

zenzüchtung 33, (1954) 363. — 12. ROEMER: Roggen. Handbuch f. Pflanzenzüchtung. — 13. RUNDFELDT: Die Ausnutzung des Heterosiseffektes in der Maiszüchtung. Zeitschrift f. Pflanzenzüchtung 31, 226—260 (1952) — 14. WELLENSIEK: Neue Methoden der Roggenzüchtung Zaa-

zaad en Pootgoed 2, Nr. 7, 9—11, Nr. 8, 8—11 (1940). — 15. WELLENSIEK: Vegetative vermeerdering en plantenveredeling, speciaal bij rogge. Landbouwk. Tijdschr. 54, 422—436 (1952). — 16. WERNER: Aufzeichnungen u. Protokolle 1943 (unveröffentlicht).

(Aus dem Institut für Vererbungs- und Züchtungsforschung, Berlin-Dahlem)

Über die Methoden zur Untersuchung der Wirkungsweise quantitativer Gene

Von GÜNTER WRICKE*

Mit 1 Textabbildung

Einleitung

Die Analyse qualitativ vererbter Merkmale, wie z. B. die von Blütenmerkmalen im Pflanzenreich, Fellausfärbungen im Tierreich oder vielen morphologischen Eigenschaften der Organismen, bietet bei dem heutigen Stand der Vererbungswissenschaft keine grundsätzlichen Schwierigkeiten mehr. Wenn auch neben den einfachsten dominanten und intermediären Erbgängen durch das Zusammenwirken mehrerer Faktoren kompliziertere oder auch durch Manifestationsschwankungen variable Zahlenverhältnisse auftreten, so kann man bei solchen qualitativen Merkmalen in spaltenden Generationen doch immer bestimmte Gruppen oder Klassen bilden, in die sämtliche Individuen einzuordnen sind. Bei einer genügend großen Zahl von Nachkommen und der Möglichkeit der Prüfung späterer Generationen ist eine Klärung des Erbganges auch in diesen Fällen fast immer möglich.

Anders ist es dagegen bei den sogenannten „polygenen“ Merkmalen, zu denen auch die meisten wirtschaftlich wichtigen Eigenschaften der landwirtschaftlichen und vieler gärtnerischer Kulturpflanzen gehören. Ihre Ausprägung ist abhängig von einer Vielzahl von Genen, die als sogenannte „Polygene“ den alternative Eigenschaften bedingenden „Majorgenen“ gegenübergestellt werden (MATHER, 1949 a). Während die letzteren immer einen deutlich sichtbaren Effekt haben und in der Form der Subletal- und Letalfaktoren oft über Leben und Tod des Organismus entscheiden, ist die Wirkung der einzelnen Polygene nicht von so entscheidender Bedeutung für die Eigenschaftsausprägung oder gar das Fortbestehen eines Individuums. Im Gegenteil, die Glieder eines Polygensystems können sich in gewissem Sinne in ihrer Wirkung vertreten. Zwei durchaus verschiedene Genotypen können in bezug auf eine quantitative Eigenschaft leistungsmäßig gleich veranlagt sein, innerhalb gewisser Zufallsschwankungen phänotypisch also die gleiche Merkmalsausprägung zeigen. Die Wirkung der einzelnen Polygene ist gering. In ihrer Gesamtheit vermögen sie jedoch beträchtliche Unterschiede in der Merkmalsausprägung auszuüben, da ihre Zahl eben groß ist. Die Anzahl der Genotypen in der F_2 -Generation nach Kreuzung zweier in bezug auf solche quantitativen Eigenschaften unterschiedlicher homozygoter Elterrrassen ist daher auch sehr groß. Theoretisch sind bei vollkommen freier Spaltung 3^n verschiedene Genotypen zu erwarten, wenn n die Zahl der spaltenden Gene ist. Bei intermediärer Merkmalsausprägung findet man

ebensoviele Phänotypen, bei voller Dominanz verringert sich die Zahl auf 2^n , sie ist aber immer noch auch bei nur kleinen Zahlen spaltender Gene recht beträchtlich. Die phänotypischen Differenzen zwischen zwei Genotypen, die sich nur in einem oder wenigen Leistungsgenen unterscheiden, werden daher auch sehr klein sein, und schon geringe umweltbedingte Schwankungen können sie leicht verwischen. Nun sind die meisten quantitativen Eigenschaften sehr stark von Umweltverhältnissen abhängig. Man findet daher in der Tat bei der Untersuchung solcher Eigenschaften nach Kreuzung zweier homozygoter Linien in der F_2 nicht eine entsprechend der Zahl der spaltenden Gene mehr oder weniger große Zahl von gut zu trennenden Klassen, die den einzelnen Genotypen entsprechen, sondern eine kontinuierliche Verteilung. Die Grenzen zwischen den einzelnen Genotypen sind durch die umweltbedingte Variation vollkommen verwischt, so daß zwei Pflanzen mit dem gleichen Meßwert durchaus nicht vom gleichen Genotyp zu sein brauchen oder genotypisch gleiche Pflanzen verschiedene Meßwerte zeigen können. Eine Einteilung in Klassen ist zwar rein theoretisch möglich, aber genetisch nicht sinnvoll. Methoden, die bei der Genanalyse qualitativ vererbter Merkmale zum Ziel führen, sind also bei quantitativen Eigenschaften nicht anwendbar, weil die Wirkung eines einzelnen Polygens nicht getrennt von der der anderen Polygene analysiert werden kann. Kontinuierliche Verteilung, wie man sie bei der Untersuchung quantitativer Merkmale findet, kann man nur durch statistische Maßzahlen analysieren, die für derartige Häufigkeitsverteilungen herausgearbeitet worden sind, also z. B. durch Mittelwerte, Varianzen, Kovarianzen oder auch statistische Maßzahlen höheren Grades. Das bedeutet aber gegenüber den klassischen mendelistischen Methoden zur Untersuchung der Erbvorgänge alternativer Merkmale eine ganz andere Betrachtungsweise und für den die Vererbung solcher quantitativen Eigenschaften Untersuchenden selbst eine gewisse Umstellung. Eine statistische Analyse der Wirkungsweise der Polygene führt eben nur zur Schätzung der durchschnittlichen Eigenschaften der Erbeinheiten und kann immer nur innerhalb gewisser Fehlergrenzen angegeben werden. Für die Anwendung dieser statistischen Methoden müssen daher oft stark einschränkende Voraussetzungen gemacht werden, deren Gültigkeit in vielen Fällen sehr schwer nachzuprüfen ist. In der englischsprachigen Literatur sind bereits mehrere derartige Methoden entwickelt worden, um drei sehr wichtige Eigenschaften von Polygenen zu untersuchen: nämlich ihr Zusammenwir-

* Herrn Prof. H. KAPPERT zum 65. Geburtstag gewidmet

ken, ihre Dominanzverhältnisse und ihre Zahl. Da diese Methoden in der deutschen Fachliteratur bisher wenig bekannt geworden sind, soll hier eine kurze Zusammenfassung der wichtigsten von ihnen gegeben werden. Die vorliegende Arbeit kann aber bei der Fülle der bereits vorhandenen Literatur in dem vorliegenden Umfang nicht als eine ausführliche Darstellung und eine eingehende Diskussion angesehen werden. Auf die Ableitung vieler Formeln muß daher verzichtet werden. Für die Anwendung der hier darzustellenden Methoden wird daher in den meisten Fällen ein Studium der Originalliteratur unumgänglich sein. Wenn also die vorliegende Arbeit einerseits nicht als Sammelreferat gelten kann und andererseits oft nur Hinweise geben wird, so glaubt der Verfasser doch, daß sie manchem, vor allem in der praktischen Züchtung Tätigen, wertvolle Anregung geben kann.

Das Zusammenwirken der Polygene

Nähere Aussagen über das Zusammenwirken oder die Wechselwirkung der Polygene zu machen, stößt deshalb auf Schwierigkeiten, weil die Wirkung eines einzelnen Gens, wie schon oben ausgeführt wurde, nicht analysiert werden kann. Als Arbeitshypothese diene daher zunächst die von NILSSON-EHLE (1910 u. 1911) und EAST (1910) geäußerte Ansicht, daß die Wirkung der eine kontinuierliche Variation bedingenden Gene additiv intermediär ist, das heißt, jedes Gen bewirkt, daß der Wert eines Genotyps um einen bestimmten Betrag erhöht wird, und daß die Homozygoten eine doppelt so große Wirkung haben wie die Heterozygoten. Jeder Faktor wirkt also unabhängig von allen übrigen Genen auf die Eigenschaftsausprägung.

Die experimentellen Ergebnisse einiger älterer Arbeiten stimmten auch gut mit dieser Annahme überein. Den überzeugendsten Beweis dafür, daß Gene eine derartige intermediäre, additive Wirkung haben können, brachte NILSSON-EHLE (1909 u. 1911) in seinen bekannten Untersuchungen über die Kornfarbe des Weizens und die Spelzenfarbe des Hafers. Die Zahl der beteiligten Gene war hier noch so klein (2 bzw. 3), daß die Wirkung der einzelnen Faktoren phänotypisch noch erkennbar war. Daher möchten DARLINGTON und MATHER (1950) diese Gene auch noch nicht als Polygene bezeichnen, obwohl sie die für diese geforderten anderen beiden Eigenschaften der ähnlichen und kumulativen Wirkung besitzen. Der Unterschied zu den tatsächlichen Polygenen, wie sie DARLINGTON und MATHER (1950) verstehen, ist aber, wie man ohne weiteres zugeben muß, nur ein gradueller. Die Wirkung dieser ist eben schon zu klein gegenüber der durch die Umwelt bedingten Variabilität, so daß man sie nur noch „en masse“ mit Hilfe statistischer Methoden und nicht mehr einzeln studieren kann.

Nicht viel später, nach den Arbeiten EASTS und NILSSON-EHLES, fand man jedoch, daß andere Ergebnisse besser erklärt werden konnten, wenn man eine multiplikative (oder geometrische) Wirkung der Polygene zugrundelegte. Als Beispiel für diese Art der Wirkung soll hier nur auf die Arbeit von MACARTHUR und BUTLER (1939) über die Vererbung des Fruchtgewichts bei Tomaten hingewiesen werden. An einer großen Zahl von Kreuzungen zwischen groß- und kleinfrüchtigen Tomaten wird gezeigt, daß der Wert der F_1 in den meisten Fällen mehr dem geometrischen Mittel

der Fruchtgewichte der beiden Eltern als dem des arithmetischen Mittels gleicht. Auch der Wert der F_2 liegt meistens näher dem des Elters mit geringerem Ertrag. Diese beiden Tatsachen und auch die positiv schiefe Verteilung der F_2 -Nachkommenschaften sprechen nun für eine multiplikative Genwirkung. Das soll an dem folgenden Schema (Tab. 1a) gezeigt werden. Es werden insgesamt 3 Faktoren (*Aa, Bb, Cc*) angenommen. Der *aabbcc*-Genotyp hat den Wert 3. Jeder mit einem großen Buchstaben bezeichnete Faktor fügt nun nicht, wie bei additiver Genwirkung, einen bestimmten Betrag zu diesem Wert hinzu, sondern multipliziert ihn mit 2. Der Genotyp *Aabbcc* hat also den Wert $3 \cdot 2 = 6$. Ersetzt man in ihm ein *b*- durch ein *B*-Allel, so würde dieser Genotyp den Wert $3 \cdot 4 = 12$ erhalten. Denselben Wert hat aber auch *AAbbcc*, nämlich 3 (Grundwert) $\cdot 2 \cdot 2 = 12$. Der Wert des F_1 -Genotyps *AaBbCc* aus der Kreuzung *AABBCC* \times *aabbcc* ist demnach $3 \cdot 2 \cdot 2 \cdot 2 = 24$. Ein Vergleich mit den beiden Elterwerten von 192 und 3 (Tabb. 1a) zeigt, daß er dem Elter mit dem kleineren Wert näher liegt und gleich ihrem geometrischen Mittel ist. Ferner liegt der F_2 -Mittelwert von 34 näher am geometrischen Mittel ($\sqrt{3 \cdot 192} = 24$) als am arithmetischen Mittel (97,5), und

Tabelle 1a. Schema einer trifaktoriellen, multiplikativen Genwirkung.

