Zeitschrift: Candollea : journal international de botanique systématique = international journal of systematic botany

Herausgeber: Conservatoire et Jardin botaniques de la Ville de Genève

Band: 36 (1981)

Heft: 2

Artikel: Etude biosystématique du Silene vulgaris s.l. (Caryophyllaceae) dans le domaine alpin : méthodologie

Autor: Aeschimann, David / Mascherpa, Jean-Michel / Bocquet, Gilbert

DOI: https://doi.org/10.5169/seals-880071

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 06.07.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Etude biosystématique du Silene vulgaris s.l. (Caryophyllaceae) dans le domaine alpin. Méthodologie

DAVID AESCHIMANN
JEAN-MICHEL MASCHERPA
&
GILBERT BOCQUET

RÉSUMÉ

AESCHIMANN, D., J.-M. MASCHERPA & G. BOCQUET (1981). Etude biosystématique du Silene vulgaris s.l. (Caryophyllaceae) dans le domaine alpin. Méthodologie. *Candollea* 36: 379-396. En français, résumé anglais.

Le *Silene vulgaris* s.l., taxon très polymorphe, a depuis longtemps intéressé les systématiciens; il les a également déroutés. De nombreux points restent en suspens, sur la voie d'une solution taxonomique entièrement satisfaisante et utilisable dans les flores. Notre travail porte sur 57 caractères morphologiques et un nombre de spécimens voisin de 3000. Pour le traitement d'une telle masse de données, nous avons choisi le recours aux méthodes de taxonomie numérique. Selon une étude préliminaire, les caractères relatifs à la surface foliaire paraissent importants: nous utilisons par conséquent un analyseur de surface automatique à microprocesseur. Une discussion des caractères retenus et l'organigramme général sont présentés. Cet article est le premier d'une série sur cette étude biosystématique.

ABSTRACT

AESCHIMANN, D., J.-M. MASCHERPA & G. BOCQUET (1981). Biosystematical studies on Silene vulgaris s.l. (Caryophyllaceae) in the alpine region. Methodology. *Candollea* 36: 379-396. In French, English abstract.

Silene vulgaris s.l., a very polymorphic taxon, has aroused the interest of systematists for a long time; it has also baffled them. Numerous topics remain unresolved on the way to an entirely satisfactory and usable taxonomic solution for floras. Our work relates to 57 morphologic characters and approximately 3000 specimens. To deal with such a bulk of information we chose the method of numerical

CODEN: CNDLAR
ISSN: 0373-2967
36(2) 379 (1981)

© CONSERVATOIRE ET JARDIN BOTANIQUES DE GENÈVE 1981

taxonomy. According to a preliminary study, the characters related to the foliar surface seem important: we therefore used an automatic microprocessor surface analyser. A discussion of the characters selected is presented, together with the general operating chart. This article is the first of a series on this biosystematic study.

Introduction

Les études systématiques du *Silene vulgaris* s.l., au sens du "Flora europaea" (CHATER & WALTERS, 1964), n'ont pas à ce jour permis de proposer des solutions entièrement satisfaisantes. Ce groupe extrêmement dynamique, tant sur le plan purement écologique qu'au niveau évolutif, montre un polymorphisme déroutant, dont nous avons analysé quelques unes des causes dans un récent travail (AESCHIMANN & BOCQUET, 1980b).

Une étude préliminaire (AESCHIMANN, 1979) portait sur l'examen classique de matériel d'herbier provenant de l'ensemble de l'arc alpin, ainsi que sur un premier sondage biométrique. L'importance du rapport longueur/largeur de la feuille a notamment été démontrée. Une esquisse de solution taxonomique et chorologique a été proposée.

De manière à dégager les unités taxonomiques de rang supérieur à la masse des formes locales et de toute la variabilité écotypique, nos propres récoltes et observations sur le terrain couvrent en détail le domaine étudié. En outre, le matériel de plusieurs herbiers suisses et étrangers complète nos observations.

Une telle étude de détail porte ainsi sur un nombre de spécimens qui avoisinera 3000. Le traitement de ces données, et la nécessité d'aboutir à une solution intégrant toutes les informations disponibles, nous a naturellement conduit vers l'utilisation des ressources modernes de l'informatique et de l'analyse des données. Leur application au problème du *Silene vulgaris* s.l. constitue l'originalité et la partie principale de notre travail. HESS & LANDOLT (1967: 784) ont souligné par ailleurs la nécessité d'études de populations. Pour le domaine atlantique, GODEAU (1977) s'est aussi engagé dans cette voie. Son étude, encore limitée, ne porte que sur cinquante spécimens et douze caractères; elle ne fait pas appel à l'analyse factorielle.

Comme l'ont montré les travaux de MARSDEN-JONES & TURILL (1957), par trente années de croisements, les divers taxa présents au sein du groupe du *Silene vulgaris* s.l. sont très interfertiles. Nous l'avons constaté lors de nos études sur le terrain: les séries introgressives sont très nombreuses (ANDERSON, 1949), aussi bien au niveau très large des zones de contact de races géographiques qu'à l'échelon local, avec le mélange de deux populations voisines. C'est pourquoi nous chercherons à préciser la localisation taxonomique des spécimens intermédiaires.

La variabilité étudiée, de niveau infraspécifique, est souvent ténue; il s'agit en fait de races géographiques. Ainsi, la codification grossière de

certains caractères, comme ceux relatifs à la forme des feuilles, se serait révélée insuffisante. De manière à accroître la précision des résultats, faciliter et accélérer la prise des données portant sur un nombre élevé de spécimens, nous utilisons un analyseur de surface électronique.

Enfin, parallèlement à ces études, d'importantes séries de témoins sont cultivés en couche au Jardin botanique de la Ville de Genève.

Tous nos comptages chromosomiques ($2n = 24$, non publiés) ont confirmé l'absence de polyploïdie dans le domaine alpin. Lors de travaux ultérieurs, avec l'extension de l'aire d'étude au niveau méditerranéen, il sera indispensable d'inclure les données cytologiques (par exemple $2n = 48$ chez le *Silene vulgaris* (Moench) Garcke subsp. *macrocarpa* Turill).

Choix des caractères pour la taxonomie numérique

MARSDEN-JONES & TURILL (1957) fournissent de nombreuses données sur les caractères taxonomiques qu'ils ont utilisé pour le groupe du *Silene vulgaris* s.l. Nous avons été amenés à choisir certains d'entre eux, tout en proposant quelques autres caractères.

A l'occasion d'un travail sur les types biologiques (AESCHIMANN & BOCQUET, 1980b), nous avons montré leur nature fondamentalement adaptative, en soulignant qu'ils ne pourraient que fournir des caractères complémentaires et aider à l'identification. C'est pourquoi nous avons décidé de renoncer à leur utilisation dans le programme de taxonomie numérique. D'ailleurs, le matériel existant en herbier, très souvent incomplet à cet égard, ne permet pas toujours une appréciation objective de tels caractères.

