

Phyton (Austria)	Vol. 23	Fasc. 1	43—48	15. 2. 1983
------------------	---------	---------	-------	-------------

Aus dem Institut für Botanik der Universität Wien

Vergleich von *Arabis hirsuta* s. str. und *A. sagittata* × *ciliata* (Cruciferae) mittels der Diskriminanzanalyse

Von

Walter TITZ*)

Mit 1 Abbildung

Eingelangt am 29. Oktober 1981

Key words: Numerical taxonomy, discriminant analysis, canonical analysis, factor analysis, principal-component analysis. — *Cruciferae*, *Brassicaceae*, *Arabis hirsuta* agg.

Summary

TITZ W. 1983. Comparison of *Arabis hirsuta* s. str. and *A. sagittata* × *ciliata* (Cruciferae) by means of discriminant analysis. — *Phyton* (Austria) 23 (1): 43—48, 1 figure. — German with English summary.

Discriminant analysis was applied to the problem of phenotypical divergence between natural tetraploid *Arabis hirsuta* s. str. and artificial tetraploid *A. sagittata* × *ciliata*: essential diversity was confirmed and illustrated by a diagram. Missing character values had been replaced either by the arithmetic group mean or (in the course of a principal-component analysis) by the total mean for each character.

Zusammenfassung

TITZ W. 1983. Vergleich von *Arabis hirsuta* s. str. und *A. sagittata* × *ciliata* (Cruciferae) mittels der Diskriminanzanalyse. — *Phyton* (Austria) 23 (1): 43—48, 1 Abbildung. — Deutsch mit englischer Zusammenfassung.

Die Diskriminanzanalyse wurde auf das Problem der phänotypischen Divergenz zwischen der natürlichen Tetraploiden *Arabis hirsuta* s. str. und der artifiziellen Tetraploiden *A. sagittata* × *ciliata* angewandt: eine essentielle Verschiedenheit der beiden Sippen wurde bestätigt und diagrammatisch dar-

*) Univ.-Prof. Dr. Walter TITZ, Institut für Botanik der Universität Wien, Rennweg 14, A-1030 Wien, Österreich.

gestellt. Zuvor waren die fehlenden Merkmalswerte entweder durch das arithmetische Gruppenmittel oder — im Verlauf einer Hauptkomponentenanalyse — durch das Gesamtmittel jedes Merkmals ersetzt worden.

Einleitung und Methode

Die Diskriminanzanalyse ist das einzige multivariate statistische Verfahren, das für eine systematisch-biologische Fragestellung entwickelt wurde, und zwar im wesentlichen von FISHER 1936 am Beispiel der taxonomischen und phylogenetischen Beziehungen zwischen drei *Iris*-Arten (vgl. SOKAL 1965, REYMENT 1973). Das Grundprinzip der seither insbesondere in Verbindung mit dem Aufschwung der elektronischen Datenverarbeitung entwickelten Prozeduren der Diskriminanzanalyse (auch unter der Bezeichnung Kanonische Analyse = Canonical Analysis subsumiert) besteht in einer Transformation der Merkmale bereits klassifizierter Objekte in sogenannte Diskriminanzfunktionen, die die bestehenden Gruppenunterschiede bestmöglich ausdrücken. Der Wert der Diskriminanzfunktion eines bestimmten Individuums i , das ist die sogenannte Diskriminanzvariable x_i , ergibt sich aus den einzelnen Merkmalswerten (a_{i1} für das Merkmal 1, a_{i2} für das Merkmal 2 usw. bis a_{im} für das Merkmal m) wie folgt:

$$x_i = g_1 \cdot a_{i1} + g_2 \cdot a_{i2} + g_3 \cdot a_{i3} \dots + g_m \cdot a_{im} = \sum_{q=1}^m g_q \cdot a_{iq},$$

wobei $g_1, g_2, \dots, g_m$ merkmalspezifische Gewichtungsfaktoren sind. Die Gewichtungsfaktoren von Merkmalen mit großer Varianz zwischen den (durch die Klassifikation vorgegebenen) Gruppen, aber geringer Varianz innerhalb derselben, werden hoch sein, umgekehrt solche von Merkmalen mit geringer Zwischengruppenvarianz und hoher Innergruppenvarianz gering (Diskriminanzkriterium). Die Errechnung dieser Gewichtungsfaktoren kann auf verschiedenem Wege erfolgen (hiez zu siehe die einschlägige Literatur, z. B. FISHER 1936, SEAL 1964, CACOULOS 1973, MORRISON 1976, NIE & al. 1975, OPTIZ 1980, TITZ 1982). Im vorliegenden Fall wurde die im Programmsystem SPSS, Version 8.0 enthaltene direkte Methode (simultane Einbeziehung aller Merkmale) als Analyseweg für die Berechnung der Gewichtungsfaktoren (= discriminant function coefficients) beschritten (vgl. NIE & al. 1975).