	<i>AABBCC</i> \times <i>aabbcc</i>
	192 ($3 \cdot 2^6$) 3 (Grundwert)
F_1	<i>AaBbCc</i> 24 ($3 \cdot 2^3$)
F_2	Durchschnittlicher Wert = 34,17
	Arithmet. Mittel der Elternwerte = 97,5
	Geometrisch. „ „ „ = 24,0
	Schiefheitsziffer $S = 2,38$.

schließlich ist auch die F_2 -Verteilung positiv schief ($S = + 2,4$). Man findet aber ganz ähnliche Verhältnisse in gewissen Fällen additiver Wirkung der Gene, nämlich dann, wenn partielle Dominanz des Allels für den kleineren Wert der zu untersuchenden Eigenschaft (z. B. Fruchtgröße oder Pflanzenlänge) vorliegt. Das zeigt das in Tabelle 1b dargestellte Schema. Der voll dominante Genotyp hat den Wert 3, jedes rezessive Leistungsgen fügt im homozygoten Zustand den Wert 8 und im heterozygoten den Wert 1,5 hinzu. Alle drei für multiplikative Genwirkung genannten Kriterien sind hier annähernd erfüllt. So ist erstens der F_1 -Wert 7,5 nur wenig kleiner als das geometrische Mittel mit einem Wert von 9 ($\sqrt{3 \cdot 27}$). Eine derartig kleine Differenz wird im Experiment meistens innerhalb der Zufallsstreuung liegen. Zweitens liegt der F_2 -Wert näher am geometrischen als am arithmetischen Mittel und drittens ist die Verteilung der F_2 wieder positiv schief.

CHARLES und SMITH (1939) haben nun die Möglichkeit zur Unterscheidung zwischen diesen beiden Arten von Genwirkung eingehender untersucht. Nicht immer ist nämlich die Übereinstimmung zwischen einer multiplikativen und einer additiven Genwirkung mit Dominanz der „kleinen“ Werte so gut wie an dem eben abgeleiteten Beispiel. Sie hängt wesentlich von dem Dominanzgrad der Gene ab. Nimmt man z. B. in Tabelle 1 b vollkommene Dominanz bei sonst gleichen Verhältnissen an, so ist der Wert der F_1 , wie man sich leicht klarmachen kann, gleich dem Wert des dominanten Elters, also = 3. Die Differenz zum geometrischen Mittel ist hier beträchtlich. Es kann somit in

gewissen Fällen schon mit Hilfe des F_1 -Wertes eine Unterscheidung zwischen diesen beiden Genwirkungen getroffen werden. Weiterhin bietet der Durchschnittswert der F_2 -Generation hierfür in vielen Fällen eine Möglichkeit.

Erwartungsgemäß ist er bei multiplikativer Genwirkung nicht viel größer als der F_1 -Wert, bei additiver Genwirkung liegt er unabhängig von allen Dominanzverhältnissen genau zwischen dem F_1 -Wert und dem arithmetischen Mittel der beiden Elter-Werte. Die von CHARLES und SMITH angegebenen Formeln lauten bei multiplikativer Genwirkung

$$\bar{F}_2 = \bar{F}_1 \left[1 + \frac{1}{3^n} (\log \bar{P}_1 - \log \bar{P}_2)^2 \right], \quad (1)$$

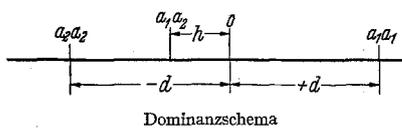
und bei additiver Wirkung

$$\bar{F}_2 = \frac{1}{4} (2\bar{F}_1 + \bar{P}_1 + \bar{P}_2), \quad (2)$$

wobei $\bar{F}_1, \bar{F}_2, \bar{P}_1$ u. \bar{P}_2 (CHARLES und SMITH benutzen hierfür die Symbole $\bar{v}_1, \bar{v}_2, \bar{v}_0$, und \bar{v}_0) die Mittelwerte der F_1, F_2 - und Parentalgenerationen sind, P_1 der Elter mit dem größeren Wert ist und n die Zahl der beteiligten Gene bedeutet.

Gleichung (1) ist eine Näherungsformel, die besonders dann gilt, wenn n , die Zahl der Gene, groß ist, und wenn die Wirkung der einzelnen Faktoren, wie es in dem Schema angenommen wurde, gleich ist. Formel (2) ist hiervon sowie von sämtlichen Dominanzverhältnissen unabhängig, wie im folgenden gezeigt werden wird.

Hat der Genotyp $a_1 a_1$ eines Genpaares $a_1 a_2$ den Wert $+d_a$, der Genotyp $a_2 a_2$ den Wert $-d_a$ und die Heterozygote $a_1 a_2$ den Wert h_a , wobei h_a positiv und negativ sein kann, so liegt der o-Punkt der Werteskala genau zwischen den Werten der beiden Homozygoten (nach MATHER, 1949a).



Entsprechend wären die Werte für die drei Genotypen eines zweiten Genpaares $+d_b$ für $b_1 b_1$, $-d_b$ für $b_2 b_2$ und h_b für $b_1 b_2$. Eine Kreuzung zwischen beiden Genotypen $a_1 a_1 b_1 b_1 \times a_2 a_2 b_2 b_2$ ergäbe in der F_1 den Genotyp $a_1 a_2 b_1 b_2$ mit einem Wert $h_a + h_b$ und eine F_2 , die folgende Zusammensetzung hat:

Genotyp	Hfk.	Wert
$a_1 a_1 b_1 b_1$	1/16	$+d_a + d_b$
$a_1 a_1 b_1 b_2$	2/16	$+d_a + h_b$
$a_1 a_2 b_1 b_1$	2/16	$+d_b + h_a$
$a_1 a_2 b_1 b_2$	4/16	$+h_a + h_b$
$a_1 a_1 b_2 b_2$	1/16	$+d_a - d_b$
$a_1 a_2 b_2 b_2$	2/16	$-d_b + h_a$
$a_2 a_2 b_1 b_1$	1/16	$-d_a + d_b$
$a_2 a_2 b_1 b_2$	2/16	$-d_a + h_b$
$a_2 a_2 b_2 b_2$	1/16	$-d_a - d_b$
S 16/16		$8/16h_a + 8/16h_b$

Neben dem Genotyp und seiner Häufigkeit steht der nach obiger Skala angenommene Wert. Als Durchschnitt für die F_2 ergibt sich nach dieser Zusammensetzung $\bar{F}_2 = 1/2 h_a + 1/2 h_b$. Dieser Wert soll nach

Tabelle 1b. Schema einer trifaktoriellen additiven Genwirkung mit Dominanz der „kleinen“ Werte

$AABBCC \times aabbcc$	
3 (Grundwert) 27	
F_1	$AaBbCc$
	7.5
F_2	Durchschnittlicher Wert = 11,25
	Arithmetisches Mittel der Elterwerte = 15
	Geometrisches „ „ „ = 9
	Schiefheitsziffer S = 1,45.

(2) gleich $1/4 (\bar{P}_1 + \bar{P}_2 + 2\bar{F}_1)$ sein. Setzt man die entsprechenden Elterwerte und den F_1 -Mittelwert in die Gleichung ein, so erhält man

$$1/4 (+d_a + d_b - d_a - d_b + 2h_a + 2h_b) = 1/2 h_a + 1/2 h_b$$

und damit den oben schon erhaltenen Durchschnittswert der F_2 . Man kann das hier dargestellte Schema auf eine beliebige Zahl von Genen ausdehnen und hat somit den Beweis für die Richtigkeit der Formel (2).

Neben dem Vergleich der Mittelwerte der F_1, F_2 - und Parentalgenerationen können auch die der Rückkreuzungsgenerationen zur Unterscheidung der additiven und multiplikativen Genwirkung herangezogen werden. Die von CHARLES und SMITH angegebenen Formeln lauten, wenn man die hier eingeführten Symbole und für den Durchschnittswert der Rückkreuzungsgeneration \bar{B} verwendet, bei geometrischer Genwirkung

$$\bar{B} = \sqrt{\bar{P} \cdot \bar{F}_2} \quad (3)$$

und bei additiver $\bar{B} = 1/2 (\bar{P} + \bar{F}_1)$. (4)

Für \bar{P} ist jeweils der Durchschnittswert des Elters einzusetzen, mit dem die F_1 zurückgekreuzt wurde.

Diese drei eben genannten Vergleiche zwischen den Durchschnittswerten der F_1, F_2 und der Rückkreuzungsgeneration und ihren Erwartungswerten bei additiver und multiplikativer Genwirkung ermöglichen in vielen Fällen schon eine Entscheidung darüber, ob die Gene mehr additiv oder multiplikativ wirken. CHARLES und SMITH (1939) geben weiterhin Ableitungen für zwei Koeffizienten, c („coefficient of genetic variability“) und t („coefficient of genetic third moment“), zu deren Berechnung auch die Varianzen und Schiefheitsmaße einer spaltenden Generation und die Varianzen der drei nichtspaltenden Generationen (F_1, P_1 und P_2) herangezogen werden. Zum näheren Studium dieser beiden Maße für die Beurteilung der Genwirkung soll hier nur auf die oben bereits mehrfach zitierte Arbeit dieser beiden Autoren verwiesen werden.

Hat man nun mit Hilfe der eben genannten Methoden festgestellt, daß die Wirkung der Faktoren weitgehend multiplikativ ist, so besteht die Möglichkeit, durch Umformung der Werte in ihre Logarithmen einen Maßstab zu erhalten, auf dem die Wirkung der Gene additiv erscheint. In dem Modellbeispiel der Tabelle 1a war für den Genotyp $AAbbcc$ ein Wert von 12, für $AABBcc$ ein Wert von 48 und für $AABBCc$ der Wert 192 festgelegt worden. Die Logarithmen der Zahlen 12, 48, 192 sind 1,07918, 1,68124 und 2,28330. Die Differenz zwischen 1,68124 und 1,07918 und zwischen 2,28330 und 1,68124 ist in beiden Fällen 0,60206. Die Abhängigkeit zwischen der Zahl der Gene und ihrer Wirkung ist also nach Verwendung einer logarithmischen Skala wieder rein linear.

Die Verteilungskurve der F_2 wäre in diesem Falle nicht mehr positiv schief, sondern symmetrisch. Durch die logarithmische Umformung wird also der rechte, obere Teil der x -Achse stärker kontrahiert als der untere.

Dem in Tab. 1a angenommenen Modellbeispiel wurde zugrunde gelegt, daß jedes mit einem großen Buchstaben bezeichnete Allel, das in einen Genotyp eingeführt wird, den Wert dieses Genotyps mit 2 multipliziert, unabhängig davon, ob das Allel bereits vorhanden ist oder nicht. Man könnte sich aber auch vorstellen, daß die Wirkung eines Allels größer oder kleiner ist, wenn der Genotyp bereits dasselbe Allel besitzt. Als Beispiel sollen folgende Zahlenreihen dienen:

Genotyp	aa	Aa	AA
Wert	2	4	12
Multiplikator	2		3
log des Wertes	0,3010	0,6021	1,0792
Genotyp	aa	Aa	AA
Wert	2	6	12
Multiplikator	3		2
log des Wertes	0,3010	0,7782	1,0792

Im ersten Beispiel ist der multiplikative Faktor in der Heterozygoten nur 2, während er durch das Hinzukommen des zweiten Allels auf 3 ansteigt. Im zweiten Beispiel ist es umgekehrt. Man könnte im ersten Fall von einer partiellen „multiplikativen Dominanz“ des a -Allels und im zweiten von einer solchen des A -Allels sprechen. In jedem Falle einer solchen multiplikativen Dominanz treffen die von CHARLES und SMITH abgeleiteten Formeln für multiplikative Genwirkung nicht mehr zu. Wohl läßt eine logarithmische Umformung die Wirkung der Gene additiv erscheinen. Jedoch findet sich jetzt auch eine partielle Dominanz, wie ein Vergleich der Logarithmen der Werte des obigen Schemas zeigt.

