- Tous les individus sont de type *AA*.
- Diversité dans la population, types *aa*, *AA*, *Aa*.

Comment modéliser l'évolution pour une petite population et prendre en compte ses fluctuations?

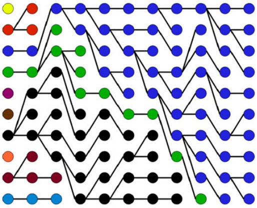
L'arbre généalogique d'une petite population

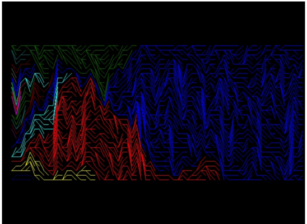
Le modèle de Wright-Fisher. (Wright, 1889-1988)

- Taille de population N constante.
- A chaque génération, chaque individu choisit son père uniformément au hasard dans la génération précédente.
- Nombre de descendants D d'un individu: **loi binomiale** $B(N, \frac{1}{N})$**.**

$$\mathbb{P}(D = k) = \binom{N}{k} \left(\frac{1}{N}\right)^k \left(1 - \frac{1}{N}\right)^{N-k}.$$

- Les effets de la dérive génétique:

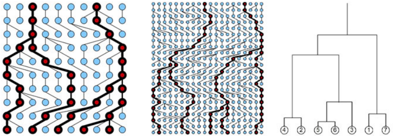




Le processus de coalescence

Quand N tend vers l'infini: la probabilité que deux individus aient le même parent: $N \times \frac{1}{N} \times \frac{1}{N} = \frac{1}{N} \to 0$.

Il faut changer l'échelle de temps: pour voir quelque chose on se place au temps $N\,t$.



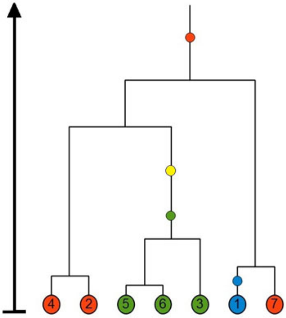
On construit quand $N \to \infty$ un objet mathématique appelé coalescent de Kingman (Kingman, 1982).

On peut remonter le temps pour trouver les lignées ancestrales d'un échantillon d'individus et leur plus récent ancêtre commun.

Objet géométrique avec des distances aléatoires.

Reconstruction de l'histoire des individus

Coalescent avec mutations:

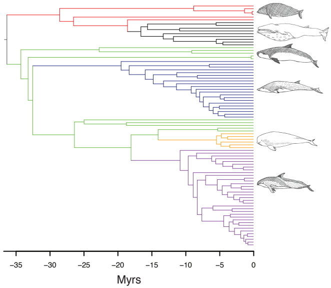


En comparant les séquences d'ADN, on peut retrouver les ancêtres communs des individus.

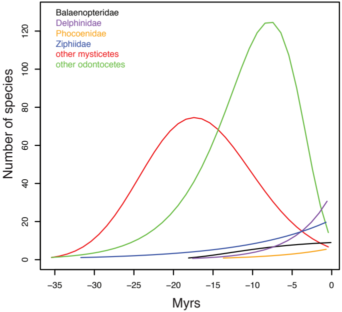
Reconstruire la biodiversité passée à partir des phylogénies moléculaires des espèces actuelles

(Morlon-Parsons-Plotkin, 2011).

a



c



Autres exemples de questions posées

- Ecologie: Les espèces cohabitent et interagissent (proie-prédateur, hôte-parasite, réseaux trophiques). Aura-t-on coexistence de toutes les espèces ou disparition de certaines d'entre elles?

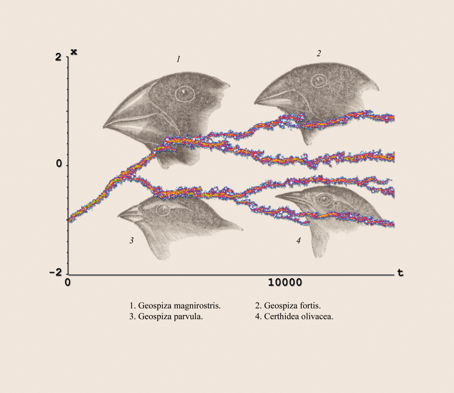

- Impact de la migration et de la fragmentation de l'habitat sur la biodiversité.

- Impact des variations de l'environnement sur la biodiversité. (Réchauffement climatique).
- Comment la reproduction sexuée se maintient-elle?

Comprendre les mécanismes de diversification des espèces: les pinsons de Darwin

- Chaque individu est caractérisé par un certain nombre de phénotypes.
  Exemple: la taille du bec à la naissance.
- **Hérédité**: chaque individu a en général le même type que son parent.
- **Mutations**: Génèrent de la variabilité dans les types.
  Exemple: type du mutant tiré suivant une loi normale autour du type de l'ancêtre.
- Tous les individus sont en **compétition** pour le partage des ressources.
  Exemple: Compétition plus grande si les oiseaux ont des tailles de bec proches et donc mangent les mêmes graines.
- **Sélection** des types les mieux adaptés.

L'évolution de chaque individu est supposée aléatoire mais quand le nombre d'individus est grand, on s'intéresse à la densité autour d'un trait donné.



(Champagnat-Ferrière-Méléard, 2011)

La dynamique de mutation - sélection peut entraîner la séparation en plusieurs espèces qui n'interagissent plus entre elles.

En conclusion...

- Les problèmes scientifiques liés à la biodiversité sont extrêmement complexes.
- Il faut des modèles mathématiques innovants.
- Ces modèles utilisent de nombreuses parties des mathématiques
  - **Modèles aléatoires**
  - **Modèles déterministes**
  - **Statistiques et méthodes numériques**
- Ils doivent permettre d'associer des quantités calculables.
- Ils permettent de développer des algorithmes de simulation (expérimentation fictive sur machine).
- A partir de données observées, on peut construire des outils statistiques permettant de prédire et quantifier différents scénarios de la biodiversité.

Merci pour votre attention