

Génétique Quantitative

Approche historique

Gregor Mendel



P. Anselm Rambousek P. Antonin Alt P. Thomas Bratranek P. Josef Lindenthal P. Gregor Mendel
P. Benedikt Fogler P. Paul Křižkovský P. Baptist Vorthey P. Cyrill Napp P. Alipius Winkelmeyer P. Wenzel Šembera

Plate III. Gregor Johann Mendel among his Fellow-Monks

Wellcome Images

Gregor Mendel, 1822-1884





Gregor Mendel est moine ; il fut également professeur de physique et de sciences naturelles au lycée de Brno. Il est passionné par les sciences, a suivi à Vienne des cours de matières diverses (mathématiques, botanique, physiologie végétale, etc) à l'institut de physique fondé par Christian Doppler.

Son mémoire de 1865 est intitulé *Expériences sur l'hybridation des plantes*. L'hybridation est couramment utilisée pour créer de nouvelles variétés ; le phénomène de la réapparition des caractères ancestraux à partir de la deuxième génération d'hybrides est bien connu et Mendel n'est pas le premier à chercher à mieux le comprendre.

Mendel procède à des expériences sur le pois (*Pisum*). Il s'intéresse principalement à des caractères discontinus (couleur, forme des graines, etc), dont il souligne l'absence de formes intermédiaires.

Loi de Mendel pour un caractère

La première génération d'hybride (F_1 en terminologie moderne) ne présente que le caractère *dominant*. À la génération suivante, le caractère *récessif* réapparaît :





	A	a
A	 A	 Aa
a	 Aa	 a

La génération F_2 présente le caractère dominant en proportion 3 : 1.

Mendel constate l'absence de variations dans la descendance des formes récessives ; il reconnaît qu'un tiers des formes dominantes a une descendance exclusivement dominante, tandis que les deux tiers restants ont une descendance où la variation s'observe à nouveau en proportion 3 : 1.

Loi de Mendel pour un caractère

La première génération d'hybride (F_1 en terminologie moderne) ne présente que le caractère *dominant*. À la génération suivante, le caractère *récessif* réapparaît :

	A	a
A	 A	 Aa
a	 Aa	 a

















Mendel comprend que cette proportion observée 3 : 1 est en fait une proportion 1 : 2 : 1 de formes (en notation moderne) AA , Aa et aa ; il l'exprime par l'expression formelle

$$A + 2Aa + a.$$

Il est à même de calculer les proportions attendues après un nombre quelconque de générations.

Loi de Mendel pour deux caractères

Mendel considère des caractères non génétiquement liés – a-t-il écarté les résultats qui ne « cadraient » pas ?

	AB	Ab	aB	ab
AB	 AB	 ABb	 AaB	 AaBb
Ab	 ABb	 Ab	 AaBb	 Aab
aB	 AaB	 AaBb	 aB	 aBb
ab	 AaBb	 Aab	 aBb	 ab

Le résultat est élucidé en combinant les deux expressions $A + 2Aa + a$ et $B + 2Bb + b$ en

$$AB + Ab + aB + ab + 2ABb + 2aBb + 2AaB + 2Aab + 4AaBb.$$

Mécanisme proposé

Mendel propose que les gamètes (*Keimzellen* ou *Pollenzellen*) soient de forme A ou a ; les hybrides Aa émettent les deux types de gamètes dans des proportions 1 : 1, et les gamètes s'apparient au hasard.

Pour illustrer ce mécanisme, Mendel écrit

$$\frac{A}{A} + \frac{A}{a} + \frac{a}{A} + \frac{a}{a} = A + 2Aa + a,$$

que nous écrivions aujourd'hui

$$AA + Aa + aA + aa = AA + 2Aa + aa.$$

Une hypothèse analogue est formulée pour le cas de plusieurs caractères (hypothèse qui n'est valable que dans le cas d'absence de liaison génétique).

Ces hypothèses sont testées par des expériences de la forme *backcross*.

Caractères continus

Les résultats de Mendel concernent des caractères discontinus.

