

Synthetic Taxonomy of *Rosa* Races Using ACT-STATIS

Olivier Raymond*, Jean-Louis Fiasson and Maurice Jay

Laboratoire de Biologie Micromoléculaire et Phytochimie, Université Claude Bernard Lyon 1,
43, boulevard du 11 novembre 1918, F-69622 Villeurbanne Cedex.

E-mail: phytochi@biomserv.univ-lyon1.fr

* Author for correspondence and reprint requests

Z. Naturforsch. **55c**, 399–409 (2000); received January 24/February 28, 2000

Rosa, ACT-STATIS, Taxonomic Congruence

Fifteen *Rosa* cultivated races were described by means of phenotypic frequencies (11 tables). Two groups of correlated contingency tables were identified by ACT-STATIS (Analyse Conjointe de Tableaux – Structuration de Tableaux à Trois Indices de la Statistique) inter-structure analysis. Three data sets appeared to be independent from the others. Typologies of races were obtained after ACT-STATIS compromise analyses for the two groups of correlated tables, and after Principal Component Analyses for the independent data sets. Each typology was original and variously influenced by genealogical structure, mutation or artificial selection pressures. A weighted synthesis was attempted in order to build a taxonomy of races taking into account these diversity factors. The good agreement between the resulting classification and the assumptions about the history of *Rosa* domestication advocated for a wider utilization of ACT-STATIS and RV coefficient when the relationships between individuals or populations have to be studied on the basis of their similarities.

Introduction

Any taxon resulting from “descent with modification”, its characteristics under a given descriptor reflect at different degrees its ascendency, adaptations (or responses to artificial selection) and mutation (Starck, 1998). Thus, the similarity between 2 taxa could be high or low according to the sensitivity of the descriptor to these various factors of diversity. Since a meta-analysis combines a number of descriptors, phenetic taxonomy (based on the assessment of similarities) as well as phylogenetic reconstruction were confronted with the problem of detecting discrepancies between several independent typologies of taxa, and of possibly retrieving a significant synthetic information. Within the phylogenetic framework, computation and validation of a consensus between several phylogenetic trees are well documented: strict consensus, Adam’s consensus or majority rule consensus could be obtained and supported by means of bootstrapping or jack-knifing. Methods for assessing the degree of incongruence between cladograms have been developed (Normark and Lanteri, 1998; Fu and Murphy, 1999). Furthermore, phylogenetic ambiguities could also be taken into account by net-

work representations (Bandelt, 1994). The limits of these methods, their advantages and the condition of their applications were often discussed and the debate between separate analyses vs. simultaneous analysis strategies is controversial (Mickel, 1978; Miyamoto, 1985; Tehler, 1995; Nixon and Carpenter, 1996). On the opposite, within phenetics the question of computing a compromise taxonomy from a set of possibly incongruent typologies remained marginal. We found no trace of a debate among taxonomists, data being generally combined prior to any analysis under the paradigm of “overall similarity”. Still, the assessment of the independence between characters, the choice between separate analyses and simultaneous analysis of potentially incongruent data matrices has the same theoretical importance here as in cladistic analyses. Now, in other research fields, numerous methods were built up to compare and summarize several matrices. Canonical analysis (Hotelling, 1936) and co-inertia analysis were developed to compare two data sets. Further developments allowed multiple comparisons: Generalized canonical analysis (Carroll, 1968; Kettenring, 1971), Multiple factor analysis (MFA) (Escofier and Pagès, 1990) and “Analyse conjointe de tab-

Race	Section
Alba	<i>Caninae</i>
Bourbon	<i>Chinenses</i>
China	—
Hybrids Perpetual	—
Noisette	—
Tea	—
Tea Hybrids	—
Rugosa Hybrids	<i>Cinnamomeae</i>
Centifolia	<i>Gallicanae</i>
Gallica	—
Moss	—
Portland	—
Floribunda	<i>Synstylae</i>
Polyantha	—
Wichuraiana Hybrids	—

a common *Chinenses* × *Gallicanae* ancestor, from which they derive by back-crossing with *Chinenses* material and / or mutation. Hybrids Perpetual cultivars, combining *Gallicanae* hardiness to *Chinenses* recurrent blossoming, were obtained by crosses between Portland, Bourbon and China Hybrids groups. Noisette, Tea and Tea Hybrids cultivars were also placed into this botanical section, despite their complex origin, involving *Chinenses*, *Synstylae* and *Gallicanae* sections.

Synstylae groups: Wichuraiana Hybrids, Polyantha and Floribunda cultivars originate from *R. wichuraiana*, *R. multiflora* or *R. moschata*. According to Arisumi (1963) the contributions of *Synstylae* species to these groups were important. Floribunda, Polyantha and Tea Hybrids share overlapping genealogies.