Au cours de nos nombreuses campagnes de récolte, nous avons constaté la faible valeur systématique des caractères tirés de la forme des pétales, de leurs échancrures, de leur éventulle coloration rose (anthocyanes) et du développement plus ou moins grand de la coronule. Tous ces caractères montrent une extrême variabilité, souvent même à l'intérieur de populations relativement homogènes.

La période de floraison ne fait pas partie de la liste des caractères adoptés pour l'analyse numérique. Il en est toutefois tenu compte au niveau de la date de récolte du spécimen, que l'on mettra en rapport avec l'état phénologique, relevé et codifié.

Pour des raisons pratiques, les caractères tirés de la période de germination et de la couleur des graines immatures ne sont pas utilisables au cours d'une étude comme la nôtre. D'après MARSDEN-JONES & TURILL (1957), ces caractères sont surtout importants pour la distinction *Silene vulgaris*/*Silene maritima*.

La morphologie de la graine, particulièrement des fines ornementations de sa surface, ne sera pas intégrée à la liste générale au cours de ce travail. Nous envisageons en effet de traiter séparément ce sujet, encore mal connu pour le groupe du *Silene vulgaris* s.l. Une série de travaux au microscope électronique à balayage est actuellement en cours.

Lors de notre étude préliminaire (AESCHIMANN, 1979), ainsi qu'au cours d'un examen récent de matériel, nous avons montré et constaté l'importance systématique des caractères tirés de la forme des feuilles, de leur pilosité (ainsi que celle de la tige), de l'inflorescence (surtout le nombre de fleurs) et de la pigmentation du calice. Tous ces caractères constituent une part importante de la liste retenue.

Enfin, la tendance ou non à la zygomorphie de la fleur est connue comme importante dans la distinction entre le *Silene maritima* (actinomorphe) et le *Silene vulgaris* (plutôt zygomorphe). D'autre part, des caractères comme la présence de racines caulinaires, ou ceux tirés de la taille et la forme de la capsule permettent par exemple de distinguer le *Silene vulgaris* subsp. *macrocarpa,* taxon méditerranéen (voir MELZHEIMER & DAMBOLDT, 1973). Notre choix de caractères permet ainsi non seulement une étude dans le domaine alpin, mais également des comparaisons avec les régions atlantiques et méditerranéennes. Toutefois, il n'est pas exclu qu'au cours d'un travail ultérieur plus vaste, incluant intégralement ces deux dernières régions, quelques adjonctions doivent être portées à la liste retenue ci-dessous.

		Sigles
	RACINE	
Caractère 1:	racines caulinaires (voir AESCHIMANN & BOCQUET, 1980a); *2 = absentes, 3 = adventives, 4 = supplémentaires*	R01CAUL
	TIGE	
Caractère 2:	longueur maximale des rameaux annuels; *en cm, arrondi* . .	T02LONG
Caractère 3:	diamètre aux 2^{me} à 3^{me} internœuds sous l'inflorescence; *en mm*	T03DIAM
Caractère 4:	(caractère 2/caractère 3) x 10; *valeur pure*	T0402/03
Caractère 5:	longueur des poils; *2 = nuls, 3 = papilles, 4 = courts (hérissés), 5 = longs (soyeux)*	T05LPIL
Caractère 6:	densité de la pilosité; *2 = nulle, 3 = éparse, 4 = dense, 5 = très dense*	T06DEPIL
Caractère 7:	forme des poils; *2 = cylindriques, 4 = coniques*	T07DIPIL
Caractère 8:	longueur du 1er internœud sous l'inflorescence; *en mm* . .	T08PINT
Caractère 9:	longueur du 2^{me} internœud; *en mm*	T09SINT
Caractère 10:	longueur du 3^{me} internœud; *en mm*	T10TINT
Caractère 11:	caractère 8/caractère 10; *valeur pure*	T1108/10
Caractère 12:	caractère 2—(somme des caractères 8, 9, 10)/10; *en cm, arrondi*	T12BASE
	FEUILLE (aux 2^{me} à 3^{me} nœuds sous l'inflorescence)	
Caractère 13:	longueur; *en mm*	F13LONG
Caractère 14:	largeur maximale; *en mm*	F14LARG
Caractère 15:	caractère 13/caractère 14; *valeur pure*	F1513/14
Caractère 16:	position de la largeur maximale sur la longueur; *valeur entre 0 et 1*	F16LA,LO

Caractère 17:	périmètre de la surface; *en mm*	F17PERI
Caractère 18:	aire de la surface; *en mm*^{2}	F18SURF
Caractère 19:	caractère 18/caractère 17; *en mm*	F1918/17
Caractère 20:	feuille très épaisse?; *2 = oui, 4 = non*	F20EPAI
Caractère 21:	longueur des poils; *2 = nuls, 3 = papilles, 4 = courts (hérissés), 5 = longs (soyeux)*	F21LPIL
Caractère 22:	densité de la pilosité; *2 = nulle, 3 = éparse, 4 = dense, 5 = très dense*	F22DEPIL
Caractère 23:	forme des poils; *2 = cylindriques, 4 = coniques*	F23DIPIL
Caractère 24:	très fines rides de surface; *2 = absentes, 3 = éparses, 4 = abondantes*	F24RIDES
Caractère 25:	marge, bord décoloré de la feuille; *2 = absente, 3 = étroite, 4 = bien marquée*	F25MARGE
Caractère 26:	longueur des poils de la marge; *2 = nuls, 3 = papilles, 4 = courts (hérissés), 5 = longs (soyeux)*	F26LPIM
Caractère 27:	densité de la pilosité de la marge; *2 = nulle, 3 = éparse, 4 = dense, 5 = très dense*	F27DEPIM
Caractère 28:	forme des poils de la marge; *2 = cylindriques, 4 = coniques*	F28DIPIM
	INFLORESCENCE	
Caractère 29:	premières bractées scarieuses?; *2 = oui, 4 = non*	I29BRSCA
Caractère 30:	longueur des premières bractées; *en mm*	I30LBRA
Caractère 31:	longueur du pédoncule de la fleur centrale du dichasium; *en mm*	I31PPED
Caractère 32:	moyenne des longueurs des deux premiers rameaux du dichasium; *en mm*	I32PRAM
Caractère 33:	caractère 8/caractère 31; *valeur pure*	I3308/31
Caractère 34:	caractère 8/caractère 32; *valeur pure*	I3408/32
Caractère 35:	caractère 32/caractère 31; *valeur pure*	I3532/31
Caractère 36:	nombre de fleurs par inflorescence; *valeur entière* . . .	I36NBFL
Caractère 37:	tendance à la zygomorphie de la fleur?; *2 = oui, 4 = non* . .	I37ZYGOM
	CALICE (état mûr, 1re à 3^{me} fleurs du dichasium)	
Caractère 38:	longueur; *en mm*	CL38LONG
Caractère 39:	diamètre maximal; *en mm*	CL39DIAM
Caractère 40:	caractère 38/caractère 39; *valeur pure*	CL403839
Caractère 41:	position du diamètre maximal sur la longueur; *2 = calice rectiligne, 3 = dans le tiers inférieur, 4 = dans le tiers intermédiaire, 5 = dans le tiers supérieur*	CL41DILO
Caractère 42:	base des dents; *en mm*	CL42BASD
Caractère 43:	hauteur des dents; *en mm*	CL43HAUD
Caractère 44:	caractère 42/caractère 43; *valeur pure*	CL444243