Nach den bisherigen Erfahrungen, die bei der Untersuchung quantitativer Merkmale gemacht worden sind, liegt kein Grund vor, nicht auch eine solche Wirkung anzunehmen. Darüber hinaus muß man nach den heutigen wissenschaftlichen Ergebnissen über die Vererbung quantitativer Merkmale feststellen, daß man durchaus nicht nur die beiden bisher betrachteten Arten der additiven und multiplikativen Genwirkung als allein mögliche ansehen darf (vgl. auch KAPPERT, 1953, S. 37). MATHER (1949a) führt in seinem Buch „Biometrical Genetics“ Beispiele an, wo auch die logarithmische Transformation der Meßwerte nicht zu einem Maßstab führt, auf dem die Wirkung der Gene einigermaßen additiv ist. So ergaben Versuche von POWERS (zitiert nach MATHER, 1949a) über die Vererbung der Fruchtgewichte bei Tomaten, daß die stärkere Kontraktion des oberen Teils des Maßstabes durch Benutzung der Logarithmen der Werte noch nicht genügte, sondern daß hier offenbar eine noch größere Kontraktion nötig ist.

Im gewissen Sinne entgegengesetzt zur multiplikativen Wirkung ist die Wechselwirkung der Gene die RASMUSSEN (1933—34) in seiner Interaktionstheorie annimmt. Der Grundgedanke dieser Theorie ist der, daß die Wirkung eines Faktors auf den Genotyp abhängig von allen anderen vorhandenen Faktoren und um so

kleiner ist, je mehr Leistungsfaktoren schon anwesend sind in dem betreffenden Genotyp. Ein mathematischer Ausdruck für diese Wechselwirkung zwischen Genen soll in folgendem entwickelt werden. Nimmt man an, daß der durchschnittliche Wert eines Faktors a ist und daß der Wert k den Grad der Wechselwirkung (interaktion) zwischen den Genen angibt (er soll von 0—1 schwanken), so ist die Zunahme, die jeder Faktor bewirkt, durch eine geometrische Reihe gegeben:

Zahl der Faktoren	0	1	2	3	n
Zunahme in der Wirkung	0	ak^0	ak^1	ak^2		ak^{n-1}

Der mathematische Ausdruck für den Wert eines beliebigen Genotyps ist dann die Summenformel der geometrischen Reihe $y = a \frac{k^x - 1}{k - 1}$. x ist die Anzahl der vorhandenen Leistungsgene.

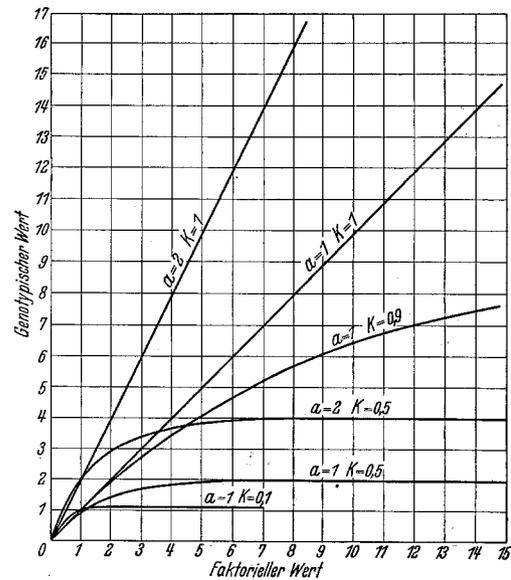


Abb. 1. Regression des „genotypischen Wertes“ auf den „faktoriellen Wert“ für einige Werte von a und k . (Nach RASMUSSEN, 1933—34, S. 254.)

Die von RASMUSSEN (S. 254) übernommene graphische Darstellung gibt die Regressionslinien des „genotypischen Wertes“ auf den bei rein additiver intermediärer Wirkung der Gene zu erwartenden Wert x , den RASMUSSEN als faktoriellen Wert bezeichnet, für einige Werte von a und k an. Die Grenzwerte für k von 0 und 1 haben beide einen biologischen Sinn. Für $k = 1$ ist die Wirkung der Gene additiv, die Regressionslinie wird zu einer Geraden (siehe Abb. 1). Ist $k = 0$, so bedeutet das, daß jedes Gen, das hinzugefügt wird, wenn schon ein Leistungsgen vorhanden ist, den Wert Null hat.

Aus der Abbildung erkennt man, daß der Wert k die Krümmung der Kurve bestimmt. Je kleiner er ist, desto früher neigt sich die Kurve, um sich einem Maximalwert asymptotisch zu nähern, was auch besagt, daß die Wirkung eines Leistungsgens, das in einem Genotyp, der schon ein paar Leistungsgene hat, eingeführt wird, um so kleiner ist, je kleiner k ist. Das veranschaulicht auch die ebenfalls von RASMUSSEN übernommene Tabelle 2.

Bei den bisher betrachteten Möglichkeiten über das Zusammenwirken der Polygene hatte jedes Leistungsgen, das in einen Genotyp eingeführt wurde, immer noch eine selbständige, erhöhende Wirkung auf die

Eigenschaftsausprägung. Man könnte sich nun aber auch vorstellen, daß es analog zu den bei Majorgenen bekannten Fällen der komplementären und epistatischen¹ Genwirkung auch bei quantitativen Merkmalen Gene gibt, die nur bei Anwesenheit anderer Gene zur Wirkung kommen. Eine derartige Wirkungsweise würde die Zahl der möglichen Phänotypen stark reduzieren.

Nimmt man an, daß drei komplementäre Faktoren erst eine Merkmalsausprägung bedingen, wie es bei-

Tabelle 2. (Aus RASMUSSEN, 1933/34, S. 253) Genotypische Werte für verschiedene Werte für a und k nach der Formel

$$y = a \frac{k^x - 1}{k - 1}$$

Faktorieller Wert	Genotypischer Wert							
	a=1 k=1	a=2 k=1	a=1 k=0,9	a=2 k=0,9	a=1 k=0,5	a=2 k=0,5	a=1 k=0,1	a=2 k=0,1
0	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
1	1,00	2,00	1,00	2,00	1,00	2,00	1,00	2,00
2	2,00	4,00	1,90	3,80	1,50	3,00	1,10	2,20
3	3,00	6,00	2,71	5,42	1,75	3,50	1,11	2,22
4	4,00	8,00	3,44	6,88	1,88	3,75	1,11	2,22
5	5,00	10,00	4,10	8,19	1,94	3,88	1,11	2,22
6	6,00	12,00	4,69	9,37	1,97	3,94	1,11	2,22
7	7,00	14,00	5,22	10,43	1,98	3,97	1,11	2,22
8	8,00	16,00	5,65	11,30	1,99	3,98	1,11	2,22
9	9,00	18,00	6,03	12,07	2,00	3,99	1,11	2,22
10	10,00	20,00	6,38	12,77	2,00	4,00	1,11	2,22
11	11,00	22,00	6,70	13,39	2,00	4,00	1,11	2,22
12	12,00	24,00	6,98	13,96	2,00	4,00	1,11	2,22
13	13,00	26,00	7,23	14,46	2,00	4,00	1,11	2,22
14	14,00	28,00	7,46	14,92	2,00	4,00	1,11	2,22
15	15,00	30,00	7,66	15,33	2,00	4,00	1,11	2,22
∞	∞	∞	10,00	20,00	2,00	4,00	1,11	2,22

spielsweise von KAPPERT (1953) bei der Vererbung der Sproß- und Blattbehaarung der Levkojen gefunden wurde, so ergäbe die Nachkommenschaft einer trihybriden F₁ bei voller Dominanz eine Aufspaltung in 2 Phänotypen, nämlich in 37 behaarte und 27 unbehaarte Pflanzen. Bei jeder anderen Art kumulativer Genwirkung erhält man dagegen bei voller Dominanz 8 verschiedene Phänotypen und bei intermediärer Vererbung sogar 27. Ähnlich, wenn auch nicht in dem Maße, wird die Anzahl der Phänotypen bei epistatischer Wirkung reduziert. Würde der Vererbungsmechanismus quantitativer Leistungseigenschaften weitgehend auf dieser Art des Zusammenwirkens beruhen, so könnte man in der F₂ zweier Eltern, die sich nur in wenig Genen unterscheiden, in gewissen Fällen eine diskontinuierliche Verteilung erwarten. In solchen Fällen wäre eine Selektion auf derartige vererbte Merkmale wesentlich erleichtert, und bei den mehr oder weniger durchgezüchteten landwirtschaftlichen Kulturpflanzen ist eventuell in bezug auf solche Gene schon eine Homozygotie erreicht.

Wie weit eine derartige komplementäre und epistatische Genwirkung bei quantitativen Eigenschaften verbreitet ist, läßt sich nicht leicht entscheiden. Die statistische Analyse wird durch diese Arten der Wechselwirkung jedenfalls erheblich kompliziert. In den

¹ Unter dem Terminus „Epistasie“ werden in der englischsprachigen Literatur ausgehend von WRIGHT und ihm folgend von vielen anderen häufig sämtliche Arten von Wechselwirkungen zwischen nicht allelen Genen zusammengefaßt. Hier soll unter Epistasie nur die Erscheinung verstanden werden, daß ein Gen die Wirkung eines anderen vollkommen verhindert (vgl. KAPPERT, S. 41).

nun zu besprechenden Methoden ist daher immer die Annahme gemacht worden, daß die Wirkung der Polygene weitgehend kumulativ ist, und daß epistatische und komplementäre Genwirkungen, wenn sie vorhanden sind, im Verhältnis zu den übrigen kumulativ wirkenden Genen nicht ins Gewicht fallen und somit die Ergebnisse nicht wesentlich beeinflussen.

Zusammenfassend läßt sich sagen, daß die Fälle einer einfachen additiven oder multiplikativen Genwirkung von Polygensystemen nicht verallgemeinert werden dürfen. Vielmehr muß man mit einer Vielfalt von Möglichkeiten des Zusammenwirkens der die quantitativen Eigenschaften bedingenden Gene rechnen. Es ist daher angebracht, sich vor jeder statistischen Analyse quantitativ vererbter Merkmale über das Zusammenwirken der Gene durch die hier angegebenen Methoden oder durch eine andere als die logarithmische Umformung (z. B. durch Verwendung von Antilogarithmen, Wurzeln, Potenzen oder irgend einer anderen Funktion) ein annäherndes Bild zu verschaffen. Die letzte Forderung, nämlich die Umformung der Werte in einen Maßstab, auf dem die Wirkung der Gene additiv erscheint, betont besonders MATHER (1949a), der sie für die Anwendung seiner Methoden zur Schätzung der durchschnittlichen Dominanz der Gene, der Genzahl und der Berechnung von Kopplungsverhältnissen zwischen Polygenen überhaupt zur Voraussetzung macht. Bei Besprechung dieser Methoden wird daher hierauf noch näher einzugehen sein.

Die Schätzung der durchschnittlichen Dominanz

Die allgemein in Genetik und Züchtung bekannte Tatsache des Luxurierens der Bastarde homozygoter Eltern, der Heterosis, kann auf zweierlei Weise erklärt werden. Einmal ist es möglich, daß der höhere Wert der F₁ durch partielle oder vollständige Dominanz der Leistungsgene und Kopplung von + und - Genen bedingt wird, zum anderen kann die Heterozygote per se größere Leistung als jede der beiden Homozygoten bedingen. In diesem Fall spricht man von Superdominanz der Gene.

Die Konsequenzen, die sich aus diesen beiden Erklärsmöglichkeiten für die Züchtung ergeben, sind, wie leicht einzusehen ist, grundverschieden. Im ersten Fall ist es generell möglich, Austauschtypen zu finden, in denen die höhere Leistung der F₁ fixiert ist. Das ist aber ausgeschlossen, wenn die Heterozygoten höhere Werte zeigen als die beiden Homozygoten. Im Hinblick auf die Analyse quantitativer Merkmale wird diese bezüglich der züchterischen Konsequenzen grundsätzliche Verschiedenheit jedoch eingengt „auf die quantitative Frage nach der normalen Stellung der Heterozygoten eines Genpaares in Beziehung zu ihren Homozygoten“ (FISHER, IMMER u. TEDIN, 1932 S. 112), d. h. ganz allgemein gesprochen, auf die Frage nach dem Dominanzgrad der Gene. Die Berechnung der durchschnittlichen Dominanzverhältnisse von Polygensystemen in der Analyse quantitativer Vererbung ist also nicht nur theoretisch eine der interessantesten und wichtigsten Fragen, sondern sie kann auch für die

Wahl der anzuwendenden Zuchtmethoden von ausschlaggebender Bedeutung sein.