Da es sich im hier publizierten Anwendungsbeispiel um den Vergleich von zwei Gruppen handelt, gibt es nur je einen Gewichtungsfaktor pro Merkmal und eine Diskriminanzvariable pro Objekt. In diesem einfachsten Fall können die Häufigkeiten der Diskriminanzvariablen in den analysierten Gruppen in der Form eines Histogramms dargestellt werden (Abb. 1).

Wir haben die Diskriminanzanalyse auf die Frage angewendet, ob die synthetische Allotetraploide *Arabis sagittata* (BERTOL.) DC. $\times$ *ciliata* CLAIRV. genotypisch bzw. phänotypisch mit der natürlichen Tetraploiden *Arabis hirsuta* (L.) SCOP. s. str. übereinstimmt (TITZ 1979 a, b). Die Mittelwerte und Spannweiten der Meßwerte der 15 analysierten Merkmale von

37 *A. hirsuta*- und 30 *A. sagittata* $\times$ *ciliata*-Pflanzen sind bereits bei TITZ 1979a in Tabelle 4 veröffentlicht, ebenso die Ergebnisse der für jedes Merkmal durchgeführten t-Tests. Dabei ergab sich, daß bei 12 von den 15 Merkmalen ein signifikanter Unterschied zwischen den beiden Sippen besteht. Dies deutet ebenso wie die Kreuzungsversuche auf einen beträchtlichen phänotypischen bzw. genotypischen Unterschied hin (TITZ 1979a, b). Die Durchführung von t-Tests für die Einzelmerkmale hat den Nachteil, daß sowohl die Korrelation zwischen den Merkmalen außer Betracht bleibt als auch nur auf die einzelnen Merkmale bezogene Aussagen möglich sind, nicht aber Schlüsse auf die Pflanzenindividuen oder Sippen als ganze. Multivariate Methoden, wie etwa die Diskriminanzanalyse, hingegen betrachten simultan alle untersuchten Merkmale einer großen Zahl von Individuen.

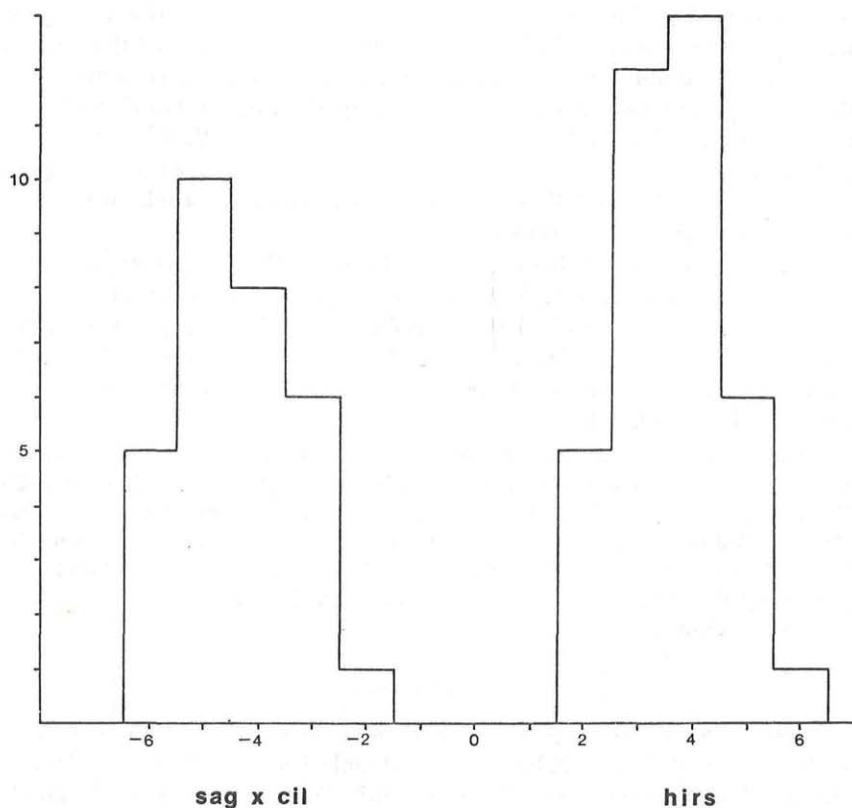


Abb. 1. Histogramm der Diskriminanzvariablen (Abszisse) von *Arabis sagittata* $\times$ *ciliata* (sag $\times$ cil) und *A. hirsuta* s. str. (hirs); Häufigkeiten auf der Ordinate. Fehlende Merkmalswerte waren durch den arithmetischen Mittelwert des jeweiligen Merkmals der Gruppe (= Sippe) ersetzt worden