Others: Rugosa Hybrids and Alba cultivars were respectively classified within the *Cinnamomeae* and *Caninae* sections. However, Alba cultivars are supposed to originate from an initial cross between *R. gallica* and *R. canina*.

Data analysis

The descriptors used are summarized in Table II. Ten of them referred to morphological characteristics. Floral flavonol heterosides chemotypes were deduced from HPLC chromatograms of MeOH:EtOH extracts (Biolley *et al.*, 1994).

Eleven contingency tables, crossing biological groups and modalities for each descriptor, contained absolute frequencies and were further encoded in a fuzzy way. The inter-structure analysis implemented in ACT-STATIS was performed in order to compute RV correlation coefficients be-

tween all the data sets (Robert and Escoufier, 1976). After identification of the correlated typologies of biological groups, sharing high RV coefficients, compromise maps were computed according to the ACT-STATIS procedure, for each set of correlated structures. Finally, separate analyses (Principal Component Analysis) were performed for original typologies, *i.e.* those exhibiting low RV coefficients with all the other data sets.

All computations were performed with ADE-4 (Analyse des données écologiques) freeware, developed at Lyon University (Thioulouse *et al.*, 1997).

Results and Discussion

Correlations between data sets

Two groups of correlated data sets were recognized by RV coefficients (Table III). RV correlations were generally high within the group A of descriptors: pilosity of the branches, bristles, prickles quantity, shape of the prickles, width of the prickles and pilosity of the leaflets, with a maximum of 0.88 between prickles quantity and shape of the prickles. Statistical association between loci was probably involved in the strong correlation between the shape and the quantity of the prickles. A common ontogenic determinism could also explain the observed correlation between the presence of bristles on the branches and the leaflets. In the same way, shape of the corolla and petal number were closely connected (group B) reflecting the structural dependance between the shape of the corolla and the number of petals : quartered and rolled corollas could only be observed when flowers displayed numerous petals. All these situa-

Table II. Studied characters and coding of the modalities. Flavonol chemotype A : kaempferol 3-rhamnoside ; flavonol chemotype B : kaempferol 3-arabinoside and 3-xyloside; flavonol chemotype C : quercetin glycosides; flavonol chemotype D : kaempferol and quercetin 3-sophorosides.

Character	Modalities					
	1	2	3	4	5	6
1-Branch/pilosity	null or weak	sparse	abundant			
2-Branch/bristles	absent	present				
3-Prickles/quantity	small	medium	high			
4-Prickles/shape	very concave	concave	straight	convex		
5-Prickles/width	very Small	small	medium	large	very large	
6-Leaflets/pilosity	glabrous	hairy lower side				
7-Flower/shape of the corolla	flat	cup	rolled up	globular	quartered	helical
8-Flower/petal number	5	<20	>20	unvisible stamens		
9-Inflorescence/flower number	1	2 to 5	6 to 10	>10		
10-Flower color	white	rose	fuchsia	bright red	yellow	orange
11-Chemotype/flavonol glycosides	A	B	C	D		

Table III. RV coefficients between frequency tables.

	1-Pilosity of the branches	2-Bristles	3-Prickles quantity	4-Shape of the prickles	5-Width of the prickles	6-Pilosity of the leaflets	7-Shape of the corolla	8-Petal quantity	9-Flower number	10-Flower color	11-Flavonol chemotype
1-Pilosity of the branches	1.00	.	.	.	.	.	.	.	.	.	.
2-Bristles	0.48	1.00	.	.	.	.	.	.	.	.	.
3-Prickles quantity	0.67	0.69	1.00	.	.	.	.	.	.	.	.
4-Shape of the prickles	0.75	0.55	0.88	1.00	.	.	.	.	.	.	.
5-Width of the prickles	0.69	0.14	0.37	0.63	1.00	.	.	.	.	.	.
6-Pilosity of the leaflets	0.77	0.21	0.52	0.56	0.55	1.00	.	.	.	.	.
7-Shape of the corolla	0.47	0.12	0.22	0.32	0.42	0.44	1.00	.	.	.	.
8-Petal quantity	0.33	0.01	0.18	0.31	0.42	0.28	0.80	1.00	.	.	.
9-Flower number	0.26	0.06	0.22	0.26	0.23	0.17	0.29	0.34	1.00	.	.
10-Flower color	0.20	0.05	0.13	0.14	0.17	0.14	0.20	0.11	0.48	1.00	.
11-Flavonol chemotype	0.32	0.16	0.34	0.29	0.22	0.46	0.25	0.19	0.28	0.50	1.00

tions could be referred as methodological biases inducing false estimation of the distances between taxa.