Im folgenden sollen daher einige, in der Literatur angegebene Methoden aufgeführt werden, die sich mit der Schätzung der durchschnittlichen Dominanz von Polygenen befassen. Wie bereits im ersten Kapitel dieser Arbeit gezeigt wurde, kann eine schiefe Verteilung spaltender Generationen nicht nur durch multiplikative Wirkung der Gene, sondern auch durch additive dominante Genwirkung verursacht werden. Die positiv schiefe Verteilung der F_2 in Tabelle 1b wurde durch Dominanz der Gene mit dem kleineren Wert verursacht. Ganz entsprechend bedingt Dominanz der Gene mit dem größeren Wert eine negativ schiefe Verteilung der spaltenden Generationen, also eine Verteilung, bei der die Klasse mit der größten Anzahl von Varianten rechts vom Mittelwert liegt. Als Maß für die Schiefheit kann der Wert der mittleren dritten Potenz der Abweichung vom Mittel benutzt werden, der als statistische Maßzahl dritten Grades bekannt ist und mit k_3 bezeichnet wird.

$$k_3 = \frac{n}{(n-1)(n-2)} S(x - \bar{x})^3$$

Diese Maßzahl und das mittlere Produkt (oder die Kovarianz) des Mittelwertes (k_1) und der Varianz (k_2) von F_3 -Familien benutzen FISHER, IMMER u. TEDIN (1932) zur Berechnung der durchschnittlichen Dominanz. Nimmt man an, wie es schon bei der Ableitung der Formel (2) in dieser Arbeit getan worden ist, daß die Werte der drei Genotypen eines Genpaares $+d$ für $a_1 a_1$, $-d$ für $a_2 a_2$ und h für $a_1 a_2$ sind und daß sich die Eltern nur in diesem Gen unterscheiden, so beträgt der mittlere k_3 -Wert aller vier F_3 -Nachkommenschaften $(-3/8) d^2 h$. Ebenso groß ist der k_3 -Wert der Mittelwerte der F_3 -Nachkommenschaften. Die Kovarianz aus Mittelwert und Varianz der einzelnen F_3 -Nachkommenschaften beträgt $1/32 h (2d^2 + h^2)$. Aus diesen drei Werten kann man nun berechnen, ob h größer, gleich groß oder kleiner als $+d$ ist, mit anderen Worten, ob also Superdominanz, absolute Dominanz oder partielle Dominanz vorliegt. Multipliziert man nämlich die Kovarianz mit 4, so ist der Wert dann größer als $(+3/8) d^2 h$, wenn h größer ist als $+d$. Er ist gleich diesem Wert, wenn $h = +d$ und kleiner, wenn h kleiner als $+d$, aber noch positiv ist. Bei additiver Wirkung der Gene bleiben diese Gesetzmäßigkeiten bei jeder beliebigen Anzahl von Genen erhalten. Ein Vergleich der vierfachen Kovarianz zwischen Mittelwert und Varianz der verschiedenen F_3 -Familien mit dem mit umgekehrtem Vorzeichen versehenen durchschnittlichen k_3 -Wert der F_3 -Familien oder dem mit umgekehrtem Vorzeichen versehenen k_3 -Wert der Mittelwerte der F_3 -Nachkommenschaften macht also eine Aussage über die Stellung der Heterozygoten möglich. Die im Experiment gefundenen Werte sind naturgemäß mit einem Fehler behaftet. Ein Signifikanztest ist mit Hilfe der von FISHER, IMMER u. TEDIN (1932) angegebenen Formeln für die Varianz dieser Werte möglich. Sie beträgt für den mittleren Wert von k_3 für die einzelnen F_3 -Nachkommenschaften $[1/(n-1)n] S(k_3 - \bar{k}_3)^2$, wo n die Zahl der insgesamt berechneten k_3 -Werte ist; für k_3 der Mittelwerte gleich $[6n/(n-1)(n-2)] k_3^2$ und für die Kovarianz aus Mittelwert und Varianz der einzelnen F_3 -Nachkommenschaften

$(1/n-1) [V(k_1) V(k_2) + V(k_1 k_2)]$. $V(k_1)$ und $V(k_2)$ sind die Varianzen der Mittelwerte und die der Varianzen der einzelnen F_3 -Familien, und $V(k_1 k_2)$ ist ihre Kovarianz. FISHER, IMMER u. TEDIN haben einige Beispiele aus der Literatur nach ihrer Methode durchgerechnet, die hier nicht im einzelnen aufgeführt werden können.

Weitere Möglichkeiten zur Schätzung der Dominanzverhältnisse von Polygenen werden von GRIFFING (1950) beschrieben, von denen sich die erste der Regression zwischen F_1 - und Elterwerten bedient. Diese Methode geht in ihren Anfängen auf Arbeiten von HULL zurück. Ihr besonderer Vorteil besteht darin, daß bei ihr nur Meßwerte der drei nichtspaltenden Generationen für die Analyse benutzt werden.¹ Dadurch ist die Zahl der Fehlerquellen stark reduziert. Zweitens braucht man für genetisch nichtspaltende Generationen eine wesentlich geringere Zahl von Individuen, um den genotypischen Wert mit einiger Sicherheit abschätzen zu können, als für solche Generationen, deren Variabilität durch ihre Aufspaltung in erblich verschiedene Typen erhöht wird. Nicht zuletzt erlaubt eine Analyse, die, wie es diese Methode erfordert, nicht nur zwei Ausgangseltern, sondern eine größere Anzahl für ihre Untersuchungen verwendet, allgemein gültigere Aussagen über die Vererbung der betreffenden Eigenschaft zu machen. Ihr Nachteil liegt naturgemäß in der Tatsache, daß eben mehr als nur 2 Elterlinien für die Analyse nötig sind. Kopplung von Polygenen hat auf den Wert dieser Methode keinen Einfluß, wenn man voraussetzt, daß sich bei den verwendeten Eltern die Repulsions- und Attraktionsphasen das Gleichgewicht halten.

Nehmen wir an, wir hätten verschiedene Linien eines Selbstbefruchtlers oder Inzuchtlinien eines Fremdbefruchtlers, so können wir aus diesen homozygoten Eltern $\binom{n}{2} = \frac{n!}{2!(n-2)!}$ Kreuzungen herstellen.

Jeder bestimmte Elter kann mit allen übrigen $(n-1)$ Eltern gekreuzt werden. $(n-1) F_1$ -Generationen haben also immer einen gemeinsamen Elter. Innerhalb dieser Kreuzungsgruppe (constant parent group), in der in allen Kreuzungen ein gleicher Elter vorhanden ist, kann man die Regression der F_1 -Werte auf die Werte des jeweils verschiedenen Elters berechnen. Zeigen die so berechneten Regressionskoeffizienten aller Kreuzungsgruppen, wenn man sie nach der Größe des konstanten Elters ordnet, einen abnehmenden Trend, so liegt Dominanz der Gene mit dem größeren Wert der Eigenschaft vor. Nimmt der Trend der Regressionslinie zu, so muß man zuerst untersuchen, ob die Genwirkung additiv oder multiplikativ ist. Hierfür hat man folgende Möglichkeiten:

1. Steigen die größeren Regressionswerte stark über 1 an und wird der Trend durch Verwendung der Logarithmen der Werte stark reduziert, so deutet das auf multiplikative Genwirkung.

2. Wie allgemein aus der Regressionsrechnung bekannt ist, kann man die Summe der Abweichungsquadrate von \bar{y} , dem Mittelwert der relativen Eigenschaft, aufteilen in den Teil, der durch die Abweichung

¹ Nach Drucklegung der vorliegenden Arbeit wurde mir erst eine neuere Veröffentlichung von JINKS (Genetics 39, 767-788, 1954) bekannt, in der für die Berechnung der durchschnittlichen Dominanz ebenfalls nur die Elter- und F_1 -Werte verwandt werden und auf die hier noch hingewiesen werden soll.

der Werte von der Regressionsgeraden und in den restlichen Teil, der durch die Abweichung der Regressionslinie selbst vom Mittelwert \bar{y} verursacht wird (siehe SNEDECOR, 1950, S. 127). Diesen Mittelwert kann man sich graphisch als eine durch den Wert für \bar{y} gehende, zur x -Achse parallele Linie vorstellen. Bezeichnen x und y die Abweichung der gefundenen Werte von ihrem Mittel, so ist der mathematische Ausdruck für den Teil der Varianz, der durch die Abweichung von der Regressionsgeraden bedingt wird, gleich $\frac{(Sxy)^2}{Sx^2}$. Wird nun dieser Teil der Varianz der y -, in diesem Fall also der F_1 -Werte, nach Verwendung der Logarithmen stark reduziert, so zeigt das ebenfalls multiplikative Genwirkung an.

3. Eine dritte Möglichkeit besteht darin, den von GRIFFING (1950) als „potence value“ bezeichneten Ausdruck $h_p = \frac{F_1 - M_p}{P - M_p}$ zu berechnen. F_1 bezeichnet den Wert der F_1 , M_p das arithmetische Mittel der Eltern (Mittelerterwert) und P den Elter, der den positiven oder negativen extremen Wert der Eigenschaft darstellt, d. h. also nur Leistungsgene oder nur ihre Allele besitzt. Wenn man einen solchen Elter zur Verfügung hat, was natürlich in den seltensten Fällen möglich sein wird, und die Genwirkung rein additiv ist, so gibt der h_p -Wert den genauen durchschnittlichen Dominanzgrad an. Da die eben genannte Bedingung, daß man einen extremen Elter zur Verfügung hat, selten erfüllt sein wird, kann dieser Wert meistens nur ein roher Schätzwert sein. Zu der hier in Frage stehenden Prüfung auf additive oder multiplikative Genwirkung kann aber die mehr oder weniger große Gleichmäßigkeit der h_p -Werte herangezogen werden. Sind diese Werte, die man leicht für alle vorhandenen F_1 -Generationen berechnen kann, nach Benutzung der Logarithmen bedeutend gleichmäßiger geworden, so ist auch das ein Hinweis auf eine multiplikative Wirkung der Gene.

Findet man nun keines der drei Kriterien für multiplikative Genwirkung erfüllt, so kann man bei zunehmendem Trend der Regressionswerte eine weitgehend additive Genwirkung mit Dominanz der Gene mit geringer Eigenschaftsausprägung annehmen (GRIFFING bezeichnet diese auch als negative Dominanz gegenüber der positiven Dominanz der Leistungsgene). Wird dagegen der Trend der Regressionskoeffizienten nach Verwendung der Logarithmen stark reduziert, verringert sich weiter der Teil der Varianz der F_1 -Werte, der durch die Abweichung von den Regressionsgeraden bedingt wird, und werden auch die h_p -Werte einheitlicher, so wird man in der Analyse von den logarithmischen Werten ausgehen.

Man wird, wie sich aus der eingangs erwähnten Vielfalt des Zusammenwirkens von Polygenen ergibt, auch nach logarithmischer Transformation nicht immer vollkommene Übereinstimmung mit den Erwartungen additiver Genwirkung finden. Es bleibt den Versuchsanstellern dann überlassen, nach einem besseren Maßstab zu suchen (siehe auch S. 269).

Bis zu einem gewissen Grade ist es möglich, neben der Richtung der Dominanz (positiv oder negativ) auch den Dominanzgrad aus den Regressionskoeffizienten selbst und der Regression dieser bereits errechneten Regressionskoeffizienten auf die entsprechenden „konstanten Elterwerte“ annähernd abzuschätzen.

