Quantitative population-epigenetics in screening and development of regulator-active compounds

Quantitative Populations-Epigenetik beim Screening und der Entwicklung von regulatorisch aktiven Substanzen

Reinhold Stauß
Ritzebeker Weg 13, D-24222 Schwentinental Ortsteil Klausdorf
Reinhold.Stauss@googlemail.com

DOI: 10.5073/jka.2012.434.074

Summary

Quantitative (Population-) Epigenetics describes the variability observed in characters due to factors in the environment induced primarily by the gamut of chemical compounds screened for regulatory efficacy.

Likewise, index selection based on the statistical Quantitative (Population-) Epigenetics theory can be used to improve efficiency in screening compounds for their potential to enhance quantitative characters such as yield, stability and resistance to unfavourable environmental influences (e.g., water stress, cold temperatures or disease resistance).

1. Optimal screening efficiency of regulator-active compounds is obtained with
2. high environmental variability,
3. low heritability (characters for which the genotype sets a wide 'norm of reaction' on environmental influences),
4. high correlation between characters under indirect selection and intensity of selection is shown.

Keywords: Biologically-active chemical compounds, efficacy value, environmental variation, genetic analogues, genotype, heritability, index selection, phenotype, quantitative characters, screening efficiency

Zusammenfassung

Quantitative (Populations-) Epigenetik beschreibt die durch regulatorische Substanzen als Umweltfaktoren bewirkte Variabilität von Merkmalen.

Die Anwendung von Indexselektion auf Grundlage der statistischen Quantitativen (Populations-) Epigenetik-Theorie zur Erhöhung der Screeningeffizienz von chemischen Substanzen bei quantitativen landwirtschaftlichen Merkmalen wie zum Beispiel Ertrag, Standfestigkeit oder Erhöhung der Widerstandsfähigkeit gegen negative Umweltfaktoren (z. B. Wassermangel, Frost oder Krankheiten), wird dargestellt.

1. Optimale Screeningeffizienz bei regulatorisch wirksamen Substanzen wird erzielt
2. bei großer Umweltvariabilität,
3. bei niedriger Heritabilität (Merkmale, bei denen der Genotyp eine große Reaktionsbreite ('norm of reaction') für die Modifikation durch die Umwelt zulässt),
4. bei hoher Korrelation zwischen Hilfsmerkmal und Zielmerkmal bei indirekter Selektion und mit hoher Selektionsintensität.

Stichwörter: Biologisch aktive chemische Substanzen, genetische Analogie, Genotyp, Heritabilität, Phänotyp, quantitative Merkmale, Screeningeffizienz, umweltbedingte Variabilität, Wirkungswert

1. Introduction

The potential for using plant-growth regulators in agriculture is far-reaching. Yet, despite extensive research on plant-growth regulators, only a few chemical compounds have achieved practicability. Lack of success may be attributed to two factors: up to now, either chemists have not devised such biologically active compounds or screening and development procedures are not adequately sensitive to detect effects on quantitatively inherited characters.

This publication propounds to render screening more efficient by taking into account laws of

inheritance. The argument undertakes to show that statistical (epi-)genetic theory as a basis for developing regulator screening methods may be appropriated with the same facility as is done in plant and animal breeding schemes. The research discipline and the treatment subject are the same for both the breeder and the investigator of regulatory agents, save each treats different sides of the same coin (organism). The breeder endeavours to improve the genotype - for him environments are 'fixed' effects; the chemical researcher is not able to augment the genotype - one strives to intervene in the environment by effecting a specific phenotypic expression with a chemical compound within the 'norm of reaction' inherent in the genotype.

Likewise index selection based on statistical epigenetic theory can be used to improve efficiency in screening compounds for their potential to enhance quantitative characters such as yield, stability and resistance to unfavourable environmental influences (e.g., water stress, cold temperatures, disease resistance) - as well indeed, for potential in pharmacological intervention.