As typologies were highly congruent within each group, a compromise typology allowed to avoid redundancy. On the contrary, the following data sets were not related to any other ones: number of flowers / inflorescence, floral colour and floral flavonol chemotype. Their underlying structures being original, each of those 3 sets was analysed separately.

Compromise typology of group A

Compromise analysis of tables pilosity of the branches, bristles, prickles quantity, shape of the prickles, width of the prickles and pilosity of the leaflets supplied a bi-dimensional map, displaying 86.3% of total inertia (Fig. 1A). *Rugosa* Hybrids and *Moss Centifolia* groups were defined by an abundant pilosity of the branches and presence of bristles. *Alba*, *Gallica*, *Centifolia* and *Portland* groups were defined by high frequencies of medium to high prickles quantity as well as by high

frequencies of small width and straight shape of the prickles. On the contrary, *Chinenses* and *Synstylae* races shared high frequencies of low prickles quantity, medium to large width and curved shape of the prickles.

Sections *Chinenses* and *Synstylae* were opposed to section *Gallicanae* along the F1 axis, while section *Cinnamomeae* was taken into account by F2 axis. Whereas the coordinate of *Moss Centifolia* on the F1 axis was coherent with its botanical classification in section *Gallicanae*, its position along the F2 axis emphasized the originality of this mutant group. Since classical taxonomy was based on other criteria than those used here, e.g. type of inflorescence, form of stipules and length of styles (Rehder, 1940), the observed congruencies between the compromise map of group A and the botanical classification support the value of prickles and pilosity as taxonomic markers, and have not been documented as yet.

Compromise typology of group B

Compromise analysis of tables shape of the corolla and petal number resulted in a bi-dimen-

