0 Trop beau pour être vrai

histoires de statistique

Trop beau pour être vrai

les expériences de Mendel

S'il y a bien un grand nom dans l'histoire des sciences, c'est celui de Mendel. Ses expériences sur les pois sont le fondement de la génétique moderne.

1 Gregor (Johann) Mendel (1822–1884)

Gregor (Johann) Mendel (1822–1884)

Gregor Mendel s'appelait Johann à la naissance. Il avait pris le prénom de Gregor quand il s'était fait moine, et il s'était fait moine, surtout pour des raisons économiques : ça lui avait permis de suivre, à Vienne, les études que sa famille n'aurait pas pu lui payer.

2 la communauté du Königskloster en 1866

la communauté du Königskloster en 1866

Gregor Mendel (1822–1884)

Il était entré dans la communauté du Königskloster, et avait gravi les échelons, jusqu'à en devenir prieur, directeur si vous préférez. Vous voyez ici une photo de la communauté. Mendel est le troisième à partir de la droite. Il est debout et il tient à la main une tige de ces fameux pois, qu'il venait de passer une dizaine d'années à faire pousser.

3 Le jardin du Königskloster à Brünn

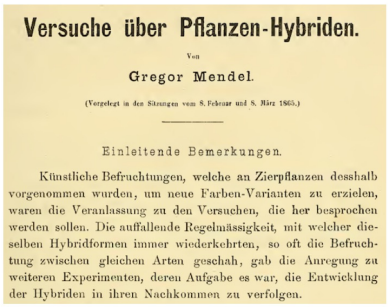
Le jardin du Königskloster à Brünn
Gregor Mendel (1822–1884)

Il les faisait pousser dans le jardin du couvent, le Königskloster, qui se situe dans la ville de Brünn. C'est actuellement la seconde ville de la République tchèque, et elle s'appelle Brno.

Dans cette ville de Brünn, il y avait une société savante, qui publiait un journal. Dans ce journal est paru en 1865 l'article de Mendel.

4 Versuche über Pflanzen-Hybriden (1865)

Versuche über Pflanzen-Hybriden (1865)
Gregor Mendel (1822–1884)



Versuche über Pflanzen-Hybriden.

Von

Gregor Mendel.

(Vorgelegt in den Sitzungen vom 8. Februar und 8. März 1865.)

Einleitende Bemerkungen.

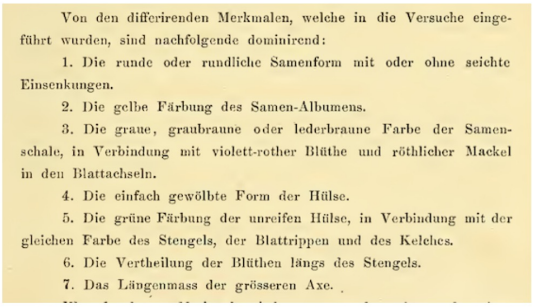
Künstliche Befruchtungen, welche an Zierpflanzen desshalb vorgenommen wurden, um neue Farben-Varianten zu erzielen, waren die Veranlassung zu den Versuchen, die hier besprochen werden sollen. Die auffallende Regelmässigkeit, mit welcher dieselben Hybridformen immer wiederkehrten, so oft die Befruchtung zwischen gleichen Arten geschah, gab die Anregung zu weiteren Experimenten, deren Aufgabe es war, die Entwicklung der Hybriden in ihren Nachkommen zu verfolgen.

« Expériences sur les hybrides de plantes ». C'est maintenant un des articles les plus célèbres de l'histoire, mais il est resté totalement ignoré pendant 35 ans, jusqu'après la mort de Mendel, qui n'a jamais rien su de sa notoriété à venir.

Sans doute était-ce trop tôt. Pourtant maintenant, la théorie nous semble relativement simple. Surtout, Mendel l'avait expliquée de façon très claire.

5 sept caractères dominant/récessif

sept caractères dominant/récessif
Mendel, Versuche über Pflanzen-Hybriden (1865)



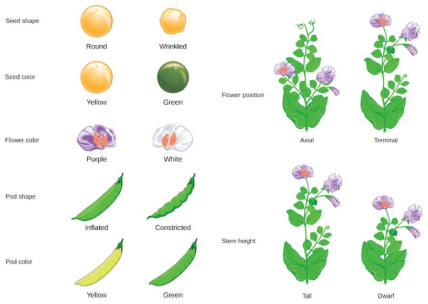
Von den differirenden Merkmalen, welche in die Versuche eingeführt wurden, sind nachfolgende dominirend:

1. Die runde oder rundliche Samenform mit oder ohne seichte Einsenkungen.
2. Die gelbe Färbung des Samen-Albumens.
3. Die graue, graubraune oder lederbraune Farbe der Samenschale, in Verbindung mit violett-rother Blüthe und röthlicher Mackel in den Blattachseln.
4. Die einfach gewölbte Form der Hülse.
5. Die grüne Färbung der unreifen Hülse, in Verbindung mit der gleichen Farbe des Stengels, der Blattrippen und des Kelches.
6. Die Vertheilung der Blüthen längs des Stengels.
7. Das Längenmass der grösseren Axe.