Il a cependant également mené des expériences sur les haricots (*Phaesusolus*), s'intéressant notamment à la couleur des fleurs. Il est quelque peu dérouté par les résultats obtenus : un continuum de variation, de trop rares retours à la forme récessive...

Il note cependant : *Ces résultats énigmatiques, cependant, peuvent probablement s'expliquer par les lois qui régissent Pisum si nous supposons que la couleur des fleurs et des graines de Ph. multiflorus est la combinaison de deux couleurs entièrement indépendantes ou plus, qui se comportent individuellement comme n'importe quel autre caractère constant de la plante.*

👉 Il me semble que c'est une esquisse assez lucide du modèle polygénique !

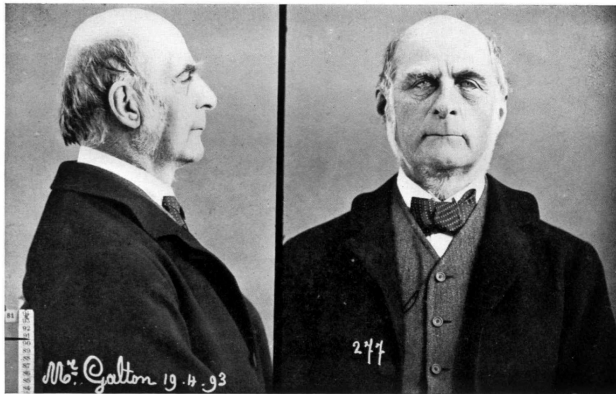
Les travaux de Mendel sont passés inaperçus de son vivant. Le fait qu'ils aient été présentés comme des travaux sur l'hybridation et non sur l'hérédité a pu y contribuer.

Mendel n'était pas un scientifique professionnel ; il est devenu supérieur de son couvent en 1868 et ne pouvait consacrer son temps à défendre ses travaux. Il continuera à s'intéresser à l'horticulture et l'apiculture, et publiera également des relevés météorologiques détaillés.

Son mémoire fut redécouvert en 1900, en même temps que Correns, de Vries, et von Tschermak réalisaient des expériences similaires.

Plus tard (1936) Fisher montrera que les résultats rapportés par Mendel sont trop beaux pour être vrais : consciemment ou non, Mendel a biaisé les chiffres pour se rapprocher des proportions théoriques.

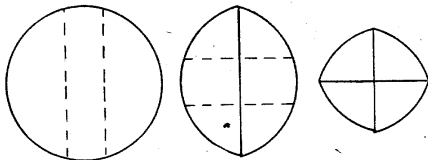
Francis Galton (Sir)



Francis Galton (1822-1911)

Cousin éloigné de Darwin, génie précoce, Galton fera une dépression à l'université quand il ne se classera que second en mathématiques.

Touche-à-tout, il participera à plusieurs voyages d'exploration en Afrique, y réalisa notamment des relevés géographiques. Il publiera dans Nature (1906) une méthode pour couper un gâteau de façon scientifique afin d'éviter que l'entame ne sèche.



Broken straight lines show intended cuts. Ordinary straight lines show the cuts that have been made. The segments are kept in apposition by a common elastic band that encloses the whole. In the above figures about one-third of the area of the original disc is removed by each of the two successive operations.

Mais c'est comme pionnier de la biométrie et des statistiques et comme fondateur de l'eugénisme qu'il reste dans l'histoire.

Big data à l'ère victorienne

La révolution industrielle est le moment d'une collecte de données chiffrées sans précédent (chemins de fers, commerce, etc) et du développement des méthodes pour les analyser.

C'est dans ce contexte que naît la biométrie. Le statisticien Quetelet (1796-1874) remarque le premier que la taille et d'autres mesures suivent une distribution comparable à la « loi des erreurs » ou loi normale. Il en tirera l'idée d'un « homme-moyenne » parfait...

Galton est obsédé par les chiffres. Il mesure tout, et note même la beauté des femmes qu'il croise (trois notes) pour dresser une carte de beauté des îles britanniques (malheureusement non publiée). Il est fatalement attiré par la biométrie, dont il est un des fondateurs !

L'hérédité

Son intérêt pour la théorie de la descendance avec modification de Darwin se mêle à un intérêt personnel pour l'hérédité.