Ecological and Evolutionary Epigenetics is a new field of frontier research at the intersection between molecular genetics and evolutionary ecology. The term 'Epigenetics' has been used only since about ten years. The statistical Quantitative (Population-) Epigenetics theory was published with "Genetic analogues in chemical screening" in 1992 (STAUSS, 1992).

2. Quantitative Variation and the Concept of Heritability

Concerning the inheritance of quantitative characters, it is manifest that the response in the environment is not known precisely. Nonetheless, the measure of the action of an agent ('efficacy value' [Bc]) is determined as the average mean of verifiable attributes (phenotypes) for a chemical effect taken over all random samplings in the environment. (It is understood that a given 'efficacy value' [Bc] is contingent upon described dosage).

"In a strict sense, the question of whether a characteristic is hereditary or environmental has no meaning. The genes cannot cause a character to develop unless they have the proper environment, and, conversely, no amount of manipulation of the environment will cause a characteristic to develop unless the necessary genes are present. Nevertheless, we must recognize that the variability observed in some characters is caused primarily by differences in the genes carried by different individuals and that the variability in other characters is due primarily to differences in the environments to which individuals have been exposed." (ALLARD, 1960).

Therefore the prerequisite for an efficient screening of chemical compounds which could influence a designated quantitative character is the variability of this character induced by given environmental factors.

Quantitative Variation: $P = E + G$

(where *P* = phenotypic value; *E* = response of the environment and/or a chemical compound; *G* = deviation due to different genotypes).

Heritability:

The proportion of the total phenotypic variance (σ_P^2) contingent upon *genetic* differentiation is a measure of the amount of genetic variability (σ_G^2) of the total variance for a character in question (heritability in the broad sense [h^2]).

By transformation of the equation of the heritability definition

$$h^2 = \frac{\sigma_G^2}{\sigma_P^2} = \frac{\sigma_G^2}{\sigma_G^2 + \sigma_E^2},$$

the following expression is derived for $1-h^2$:

$$1-h^2=\frac{\sigma_E^2}{\sigma_P^2}.$$

The proportion of the total phenotypic variance (σ_P^2) that is due to *environmental* effects is a measure of the extent of environmental variability (σ_E^2) of the total variance for a character under observation (one minus heritability in the broad sense [$1-h^2$]).

When all variation is due to genetic cause or if no environmental variance obtains, $h^2=1$ or 100 % expressed as a percentage; as the environmental element of variance increases, h^2 decreases.

Environmental variance:

The main components of environmental variance are 1) 'chemical deviation', $\sigma_{E_c}^2$ (response-effect deviations for an applied agent accruing to all genotypes and random environmental integrants), 2) 'random environmental deviation', $\sigma_{E_r}^2$ (deviation effected by environmental constituents) and 3) 'interactional deviation', $\sigma_{E_i}^2$ (deviation resulting from interaction between a chemical compound and environmental factors).

The 'efficacy value' [Bc], by definition, is not correlated to dominance or epistatic effects. Furthermore, it is assumed that there is no correlation due to effects between genotypes and environmental factors and that there is no association between an applied agent and the environment.

Conclusion:

The smaller the heritability the greater is the facility to influence a character by the environment or by a biologically active chemical agent as an integrant of the environment.

3. The Significance of the Heritability Factor

The relationship between the breeding value or the 'efficacy value' [Bc] of a chemical compound and a phenotype may be delineated by means of appropriating biometric expressions: covariance, correlation and regression.

The *correlation* between the 'efficacy value' [Bc] and a phenotypic value is:

$$\rho_{EP}=\frac{\mathrm{cov}(E,P)}{\sigma_E\sigma_P}=\frac{\sigma_E^2}{\sigma_E\sigma_P}=\frac{\sigma_E}{\sigma_P}=\sqrt{1-h^2}.$$

The *regression* between the 'efficacy value' [Bc] and a phenotypic value is then:

$$\beta_{EP}=\frac{\mathrm{cov}(E,P)}{\sigma_P^2}=\frac{\sigma_E^2}{\sigma_P^2}=1-h^2.$$

Conclusions:

1. The regression of the 'efficacy value' [Bc] of a chemical compound to a phenotypic value is equal to one minus heritability ($1-h^2$).
2. The correlation between the 'efficacy value' [Bc] of a compound and a phenotypic value is the square root of one minus heritability ($\sqrt{1-h^2}$).