Mendel annonce qu'il va considérer sept caractères de la plante. Il y a la forme et la couleur des graines, la couleur des fleurs, la forme et la couleur des gousses, la position des fleurs, et la taille de la plante. Chacun des sept caractères a une forme dominante et une forme récessive.

6 sept caractères dominant/récessif

sept caractères dominant/récessif
Mendel, Versuche über Pflanzen-Hybriden (1865)



Sur cette image vous voyez les sept caractères, avec la version dominante à gauche, récessive à droite. La forme des graines est ronde (dominant) ou ridée (récessif), la couleur des graines est jaune (dominant) ou verte (récessif), la couleur des fleurs est pourpre (dominant) ou blanc (récessif) etc.

7 de vraies données !

Quand j'ai découvert le contenu de cet article, je me suis dit : mais c'est génial ! J'enseigne la statistique à des biologistes, je fais de mon mieux pour rendre la discipline crédible, et voilà que ce monument de la littérature scientifique, l'article fondateur de la génétique, contient une profusion de données expérimentales, de quoi faire une quantité de tests statistiques. Et, cerise sur le gâteau, ces tests sont extrêmement faciles à expliquer et à réaliser.

Alors je me suis mis à préparer mon cours.

de vraies données !

Mendel, Versuche über Pflanzen-Hybriden (1865)

1. Versuch.			2. Versuch.	
Gestalt der Samen.			Färbung des Albumens.	
Pflanze	rund	kantig	gelb	grün
1	45	12	25	11
2	27	8	32	7
3	24	7	14	5
4	19	10	70	27
5	32	11	24	13
6	26	6	20	6
7	88	24	32	13
8	22	10	44	9
9	28	6	50	14
10	25	7	44	18

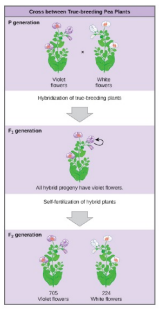
8 génération $[F_2]$

Que dit la théorie de Mendel ?

Prenons l'exemple de la couleur des fleurs : pourpre dominant, blanc récessif. Si on croise une plante à fleurs pourpres avec une plante à fleurs blanches, à la première génération que Mendel note F1, on n'obtient que des fleurs pourpres. Mais à la génération suivante, notée F2, on doit obtenir théoriquement trois quarts de fleurs pourpres, un quart de fleurs blanches. Or les chiffres donnés par Mendel sont 705 fleurs pourpres, 224 fleurs blanches, pour un total de 929 plantes. Est-ce que 705 sur 929 est significativement différent de trois quarts ?

génération $[F_2]$

Mendel, Versuche über Pflanzen-Hybriden (1865)



9 génération $[F_2]$

Il suffit de demander à R ! on rentre x=705, parmi n=929, pour une hypothèse nulle de p=3/4, dans la fonction prop.test, et la réponse tombe. 0.75 est bien dans l'intervalle de confiance, la p-valeur vaut 0.5571 ce qui est beaucoup plus grand que le seuil de 5 pour cent, donc tout va bien : l'observation 705 sur 929 est bien compatible avec la théorie 3/4.

Et puisqu'il y a 7 caractères, tant vaut faire le calcul pour tous.

génération $[F_2]$

test de proportion

```
> prop.test(x=705,n=929,p=3/4)

1-sample proportions test with continuity correction

data:  705 out of 929, null probability 3/4
X-squared = 0.34482, df = 1, p-value = 0.5571
alternative hypothesis: true p is not equal to 0.75
95 percent confidence interval:
 0.7297867 0.7858006
sample estimates:
       p
0.7588805
```

Conclusion : 705/929 compatible avec $p = 3/4$

10 génération $[F_2]$

Voici les données pour les 7 caractères, avec le nombre de dominants dans la première colonne, le nombre total d'observations dans la seconde, et la p-valeur du test dans la troisième.

Quand j'ai vu ces résultats, je me suis dit : ah là j'ai un problème ! Il faut que j'enseigne que la p-valeur sous l'hypothèse nulle, c'est-à-dire si Mendel a raison, est uniformément distribuée entre 0 et 1. C'est-à-dire qu'une fois sur deux elle doit être plus petite que un demi. Or regardez la colonne de droite. La plus petite valeur est 0.457. Si j'avais tiré des nombres au hasard entre 0 et 1 je n'aurais eu qu'une chance sur cent de n'obtenir que des valeurs aussi grandes. C'est un peu bizarre.