Caractère 45:	intensité de la pigmentation anthocyanique; *2 = faible, 3 = moyennement intense, 4 = intense*	CL45IANT
Caractère 46:	répartition (contraste) de la pigmentation anthocyanique; *2 = diffuse, 3 = moyennement concentrée sur les veines, 4 = étroitement concentrée sur les veines*	CL46RANT
Caractère 47:	intensité de la pigmentation chlorophyllienne; *2 = faible, 3 = moyennement intense, 4 = intense*	CL47ICHL
Caractère 48:	marge blanche dépourvue d'anthocyanes, à la jointure des dents; *2 = absente, 3 = étroite, 4 = marquée*	CL48MABL
	CARPOPHORE (état mûr, 1re à 3me fleurs du dichasium)	
Caractère 49:	longueur; *en mm*	CR49LONG
Caractère 50:	caractère 52/caractère 49; *valeur pure*	CR505249
Caractère 51:	caractère 38/caractère 49; *valeur pure*	CR513849
	CAPSULE (état mûr, 1re à 3me fleurs du dichasium)	
Caractère 52:	longueur; *en mm*	CP52LONG
Caractère 53:	diamètre maximal; *en mm*	CP53DIAM
Caractère 54:	caractère 52/caractère 53; *valeur pure*	CP545253
Caractère 55:	diamètre d'ouverture; *en mm*	CP55DIOU
Caractère 56:	rétrécissement en bec à l'extrémité; *2 = nul, 3 = modéré, 4 = marqué*	CP56RETR
Caractère 57:	dents érigées?; *2 = oui, 4 = non (réfléchies)*	CP57DEER

Liste des caractères retenus

La matrice des données comprend 57 caractères relevés sur la racine, la tige, la feuille, l'inflorescence, le calice, le carpophore et la capsule. Vingt-deux d'entre eux sont codés à plusieurs états; 22 sont mesurés, ou se voient attribuer des nombres purs; 13 sont des rapports calculés sur la base des caractères mesurés. Nous sommes donc en présence d'une matrice de données hétérogènes, dont nous savons que le traitement est toujours délicat. C'est pourquoi, dans ce cas précis, nous avons adopté une stratégie légèrement différente de celle proposée précédemment (MASCHERPA, 1976; MARÉCHAL & al., 1978).

Dans la liste ci-dessus, pour les caractères codés, nous adoptons toujours la convention suivante:

— code 0 = information absente;

— code 1 = information présente, mais dans un état ne permettant pas la codification.

Un certain nombre de caractères sont illustrés aux figures 1, 2 et 3, les numéros reportent à la liste.

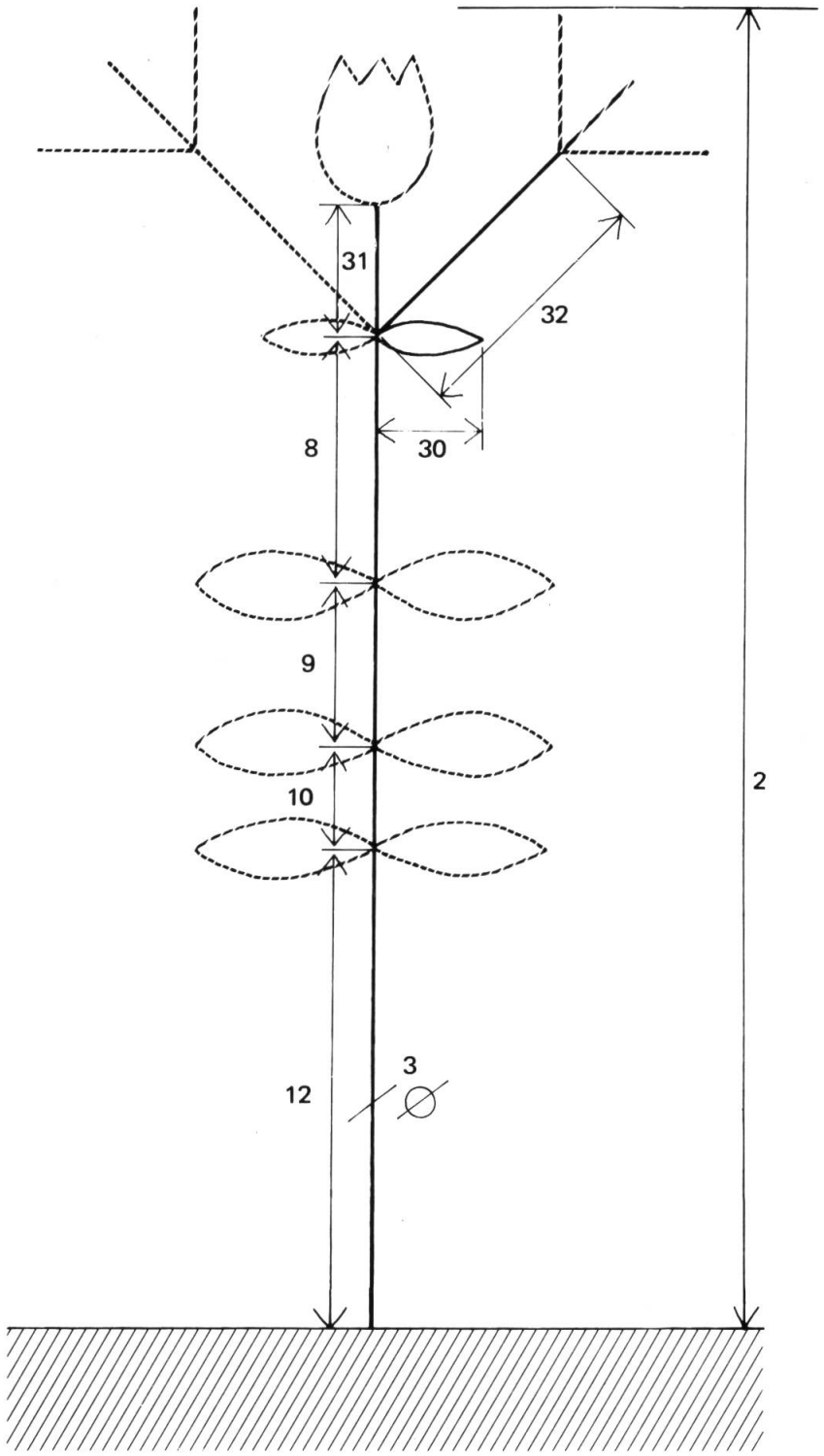

Fig. 1. — Caractères relatifs aux dimensions générales de la plante (2 à 32).