GRIFFING (1950) zeigt in seiner erwähnten Arbeit weiter, daß man mit Hilfe der Komponenten der Varianz der F_1 -Werte innerhalb einer Kreuzungsgruppe eine exaktere Schätzung der Dominanz vornehmen kann. Die Varianz der F_1 -Werte kann aufgeteilt werden in die umweltbedingte Varianz und in die durch die verschiedenen F_1 -Genotypen einer Kreuzungsgruppe bedingte erbliche Varianz. Die letzte setzt sich zusammen aus dem Teil, der durch die d -Komponente verursacht wird (siehe S. 264), also abhängt von den Differenzen zwischen den Homozygoten der beteiligten Genpaare und dem Teil, der durch die h -Werte bedingt wird, und somit durch die Differenz zwischen Heterozygoten und Durchschnittswert der beiden Homozygoten entsteht.

Der Dominanzgrad ist dann $\sqrt{\frac{\sigma_h^2}{\sigma_d^2}} \cdot \sigma_h^2$ ist aber, wie GRIFFING zeigen kann, gleich der Varianz der Differenzen zwischen den F_1 - und Mittelerterwerten minus $5/4$ der umweltbedingten Varianz, σ_d^2 ist gleich der Varianz der jeweils verschiedenen Eltern einer Kreuzungsgruppe minus umweltbedingter Varianz. Die umweltbedingte Varianz kann direkt aus dem in Wiederholungen angelegten Versuch erhalten werden.

ГОРОХ (1953) hat nach diesen von GRIFFING angegebenen Methoden einige quantitative Eigenschaften von *Solanum melongena* untersucht. Insgesamt verwendet er 7 Sippen für seine Berechnungen. Er findet z. B. für den Längenbreitenindex der Frucht additive Genwirkung mit geringer negativer Dominanz, für das Fruchtgewicht multiplikative Genwirkung und für die Blütezeit additive Genwirkung mit negativer Dominanz. Der Dominanzgrad variierte in diesem Fall stark zwischen den einzelnen Kreuzungsgruppen.

Vor der Besprechung der von MATHER entwickelten Methoden zur Analyse quantitativer Genwirkungen, die in seinem 1949 erschienenen Buch „Biometrical Genetics“ eine umfassende Darstellung erfahren haben, soll noch einmal auf die Frage der Anwendung der geeigneten Skala für die den Berechnungen zugrundeliegenden Messungen näher eingegangen werden. MATHER betont immer wieder mit Nachdruck, daß vor jeder Untersuchung quantitativ wirkender Gene mit Hilfe von Varianzen und Kovarianzen, wie er sie durchführt, ein Maßstab gefunden werden sollte, auf dem die Wirkung der die Variation bedingenden Faktoren, also die der Gene und die der Umweltfaktoren, weitgehend additiv ist und somit zwischen ihnen, keine Wechselwirkung besteht (MATHER, 1946, MATHER 1949a, MATHER 1949b). Je weniger Wechselwirkungen oder Interaktionen zwischen diesen Faktoren auftreten, desto besser kann die Analyse durchgeführt werden und um so genauer werden die Aussagen, die man machen kann. Eine der Interaktionen ist in der multiplikativen Wirkung der Gene schon aufgezeigt worden. Jedes Gen ist in seiner Wirkung von den bereits vorhandenen Leistungsgenen abhängig, da es nicht wie bei additiver Genwirkung einen bestimmten Betrag hinzufügt, sondern den Wert eines Genotyps mit einem bestimmten Faktor multipliziert, seine verbessernde Wirkung also letztlich von den bereits vorhandenen Leistungsgenen abhängt.

Aber auch die Dominanz kann man sich in der biometrischen Analyse als eine Art Wechselwirkung vorstellen, nämlich als eine solche zwischen den Allelen

eines Genpaares. Da diese aber nun ein Hauptgegenstand der Untersuchung ist und in den Formeln volle Berücksichtigung findet, braucht und soll sie durch Transformation der Meßwerte in irgend einen anderen Maßstab nicht verschwinden. Wichtig bei der Auswahl eines geeigneten Maßstabes und deswegen voll zu berücksichtigen ist aber die Interaktion zwischen Umweltfaktoren und Genotyp, die dadurch entsteht, daß gleiche Umweltfaktoren nicht alle Genotypen gleichmäßig beeinflussen. Solche Interaktionen können durch einen Vergleich der Varianzen verschiedener Genotypen getestet werden. In den beiden Elterngenerationen sowie in der F_1 -Generation hat man vollkommen erbgleiche Generationen zur Verfügung, deren Varianz rein umweltbedingt sein muß. Die Varianzen dieser drei nichtspaltenden Generationen müßten also innerhalb der Fehlergrenzen gleich sein, wenn die Umweltfaktoren alle drei Genotypen gleichmäßig beeinflussen. Eine Prüfung auf Homogenität dieser Varianzen gibt also eine Auskunft über das Vorhandensein von Interaktionen zwischen Umwelt und Genotyp. Für den Test auf additive Wirkung der Gene benutzt MATHER (1949a) die Mittelwerte der Rückkreuzungs-, F_2 - und F_3 -Generationen:

$$\bar{B}_1 = \frac{1}{2} (\bar{P}_1 + \bar{F}_1) \quad (4)$$

$$\bar{B}_2 = \frac{1}{2} (\bar{P}_2 + \bar{F}_1) \quad (4)$$

$$\bar{F}_2 = \frac{1}{4} (\bar{P}_1 + \bar{P}_2 + 2 \bar{F}_1) \quad (2)$$

$$\bar{F}_3 = \frac{1}{4} (\bar{P}_1 + \bar{P}_2 + 2 \bar{F}_2) \quad (5)$$

Die Formeln für die Rückkreuzungs- und F_2 -Generationen sind auch schon von CHARLES und SMITH (1939) für die Unterscheidung zur multiplikativen Genwirkung benutzt worden. Für die Formel (2) ist auf Seite 264 die Ableitung gegeben worden. Ebenso leicht läßt sich der Beweis für die Formeln (4) und (5) erbringen. Aus den obengenannten Gleichungen ergeben sich weiterhin folgende Beziehungen:

$$\bar{F}_2 = \frac{1}{2} (\bar{B}_1 + \bar{B}_2) \quad (6)$$

$$\bar{F}_3 = \frac{1}{8} (3 \bar{P}_1 + 3 \bar{P}_2 + 2 \bar{F}_1) \quad (7)$$

Sind die durch diese Formeln angegebenen Bedingungen bei Benutzung der in g, cm usw. gemessenen Werte nicht erfüllt, so schlägt MATHER vor, eine Umwandlung der Werte in ihre Logarithmen, Potenzen, Antilogarithmen oder auf irgendeine andere Art vorzunehmen.

COPP und WRIGHT (1952), die die Vererbung des durchschnittlichen Korngewichtes in einer Weizenkreuzung untersuchten, gingen dabei folgendermaßen vor:

Nahmen sie keine Umformung der Werte vor, so wichen die Werte für

$$A = 4 \bar{F}_2 - 2 \bar{F}_1 - \bar{P}_1 - \bar{P}_2 \quad \text{und}$$

$$B = 8 \bar{F}_3 - 2 \bar{F}_1 - 3 \bar{P}_1 - 3 \bar{P}_2$$

signifikant von ihrem Erwartungswert, der nach Formel (2) bzw. (7) gleich 0 ist, ab. Es wurden negative Werte erhalten. Neben diesen negativen Werten von A und B zeigte auch die Ungleichmäßigkeit der Varianzen für die Eltern- und F_1 -Generationen, daß die in mg gemessenen Werte für das Korngewicht nicht geeignet waren für die Analyse. Auch die Benutzung der Wurzeln konnte die

Kriterien A und B nicht ganz erfüllen. Es wurde daher zu einer modifizierten logarithmischen Transformation übergegangen:

$G = \log(g - a)$. g ist das ursprüngliche Korngewicht in mg, G das entsprechende Gewicht auf der neuen erst noch zu findenden Skala und a eine Konstante, die bestimmt werden sollte. Es wurde festgestellt, daß der Wert für A das Vorzeichen bei einem Wert für a zwischen 25 und 30 und der Wert B für a zwischen 20 und 25 ändert. So schien die Funktion $G = \log(g - 25)$ eine geeignete Transformation für die Werte darzustellen. Es zeigte sich jedoch, daß bei Benutzung dieser Transformation die Varianz des Elters mit niedrigerem Korngewicht gesichert größer war als die des anderen Elters. Zur Angleichung dieser beiden Varianzen mußte also eine weniger wirksame Transformation benutzt werden. Diese Forderung wurde zwar bei Verwendung des Ausdrucks $\log(g - 10)$ erfüllt; hier waren aber die Werte für A und B wieder von o verschieden. Der Ausdruck $\log^2(g - 25)$ zeigte sich schon als eine günstigere Kompromißlösung. Eine weitere Prüfung ergab, daß der Ausdruck $\log 5/2(g - 25)$ den Forderungen am besten genügte.

Die eben dargestellten Bemühungen COPPS und WRIGHTS, für ihre Untersuchungen einen geeigneten Umwandlungsmodul zu finden, zeigen, daß diese von MATHER nachdrücklich betonte Forderung nicht immer leicht zu erfüllen ist. In den meisten Fällen wird man kaum eine derart ideale Skala finden, auf der alle die Variabilität beeinflussenden Faktoren eine vollkommene additive Wirkung zeigen. Vielmehr muß man sehr oft einen Kompromiß schließen und einen Modul benutzen, der den kleinsten möglichen Rest an Interaktionen übrig läßt. Auch wird man darauf achten müssen, daß der rechnerische Aufwand zum Auffinden einer noch besseren Transformation in einem vernünftigen Verhältnis zu der erreichbaren Verbesserung steht.

Die Methoden von MATHER zur Schätzung der durchschnittlichen Dominanz unterscheiden sich von den von GRIFFING angegebenen grundsätzlich dadurch, daß MATHER die Varianzen und Kovarianzen spaltender Generationen für seine biometrische Analyse benutzt.

Geht man wieder von den 3 Genotypen a_1a_1 , a_1a_2 und a_2a_2 eines Genpaares mit den Werten $+d_a$, h_a und $-d_a$ aus (siehe Seite 264), so kann man die Varianz der F_2 in d_a und h_a ausdrücken und sie somit in diese beiden Komponenten trennen:

Die F_2 besteht aus 1 a_1a_1 mit dem Wert $+d_a$, 2 a_1a_2 mit dem Wert h_a und 1 a_2a_2 mit dem Wert $-d_a$. Die Summe der Abweichungsquadrate dieser 4 Genotypen vom Mittelwert 0 ist $d_a^2 + 2h_a^2 + d_a^2 = 2d_a^2 + 2h_a^2$, das durchschnittliche Abweichungsquadrat von diesem Wert also $1/2 d_a^2 + 1/2 h_a^2$. Der F_2 -Mittelwert ist $h/2$, und somit die Varianz der $F_2 = 1/2 d_a^2 + 1/2 h_a^2 - h_a^2/4 = 1/2 d_a^2 + 1/4 h_a^2$.