Il mène des expériences pour tester la théorie des gemmules de Darwin ; il élabore ensuite sa propre théorie, postulant l'existence du *stirp*, un ensemble d'éléments hérités des parents, transmis aux enfants, dont une partie seulement s'« exprime » chez un individu donné.

Il abandonnera progressivement toute référence explicite à cette théorie, pour se concentrer sur la mathématisation de l'hérédité.

Galton veut expliquer par une loi unique l'hérédité des traits continus (la taille), et celle des traits discontinus (couleur des yeux, couleur du pelage des bassets...), et en particulier pour ces derniers la réapparition de caractères ancestraux (qu'il appelle *atavisme*).

La taille et la régression vers la moyenne

TABLE I.
NUMBER OF ADULT CHILDREN OF VARIOUS STATURES BORN OF 205 MID-PARENTS OF VARIOUS STATURES.
(All Female heights have been multiplied by 1.08).

Heights of the Mid-parents in inches.	Heights of the Adult Children.														Total Number of		Medians.
	Below	62.2	63.2	64.2	65.2	66.2	67.2	68.2	69.2	70.2	71.2	72.2	73.2	Above	Adult Children.	Mid-parents.	
Above	1	3	..	4	5	..
72.5	1	2	1	2	7	2	4	19	6	72.2
71.5	1	3	4	3	5	10	4	9	2	2	43	11	69.9
70.5	1	..	1	..	1	1	3	12	18	14	7	4	3	3	68	22	69.5
69.5	1	16	4	17	27	20	33	25	20	11	4	5	183	41	68.9
68.5	1	..	7	11	16	25	31	34	48	21	18	4	3	..	219	49	68.2
67.5	..	3	5	14	15	36	38	28	38	19	11	4	211	33	67.6
66.5	..	3	3	5	2	17	17	14	13	4	78	20	67.2
65.5	1	..	9	5	7	11	11	7	7	5	2	1	66	12	66.7
64.5	1	1	4	4	1	5	5	..	2	23	5	65.8
Below ..	1	..	2	4	1	2	2	1	1	14	1	..
Totals ..	5	7	32	59	48	117	138	120	167	99	64	41	17	14	928	205	..
Medians	66.3	67.8	67.9	67.7	67.9	68.3	68.5	69.0	69.0	70.0

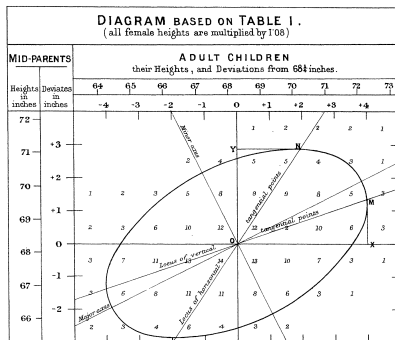
Galton remarque que la distribution jointe de la taille des enfants et de la moyenne de la taille des deux parents suit une loi normale bivariée.

La taille des femmes est multipliée par 1,08 pour obtenir une distribution approximativement identique à celle de la taille des hommes.

Galton 1886, Régression towards mediocrity in hereditary stature.

Galton 1889, Inheritance in man.

La taille et la régression vers la moyenne



Galton remarque que la distribution jointe de la taille des enfants et de la moyenne de la taille des deux parents suit une loi normale bivariée.

La taille des femmes est multipliée par 1,08 pour obtenir une distribution approximativement identique à celle de la taille des hommes.

Galton 1886, Régression towards mediocrity in hereditary stature.

Galton 1889, Inheritance in man.

La taille et la régression vers la moyenne

Galton met en l'évidence une corrélation positive entre la taille des parents et celle des enfants.

Une conséquence de cette distribution est que l'écart $Y - \mu$ entre la taille Y d'un enfant et la taille moyenne μ est, en moyenne, égal au deux-tiers de l'écart entre la taille moyenne M de ses parents et μ :

$$Y - \mu = \frac{2}{3}(M - \mu)$$

d'où le terme de « régression vers la moyenne ».