The importance of using one minus heritability as a regression is in estimating the 'efficacy value' [Bc] of a chemical compound with the help of a phenotypic value:

$$\hat{Bc} = \hat{E} = \beta_{EP}(P-\mu) = (1-h^2)(P-\mu).$$

To see how accurate this estimation is, the correlation between the estimated and the true values can be calculated thusly:

$$\rho_{Bc,\hat{Bc}} = \rho_{E\hat{E}} = \frac{\operatorname{cov}\left(E,\hat{E}\right)}{\sigma_E \sigma_{\hat{E}}} \qquad (3.1)$$

with

$$\hat{\mathrm{E}} = \beta_{\mathrm{EP}}(\mathrm{P}-\mu);$$

so that

$$\operatorname{cov}(\mathrm{E},\hat{\mathrm{E}}) = \operatorname{cov}\left[\mathrm{E},\beta_{\mathrm{EP}}(\mathrm{P}-\mu)\right] = \beta_{\mathrm{EP}}\operatorname{cov}(\mathrm{E},\mathrm{P}) \qquad (3.2)$$

and

$$\sigma_{\hat{\mathrm{E}}} = \sigma_{\beta_{\mathrm{EP}}}(\mathrm{P}-\mu) = \beta_{\mathrm{EP}}\sigma_{\mathrm{P}}. \qquad (3.3)$$

Applying (3.2) and (3.3) to (3.1), the reliability of an estimated 'efficacy value' [Bc] may be computed

$$\begin{aligned}\rho_{Bc,\hat{Bc}} &= \rho_{E\hat{E}} \\ &= \frac{\beta_{EP}\operatorname{cov}(E,P)}{\sigma_E \beta_{EP}\sigma_P} \\ &= \frac{\operatorname{cov}(E,P)}{\sigma_E \sigma_P} \\ &= \sigma_{EP} \\ &= \sqrt{1-h^2}.\end{aligned}$$

Conclusion:

The square root of one minus heritability is a reliability measure of an 'efficacy value' [Bc] estimation of a chemical compound as the function of a phenotypic value.

4. Correlation of Characters

The relationship (correlated variation) between two or more characters can be quantified by means of the biometric expressions

3. covariance and
4. correlation.

The methodology must take into account correlation between characters -- 'genetic character correlation' and/or 'reciprocal response' to mutual environmental factors.

By this correlation it is inferred that a modification in one character induces alteration in another.

This can be used,

1. in doing indirect selection, and
2. in doing direct selection with the intention of enhancing two or more characters (albeit at times negatively correlated).

Taking P_1 and P_2 as phenotypic values of an organism,

$$P_1 = \mu_1 + A_1 + E_1^{'} \text{ and } P_2 = \mu_2 + A_2 + E_2^{'}$$

where A_1 and A_2 express the breeding values of two characters; $E_1^{'}$ and $E_2^{'}$ denote the respective sums of factors in an 'environmental Syndrome', plus the effects of 1) dominance, 2) epistasis and 3) genotypic-environmental interaction.