Alors comme j'ai mauvais esprit, j'ai pris tous les résultats expérimentaux donnés par Mendel dans son article, et j'ai calculé toutes les p-valeurs possibles. J'en ai trouvé 66.

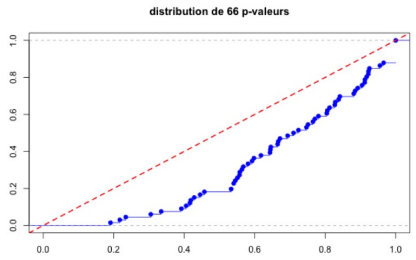
génération $[F_2]$
test de proportion

dominant	total	proportion	p-valeur
5474	7324	0.747	0.618
6022	8020	0.751	0.913
705	929	0.759	0.557
882	1181	0.747	0.827
428	580	0.738	0.533
651	858	0.759	0.581
787	1064	0.740	0.457
14949	19959	0.749	0.747

11 distribution de 66 p-valeurs

Voici leur fonction de répartition. La théorie dit que les points bleus devraient être proches de la ligne pointillée rouge. Comme vous le constatez, ce n'est pas vraiment le cas.

distribution de 66 p-valeurs
Mendel, Versuche über Pflanzen-Hybriden (1865)



12 distribution de 66 p-valeurs

Ces 66 p-valeurs qui devraient être uniformément réparties entre 0 et 1, en fait la plus petite est 0.19, en moyenne elles valent 0.70. Il devrait y en avoir 50% plus grandes que 0.5, il y en a 82%. Il devrait y en avoir 10% plus grandes que 0.9, il y en a 26%. Le test de comparaison des distributions donne une p-valeur inférieure à 10^{-7}. Non décidément, les p-valeurs ne sont pas uniformément distribuées.

Elles sont trop grandes, les résultats expérimentaux de Mendel sont trop proches de sa théorie, trop beaux pour être vrais. À ce stade là j'étais bien embêté. Alors j'ai cherché à en savoir plus sur l'article. Et vous imaginez bien que, si je me permets de vous raconter cette histoire, c'est que d'autres plus autorisés que moi avaient déjà fait les mêmes constatations, bien avant moi.

Est-il possible que les données aient été inventées ou trafiquées ? Je ne le crois pas. Simplement, Mendel n'était pas dans une démarche de vérification statistique. On en trouve des indications dans l'article.

distribution de 66 p-valeurs
Mendel, Versuche über Pflanzen-Hybriden (1865)

- minimum : 0.19
- moyenne : 0.70
- proportion > 0.5 : 82%
- proportion > 0.9 : 26%
- `ks.test(pp,punif,alternative="less")`
  p-valeur : 8.4×10^{-8}

13 pourquoi ?

« L'Expérience 5, qui fait apparaître l'écart le plus grand, a été répétée, et alors, au lieu du rapport de 60 à 40, celui de 65 à 35 a été obtenu. »

C'est un exemple où la théorie doit donner un rapport deux tiers un tiers. Il se trouve que 60 contre 40 n'est pas significativement différent de ce rapport, mais il n'est pas assez proche au goût de Mendel. Alors il recommence pour trouver 65 contre 35 qui est presque parfait.

Voici un autre exemple.

« Les rapports 10, 19, 43, 78 sont si proches des rapports 10, 20, 40, 80, ou 1, 2, 4, 8, que ces derniers représentent indubitablement la vraie valeur. »

Clairement, Mendel ne s'attend pas à des fluctuations statistiques, mais à une adéquation la plus parfaite possible avec sa théorie.

pourquoi ?
Mendel, Versuche über Pflanzen-Hybriden (1865)

L'Expérience 5, qui fait apparaître l'écart le plus grand, a été répétée, et alors, au lieu du rapport de 60 à 40, celui de 65 à 35 a été obtenu.

[...]

Les rapports 10, 19, 43, 78 sont si proches des rapports 10, 20, 40, 80, ou 1, 2, 4, 8, que ces derniers représentent indubitablement la vraie valeur.

14 Karl Pearson (1857–1936)

Il se trouve que le moment où l'article de Mendel a été redécouvert, c'est-à-dire juste avant 1900, était un grand moment pour la statistique anglaise. Karl Pearson régnait sans partage sur la biométrie.

Karl Pearson (1857–1936)

15 On the Criterion that a given System of Deviations (1900)

Et justement en 1900, au moment où l'article de Mendel était peut-être le plus discuté, Pearson publie un article. Vous voyez un titre à rallonge qui ne dit pas du tout de quoi il s'agit. C'est la généralisation du test de proportion à des données multinomiales, en clair, le test du chi-deux. Précisément ce qu'il faut faire pour vérifier les affirmations de Mendel.