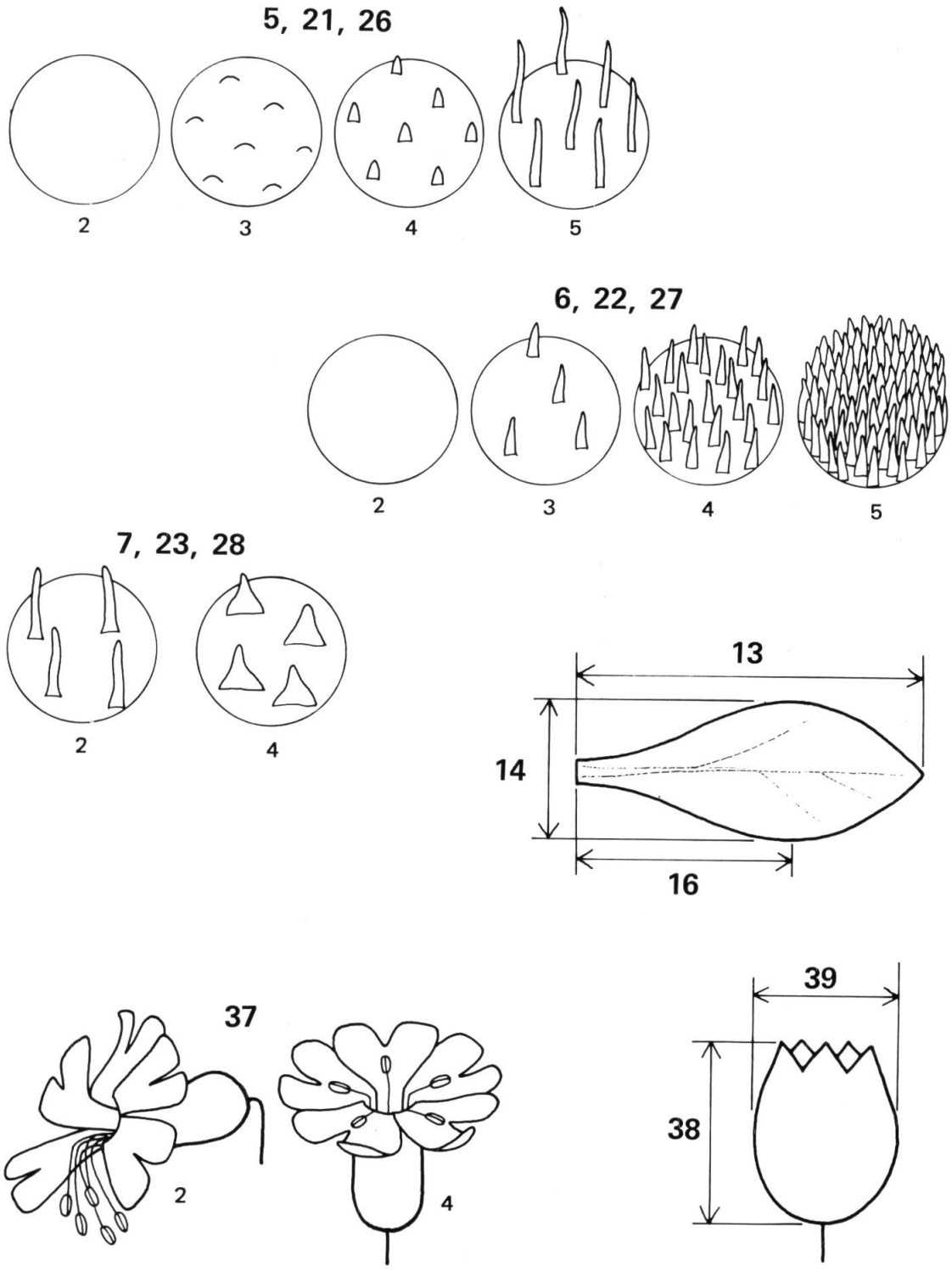

Fig. 2. — Caractères relatifs à la pilosité, à la surface foliaire et à la morphologie florale (5 à 39). Les petits chiffres correspondent au codage des différents états.