Il s'intéresse aussi à des relations de parentés plus lointaines. Il proposera plusieurs lois mathématiques pour régir les corrélations constatées. Toutes ses tentatives supposent que la loi est la même pour tous les traits étudiés. Pour les traits discontinus, il utilise ces mêmes lois pour calculer les probabilités des diverses catégories d'un même trait...

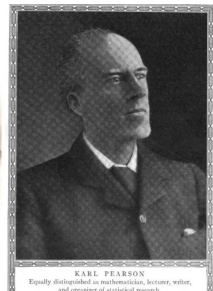
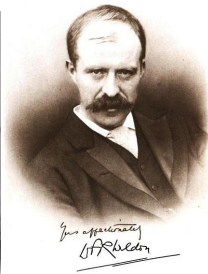
Les lois proposées par Galton sont fausses et oubliées.

Il reste néanmoins un pionnier de la biométrie et des statistiques descriptives, attirant le premier l'attention sur l'intérêt d'analyser ce qui ne s'appelait pas encore les « corrélations ».

Son rôle comme fondateur et promoteur actif de l'eugénisme nuit certainement quelque peu à sa mémoire.

It must be introduced into the national conscience, like a new religion. It has, indeed, strong claims to become an orthodox religious tenet of the future, for eugenics co-operate with the workings of nature by securing that humanity shall be represented by the fittest races.

La controversie



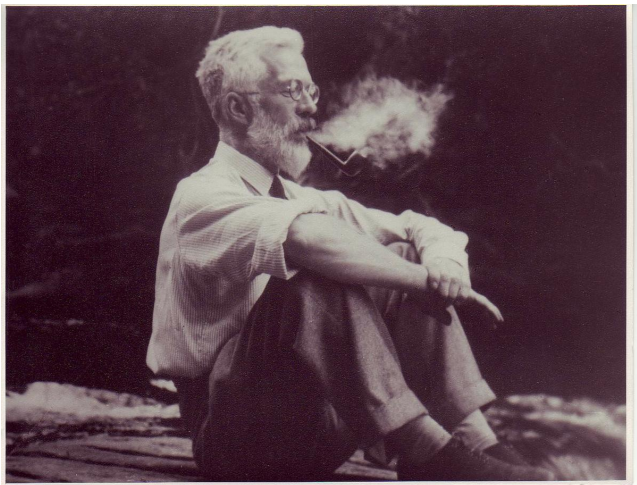
La controverse entre biométriciens et mendéliens

Les biométriciens, représentés essentiellement par Weldon et Pearson, rejettent la théorie de l'hérédité mendélienne qu'ils jugent incapable d'expliquer l'hérédité des traits continus.

De son côté, Bateson se fait le champion du mendélisme, attaquant avec vigueur ses opposants,

Les protagonistes s'opposent également sur les mécanismes de spéciation ; et les généticiens (Bateson mais aussi Morgan) sont également très critiques à l'égard des théories eugénistes...

Ronald Aylmer Fisher



R.A. Fisher (1890-1962)

C'est à Fisher qu'échoit l'honneur de réaliser la synthèse, dans un article de 1918 (de lecture ardue) intitulé *The correlation between relatives on the supposition of Mendelian inheritance*.

Mendel suppose qu'un grand nombre de facteurs mendéliens participent de façon additive à la construction d'un phénotype. Celui-ci doit alors suivre une loi approximativement normale.

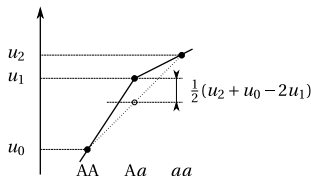
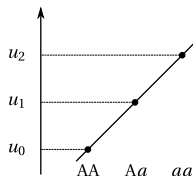
Fisher calcule les corrélations attendues entre apparentés sous ce modèle et montre qu'elles sont compatibles avec les corrélations observées par les biométriciens.

La solution de Fisher (revisitée)

Considérons un facteur à deux allèles A et a dont la contribution à un trait est donnée par

$$g = \begin{cases} u_0 & \text{si le génotype est } AA \\ u_1 & \text{si le génotype est } Aa \\ u_2 & \text{si le génotype est } aa \end{cases}$$

Si $u_2 - u_1 = u_1 - u_0$ on a un effet additif. L'écart à l'additivité peut être mesuré par $u_0 + u_2 - 2u_1$.