On y trouve le théorème principal, y compris avec la notation chi-deux, on y trouve même un algorithme de calcul de la loi du chi-deux, et on y trouve aussi des exemples.

On the Criterion that a given System of Deviations (1900)
Karl Pearson (1857–1936)

X. *On the Criterion that a given System of Deviations from the Probable in the Case of a Correlated System of Variables is such that it can be reasonably supposed to have arisen from Random Sampling.* By KARL PEARSON, *F.R.S., University College, London*.*

THE object of this paper is to investigate a criterion of the probability on any theory of an observed system of errors, and to apply it to the determination of goodness of fit in the case of frequency curves.

16 On the Criterion that a given System of Deviations (1900)

Admirez quand même le recueil de ces données-là : un lancer de douze dés a été effectué 26 306 fois. Ce qui a été relevé c'est le nombre de dés parmi les douze qui étaient à cinq ou six.

On the Criterion that a given System of Deviations (1900)
Karl Pearson (1857–1936)

(6) *Frequency known or supposed known a priori.*
Illustration I.
The following data are due to Professor W. F. R. Weldon, F.R.S., and give the observed frequency of dice with 5 or 6 points when a cast of twelve dice was made 26,306 times :—

No. of Dice in Cast with 5 or 6 Points.	Observed Frequency, m'.	Theoretical Frequency, m.	Deviation, e.
0	185	203	− 18
1	1149	1217	− 68
2	3265	3345	− 80
3	5475	5576	− 101
4	6114	6273	− 159
5	5194	5018	+176
6	3067	2927	+140
7	1331	1254	+ 77
8	403	392	+ 11
9	105	87	+ 18
10	14	13	+ 1
11	4	1	+ 3
12	0	0	0
	26306	26306	

17 Walter Frank Raphael Weldon (1860–1906)

Comme il est dit au-dessus, le courageux lanceur de dés était Frank Weldon. A l'origine c'était un zoologiste, converti à la biométrie par sa collaboration avec Pearson. Ce n'est pas pour diminuer son mérite de lanceur, mais il a réalisé une grande partie de ses travaux en collaboration avec son épouse, qui est rarement citée. J'ignore quel rôle elle a joué dans les 26 306 lancers.

Cela n'enlève rien à l'appréciation de Pearson, qui a dit : « Weldon était par nature un poète, et ce sont eux qui donnent le meilleur à la science, car ils donnent des idées ».

Walter Frank Raphael Weldon (1860–1906)

18 Mendel's laws of alternative inheritance in Peas

En 1901, Weldon saisit l'occasion du test du chi-deux de son ami Pearson pour écrire un article sur les lois de Mendel. Il y reprend les arguments de Mendel, mais en plus il applique aux résultats expérimentaux, des tests statistiques.

Mendel's laws of alternative inheritance in Peas
Raphael Weldon (1860–1906)

MENDEL'S LAWS OF ALTERNATIVE INHERITANCE IN PEAS.

By W. F. R. WELDON.

[Received December 9, 1901.]

If two plants, or two animals, of different characters be allowed to breed together, the parental characters may affect the offspring in any one of three ways. In the most usual case, the characters by which the parents differ may appear so intimately blended in the offspring that each young animal or plant appears intermediate in character between its parents, but it is not generally possible for us at present to resolve its body into separate elements, some of which resemble one parent, and some the other. In other cases, however, the body of the young is easily divisible into regions, in some of which the character of one parent is presented in a recognisable and often apparently unaltered state, while the rest of the body presents a similar resemblance to the other parent. In the cases of the third class the body of the offspring may entirely resemble that of one parent, the characters of the other being apparently unrepresented. While it is perfectly possible and indeed probable that the difference between these three forms of inheritance is only one of degree, it is still convenient to discuss them separately.

19 such test as may be possible

« Il est clairement important de tester ces remarquables affirmations par une étude soigneuse des résultats numériques, et en appliquant tous les tests possibles. Il me semble, que pour avoir négligé ces précautions, quelques auteurs ont ignoré la manière merveilleusement consistante, dont les résultats de Mendel s'accordent avec sa théorie. »

Mmhh : merveilleusement consistante hein ? Weldon fait les calculs du test de proportion et insiste.

such test as may be possible
Weldon, Mendel's laws of alternative inheritance (1902)

It is clearly important to test these remarkable statements by a careful study of the numerical results, and by the application of such tests as may be possible. It seems to me that by neglecting these precautions some writers have been led to overlook the wonderfully consistent way in which Mendel's results agree with his theory,...

20 These results accord so remarkably

« Un écart entre l'hypothèse et les observations supérieur ou égal à l'erreur probable, ne se produit que deux fois sur 7, et des écarts beaucoup plus grands que l'erreur probable ne se produisent jamais.