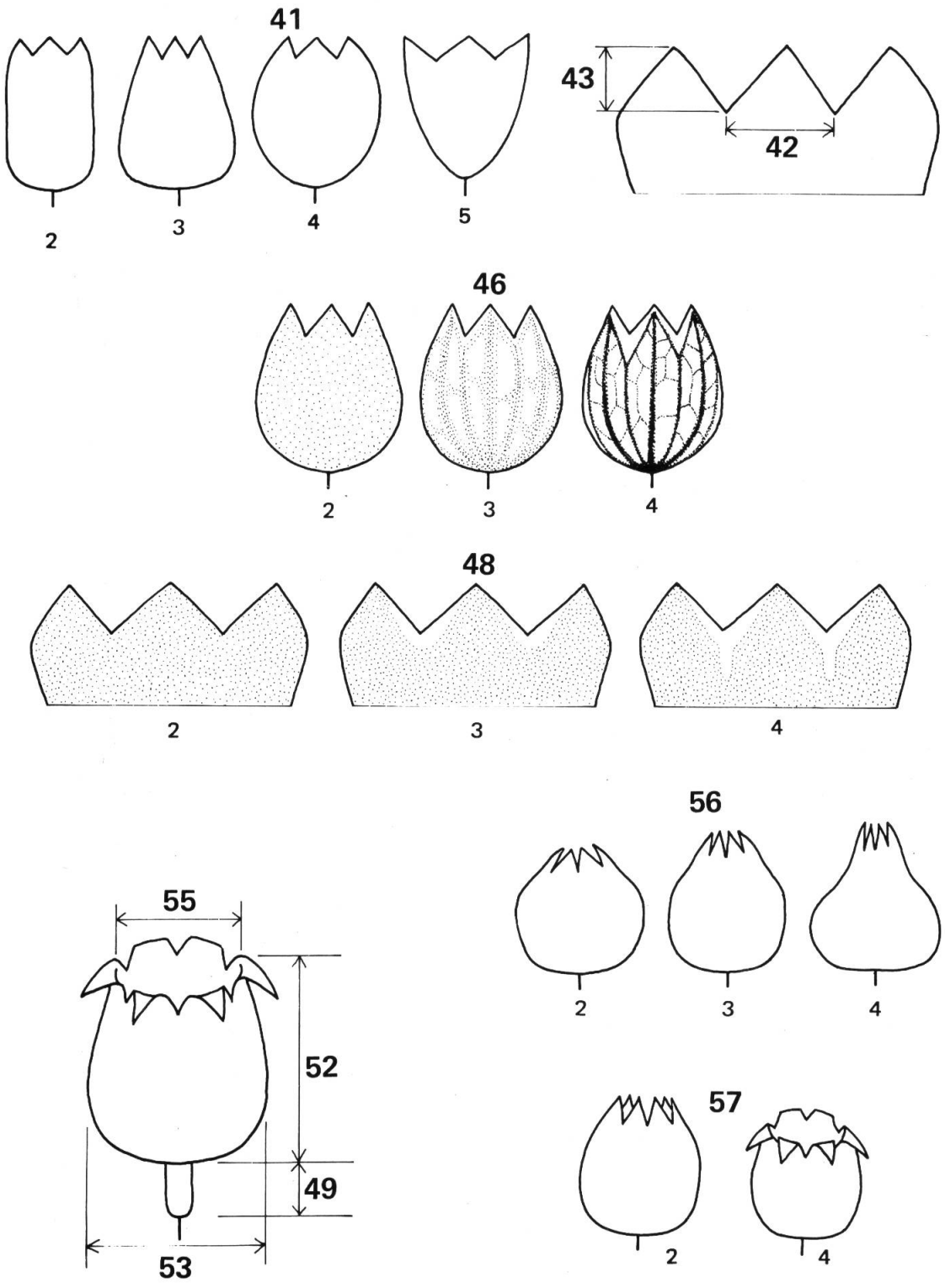

Fig. 3. — Caractères relatifs au calice (forme et pigmentation) et à la capsule (41 à 57). Les petits chiffres correspondent au codage des différents états.

Matériel électronique utilisé

Les mesures sont effectuées au moyen d'un analyseur de surface automatique "Videoplan" de marque Kontron (fig. 4). Il est composé d'une unité centrale Microsoft de 64 Kbytes basée sur un microprocesseur 8 bits Z80 de Zilog; d'une unité d'enregistrement de données à deux disquettes souples de 5″ (double densité) de 140 Kbytes chacune; d'un printer-plotter à impression thermique; d'une tablette de digitalisation de 28 × 26 cm; d'un clavier et d'un display-video. Outre les classiques interpréteur BASIC, compilateur FORTRAN et éditeur de texte, le logiciel fourni par Kontron est composé d'une série de programmes permettant la saisie automatique des données, et leur traitement statistique immédiat: analyse de population, histogrammes, analyse de variances, test de T, etc.

L'analyse numérique est réalisée par le Service d'informatique du Centre de botanique (Genève). Le traitement des données est effectué sur l'ordinateur UNIVAC 1100/60 du Centre universitaire d'informatique de Genève.

Dans la configuration actuelle, nous ne possédons pas l'interface nécessaire entre le Videoplan et l'ordinateur UNIVAC. Un passage par un encodage des données est donc encore nécessaire. Nous avons établi un

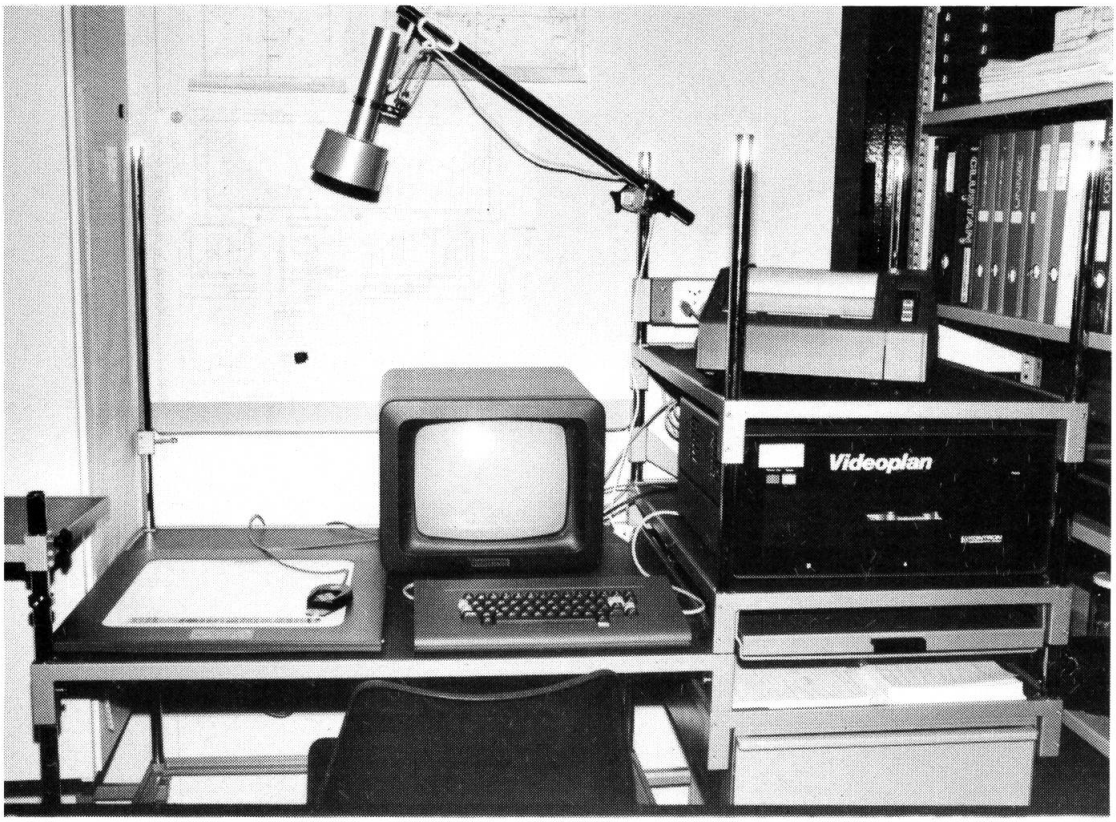

Fig. 4. — Analyseur de surface automatique "Videoplan" au Conservatoire botanique de la Ville de Genève.

bordereau de prise de données pour les travaux de taxonomie numérique (fig. 5). Il comprend deux parties distinctes: au recto, les information nomenclaturales, stationnelles, etc., relevées classiquement sur les étiquettes d'herbier. Cette partie est commune à tous les chercheurs de l'institut. Au verso, chaque chercheur peut définir à son gré les différents caractères qu'il compte utiliser. De cette manière, une banque de données taxonomiques peut être réalisée au long des divers travaux, tout en conservant à chacun d'eux son originalité.

Méthodes et programmes

Comme nous l'avons signalé plus haut, la matrice originale des données n'est pas homogène, c'est-à-dire que nous avons en présence, simultanément, des variables quantitatives et qualitatives, en parts à peu près égales. De plus, les spécimens utilisés pour l'analyse appartiennent à des populations différentes, de sous-espèces différentes. Il doit donc "a priori" exister une "structure" de ces données et c'est précisément elle que nous voulons mettre en évidence. Selon la terminologie de DAGNELIE (1977), nous sommes en présence d'une matrice rectangulaire, de structure (k_1, k_2). A ce stade de la réflexion, nous sommes convaincus, avec quelques autres auteurs (DAGNELIE, 1981, communication personnelle), qu'il n'existe pas de solution totalement satisfaisante pour résoudre ce problème dans son entièreté. Il nous faudra à chaque pas utiliser des méthodes robustes, ayant fait leurs preuves dans de nombreux domaines différents, ainsi que des logiciels confirmés.