Zeigen die Gene keine Wechselwirkung und sind sie nicht gekoppelt, so ist die totale erbliche Varianz der F_1 gleich der Summe aller $1/2 d_a^2 + 1/4 h_a^2$, also $1/2 S(d_a^2) + 1/4 S(h_a^2)$. Der erste Teil dieser Varianz stellt den erblichen, fixierbaren Anteil, der zweite den nicht fixierbaren, durch die Werte der Heterozygoten bedingten Anteil dar. Für $S(d_a^2)$ und $S(h_a^2)$ schreibt MATHER auch D und H . Bezeichnet man die umweltbedingte Varianz der F_2 mit E (environment), so ist die totale Varianz einer F_2 -Generation

$$V_{F_2} = 1/2 D + 1/4 H + E. \quad (8)$$

Entsprechend einfach ist die Ableitung der nachstehend aufgeführten Formeln (MATHER, 1949 a, S.56):

$$V_{\bar{F}_3} = \text{Varianz der Mittelwerte der } F_3\text{-Nachkommenschaften} \quad \frac{1}{2} D + \frac{1}{16} H + E$$

$$W_{F_2/F_3} = \text{Kovarianz aus den } F_3\text{-Mittelwerten und den } F_2\text{-Elterwerten} \quad \frac{1}{2} D + \frac{1}{8} H$$

$$\bar{V}_{F_3} = \text{Durchschnittliche Varianz der } F_3\text{-Nachkommenschaften} \quad \frac{1}{4} D + \frac{1}{8} H + E$$

$$V_{\bar{Bip}} = \text{Varianz der Mittelwerte von Biparental-Nachkommenschaften}^1 \quad \frac{1}{4} D + \frac{1}{16} H + E$$

$$W_{F_2/Bip} = \text{Kovarianz der Mittelwerte der Biparental-Nachkommenschaften und ihrer } F_2\text{-Elterwerte} \quad \frac{1}{4} D$$

$$\bar{V}_{Bip} = \text{Durchschnittliche Varianz der Biparental-Nachkommenschaften} \quad \frac{1}{4} D + \frac{3}{16} H + E$$

$$V_{\bar{Mat}} = \text{Varianz der Mittelwerte der Maternal-Nachkommenschaften}^1 \quad \frac{1}{8} D + E$$

$$\bar{V}_{Mat} = \text{Durchschnittliche Varianz der Maternal-Nachkommenschaften} \quad \frac{3}{8} D + \frac{1}{4} H + E$$

Hat man nun z. B. im Experiment die Varianz der F_2 , die durchschnittliche Varianz der F_3 -Nachkommenschaften, die Varianz der F_3 -Mittelwerte, die Kovarianz aus den F_2 -Werten und den Mittelwerten der F_3 -Nachkommenschaften und aus den Varianzen der nichtspaltenden Generationen (P_1 , P_2 , u. F_1) die Werte für E_1 und E_2 ermittelt, so erhält man folgendes Gleichungssystem:

$$\begin{aligned} V_{F_2} &= \frac{1}{2} D + \frac{1}{4} H + E_1 = a \\ V_{\bar{F}_3} &= \frac{1}{2} D + \frac{1}{16} H + E_2 = b \\ W_{F_2/F_3} &= \frac{1}{2} D + \frac{1}{8} H = c \\ \bar{V}_{F_3} &= \frac{1}{4} D + \frac{1}{8} H + E_1 = d \\ &E_1 = e \\ &E_2 = f \end{aligned}$$

Wie alle Meßwerte sind auch die im Versuch ermittelten Werte für V_{F_2} , usw., die hier mit a, b, c, d, e und f bezeichnet sind, mit einem Fehler behaftet. Zur Schätzung ihrer wahrscheinlichsten Werte benutzt MATHER die Methode der kleinsten Quadrate. Nach dieser Methode ergibt sich die Forderung, alle Gleichungen der Reihe nach mit den Koeffizienten der verschiedenen Unbekannten D, H, E_1 und E_2 zu multiplizieren und die jeweils nach Multiplikation der Koeffizienten einer Unbekannten erhaltenen Gleichungen dann zu addieren (vgl. auch WRICKE, 1955).

¹ Biparental-Nachkommenschaften sind Nachkommenschaften aus der Kreuzung zweier zufällig aus der F_2 ausgewählter Pflanzen.

Maternal-Nachkommenschaften erhält man aus Bestäubung einer F_2 -Pflanze durch die gesamte F_2 -Generation.

Nach der Multiplikation mit den Koeffizienten von D ergäbe sich also:

$$\begin{aligned} \frac{1}{4} D + \frac{1}{8} H + \frac{1}{2} E_1 &= \frac{a}{2} \\ \frac{1}{4} D + \frac{1}{32} H + \frac{1}{2} E_2 &= \frac{b}{2} \\ \frac{1}{4} D + \frac{1}{16} H &= \frac{c}{2} \\ \frac{1}{16} D + \frac{1}{32} H + \frac{1}{4} E_1 &= \frac{d}{4} \\ \hline \frac{13}{16} D + \frac{1}{4} H + \frac{3}{4} E_1 + \frac{1}{2} E_2 &= \frac{1}{2} (a + b + c) + \frac{1}{4} d = r \end{aligned}$$

In Dezimalbrüchen geschrieben lautet die Gleichung:

$$0,8125 D + 0,25 H + 0,75 E_1 + 0,5 E_2 = r$$

Die Gleichungen, die die Unbekannte, mit deren Koeffizienten multipliziert wird, nicht enthalten, werden nicht benutzt. In diesem Fall waren es also

$$E_1 = e \quad \text{und} \quad E_2 = f.$$

Ganz entsprechend kann man nun alle Gleichungen mit dem Koeffizienten von H multiplizieren und erhält die zweite Formel:

$$0,25 D + 0,097656 H + 0,375 E_1 + 0,0625 E_2 = s.$$

Genau so erhält man nach Multiplikation mit den Koeffizienten von E_1 und E_2 und anschließender Addition zwei weitere Gleichungen:

$$0,75 D + 0,375 H + 3,0 E_1 = t$$

$$0,5 D + 0,0625 H + 2,0 E_2 = u$$

Man erhält also insgesamt soviel Gleichungen wie Unbekannte vorhanden sind. Nachfolgend sind die Gleichungen für das vorliegende Beispiel noch einmal zusammen aufgeführt:

$$0,812500 D + 0,250000 H + 0,750000 E_1 + 0,500000 E_2 = r$$

$$0,250000 D + 0,097656 H + 0,375000 E_1 + 0,062500 E_2 = s$$

$$0,750000 D + 0,375000 H + 3,000000 E_1 = t$$

$$0,500000 D + 0,062500 H + 2,000000 E_2 = u$$

Die einfachste Lösung dieses neuen Systems von Gleichungen besteht darin, es in ein Diagonalsystem zu verwandeln und dann von unten her aufzurollen¹.

¹ Das sei an einem einfachen Beispiel kurz dargestellt: Hätte man die drei Gleichungen

$$\begin{aligned} 2X_1 + 3X_2 + X_3 &= 17 \\ X_1 - X_2 + 2X_3 &= 7 \\ 2X_1 + 2X_2 - 3X_3 &= -2 \end{aligned}$$

zu lösen, so würde man zweckmäßig so vorgehen, daß man im 1. Schritt die untere Gleichung mit $+1$, die mittlere mit $+2$ multipliziert und beide von der ersten abzieht:

$$\begin{aligned} 2X_1 + 3X_2 + X_3 &= 17 \\ 5X_2 - 3X_3 &= 3 \\ X_2 + 4X_3 &= 19 \end{aligned}$$

Multipliziert man jetzt die letzte Gleichung mit $+5$ und zieht sie von der zweiten ab, so erhält man das folgende Diagonalsystem, das einfach von unten her zu lösen ist:

$$\begin{aligned} 2X_1 + 3X_2 + X_3 &= 17 \\ 5X_2 - 3X_3 &= 3 \\ -23X_3 &= -92 \\ X_3 = 4, X_2 = 3, X_1 = 2. \end{aligned}$$

Man erhält so die Werte für D, H, E_1 und E_2 , und der Ausdruck $\sqrt{\frac{H}{D}} = \frac{h}{d}$ ist ein Maß für die durchschnittliche Dominanz der Polygene.

MATHER (1949 a) geht etwas anders vor: Er setzt, wie es FISHER (1950) für die Analyse multipler Regressionen vorschlägt, für die im Experiment erhaltenen Werte auf der rechten Seite der Gleichungen sukzessive die Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

ein und löst zuerst diese Ersatzsysteme.

Die Lösung dieser Ersatzsysteme gibt eine Matrix von Koeffizienten, mit deren Hilfe die Werte für D, H und E leicht zu berechnen sind.

Der große Vorteil der Berechnung der Matrix besteht darin, daß diese für alle je durchzuführenden Versuche gilt, in denen die Unbekannten D, H und E aus denselben Varianzen und Kovarianzen, also aus $V_{F_1}, V_{F_2}, W_{F_1/F_2}$ und \bar{V}_{F_3} , errechnet werden sollen, da sie vollkommen von den im Versuch erhaltenen Werten unabhängig ist. Wenn man also einmal diese Matrix errechnet hat, so ist die Berechnung von D, H und E sehr einfach. Für eine Reihe von Versuchen sind die Matrizen bereits von MATHER (1949 a) berechnet worden. Es besteht daher für jeden Versuchsansteller die Möglichkeit, bei einem gleichartig angelegten Versuch auch ohne spezielle Kenntnisse in der Handhabung derartiger Matrizen die Berechnung der Werte für D, H und E für seinen Versuch schnell durchzuführen. Weiter sei in diesem Zusammenhang auf die bereits erwähnte Arbeit von COPP und WRIGHT verwiesen.

In der Ableitung der oben genannten Formeln für die Varianzen und Kovarianzen wurde vorausgesetzt, daß keine Kopplung zwischen den Genen vorliegt. Für den Fall der Kopplung ändern sich die Varianzen von Generation zu Generation. Die Mittelwerte werden dagegen hierdurch nicht beeinflusst. Eine Prüfung auf Kopplung kann man nun so durchführen, daß die Werte für D, H, E_1 und E_2 noch einmal berechnet werden, wenn man die Varianz einer spaltenden Generation wegläßt oder die einer anderen hinzunimmt und die Homogenität der in beiden Fällen für D, H, E_1 und E_2 erhaltenen Werte prüft. Sind sie inhomogen, so schließt MATHER auf Kopplung der Gene.

Für den hier besprochenen Fall könnte man die Berechnung für D, H, E_1 und E_2 noch einmal unter Ausschluß des Wertes für \bar{V}_{F_3} durchführen. Die Varianz der F_3 -Mittelwerte, V_{F_3} , ist mit der F_2 -Varianz auch bei Kopplung vergleichbar, da, wie eben bereits ausgeführt, die Mittelwerte durch Kopplung nicht verändert werden. Dasselbe gilt für die Kovarianz zwischen den F_2 -Werten und den Mittelwerten ihrer F_3 -Nachkommenschaften, W_{F_2/F_3} , zu deren Berechnung nur die Mittelwerte der F_3 -Generation benutzt werden.

Alle bisher besprochenen Methoden setzten voraus, daß die in der Analyse benutzten Eltern weitgehend homozygot sind. Sie eignen sich also besonders für Selbstbe-

fruchter und solche Fremdbefruchter, von denen leicht homozygote Inzuchtlinien herzustellen sind. MATHER (1949 a) gibt auch ein Verfahren zur Schätzung der Dominanz in fremdbefruchtenden Populationen an, zu dem nur die Korrelationskoeffizienten zwischen Vollgeschwistern und die zwischen Eltern und Nachkommen benötigt werden.

Unter der Annahme, daß die Häufigkeit der Allele der einzelnen Polygene in der Population gleich ist und daß die Wirkung der Gene additiv ist, kann man zeigen, daß die Varianz einer beliebigen Generation einer fremdbefruchtenden Population wie bei Selbstbefruchtern gleich $1/2 D + 1/4 H + E$ ist, daß ferner die Kovarianz zwischen Vollgeschwistern $W_{S/S} = 1/4 D + 1/16 H$ und die Kovarianz zwischen Eltern und Nachkommen $W_{P/O} = 1/4 D$ ist. Der Korrelationskoeffizient zwischen Eltern und Nachkommen ist nun $r_{P/O} = \frac{W_{P/O}}{\sqrt{V_P \cdot V_O}}$. Da man $V_P = V_O$ setzen kann, so vereinfacht sich die Formel zu

$$r_{P/O} = \frac{W_{P/O}}{V} = \frac{1/4 D}{1/2 D + 1/4 H + E}$$

Ganz entsprechend ist die Korrelation zwischen Vollgeschwistern

$$r_{S/S} = \frac{W_{S/S}}{V} = \frac{1/4 D + 1/16 H}{1/2 D + 1/4 H + E}$$

Die Kenntnis dieser beiden Korrelationskoeffizienten genügt also zur Schätzung des Verhältnisses zwischen H und D sowie zwischen E und D .

$$\frac{H}{D} = \frac{4(r_{S/S} - r_{P/O})}{r_{P/O}}, \quad \frac{E}{D} = \frac{1 - 2r_{P/O}}{4r_{P/O}}$$

Als letzte Methode zur Schätzung der Dominanz der Polygene soll noch kurz ein Verfahren von COMSTOCK und ROBINSON (1948) erwähnt werden. Diese Autoren benutzen für ihre Untersuchungen Biparentalnachkommenschaften, also Kreuzungsnachkommenschaften aus F_2 -Pflanzen und Rückkreuzungsgenerationen.