On peut décomposer g en une composante additive et une composante non-additive indépendantes.

La solution de Fisher (revisitée)

On note p et q les fréquences des allèles A et a .

On définit deux variables centrées et réduites (sous l'équilibre de Hardy-Weinberg), X_a et X_d , par

$$X_a = \begin{cases} \frac{1}{\sqrt{2pq}}(0 - 2q) & \text{si } AA \\ \frac{1}{\sqrt{2pq}}(1 - 2q) & \text{si } Aa \\ \frac{1}{\sqrt{2pq}}(2 - 2q) & \text{si } aa \end{cases} \quad \text{et} \quad X_d = \begin{cases} \frac{q}{p} & \text{si } AA \\ -1 & \text{si } Aa \\ \frac{p}{q} & \text{si } aa \end{cases}$$

On a $E(X_a) = E(X_d) = 0$, $\text{var}(X_a) = \text{var}(X_d) = 1$, et $\text{cov}(X_a, X_d) = 0$.

On note que les variables de la forme $g = \mu + \alpha X_a$ sont les variables avec des effets additifs ($u_2 - u_1 = u_1 - u_0 = \frac{\alpha}{\sqrt{2pq}}$).

La solution de Fisher (revisitée)

Dans le cas général, on a

$$g = \mu + \alpha X_a + \beta X_d$$

avec

$$\mu = E(g) = p^2 u_0 + 2pq u_1 + q^2 u_2$$

$$\alpha = \sqrt{2pq}(p(u_1 - u_0) + q(u_2 - u_1))$$

$$\beta = pq(u_0 + u_2 - 2u_1)$$

La composante additive est αX_a et la composante non-additive, dite « de dominance », est βX_d .

On a $\text{var}(g) = \alpha^2 + \beta^2$: α^2 représentant la variance additive, et β^2 la variance due à l'écart éventuel à l'additivité.

La solution de Fisher (revisitée)

On décompose un phénotype P en

$$P = G + E$$

où la composante génétique $G = g_1 + \dots + g_n$ est la somme d'un grand nombre de termes indépendants de cette forme, et E est une composante environnementale, supposée normale et indépendante de G .

On peut décomposer ainsi G à son tour en deux termes

$$G = A + D$$

avec $\text{var}(A) = \sigma_A^2 = \alpha_1^2 + \dots + \alpha_n^2$, $\text{var}(D) = \sigma_D^2 = \beta_1^2 + \dots + \beta_n^2$ et $\text{var}(G) = \sigma_A^2 + \sigma_D^2$.

La solution de Fisher (revisitée)

En supposant que les locus impliqués ne sont pas génétiquement liés, la covariance entre deux apparentés est

$$\text{cov}(G_1, G_2) = 2\phi\sigma_A^2 + \mathbb{P}(IBD = 2)\sigma_D^2.$$

Pour deux germains :

$$\text{cov}(G_1, G_2) = \frac{1}{2}\sigma_A^2 + \frac{1}{4}\sigma_D^2$$

Parent et enfant :

$$\text{cov}(G_1, G_2) = \frac{1}{2}\sigma_A^2$$

Grand-parent et enfant :

$$\text{cov}(G_1, G_2) = \frac{1}{4}\sigma_A^2$$

La solution de Fisher (revisitée)

Mis à part le cas des germains (et des doubles cousins germains), dans la plupart des relations couramment considérées en biométrie on a $\mathbb{P}(IBD = 2) = 0$ et

$$\text{cov}(G_1, G_2) = 2\phi\sigma_A^2,$$

d'où, en supposant les composantes environnementales indépendantes, la corrélation entre les phénotypes

$$\text{cor}(P_1, P_2) = 2\phi \frac{\sigma_A^2}{\sigma_A^2 + \sigma_D^2 + \sigma_E^2} = 2\phi h^2.$$

L'*héritabilité restreinte* h^2 est la proportion de variance expliquée par les effets additifs.

Les estimations usuelles se font sur des données familiales.