Voyons d'abord ce que nous attendons d'une analyse numérique: aide à la prise de données, puis au stockage et à la recherche d'information, estimation de la distribution de chaque caractère, réduction et orthogonalisation de l'hyperespace des données, ordination, recherche de structure et classification. Il est évident que dans tout travail systématique, l'important est plus de connaître les relations qui existent entre les individus qu'entre les variables. Nous sommes donc amenés à choisir, parmi les différentes méthodes proposées, les techniques d'analyse factorielle, d'hiérarchisation, discriminantes. Nous adoptons l'algorithme général présenté à la figure 6. Nous ne reviendrons pas sur les points 1 à 3 qui ont fait l'objet des discussions précédentes (le stockage et la recherche d'information sont réalisés par le logiciel INFOL 2; CHENAIS & al., 1980). Pour le point 4, l'utilisation de logiciels, classiques pour un centre d'informatique, est évidente: SPSS (NIE & al., 1975), UNIVAC-STATPACK, OSIRIS. Quant au point 5! Comme nous l'avons vu, la matrice originale comprend à part égale des données qualitatives et quantitatives. Nous aurions pu choisir de traiter séparément chaque type. Mais, ayant toujours été convaincus que la grande originalité de l'analyse numérique est d'intégrer en une seule fois tous les caractères utilisés, nous avons préféré transformer les données quantitatives:

BORDEREAU TAXONOMIQUE — Numéro *1* 050802 — Détermination provisoire *2* S. vulgaris ssp. prostrata

Genre	*3*	Silene L.
Sous-genre Section	*4*	
Espèce	*5*	vulgaris (Moench) Garcke
Sous-espèce	*6*	prostrata (Gaudin) Chater & Walters
Variété	*7*	
Determinavit Date	*8*	Aeschimann, D. 1980
Syn. porté sur étiquette	*9*	
Famille	*10*	Caryophyllaceae — Collecteur *11* Aeschimann, D. — N° collecteur *12* 2274 — Date récolte *13* 12. 8. 1980
Pays	*14*	Italie — Province *15* Val d' Aoste
Localité	*16*	Alpe Thoules — Coordonnées *17* P* 45.54 N 7.20 E * — Altitude *18* 2500
Localité selon étiquette	*19*	500 m NW de Alpe Thoules, au fond du Val d' Ollomont.
Population	*20*	Homogène · Abondant · Fréquent
Biologie	*21*	Hermaphrodite / Pas de coronule.
Substrat	*22*	Micaschiste.
Station	*23*	Eboulis naturel, un peu mobile, non actif (pas d'alimentation), aux éléments moyens, pente 50 %, exp. SE.
Institution	*24*	G — Type *25*
Sources	*26*	
Remarques	*27*	Echantillon géologique récolté (Nos 2274 - 2279).
Nb d'étiquettes	*28*	2 — *29*

22/23

* 30 *	8	* 31 *	2	* 32 *	10	* 33 *	1.1	* 34 *		* 35 *	4	* 36 *	4	* 37 *	6	* 38 *	30.7	* 39 *	14.1
* 40 *	11.0	* 41 *	19.3	* 42 *	8.7	* 43 *		* 44 *	11.0	* 45 *	47.1	* 46 *	97.1	* 47 *		* 48 *	2	* 49 *	4
* 50 *	4	* 51 *	2	* 52 *	3	* 53 *	2	* 54 *	4	* 55 *	4	* 56 *	2	* 57 *	4	* 58 *	11.7	* 59 *	30.6
* 60 *	8.3	* 61 *		* 62 *		* 63 *		* 64 *	2	* 65 *	2	* 66 *	15.1	* 67 *	12.4	* 68 *		* 69 *	4
* 70 *	5.1	* 71 *	3.2	* 72 *		* 73 *	3	* 74 *	2	* 75 *	4	* 76 *	2	* 77 *		* 78 *		* 79 *	
* 80 *		* 81 *		* 82 *		* 83 *		* 84 *		* 85 *		* 86 *		* 87 *		* 88 *		* 89 *	
* 90 *		* 91 *		* 92 *		* 93 *		* 94 *		* 95 *		* 96 *		* 97 *		* 98 *		* 99 *	
* 100 *		* 101 *		* 102 *		* 103 *		* 104 *		* 105 *		* 106 *		* 107 *		* 108 *		* 109 *	
* 110 *		* 111 *		* 112 *		* 113 *		* 114 *		* 115 *		* 116 *		* 117 *		* 118 *		* 119 *	
* 120 *		* 121 *		* 122 *		* 123 *		* 124 *		* 125 *		* 126 *		* 127 *		* 128 *		* 129 *	
* 130 *		* 131 *		* 132 *		* 133 *		* 134 *		* 135 *		* 136 *		* 137 *		* 138 *		* 139 *	
* 140 *		* 141 *		* 142 *		* 143 *		* 144 *		* 145 *		* 146 *		* 147 *		* 148 *		* 149 *	
* 150 *		* 151 *		* 152 *		* 153 *		* 154 *		* 155 *		* 156 *		* 157 *		* 158 *		* 159 *	
* 160 *		* 161 *		* 162 *		* 163 *		* 164 *		* 165 *		* 166 *		* 167 *		* 168 *		* 169 *	
* 170 *		* 171 *		* 172 *		* 173 *		* 174 *		* 175 *		* 176 *		* 177 *		* 178 *		* 179 *	

Fig. 5. — Bordereau établi pour la prise de données en taxonomie numérique. En haut, le recto: informations nomenclaturales, stationnelles, etc. (items 1 à 29). En bas, le verso: caractères taxonomiques codifiés ou mesurés selon la liste propre du chercheur (items 30 à 179).