An einem einfachen Beispiel (COMSTOCK und ROBINSON, 1948) soll die Methode kurz beschrieben werden:

Aus der F_2 zweier homozygoter Ausgangseltern liest man zufällig $m \cdot n$ Pflanzen aus und kreuzt je n Mutterpflanzen mit einem der vorhandenen m Vaterpflanzen. Bringt man von jeder Kreuzungsnachkommenschaft $r \cdot k$ Pflanzen in den Versuch und von diesen je k Pflanzen in r Parzellen, so kann bei geeigneter Versuchsanlage, additive Wirkung der Gene vorausgesetzt, die Varianz zwischen den einzelnen Individuen wie folgt aufgeteilt werden:

Varianztabelle (Nach COMSTOCK u. ROBINSON 1948)

Varianzursache	F G	Varianz	Komponenten d. Varianz
Väter	$m-1$	M_1	$s^2 + ks_p^2 + rks_j^2 + nrks_m^2$
Mütter innerhalb der Väternachkommenschaft	$m(n-1)$	M_2	$s^2 + ks_p^2 + rks_j^2$
Parzellen innerhalb der Mütternachkommenschaft	$mn(r-1)$	M_3	$s^2 + ks_p^2$
Innerhalb d. Parzellen	$mnr(k-1)$	M_4	s^2

s^2 ist die Summe der umweltbedingten und genetischen Varianz innerhalb einer Parzelle, s_p^2 die Parzellenvarianz, s_j^2 die durch genetische Differenzen der verschiedenen Mutterpflanzen und s_m^2 die durch genetische Differenzen der verschiedenen Vaterpflanzen bedingte Varianz. COMSTOCK und ROBINSON können nun zeigen, daß der Ausdruck

$$a = \frac{2[(M_2 - M_3) - (M_1 - M_2)]}{M_1 - M_2} \quad (9)$$

ein Maß für die durchschnittliche Dominanz der Gene ist. a entspricht dem Wert h/d in den Formeln von FISHER und MATHER. Größere Werte für a als 1, die also Superdominanz anzeigen, können auch durch Repulsionskopplung, also durch Kopplung von Plus- und Minus-Genen, die voll oder auch nur teilweise dominant zu sein brauchen, vorgetäuscht werden. Durch eine Ausdehnung der Versuche auf spätere Generationen kann man nun wieder einen Anhaltspunkt dafür erhalten, ob Kopplung die Ursache für einen größeren Wert für a als 1 ist. Setzt man einen gewissen Austausch zwischen den in Repulsionskopplung sich befindenden + und - Genen voraus, so wird sich der Wert für die durchschnittliche Dominanz, den man aus den Experimenten mit Nachkommenschaften späterer Generationen erhält, verringern. Ist aber eine echte Superdominanz die Ursache für einen Wert von a , der größer als 1 ist, so wird man diesen auch in späteren Generationen innerhalb der zufälligen Schwankungen wieder erhalten.

Neben dem oben beschriebenen Versuch geben COMSTOCK und ROBINSON (1948) die Berechnung der durchschnittlichen Dominanz für zwei weitere Experimente an, die kurz folgendermaßen beschrieben werden sollen:

Das erste ist eine Abwandlung des oben beschriebenen Versuchs, die dann möglich ist, wenn man eine Mutterpflanze gleichzeitig mit mehreren Vaterpflanzen bestäuben kann. Jede der n zur Kreuzung ausgelesenen Mutterpflanzen wird in diesem Fall mit jedem der m Väter bestäubt, so daß man wieder $m \cdot n$ Kreuzungsnachkommenschaften erhält.

In der zweiten Versuchsanlage werden F_2 -Pflanzen als Väter mit beiden Eltern zurückgekreuzt und die Rückkreuzungsgenerationen analysiert. Der Versuch besteht also in diesem Fall aus einer Reihe von Nachkommenschaften, die paarweise einen gemeinsamen F_2 -Vater, aber zwei verschiedene Mütter, nämlich die beiden homozygoten Eltern haben. Diese beiden Eltern sind für jede der paarweise zusammengehörenden Nachkommenschaften gleich. Die Berechnung der Dominanz der Gene wird in diesen beiden Versuchen entsprechend der veränderten Versuchsanlage etwas anders durchgeführt. Die Formeln, auf die hier nicht näher eingegangen werden soll, sind bei COMSTOCK und ROBINSON (1952) angegeben.

Die Schätzung der Genzahl

Die Möglichkeit, die Genzahl eines Polygensystems abschätzen zu können, ist theoretisch von Interesse; eine annähernd genaue Kenntnis der Zahl der Gene, in denen sich zwei zu kreuzende reine Linien in bezug auf eine quantitative Eigenschaft unterscheiden, ist aber auch für den praktischen Züchter mindestens von ebenso großer Bedeutung. Sie ist genau so wichtig wie die Kenntnis über das Zusammenwirken der Gene und ihrer Dominanzverhältnisse. Erst wenn

man über diese drei Eigenschaften der Polygene mit einiger Sicherheit Aussagen machen kann, hat man die Möglichkeit, die Selektionsreserven, die in einer Kreuzung liegen, annähernd abzuschätzen. Denn diese liegen ja doch letztlich in der Größe der fixierbaren Komponente der erblichen Varianz spaltender Generationen, dem Wert D bei MATHER also, und in der Genzahl begründet.

Kann man diese beiden Werte, nämlich D und die Genzahl, mit großer Sicherheit abschätzen, so ist man — unter der Annahme gleicher d -Werte aller Gene — in der Lage, die selektiven Grenzen, also die Werte der vom Mittelwert aus nach beiden Richtungen extrem möglichen Genotypen, abzuschätzen. Ihre Differenz von diesem Mittelwert beträgt nämlich \sqrt{nD} (n = Zahl der Gene), da ja D unter der Annahme gleicher d -Werte für alle Gene gleich nd^2 ist. \sqrt{nD} ist also gleich $\pm n d$ und gibt in der positiven Lösung der Wurzel den Wert desjenigen Genotyps an, der alle + Allele besitzt und in der negativen den alle — Allele enthaltenden Genotyp. Im Züchtungsexperiment könnte man also damit ziemlich genau angeben, wie hoch der Züchtungserfolg sein wird.

Die Methoden zur Schätzung der Genzahl sind bis heute jedoch noch keineswegs befriedigend gelöst. Zu ihrer Anwendung müssen nämlich noch stärker einschränkende Voraussetzungen gemacht werden, als dies schon bei der Schätzung der durchschnittlichen Dominanzverhältnisse der Fall war. So wird durch die nachfolgende, bei WRIGHT (1934) und SEREBROWSKI (1928) angegebene Methode im günstigsten Fall nur die Mindestzahl an Genen, in denen sich die Ausgangseltern voneinander unterscheiden, geschätzt. Die am stärksten einschränkende Voraussetzung, die hier gemacht wird, ist die, daß der eine Elter nur die + Allele und der andere nur die — Allele enthalten muß. Zudem gelten die Formeln nur dann, wenn additive, intermediäre und gleiche Wirkung jedes Gens vorliegt.

Bezeichnet man die Wirkung eines Gens nach der in dieser Arbeit bereits mehrfach verwandten Symbolisierung mit d , so ist die Differenz zwischen beiden extremen Eltern $2nd$. Der genetische Teil der F_2 -Varianz ist nach (8) gleich $1/2 S(d^2) + 1/4 S(h^2)$. Unter der hier gemachten Voraussetzung der intermediären und gleichen Wirkung aller Gene vereinfacht sich diese Formel zu $nd^2/2$, da $h = 0$ ist. Der nicht genetische Anteil an der F_2 -Varianz kann durch die Varianz der Elterngeneration geschätzt werden, so daß man $nd^2/2$ aus $\sigma_{F_2}^2 - \sigma_p^2$ errechnen kann.

Bezeichnet man die Differenz zwischen den beiden Eltern mit Δ , so ist n die Zahl der Gene gleich

$$n = \frac{\Delta^2}{8(\sigma_{F_2}^2 - \sigma_p^2)} \quad (10)$$

Wie bereits erwähnt, kann man nach dieser Formel nur die Mindestzahl der Gene schätzen und würde sogar in dem Grenzfall, in dem die Eltern zwar verschiedene, aber dieselbe Anzahl von Leistungsgenen enthalten, phänotypisch also gleich sind, den Wert Null erhalten, obwohl sich die Eltern in einer mehr oder weniger großen Zahl von Genen unterscheiden. Jede Dominanz, gleich welchen Grades, bewirkt ebenfalls eine Unterschätzung der Genzahl nach dieser Formel. Die F_2 -Varianz ist dann nämlich größer als $nd^2/2$, weil in diesem Fall die durch die Abweichung der Heterozygoten vom Mittelwert der beiden Homozygoten bedingte zusätzliche erbliche Varianz, also die H -Komponente MATHERS, noch hinzukommt (siehe Formel 8).

Der Wert des Nenners in Formel (10) entspricht somit wegen der oben gemachten Voraussetzung einer intermediären Genwirkung dem Wert D in Formel (8). Hat man mit Hilfe des MATHERschen Verfahrens die Komponenten D und H der erblichen Varianz in spaltenden Generationen errechnet, so braucht man in Formel (10) also nur D in den Nenner einzusetzen, und die Formel ist jetzt auch dann anwendbar, wenn Dominanz der Gene vorliegt. MATHER betont, daß man in diesem Fall auch mit Hilfe der H -Komponente die Schätzung der Genzahl vornehmen kann, wenn man nämlich die Voraussetzung macht, daß alle Gene denselben Dominanzgrad haben. In diesem Fall ist $H = S(h^2) = nh^2$, und die Differenz zwischen dem F_1 - und dem Mittelelterwert gleich nh . Das Quadrat dieser Differenz dividiert durch H gibt jetzt wieder den Wert für n . Allerdings würde jede Ungleichheit der h -Werte der einzelnen Gene und vor allem das Vorhandensein von z. T. positiven, z. T. negativen h -Werten (also von teilweise positiver und teilweise negativer Dominanz) wieder eine zu kleine Schätzung der Genzahl ergeben.

In diesem Zusammenhang ist die oft in Züchtereisen diskutierte Frage von Interesse, wie weit man aus der Leistung der F_1 oder F_2 verschiedener Kreuzungen Hinweise darauf erhalten kann, welche dieser Kreuzungsnachkommenschaften die leistungsfähigsten Genotypen abspalten können. Hat man eine Skala gefunden, auf der die Wirkung der Gene weitgehend additiv ist, so ist bei Vorliegen von Dominanz und annähernd gleichen H -Werten für alle Gene der F_1 -Wert auf dieser Werteskala schon ein Anhaltspunkt für den Wert des besten aus dieser Kreuzung auszulesenden Genotyps. Er ist nämlich, wie eben gezeigt wurde, gleich nh und wird also um so größer sein, je mehr Gene an der Spaltung beteiligt sind. Je mehr Gendifferenzen aber zwischen den Eltern vorhanden sind, desto höher ist der Wert des leistungsfähigsten Genotyps, der aus dieser Kreuzung überhaupt ausgelesen werden kann. Das würde also bedeuten, daß von mehreren durchgeführten Kreuzungen unter der Annahme eines weitgehend gleichmäßigen h -Wertes aller an der Kreuzung beteiligten Gene diejenige Kreuzung, die den höchsten F_1 -Wert ergibt, auch die leistungsfähigsten Typen abspalten wird.

Genau so ist unter Voraussetzung der eben gemachten Bedingungen auch der F_2 -Wert ein gewisses Maß für die Selektionschancen aus einer Kreuzung. Da er aber nur noch $1/2 nh^1$ beträgt, sind vorhandene Differenzen zwischen F_2 -Werten von Kreuzungen mit verschiedenen sich unterscheidenden Genzahlen im Verhältnis zur F_1 weniger stark ausgeprägt und somit schlechter erfaßbar.