Ce résultats s'accordent si remarquablement bien avec le résumé que Mendel en fait, que s'ils étaient répétés une seconde fois, dans des conditions similaires, les chances que l'accord entre observations et hypothèses soit pire serait de 16 contre 1. »

Donc Mendel a eu de la chance sur les proportions simples. Passant ensuite aux résultats multinomiaux, Weldon dit :

These results accord so remarkably
Weldon, Mendel's laws of alternative inheritance (1902)

so that a discrepancy between the hypothesis and the observations which is equal or greater than the probable error occurs twiced out of seven times, and deviations much greater than the probable error do not occur at all. These results then accord so remarkably with Mendel's summary of them that if they were repeated a second time, under similar conditions and on a similar scale, the chance that the agreement between observations and hypothesis would be worse than that actually obtained is about 16 to 1.

21 the odds against a result as good as this

« En appliquant la méthode de Pearson, les chances d'obtenir des déviations supérieures à celles-ci sont de 0.95, en d'autres termes, si l'expérience était répétée 100 fois, on s'attend à un résultat pire 95 fois, ou les chances contre un résultat aussi bon que celui-ci sont de 20 contre 1. »

Remarquez comment Weldon enfonce le clou de trois façons différentes pour dire à quel point il est suspect d'avoir autant de bol. Il n'ose pas aller plus loin.

the odds against a result as good as this
Weldon, Mendel's laws of alternative inheritance (1902)

Applying the method of Pearson, the chance that a system will exhibit deviations as great as or greater than these from the result indicated by Mendel's hypothesis is about 0.95, or if the experiment were repeated a hundred times, we should expect to get a worse result about 95 times, or the odds against a result as good as this or better are 20 to 1.

22 suggest fruitful lines of enquiry

« En essayant de résumer les preuves sur lesquelles repose mon opinion, je n'ai aucune intention de minimiser l'importance de l'exploit de Mendel. Je souhaite simplement attirer l'attention sur une série de faits, qui me semblent suggérer de fructueuses pistes de recherche. »

Weldon avait des raisons de se montrer aussi prudent.

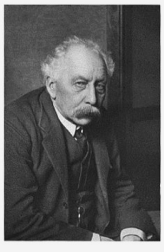
suggest fruitful lines of enquiry
Weldon, Mendel's laws of alternative inheritance (1902)

In trying to summarise the evidence on which my opinion rests, I have no wish to belittle the importance of Mendel's achievement. I wish simply to call attention to a series of facts which seem to me to suggest fruitful lines of enquiry.

23 William Bateson (1861–1926)

Il savait quelle allait être la réaction de Bateson. Ils avaient été étudiants ensemble à Cambridge, et leur inimitié datait de ce moment là. Bateson est un de ceux qui ont redécouvert l'article de Mendel. Il en avait fait en quelque sorte son cheval de bataille dans sa croisade anti-darwinienne. Pourtant Mendel et Darwin ne se connaissaient pas, et d'ailleurs leurs théories ne se contredisent pas. Mais pour Bateson, les gènes vus comme des unités discrètes par Mendel s'opposent à une évolution continue telle qu'il perçoit la théorie de Darwin. Logiquement, Bateson s'oppose aussi au cousin de Darwin, Galton et par conséquent au fils spirituel de Galton, Pearson.

William Bateson (1861–1926)



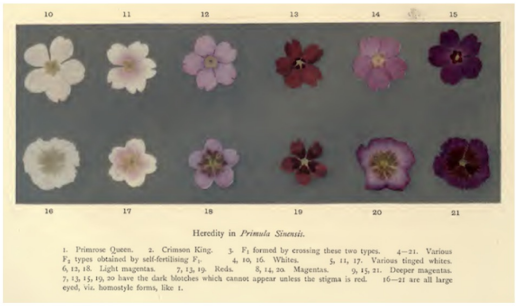
24 Mendel's principles of heredity (1909)

Mendel's principles of heredity (1909)
William Bateson (1861–1926)

Bateson écrit en 1902, et réédite en 1909, un ouvrage de vulgarisation des « principes mendéliens de l'hérédité ». Weldon est mort entre-temps, en 1906. L'ouvrage donne non seulement une biographie de Mendel et une traduction de son article, mais aussi de magnifiques illustrations en couleurs des pois de Mendel bien sûr,. . .

25 Mendel's principles of heredity (1909)

Mendel's principles of heredity (1909)
William Bateson (1861–1926)



mais aussi d'autres espèces qui montrent des caractères phénotypiques s'accordant parfaitement avec la théorie de Mendel, comme sur cette image, la primevère chinoise.

Le ton de la préface est très polémique.

26 to repell the attack

to repell the attack
Bateson, Mendel's principles of heredity (1909)

With modifications I published it separately in 1902, giving a brief summary of Mendelism as then developed, under the title *Mendel's Principles of Heredity : a Defence*. The object of that publication was to put Mendel's work before the English speaking peoples and to repell the attack which the late Professor Weldon had recently made on Mendelian methods and the conclusions drawn from them.