Phenotypic covariance may then be calculated:

$$\begin{aligned}\operatorname{cov}(P_1, P_2) &= E\left[\left(P_1 - \mu_1\right)\left(P_2 - \mu_2\right)\right] \\ &= E\left[\left(A_1 + E_1^{'}\right)\left(A_2 + E_2^{'}\right)\right] \\ &= E\left(A_1 A_2\right) + E\left(A_1 E_2^{'}\right) + E\left(E_1^{'} A_2\right) + E\left(E_1^{'} E_2^{'}\right).\end{aligned}$$

Assuming that genotypes are distributed randomly in environments,

$$E\left(A_1 E_2^{'}\right) = E\left(E_1^{'} A_2\right) = 0.$$

By substitution the formulation of phenotypic covariance is rendered

$$\operatorname{cov}\left(P_1, P_2\right) = \operatorname{cov}\left(A_1, A_2\right) + \operatorname{cov}\left(E_1^{'}, E_2^{'}\right).$$

To compare correlations of associated characters it is necessary to standardize covariances. This is carried out by the formulation

$$\begin{aligned}\rho_P^{'} &= \frac{\operatorname{cov}\left(P_1, P_2\right)}{\sigma_{P_1}\sigma_{P_2}} \\ &= \frac{\operatorname{cov}\left(A_1, A_2\right) + \operatorname{cov}\left(E_1^{'}, E_2^{'}\right)}{\sigma_{P_1}\sigma_{P_2}}\end{aligned}$$

where

$$\operatorname{cov}\left(A_1, A_2\right) = \rho_A \sigma_{A_1} \sigma_{A_2}$$

and

$$\operatorname{cov}\left(E_1^{'}, E_2^{'}\right) = \rho_{E^{'}} \sigma_{E_1^{'}} \sigma_{E_2^{'}}$$

which means

$$\frac{\sigma_{A_1}}{\sigma_{P_1}} = h_1, \frac{\sigma_{E_1^{'}}}{\sigma_{P_1}} = \sqrt{1 - h_1^2},$$

$$\frac{\sigma_{A_2}}{\sigma_{P_2}} = h_2, \frac{\sigma_{E_2^{'}}}{\sigma_{P_2}} = \sqrt{1 - h_2^2}.$$

By transformation the model is converted into

$$\rho_P = h_1 h_2 \sigma_A + \sigma_{E^{'}} \sqrt{\left(1 - h_1\right)\left(1 - h_2\right)}.$$

With the help of this formula environmental correlation may be computed as

$$\rho_{E^{'}} = \frac{\rho_P - h_1 h_2 \sigma_A}{\sqrt{\left(1 - h_1^2\right)\left(1 - h_2^2\right)}}.$$

Conclusion:

If heritability is high, phenotypic correlation is due predominantly to genotypic correlation; if it is low, phenotypic correlation is due predominantly to mutual environmental factors.

5. The Essential of Discerning Nonrandom Genetic and Environmental Effects

The 'efficacy value' [Bc], strictly, is an estimation.

Genetic effects (*A*, *D*, *Ep*) as well as environmental effects (E_r) are comprehended in the 'efficacy value' [Bc] of a compound:

1. random and nonrandom (systematic) genetic effects (depending upon occurrence in the genotype sampling) and
2. random and nonrandom environmental effects (due to e.g., seasonal or nutritional situation or experimental conditions).

This may be formulated as

$$P = \mu + G + E_c + E_r$$
$$= \mu + G_r + G_s + E_c + E_r^{'} + E_s .$$

A phenotypic value may be derived from nonrandom effects by recourse to control mean (*cm*) deviation. By 'control mean' is designated the mean (of all genotypes and all compounds) plus the sum of nonrandom experimental effects -- both genetic (G_s) and environmental (E_s):

$$cm = \mu + G_s + E_s .$$

The deviations constituted in the 'efficacy value' [Bc] then may be obtained with

$$P - cm = G_r + E_c + E_r .$$

Oftentimes there are diverse nonrandom effects to delimit (e.g., association between genotype subgroups, seasonal variation, climate chamber trial modalities). By means of the ensuing equation, any experimental modality may be 'corrected' to compensate for whatever systematic control influences (up to n genotypic and/or m environmental factors):

$$cm^{'} = \mu + G_{s_1} + \ldots + G_{s_n} + E_{s_1} + \ldots + E_{s_m} .$$

6. Selection indices

The indices may be formulated in a manner similar to the optimum index of SMITH (1936) and HAZEL (1943), that employs heritabilities as index weights (which weights correspond to weights from the optimum index if traits are uncorrelated), or according to the base index proposed by WILLIAMS (1962), which uses economic weights as index weights.