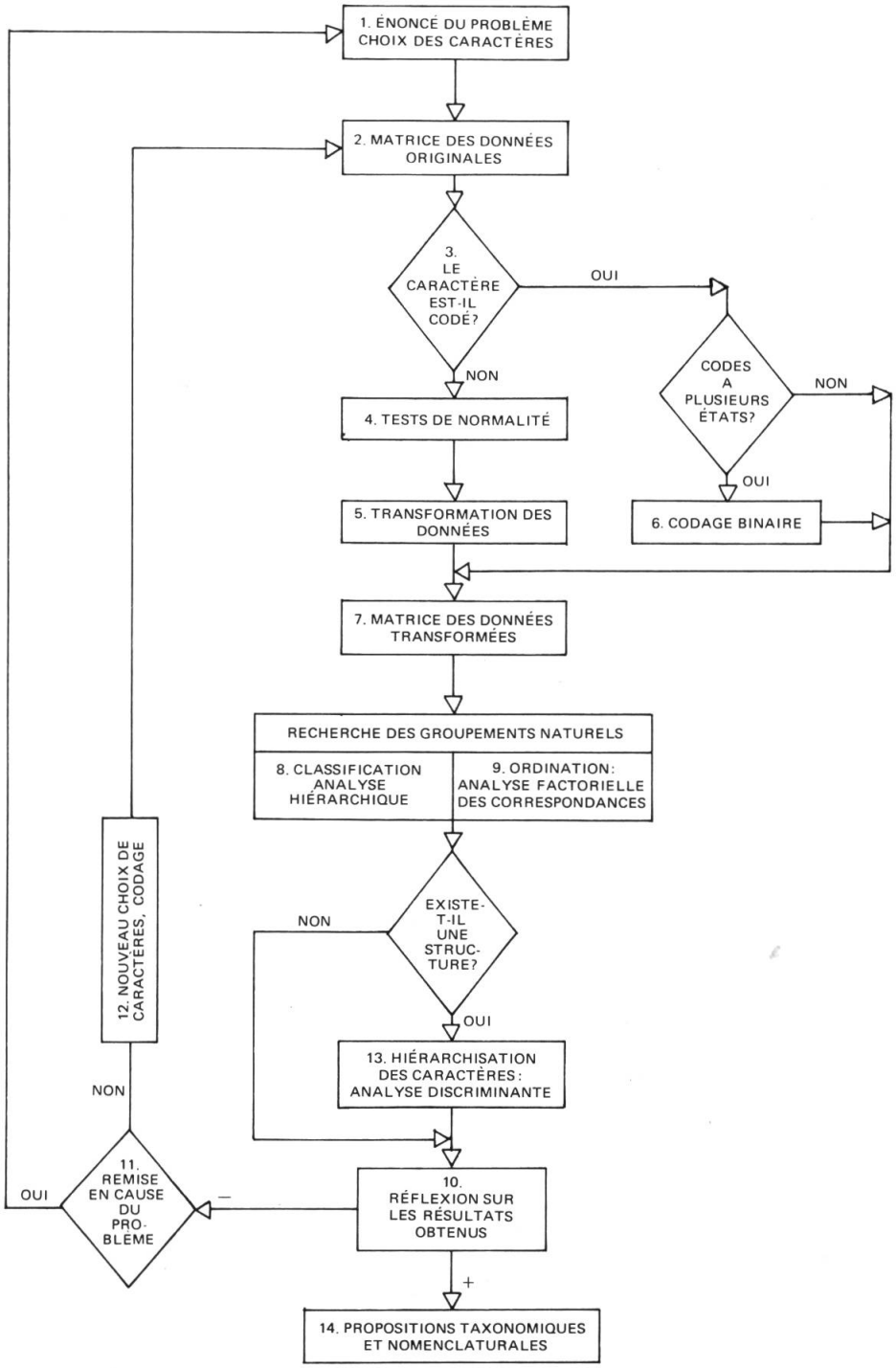

Fig. 6. — Organigramme général retenu pour notre étude.

— chaque variable quantitative x_n est transformée en deux variables x_{n1} et x_{n2}, suivant la méthode proposée par ESCOFIER (1979):

$$x_{n1} = \frac{1 - x_i}{2} \quad \text{et} \quad x_{n2} = \frac{1 + x_i}{2} \; ;$$

avec $x_{n1} + x_{n2} = 1$, où x_i est la valeur de la variable quantitative normée centrée;

— chaque donnée pure (nombre pur, rapport) est simplement normalisé selon la formule de GOWER (1971):

$$x = \frac{x - x_{min}}{(x_{max} - x_{min})}$$

— chaque caractère codé à plusieurs états est mis sous la forme disjonctive complète. Ainsi, un caractère pouvant prendre trois états différents, donne après transformation, trois caractères logiques.

Une fois ces séries de transformations effectuées, la matrice des données comprend, au point 7, 116 caractères, dont 59 logiques, 44 à 2 états et 13 purs, appartenant au sous-ensemble R^+ (0,1). Des logiciels de recodage des données originales sont proposés par FENELON & LEBEAUX (1978: 43-56), GHERMANI & al. (1977).

La recherche des groupements naturels est le nœud central de toute étude systématique. C'est à partir d'elle que se fait le découpage en taxa (ou unités taxonomiques), qui décidera de la solution taxonomique, puis de la typification et de la nomenclature. En taxonomie numérique, deux types de méthodes sont utilisables: les méthodes de classification et celles d'ordination.

Au point 8, les analyses hiérarchiques utilisées prennent pour base les logiciels CLUSTAN (WISHART, 1978) et DGRAPH5 (PARKS, 1970). Le produit final est presque toujours une matrice de similitude entre individus, calculée sur la matrice de données. Un dendrogramme permet le plus souvent de visualiser les nuages "groupements d'individus" en une solution graphique (MASCHERPA, 1976). Plusieurs indices sont utilisés pour construire la matrice de similitude: distance euclidienne, Rogers & Tanimoto, Sokal & Michener, etc.

Pour le point 9, nous avons délibérément choisi l'analyse factorielle des correspondances (ROBERT & NICOLAU, 1971), ce qui a conditionné les choix des étapes 4, 5, 6 et 7. Cette méthode, due à l'école française de BENZÉCRI (1973), présente l'avantage de permettre une visualisation simultanée des nuages des individus et des variables sur des graphiques plans des axes principaux d'inertie. Elle ne pose a priori aucune hypothèse de normalité, et ne prétend tester aucune hypothèse de départ. Comme la recherche des

corrélations entre variables n'est pas primordiale en systématique, la non-quantification des axes principaux n'est pas un désavantage par rapport aux autres analyses factorielles, de même que l'absence de possibilité de rotation des axes.

Bien que conçue au départ pour traiter des tableaux de contigence, elle a été très vite appliquée à l'étude de tableaux binaires. De nombreuses disciplines scientifiques, littéraires et sociologiques l'ont utilisé, bien que ce soit une technique relativement récente. Les "Cahiers de l'analyse des données" en donnent suffisamment de larges preuves. Dans les domaines des sciences naturelles, nous sommes souvent en présence de données hétérogènes. Très vite, les théoriciens de l'analyse factorielle des correspondances ont montré que le traitement de tels tableaux était possible, et même souvent sans codage des données quantitatives.

Ainsi par exemple: "... ne contenant que des nombres positifs, les tableaux de grandeurs additives extensives sont directement raitables par l'analyse des correspondances; ce traitement a l'avantage d'être, de par le principe d'équivalence distributionnelle, insensible aux regroupements ou subdivisions éventuels de colonnes. Il n'est donc pas besoin ici de codage ..." (BENZÉCRI, 1977: 30).