« Avec quelques modifications je le publiai séparément en 1902, donnant un bref résumé du mendélisme tel qu'on le voyait, sous le titre : « Les principes de l'hérédité de Mendel : une défense ». L'objet de cette publication était de présenter le travail de Mendel au public anglophone, et de repousser l'attaque que le regretté professeur Weldon avait récemment menée, sur les méthodes mendéliennes et leurs conclusions. »

Bateson en profite pour régler ses comptes avec Galton.

27 Galton's method failed for want of analysis

Galton's method failed for want of analysis
Bateson, Mendel's principles of heredity (1909)

Meanwhile we may note that though there was admittedly a statistical accord between Galton's theory and some facts of heredity, yet no one familiar with breeding or even with the literature of breeding could possibly accept that theory as a literal of adequate presentation of facts.

[. . .]

As we now know, Galton's method failed for want of analysis. His formula should in all probability be looked upon rather as an occasional consequence of the actual laws of heredity than in any proper sense one of those laws.

« Cependant, nous notons, tout en admettant qu'il y ait un accord statistique entre la théorie de Galton et certains faits de l'hérédité, que personne qui serait familier avec l'élevage ou même seulement sa littérature, ne pourrait accepter que cette théorie soit une représentation correcte des faits.

Comme nous le savons maintenant, la méthode de Galton a échoué faute d'analyse. Sa formule devrait être regardée plutôt comme une conséquence fortuite des lois de l'hérédité, que comme l'une de ces lois. »

28 the so-called investigations of heredity

« Sur les soi-disant recherches sur l'hérédité poursuivies par des extensions des méthodes non-analytiques de Galton, et promues par le professeur Pearson et l'école biométricienne anglaise, il n'y a pratiquement rien à dire. Que ce travail puisse dans le futur contribuer au développement de la théorie statistique ne peut pas être nié, mais dans l'application aux problèmes de l'hérédité, cet effort n'aboutit qu'à masquer un ordre, qu'il était pourtant destiné à révéler. »

the so-called investigations of heredity
Bateson, Mendel's principles of heredity (1909)

Of the so-called investigations of heredity pursued by extensions of Galton's non-analytical method and promoted by Professor Pearson and the English Biometrical school it is now scarcely necessary to speak. That such work may ultimately contribute to the development of statistical theory cannot be denied, but as applied to the problems of heredity the effort has resulted only in the concealment of that order which it was ostensibly undertaken to reveal.

29 work so unsound in construction

« Pour ceux qui dans l'avenir étudieraient cet épisode dans l'histoire de la science biologique, il apparaîtra inexplicable qu'un travail si peu solide dans sa construction, ait été reçu avec respect par le monde scientifique.

Avec la découverte de la ségrégation, il est devenu évident que les méthodes qui se dispensent d'une analyse individuelle du matériau sont inutiles. La seule alternative ouverte aux utilisateurs de ces méthodes était soit d'abandonner leurs illusions, soit de nier la vérité du mendelisme. Mais avec le passage du temps, le nombre de personnes qui ont eux-mêmes témoigné des phénomènes a augmenté de telle sorte que ces dénégations ont cessé d'être dangereuses et peuvent être considérées comme purement formelles. »

work so unsound in construction
Bateson, Mendel's principles of heredity (1909)

To those who hereafter may study this episode in the history of biological science it will appear inexplicable that work so unsound in construction should have been respectfully received by the scientific world. With the discovery of seggregation it became obvious that methods dispensing with individual analysis of the material are useless. The only alternatives open to the inventors of those methods were either to abandon their delusion or to deny the truth of Mendelism, but with the lapse of time the number of persons who have themselves witnessed the phenomena has increased so much that these denials have lost their dangerous character and may be regarded as merely formal.

30 Ronald Aylmer Fisher (1890–1962)

Weldon étant décédé, personne ne juge utile de répondre à Bateson, et la polémique se calme pour un temps. Jusqu'à l'entrée en scène de Fisher. Le problème avec Fisher est qu'on ne peut ni le traiter comme un mathématicien qui n'entend rien à la biologie, ni comme un généticien qui n'entendrait rien aux statistiques.

Il se trouve être à la fois le fondateur de la statistique mathématique, avec la théorie du maximum de vraisemblance, et le fondateur de la génétique théorique.

Ronald Aylmer Fisher (1890–1962)

31 The genetical theory of natural selection (1930)

Ce livre en particulier, « la Théorie génétique de la sélection naturelle », a été salué dès sa parution comme une avancée majeure. Fisher ne s'y oppose pas du tout à la théorie de Mendel, bien au contraire.