In der Praxis ist es freilich oft nicht möglich, an jedem Individuum alle Merkmalsbereiche zu analysieren, vor allem wenn diese verschiedenen Entwicklungsstadien angehören. Da die Diskriminanzanalyse aber vollständiger Daten bedarf, müssen fehlende Merkmalswerte ersetzt werden. Dafür ergeben sich vor allem zwei Möglichkeiten:

1. An ihre Stelle den Gruppenmittelwert des Merkmals zu setzen (z. B. den arithmetischen Mittelwert);

2. Eine andere Form einer multivariaten Analyse, nämlich eine Hauptkomponentenanalyse (principal-component analysis) vor der Diskriminanzanalyse durchzuführen. In diesem faktorenanalytischen Verfahren können primär fehlende Werte toleriert werden, bei der Berechnung der Faktorenwerte wird für sie sodann jeweils der Gesamtmittelwert des Merkmals (für alle Gruppen zusammengekommen) verwendet. Die an die Stelle der Merkmale tretenden Faktoren fassen jeweils mehrere korrelierte Merkmale in einen Faktor zusammen und verringern im Fall des Vorliegens hochkorrelierter Merkmale die damit gegebene Redundanz an Information.

Beide Prozeduren zum Ersatz der fehlenden Merkmalswerte wurden auf das vorliegende Problem angewandt; im Zuge der zweiten lief eine Hauptkomponentenanalyse (PA1 aus SPSS, Version 8.0) mit Varimax-Rotation (NIE & al. 1975). Es wurden dabei 5 Faktoren (Eigenwertkriterium: Eigenwert ≥ 1) extrahiert und diese sodann der oben genannten direkten Methode der Diskriminanzanalyse unterworfen.

Sowohl die Diskriminanzanalyse als auch die Hauptkomponentenanalyse setzen annähernde Normalverteilung der Merkmalswerte voraus, im vorliegenden Fall ergab eine Überprüfung mittels deskriptiver Statistiken (Programm BMDP2D aus dem Programmpaket BMDP-79, siehe DIXON & BROWN 1979) ein weitgehendes Zutreffen dieses Modells für die meisten Einzelmerkmale.

Die gesamte rechnerische Analyse erfolgte über das interfakultäre EDV-Zentrum der Universität Wien. Die für die Anlegung der Datenfiles und die Ersatzprozeduren für die fehlenden Merkmalswerte nötigen Programme wurden zum Teil in Zusammenarbeit mit Herrn Doz. Dr. W. TRMISCHL (Technische Universität Wien) erstellt, dem ich dafür großen Dank schulde. Dem Fonds zur Förderung der wissenschaftlichen Forschung (Projekt 4609) wird für Unterstützung gedankt.

Ergebnisse

Mit Hilfe des oben genannten Verfahrens der Hauptkomponentenanalyse werden die 15 analysierten Merkmale (siehe TRTZ 1979a, Tab. 4) in 5 Faktoren zusammengefaßt, wobei mit Faktor 1 folgende Merkmale besonders verknüpft sind (Faktorladungen über 0,6): Nr. 6, 7 und 13 (d. s. Länge und Längen-Breiten-Verhältnis des längsten Blattes der oberen Stengelhälfte sowie die maximale Schotenlänge); mit Faktor 2: Merkmale Nr. 3, 4, 6 und 8 (Anzahlen der basalen und der oberen Seitenäste sowie Län-

gen der längsten Blätter der oberen Stengelhälfte und des Gesamtstengels); mit Faktor 3: Merkmale Nr. 12, 14, 15 (Maximalwerte der Schotenstiel-länge, Griffellänge und Samenlänge); mit Faktor 4: Merkmale Nr. 9, 10, 11 (Sepalenlänge, Petalenlänge, Petalenbreite); mit Faktor 5: Merkmale Nr. 1, 2, 5 (Höhen des gesamten Hauptstengels und dessen beblätterten Teils sowie Zahl der Stengelblätter). In fast allen Fällen können die dabei zum Ausdruck kommenden Korrelationen zwischen den in einem Faktor (mit hohen Faktorladungen) vereinigten Merkmalen logisch bzw. morphogenetisch begründet werden.

Erst bei der Berechnung der Faktorenwerte für die einzelnen Pflanzen-individuen (über die von den Faktorladungen abhängigen Faktorenwerte-koeffizienten) werden für die fehlenden ursprünglichen Merkmalswerte die arithmetischen Gesamtmittelwerte der einzelnen Merkmale eingesetzt. Dies führt gegenüber dem die Unterschiede zwischen den Gruppen betonen-den Ersatz durch die Gruppenmittelwerte (s. unten) naturgemäß zu einer gewissen Nivellierung der Differenzen. Trotzdem gibt es nach dieser Prozedur in den Werten der Diskriminanzvariablen der beiden Sippen nur geringfügige Überschneidungen, für die Individuen von *Arabis sagittata* $\times$ *ciliata* rangieren sie zwischen $-4,3$ und $+0,2$, für *A. hirsuta* s. str. zwischen $-0,9$ und $+3,8$. Im fraglichen Überschneidungsbereich von $-0,9$ bis $+0,2$ liegen nur 5 Individuen (drei von *A. hirsuta* s. str., zwei von *A. sagittata* $\times$ *ciliata*).