Ou encore: "... le succès maintenant bien compris des analyses de tableaux en 0.1, mis sous forme disjonctive complète, invite à rapprocher de cette forme, par un cadage approprié, les données les plus diverses ..." (BENZÉCRI, 1977: 31).

Et pour finir, cette profession de foi: "... mais parfois aussi, notamment en biologie, des mesures hétérogènes sont tout ce dont on dispose pour décrire un phénomène important, avec quelque fidélité; un sous-ensemble homogène de mesures laissant au contraire échapper tel ou tel aspect important. Le statisticien doit alors s'efforcer de traiter l'hétérogène ..." (BENZÉCRI, 1977: 333).

Après avoir établi et vérifié la structure de la population, il est utile de pouvoir rechercher, parmi tous les caractères observés, ceux qui serviront le mieux à définir les nouveaux taxa décrits. C'est le point 13 de notre organigramme, où les analyses discriminantes sont les plus performantes. Pour notre part, nous continuerons à utiliser le programme MAHAL3 (ROMEDER, 1973), qui traite les analyses discriminantes à plusieurs groupes.

Conclusion

La fin logique de toute analyse numérique revient évidemment au taxonomiste lui-même. C'est le point 14. Les nouveaux groupes taxonomiques ayant été précisés, typification et nomenclature peuvent venir achever la recherche. Il est cependant encore une question à poser. La structure que nous venons de retrouver dans la matrice des données est-elle cohérente avec ce que nous connaissons de nos spécimens, tout ce que le taxonomiste a pu

enregistrer au cours des récoltes sur le terrain, ou de l'observation en herbier? C'est l'objet de la réflexion des points 10, 11, 12 qui amène la remise en question soit de la structure découverte, soit celle du problème taxonomique lui-même, soit celle des données et de leur codification.

Contrairement à ceux pour qui l'ordinateur est un "Deus ex Machina", à ceux pour qui tout traitement automatique est œuvre malsaine, nous pensons qu'il est un juste milieu où le taxonomiste averti peut user des techniques modernes et garder les traditions classiques.

REMERCIEMENTS

Nous remercions vivement le Fonds national suisse de la recherche scientifique (projet n° 3.571-0.79) de son généreux appui; nous avons, grâce à cette importante aide, pu entre autres faire l'acquisition de l'analyseur automatique de surface "Videoplan". Nous remercions également Line Guibentif de sa précieuse collaboration dans la réalisation des dessins.

RÉFÉRENCES BIBLIOGRAPHIQUES

AESCHIMANN, D. (1979). *Etude morphologique et phytogéographique du Silene vulgaris s.l. (Caryophyllaceae) dans le domaine alpin occidental et les territoires avoisinants.* Travail de diplôme EPF, Zürich, 151 pp. (non publié).

— & G. BOCQUET (1980a). Allorhizie et homorhizie, une reconsidération des définitions et de la terminologie. *Candollea* 35: 19-35.

— & G. BOCQUET (1980b). Les types biologiques du Silene vulgaris s.l. (Caryophyllaceae). *Candollea* 35: 451-495.

ANDERSON, E. (1949). *Introgressive hybridization.* Wiley & Sons, New York, 109 pp.

BENZÉCRI, J.-P. (1973). *Analyse des données. I. La taxinomie. II. L'analyse des correspondances.* Dunod (éd.), Paris.

CHATER, A. O. & S. M. WALTERS (1964). Silene L. *In:* TUTIN, T. G. & al. (éds.), *Flora europaea:* 1. University Press, Cambridge.

CHENAIS, J.-M., J.-P. BAUD & A. JACQUESSON (1980). *INFOL 2. Manuel de référence.* Centre universitaire d'informatique, Genève.

DAGNELIE, P. (1977). A flow chart of multivariate statistical methods. *In:* BARRA & al. (éds.), *Recent developments in statistics.* North-Holland Publishing Company.

ESCOFIER, B. (1979). Traitement simultané de variables qualitatives et quantitatives en analyse factorielle. *Les Cahiers de l'analyse des données* 4: 137-146.

FENELON, J.-P. & M.-O. LEBEAUX (1978). Recodage d'un ensemble de données hétérogènes. *In:* JAMBU, M. & M.-O. LEBEAUX (éds.), *Classification automatique pour l'analyse des données. 2. Logiciels.* Dunod, Paris.

GHERMANI, B., C. ROUX & M. ROUX (1977). Sur le codage logique des données hétérogènes: présentation de deux programmes permettant de rendre homogènes des données quelconques. *Les Cahiers de l'analyse des données* 2: 115-118.

GODEAU, M. (1977). Valeur taxinomique des Silènes armoricains du groupe vulgaris (Moench) Garcke. I. Morphologie de l'appareil végétatif et des fleurs. *Rev. Gen. Bot.* 84: 355-376.

GOWER, J. C. (1971). A general coefficient of similarity and some of its properties. *Biometrics* 27: 857-871.

HESS, H., E. LANDOLT & R. HIRZEL (1967). *Flora der Schweiz:* 1. Birkhaeuser, Basel.

MARÉCHAL, R., J.-M. MASCHERPA & F. STAINIER (1978). Etude taxonomique d'un groupe complexe d'espèces des genres Phaseolus et Vigna (Papilionaceae) sur la base de données morphologiques et polliniques, traitées par l'analyse informatique. *Boissiera* 28.

MARSDEN-JONES, E. M. & W. B. TURILL (1957). *The bladder campions (Silene maritima and Silene vulgaris).* Ray Society, London, 378 pp.

MASCHERPA, J.-M. (1976). *Applications des méthodes informatiques à la taxonomie végétale, notamment des Phaseolinae, et aux problèmes biologiques qui lui sont liés.* Thèse 1767, Université de Genève.

MELZHEIMER, V. & J. DAMBOLDT (1973). Zur Morphologie und Cytologie tetraploider Sippen von Silene vulgaris (Caryophyllaceae). *Willdenowia* 7(1): 83-100.

NIE, N. H., C. HALDAI HULL, J. G. JENKINS, K. STEINBRENNER & D. H. BENT (1975). *SPSS: Statistical package for the social sciences.* Ed. 7. Mc Graw-Hill, New York.

PARKS, J. M. (1970). FORTRAN IV Program for Q. mode Cluster analysis on distance function with printed dendrogram. *Computer Contribution* 46. D. F. Merriam (ed.), Lawrence.

ROBERT, J. & F. NICOLAU (1971). CORRESPO 1. *In:* BENZÉCRI, J.-P., *Analyse factorielle des correspondances.*

ROMEDER, J.-M. (1973). *Méthodes et programmes d'analyse discriminante.* Paris.

WISHART, D. (1978). *CLUSTAN. User manual.* Ed. 3. Program library unit, Edinburgh.

Adresse des auteurs: Conservatoire et Jardin botanique de la Ville de Genève, Case postale 60, CH-1292 Chambésy/GE.