The genetical theory of natural selection (1930)
Ronald Aylmer Fisher (1890–1962)

THE

GENETICAL THEORY OF NATURAL SELECTION

BY

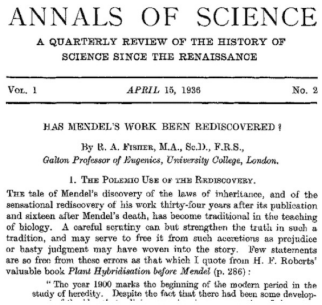
R. A. FISHER, Sc.D., F.R.S.

32 Has Mendel's work been rediscovered ? (1936)

Alors quand Fisher, six ans après son livre, publie cet article, il est lu avec attention autant par les statisticiens que par les biologistes. La revue Annals of Science, est une revue d'histoire des sciences. Le titre est une question : « Est-ce qu'on a vraiment redécouvert le travail de Mendel ? »

L'article commence fort. Regardez le titre du premier chapitre : « l'usage polémique de la redécouverte ». Fisher y reprend l'historique des publications. En particulier il décrit le rôle de Bateson.

Has Mendel's work been rediscovered ? (1936)
Ronald Aylmer Fisher (1890–1962)



ANNALS OF SCIENCE

A QUARTERLY REVIEW OF THE HISTORY OF SCIENCE SINCE THE RENAISSANCE

VOL. 1 — APRIL 15, 1936 — No. 2

HAS MENDEL'S WORK BEEN REDISCOVERED ?

By R. A. FISHER, M.A., Sc.D., F.R.S., *Galton Professor of Eugenics, University College, London.*

1. THE POLEMIC USE OF THE REDISCOVERY.

THE tale of Mendel's discovery of the laws of inheritance, and of the sensational rediscovery of his work thirty-four years after its publication and sixteen after Mendel's death, has become traditional in the teaching of biology. A careful scrutiny can but strengthen the truth in such a tradition, and may serve to free it from such accretions as prejudice or hasty judgment may have woven into the story. Few statements are so free from these errors as that which I quote from H. F. Roberts' valuable book *Plant Hybridisation before Mendel* (p. 286) :

"The year 1900 marks the beginning of the modern period in the study of heredity. Despite the fact that there had been some develop-

33 the interest of a zealous partisan

« On ne peut pas nier que l'intérêt de Bateson dans la redécouverte était celui d'un fervent partisan. On doit lui attribuer deux éléments dans la légende, qui ne semblent avoir aucun autre fondement : premièrement la croyance que l'influence de Darwin est responsable du manque d'attention pour le travail de Mendel, et deuxièmement la croyance que Mendel était hostile aux théories de Darwin, et s'imaginait que son travail les contredisait. »

Plutôt qu'à Bateson dont les arguments relèvent plus de la croyance que de la science, Fisher s'intéresse surtout à Mendel et à ses rapports avec la statistique.

the interest of a zealous partisan
Fisher, Has Mendel's work been rediscovered (1936)

It cannot be denied that Bateson's interest in the rediscovery was that of a zealous partisan. We must ascribe to him two elements in the legend which seem to have no other foundation : (1) The belief that Darwin's influence was responsible for the neglect of Mendel's work, and of all experimentation with similar aims ; and (2) the belief that Mendel was hostile to Darwin's theories, and fancied that his work controverted them.

34 had the matter caused him any anxiety

« Les études mathématiques de Mendel à Vienne, ont pu n'accorder que peu d'attention à la théorie des probabilités. Mais nous savons qu'il avait engagé des recherches à caractère statistique, par exemple en météorologie, en relation avec les tâches solaires. De sorte qu'il est à peine concevable, si le sujet lui avait causé le moindre souci, qu'il n'ait pas connu un livre ou un ami qui lui aurait permis d'examiner objectivement si oui ou non les écarts observés par rapport aux attentes, étaient conformes aux lois du hasard. »

Fisher remarque un détail extrêmement gênant.

had the matter caused him any anxiety
Fisher, Has Mendel's work been rediscovered (1936)

Mendel's mathematical studies in Vienna may have given little attention to the theory of probability ; but we know that he was engaged in other researches of a statistical character, in meteorology, and in connection with sun-spots, so that it is scarcely conceivable, had the matter caused him any anxiety, that he knew of no book or friend that would enable him to examine objectively whether or not the observed deviations from expectation conformed with the laws of chance.

35 the data have been falsified

« Une incohérence grave, et quasiment inexplicable est cependant apparue, dans le fait que, dans l'une des séries de résultats, les nombres observés sont parfaitement conformes au rapport de 2 à 1, ce que Mendel lui-même attendait, mais diffèrent notablement de ce qui aurait dû être attendu si sa théorie avait été corrigée pour la faible taille de ses lignées.