Zum anderen wird die F_2 oft deswegen besser zur Untersuchung dieser Frage geeignet sein, da von ihr schon eine größere Zahl von Pflanzen zur Verfügung steht, während das bei der F_1 oft nicht der Fall ist. Im allgemeinen werden die Aussagen, die man aus Vergleichen zwischen F_1 - und F_2 -Werten in dieser Hin-

sicht machen kann, naturgemäß um so sicherer sein, je größer der Wert für h , je größer also die Dominanz ist, da dann auch die Differenzen zwischen F_1 - und F_2 -Werten der einzelnen Kreuzungen größer werden.

Ganz ähnlich, wie eben die Berechnung der Genzahl mit Hilfe der F_2 -Varianz durchgeführt wurde, kann man statt dieser auch die Varianz der Rückkreuzungsgeneration verwenden, die nur halb so groß ist wie die der F_2 . Man erhält dann

$$n = \frac{\Delta}{4(\sigma_R^2 - \sigma_p^2)} \quad (11)$$

Neben diesen auf die Arbeiten WRIGHTS und SEREBROWSKIS zurückgehenden Schätzmethode hat PANSE (1940) eine Berechnung der Genzahlen mit Hilfe der Varianz von F_3 -Familien vorgeschlagen, die nicht zur Voraussetzung hat, daß alle Leistungsgene in dem einen Elter und ihre Allele in dem anderen vorhanden sein müssen.

Zur Ableitung der Formel sollen wieder die bereits bekannten Symbole d und h benutzt werden. Die genetische Varianz der F_2 ist gleich $1/2 D + 1/4 H$. Sie soll mit ${}_G V_{F_2}$ bezeichnet werden. Die durchschnittliche genetische Varianz der F_3 , ${}_G \bar{V}_{F_3}$, ist $1/4 D + 1/8 H$ (siehe S. 270). Bezeichnet man die genetische Varianz der F_2 mit x , so ist die der F_3 gleich $1/2 x$. Die Varianz der Varianzen der F_3 , $V({}_G V_{F_3})$, ist $1/4 x^2$.

Das kann man sich leicht folgendermaßen klarmachen: Nimmt man zunächst einmal an, daß sich die Eltern nur in einem Genpaar unterscheiden, dann besteht die F_2 aus den Genotypen AA , Aa , Aa und aa , von denen nur die beiden Heterozygoten in der F_3 spalten. Ihre Varianz ist gleich der Varianz der F_2 , also gleich x . Die 4 F_3 -Familien haben demnach in der Reihenfolge, wie die Elternpflanzen eben aufgeführt wurden, die Varianzen 0 , x , x und 0 . Die Varianz dieser Werte, also $V({}_G V_{F_3})$, ist, wie man sofort sieht, $1/4 x^2$.

Ist der Wert für d und h für alle Gene gleich, so kann man diese Formeln auf beliebig viele Gene ausdehnen. Bei n vorhandenen Genen ist $V({}_G V_{F_3})$ dann gleich $1/4 nx^2$ und ${}_G \bar{V}_{F_3} = 1/2 nx$. Durch Division von ${}_G \bar{V}_{F_3}^2$ durch $V({}_G V_{F_3})$ erhält man

$$n = \frac{{}_G \bar{V}_{F_3}^2}{V({}_G V_{F_3})} \quad (11)$$

Wenn für diese Formel auch nicht mehr die Voraussetzung einer derartigen Genverteilung notwendig ist, daß der eine Elter nur Leistungsgene und der andere die entsprechenden Allele besitzt, so müssen die d - und h -Werte für alle Gene doch noch weitgehend gleich sein, um einen annähernd richtigen Wert für n zu erhalten.

Eine weitere und sehr wesentliche Ursache für die Unterschätzung der Genzahl nach beiden Methoden ist die Kopplung der Gene. In den meisten in der Literatur angegebenen Versuchen, in denen eine Schätzung der Genzahl nach diesen Methoden vorgenommen wurde, war die errechnete Zahl auch dementsprechend niedrig. (siehe MATHER, 1949 a, S. 128). Man muß sich wohl darüber im klaren sein, daß die hier erhaltenen Zahlen nicht die tatsächlich vorhandenen Gendifferenzen der beiden Ausgangseltern anzeigen. PANSE und auch MATHER (1949) betonen das ausdrücklich und weisen darauf hin, daß die in ihren Berechnungen erhaltenen Zahlen für n in Formel (10) und (11) nicht

¹ Ganz allgemein liegt der Wert irgendeiner Filialgeneration bei additiver Genwirkung von der F_2 ab unabhängig von allen Dominanzverhältnissen genau zwischen dem Mittelelterwert und dem Wert der vorhergehenden Generation und ist unter der obigen Annahme, daß der Mittelelterwert den Wert 0 hat, gleich dem halben Wert der vorhergehenden Generation. In diesem Fall wäre also ganz allgemein $F_n = 1/2 F_{n-1}$.

die Zahl der Gene, sondern die der „effective factors“, wie sie es nennen, angeben. Welche Einheiten die „effective factors“ darstellen, ist im einzelnen schwer zu sagen. Man könnte sich vorstellen, daß es ganze Chromosomenabschnitte oder, bei sehr geringer Chiasmenzahl je Chromosom, gar ganze Chromosomen sind, die sich in der Analyse als spaltende Einheiten zeigen.

Zur weiteren Klärung dieser Frage wird man bestrebt sein müssen, erstens ein größeres Material nach diesen schon bekannten Methoden durchzuprüfen und zweitens durch eine Verbesserung der schon vorhandenen und Herausarbeitung neuer Methoden einen tieferen Einblick in diese Frage zu erhalten.

Schlußbetrachtungen

Die in der vorliegenden Arbeit dargestellten Methoden zur Untersuchung der Wirkungsweise quantitativer Gene lassen deutlich erkennen, daß man trotz zahlreicher Schwierigkeiten durchaus erfolgversprechende Wege in der Erforschung quantitativer Genvererbung mit Hilfe bestimmter statistischer Verfahren aufzeigen konnte. Sie unterstreichen aber auch die einleitend erwähnte Tatsache, daß zu ihrer Anwendung vereinfachende Voraussetzungen gemacht werden müssen, die nicht immer erfüllt sind und daher die praktische Bedeutung der entwickelten Methoden bis zu einem gewissen Grade fraglich erscheinen lassen. Es muß aber ausdrücklich betont werden, daß die hier erörterten statistischen Methoden vorläufig als die geeignetsten erscheinen und daß es zumindest fraglich ist, ob wir auf irgendeinem anderen Weg schneller zum Ziel kommen. Da das Gebiet der quantitativen Genvererbung nun aber nicht nur dem wissenschaftlich Arbeitenden ein reiches Betätigungsfeld bietet, sondern die Lösung der hier behandelten Probleme auch für die praktische Züchtung von großer Bedeutung ist, erscheint eine ausführliche Beschäftigung mit diesem Fragenkomplex durchaus notwendig zu sein.

Im praktischen Zuchtbetrieb wird ohnehin Jahr für Jahr Versuchsmaterial von einem derartigen Umfang geprüft, wie es an rein wissenschaftlichen Instituten oft arbeitsmäßig nicht möglich, für die hier kurz besprochenen statistischen Methoden aber erforderlich ist. Die Verarbeitung dieses Materials nach diesen Methoden kann dem Züchter sicher wertvolle Anhaltspunkte für die anzuwendenden Zuchtverfahren geben. Das ist zumal dann der Fall, wenn sich die Untersuchungen über mehrere Jahre erstrecken. Durch die Prüfung eines bestimmten Materials nach den hier beschriebenen Methoden über mehrere Generationen wäre dann auch eine Möglichkeit gegeben, den Einfluß von Kopplungserscheinungen auf die Größe der Werte für den Dominanzgrad und die Genzahl zu untersuchen. Die Rechenarbeit wird sich zwar gegenüber den bisher üblichen, in der praktischen Pflanzen-

züchtung angewandten Methoden beträchtlich erhöhen. Jedoch erscheint bei dem im allgemeinen relativ hohen Aufwand in der heutigen Züchtung eine gründlichere genetisch-statistische Auswertung des Materials durchaus gerechtfertigt. Darüber hinaus würde sie nicht nur dem Züchter wertvolle Hinweise geben können, sondern sicher auch zur Lösung einiger, bisher noch sehr problematischer Fragen bezüglich der quantitativen Genvererbung beitragen können.

Literatur

1. CHARLES, D. R., and H. H. SMITH: Distinguishing between two types of gene action in quantitative inheritance, *Genetics*, **24**, 34—48, (1939). — 2. COMSTOCK, R. E. and H. F. ROBINSON: The components of genetic variance in populations of biparental progenies and their use in estimating the average degree of dominance, *Biometrics* **4**, 254—266, (1948). — 3. COMSTOCK, R. E. and H. F. ROBINSON: Estimation of average dominance of genes, *Heterosis* **S. 494—516**, Iowa state college press, (1952). — 4. COPP, L. G. L. and G. M. WRIGHT: The inheritance of kernel weight in a triticum cross, *Heredity* **6**, 187—199, (1952). — 5. DARLINGTON, C. D. and K. MATHER: The elements of Genetics, Methuen u. Co. Ltd., London, (1950). — 6. EAST, E. M.: A mendelian interpretation of variation that is apparently continuous, *Amer. Nat.* **44**, 65—82, (1910). — 7. FISHER, R. A.: The correlation between relatives on the supposition of mendelian inheritance, *Trans. Roy. Soc. Edin.* **52**, 399—433, (1918). — 8. FISHER, R. A., F. R. IMMÉR and O. TÈDIN: The genetical interpretation of statistics of the third degree in the study of quantitative inheritance, *Genetics* **17**, 107—124. — 9. ГОРОХ, К.: Genetic studies on eggplant (*Solanum melongena*). I. Regression analysis of quantitative gene action, *Genetica* **26**, 445—452, (1950). — 10. GRIFFING, B.: Analysis of quantitative gene action by constant parent regression and related techniques, *Genetics* **35**, 303 bis 321, (1950). — 11. KAPPERT, H.: Die vererbungswissenschaftlichen Grundlagen der Züchtung, Verl. Parey, Berlin u. Hamburg, (1953). — 12. MACARTHUR, J. W. and L. BUTLER: Size inheritance and geometric growth processes in the tomato fruit, *Genetics* **23**, 253—268, (1939). — 13. MATHER, K.: Dominance and heterosis, *Am. Nat.* **80**, 91—96, (1946). — 14. MATHER, K.: *Biometrical Genetics*, Methuen u. Co. Ltd., London, (1949a). — 15. MATHER, K.: The genetical theory of continuous variation, Supplementary volume of *Hereditas*, **1949**, 376 bis 401, (1949b). — 16. NILSSON-EHLE, H.: Kreuzungsuntersuchungen an Hafer und Weizen, *Lunds. Univ. Arscrift*, **N. F. Avd. 2**, Bd. 5 und Bd. 7, (1909 u. 1911). — 17. PANSE, V. G.: Application of genetics to plant breeding. II. The inheritance of quantitative characters and plant breeding, *Journ. Gen.* **40**, 283—302, (1940). — 18. RASMUSSEN, J.: A contribution to the theory of quantitative character inheritance, *Hereditas* **18**, 245—261, (1933—34). — 19. SEREBROWSKY, A. S.: An analysis of the inheritance of quantitative transgressive characters, *Ztschr. Ind. Abst. Verbgsl.* **48**, 229—243, (1928). — 20. SNEDECOR, G. W.: *Statistical methods*, Iowa state college press, (1950). — 21. WRICKE, G.: Ein Fall von Superdominanz bei einer experimentell hergestellten Autotetraploiden von *Arabidopsis thaliana*, *Ztschr. Ind. Abst. Verbgsl.* (im Druck), (1955). — 22. WRIGHT, S.: The results of crosses between inbred strains of guinea pigs, differing in number digits, *Genetics* **19**, 537—551, (1934).