Erwartungsgemäß noch klarer, nämlich völlig ohne Überschneidungen fällt das Histogramm der Diskriminanzvariablen nach Ersatz der fehlenden Werte durch die Gruppenmittelwerte aus (Abb. 1): die Diskriminanz-variablen der Individuen von *Arabis sagittata* $\times$ *ciliata* liegen zwischen $-6,2$ und $-1,5$, die von *A. hirsuta* s. str. zwischen $+1,7$ und $+5,7$.

Wie oben schon erwähnt, erlaubt es die Diskriminanzanalyse zum Unterschied von univariaten statistischen Methoden, in einem Analyse-schritt alle Merkmale zu erfassen und auf Grund dessen — in einem Diagramm — eine Aussage über die Verschiedenheit zweier (oder mehrerer) Sippen zu machen. Aus der Anwendung dieser Methode auf das diskutierte Problem ergibt sich mit großer Sicherheit, daß die artifizielle Allotetraploide *Arabis sagittata* $\times$ *ciliata* phänotypisch und damit auch genotypisch von der natürlichen Tetraploiden *Arabis hirsuta* s. str. ganz essentiell verschieden ist und daher taxonomisch nicht konspezifisch mit dieser sein kann.

Das Verfahren der Diskriminanzanalyse hat sich als für die geschilderte Fragestellung als im höchsten Maße geeignet erwiesen und kann somit als wichtiges Hilfsmittel bei der Überprüfung phänotypischer Divergenzen und darauf aufbauender taxonomischer Klassifikationen angesehen werden (vgl. THORPE 1980, TITZ 1982).

Literatur

- CACOULOS T. (Ed.) 1973. Discriminant analysis and applications. — Academic Press, New York.
- DIXON W. J. & BROWN M. B. (Ed.) 1979. BMDP-79. Biomedical Computer Programs, P-Series. — Univ. of California Press, Berkeley.
- FISHER R. A. 1936. The use of multiple measurements in taxonomic problems. — Ann. Eugenics 7: 179—188.
- MORRISON D. F. 1976. Multivariate statistical methods. 2nd edition. — McGraw-Hill Kogakusha Ltd., Tokyo.
- NIE N. H., HULL C. H., JENKINS J. G., STEINBRENNER K. & BENT D. H. 1975. SPSS. Statistical package for the social sciences. 2nd edition. — McGraw-Hill Book Company, New York.
- OPITZ O. 1980. Numerische Taxonomie. — Gustav Fischer, Stuttgart.
- REYMENT R. A. 1973. The discriminant function in systematic biology. — In: CACOULOS T. (Ed.), Discriminant analysis and applications, p. 311—337. — Academic Press, New York.
- SEAL H. L. 1964. Multivariate statistical analysis for biologists. — Methuen & Co. Ltd., London.
- SOKAL R. R. 1965. Statistical methods in systematics. — Biol. Rev. 40: 337—391.
- THORPE R. S. 1980. A comparative study of ordination techniques in numerical taxonomy in relation to racial variation in the ringed snake *Natrix natrix* (L.). — Biol. J. Linn. Soc. 13: 7—40.
- TITZ W. 1979a. Kritisch-experimentelle Untersuchungen zur Resynthesierbarkeit von *Arabis hirsuta* s. str. (4x) aus *A. sagittata* (2x) und *A. ciliata* (2x). — Beitr. Biol. Pflanzen 54: 443—466.
- 1979b. Attempts to resynthesize natural allopolyploids from the predicted parents. — Eucarpia Cruciferae Newsletter 4: 30.
- 1982. Über die Anwendbarkeit biomathematischer und biostatistischer Methoden in der Systematik. — Ber. deutsch bot. Ges. 95: 149—154.

ZOBODAT - www.zobodat.at

Zoologisch-Botanische Datenbank/Zoological-Botanical Database

Digitale Literatur/Digital Literature

Zeitschrift/Journal: [Phyton, Annales Rei Botanicae, Horn](#)

Jahr/Year: 1983

Band/Volume: [23_1](#)

Autor(en)/Author(s): Titz Walter

Artikel/Article: [Vergleich von *Arabis hirsuta* s.str. und *A. sagittata* x *ciliata* \(Cruciferae\) mittels der Diskriminanzanalyse. 43-48](#)