Bien qu'aucune explication satisfaisante ne puisse être espérée, il reste une possibilité parmi d'autres qui est que Mendel ait été trompé par un des ses assistants qui ne savait que trop bien ce qui était attendu. Cette possibilité est corroborée par des indices indépendants qui montrent que les données de la plupart si ce n'est toutes les expériences ont été falsifiées de manière à s'accorder de près avec les attentes de Mendel. »

Même si Fisher faisait tout ce qu'il pouvait pour exonérer Mendel lui-même de ses responsabilités, sa conclusion était nette et sans appel.

Le dialogue de sourds entre statisticiens et généticiens a duré au moins jusqu'au centenaire de l'article de Mendel, qui a été fêté en grande pompe et à grand renfort de publications, en 1965. En particulier, une histoire de la génétique a été écrite par Sturtevant, qui tente de réfuter l'argument statistique.

the data have been falsified
Fisher, Has Mendel's work been rediscovered (1936)

A serious and almost inexplicable discrepancy has, however, appeared, in that in one series of results the numbers oberved agree excellently with the two to one ratio, which Mendel himself expected, but differ significantly from what should have been expected had his theory been corrected for the small size of his test progenies. [...] Although no explanation can be expected to be satisfactory, it remains a possibility among others that Mendel was deceived by some assistant who knew too well what was expected. This possibility is supported by independent evidence that the data of most, if not all of the experiments have been falsified so as to agree closely with Mendel's expectations.

36 a statistician will raise his eyebrows

« Il est possible de questionner l'argument selon lequel un accord est trop proche de ce qui est attendu. Si je raconte que j'ai lancé 1000 pièces et obtenu 500 piles et 500 faces, un statisticien froncera les sourcils, bien que ce résultat soit le plus probable de tous. Si je dis que j'ai obtenu 480 piles et 520 faces, le statisticien dira que c'est à peu près ce à quoi il s'attendait, bien que ce résultat soit moins probable que 500 pile 500 face.

Maintenant, si je raconte que j'ai lancé 1000 pièces dix fois et obtenu chaque fois 500 pile et 500 face, le statisticien dira que je suis un menteur, bien que ce résultat soit encore une fois le plus probable. Cet argument est dangereusement près de dire qu'aucune expérience ne peut être observée, parce que chaque résultat particulier n'a qu'une probabilité négligeable de se produire. »

Sturtevant cependant, est bien obligé de reconnaître la validité de l'analyse de Fisher. Oui les données de Mendel ont été arrangées.

a statistician will raise his eyebrows
Sturtevant, History of Genetics (1965)

[...] it is possible to question the logic of the argument that a fit is too close to expectation. If I report that I tossed 1000 coins and got exactly 500 heads and 500 tails, a statistician will raise his eyebrows, though this is the most probable exactly specified result. If I report 480 heads and 520 tails, the statistician will say that is about what one would expect – though this result is less probable than the 500:500 one. [...] If I now report that I tossed 1000 coins ten times, and got 500:500 every time, our statistician will surely conclude that I am lying, though this is the most probable result thus exactly specified. The argument comes perilously close to saying that no such experiment can be carried out, since every single exactly specified result has a vanishingly small probability of occuring.

37 how the data came to be as they are

« Il reste la question de savoir comment les données en sont venues à être ce qu'elles sont. Il existe au moins trois possibilités.

1. Il peut y avoir une tendance inconsciente à classer des individus douteux de manière à aller dans le sens attendu.

2. Il peut y avoir des familles qui semblaient aberrantes, et qui ont été omises comme probablement dues à une erreur d'expérience.

3. Quelques uns des comptes peuvent avoir été faits par des étudiants ou des assistants qui étaient conscients de ses attentes, et voulaient lui faire plaisir.

Peut-être la meilleure réponse – avec laquelle Fisher aurait été d'accord – est que, après tout, Mendel avait raison ! »

how the data came to be as they are
Sturtevant, History of Genetics (1965)

There remains the question of how the data came to be as they are. There are at least three possibilities :

1. There may have been an unconscious tendency to classify somewhat doubtful individuals in such a way as to fit the expectation.
2. There may have been some families that seemed aberrant, and that were omitted as being probably due to experimental error.
3. Some of the counts may have been made for him by students or assistants who were aware of his expectations, and wanted to please him.

Perhaps the best answer – with which Fisher would have agreed – is that, after all, Mendel was right !

38 références

Oui, évidemment, Mendel avait raison. Mais moi, pour mes cours de stats, ça m'aurait quand même arrangé que ses données ne soient pas arrangées !

références

- W. Bateson (1909) *Mendel's principles of Heredity*, Cambridge University Press
- R. A. Fisher (1936) Has Mendel's work been rediscovered ? *Annals of Science*, 1, 115–137
- A. Franklin et al. (2008) *Ending the Mendel - Fisher controversy*, University of Pittsburg Press
- A. H. Sturtevant (1965) *A History of Genetics*, Cold Spring Harbor University Press