7. Discussion

Quantitative (Population-) Epigenetics is the study of continuous traits (such as stress, height or weight) and their underlying mechanisms. It is the combined effect of the many underlying genes and epigenetical effects resulting in a continuous distribution of phenotypic values.

The main application of quantitative epigenetics to artificial and natural populations could be using the pattern of genetic variances and covariances to predict the response of the mean phenotype to biologically active chemical compounds as artificial environmental factors - from 'chemistry' to phenotype:

- Quantitative epigenetics aims to link phenotypic variation for complex traits to its underlying epigenetic basis in order to understand and predict better epigenetic architecture and changes within natural, agricultural and human populations - due to environmental factors.

- Traditionally built upon statistical abstractions of epigenetic effects (environmental, biologically active chemical compounds), the field could be used to reveal explicit links between epigenome and complex phenotypes, and could therefore serve as a focal point for bringing together many emerging areas of genetics, epigenetics, genomics, physiology, statistics, bioinformatics, and computational biology.
- This synthesis could have a large impact on the areas of evolutionary biology, selection and development of biologically active chemical compounds, and the epigenetic analysis of human disease.

The application of statistical Quantitative (Population-) Epigenetics to the selection and to the development of biologically active substances (e.g. plant growth regulators) is a fundamentally new approach in planning, evaluation and assessment of experiments.

The following considerations are offered:

1. In initial screenings, a 'random' subpopulation of genotypes should manifest moderate expression of a target character - resulting in low heritability.
2. In testing for quantitative character enhancement, initial screening should be conducted under stress-environment conditions - in order to obtain an optimum differentiation of agents (low heritability).
3. To judge constancy in performance, testing should employ a random sampling of genotypes and non-stress environments - which afford assessment of the interactions: a) agent-genotype, b) agent-environment and c) agent-genotype-environment.
4. It must be possible to identify specific genotype-environment constellations from which issue 'amplifier' interactions that intensify differentiation suitability -- as in breeding practice, where partial positive covariance of genotype-environment interaction avails to intensify differentiation suitability of location effects, or here genotype-environment effect.

Hence the objective is to identify which genotype-environment constellation is the most auspicious to make use of as a 'reference combination' to achieve optimum screening efficiency.

The statistical Quantitative (Population-) Epigenetics theory provides basic rules for experimental designs and data analysis concerning

1. experimental design: fix or random effects, size, trial conditions e.g. stress or non-stress, etc.,
2. post-experimental evaluation,
3. optimization of experimental designs and
4. quantitative description of single or multiple traits designs e.g. using selection indices.

The impetus to translate this Quantitative (Population-) Epigenetics theory into practice is weighted by a) screening sensitivity, b) time expediency, c) ease of replication, d) reliability, e) heuristic incentive, f) elimination of conjectural risks and g) financial returns.

References

ALLARD, R.W., 1960: PRINCIPLES OF PLANT BREEDING, PP. 83-88. JOHN WILEY & SONS, NEW YORK.

HAZEL, L.N., 1943: THE GENETIC BASIS FOR CONSTRUCTING SELECTION INDICES. GENETICS **28**, 476-490.

SMITH, H.F., 1936: A DISCRIMINANT FUNCTION FOR PLANT SELECTION. ANNALS OF HUMAN GENETICS **7**, 240-250.

STAUSS, R., 1992: GENETIC ANALOGUES IN CHEMICAL SCREENING. ZEITSCHRIFT FÜR PFLANZENKRANKHEITEN UND PFLANZENSCHUTZ **99**, 653-656.

WILLIAMS, J.S., 1962: THE EVALUATION OF A SELECTION INDEX. BIOMETRICS **16**, 375-393.