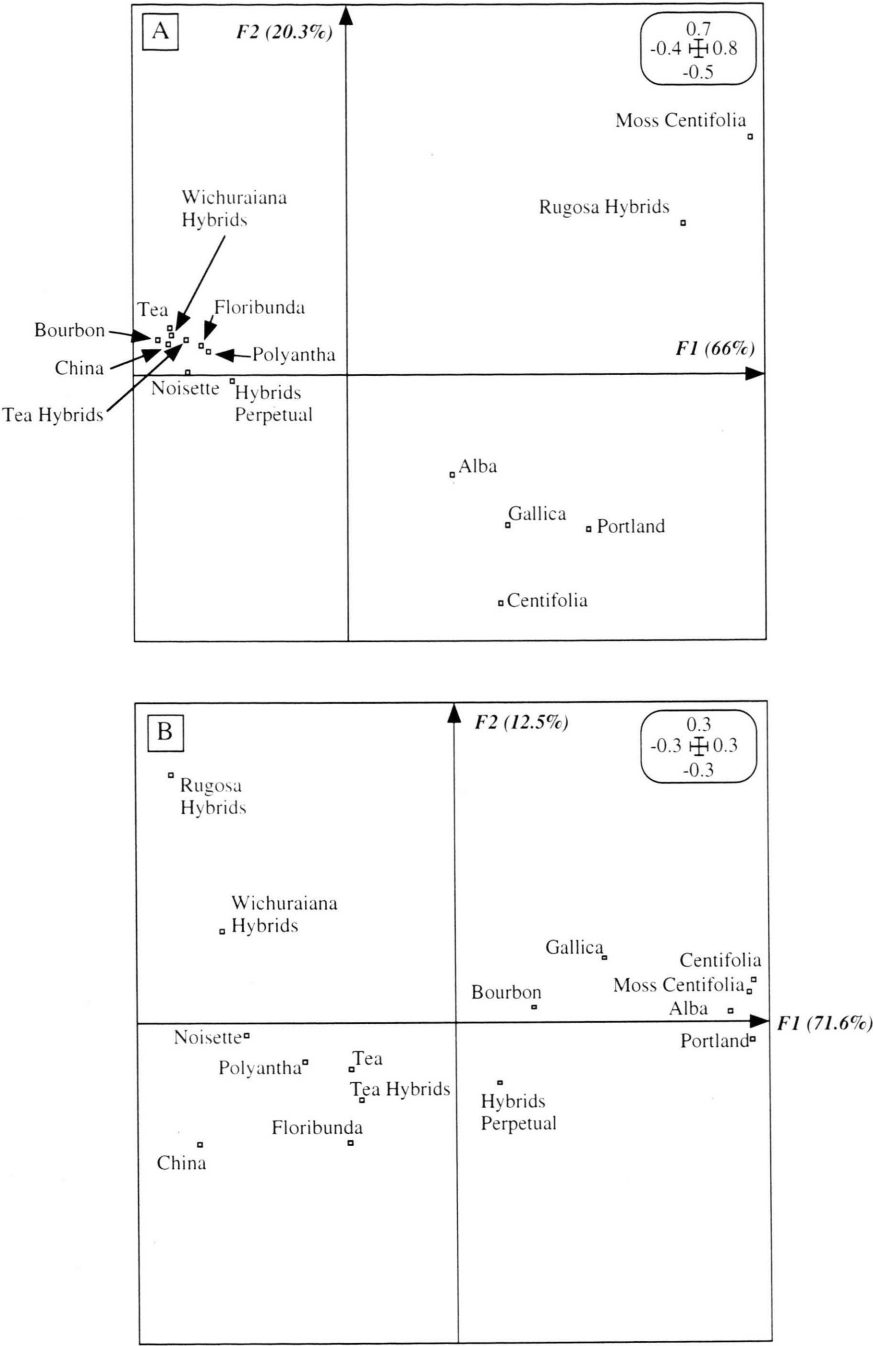


Fig. 1. A-F1*F2 map of the compromise between matrices of group A (pilosity of the branches, bristles, prickles quantity, shape of the prickles, width of the prickles, pilosity of the leaflets). B-F1*F2 map of the compromise between matrices of group B (shape of the corolla, petal quantity).

sional map dedicated to 84.1% of total inertia, revealing a complex pattern of diversity (Figure 1B). The F1 axis underlines the opposition between groups showing high frequencies of cultivars with high petal number and quartered corolla, namely Gallica, Centifolia, Moss Centifolia, Alba and Portland, vs. groups with high frequencies of cultivars with low petal number and cup corollas : China and Noisette. Groups exhibiting intermediate positions between these two extremes either contained a mix between high and low petal numbered forms – e.g. Bourbon and Hybrids Perpetual –, or included rolled-corollas cultivars –Tea Hybrids and Floribundas–. The F2 axis highlights the originality of Rugosa and Wichuraiana Hybrids, set up by high frequencies of flat corollas with low petal number, close to ancestral botanical forms.

Though the underlying botanical structure remained visible on the compromise map, tables for shape of the corolla and petal number brought informations on the hybrid origin of Bourbon and Hybrids Perpetual groups as well as on recent artificial selection trends, privileging rolled corollas over ancestral cup (China) and quartered (Gallica) forms.

Separate analyses of flower number, floral colour and flavonol chemotype

Original typologies of tables flower number, floral colour and flavonol chemotype required separate analyses (Fig. 2).

Flower number / inflorescence : races were split up into groups with (Fig. 2A):

- high frequencies of single-flowered cultivars (Tea, Tea Hybrids);
- high frequencies of cultivars with 2 to 5 flowers / inflorescence e.g. Gallica and Portland;
- high frequencies of 6 to 10 and more flowers / inflorescence, e.g. Wichuraiana Hybrids, Polyantha, Floribunda and Noisette, *i.e.* races with *Synstylae* ancestors.

Thus, flower number reveals the originality of *Synstylae*, which were not distinguished from *Chinenses* by previous compromise analyses. Moreover, Tea and Tea Hybrids illustrated another artificial selection trend, privileging single-flowered cultivars.

Floral colour did not reveal clear relationship with botanical origins of the different groups (Fig. 2B). However, the pattern of diversity of ancient races (Gallica, China, Alba and their first hybrids) was based on white, rose and *Fuchsia* colours. On the contrary, the diversity within some modern races (Noisette, Floribunda, Tea and Tea Hybrids) explored yellow and orange, indicating a recent trend of diversification.

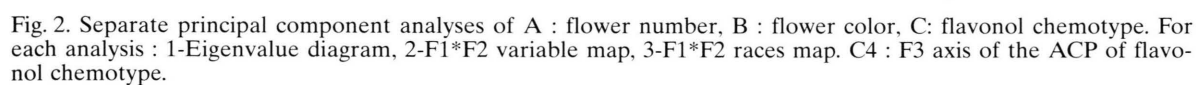
Flavonol chemotype distribution did not fit to any other structure, because of its tri-dimensional shape (Fig. 2C). Alba and Rugosa Hybrids shared high frequencies of 3-sophorosides chemotypes, known to be characteristic of *Cinnamomeae* and *Caninae* sections (Mikanagi *et al.*, 1990; Raymond *et al.*, 1995). China, Bourbon and Hybrids Perpetual shared high frequencies of quercetin based metabolisms whereas groups with *Synstylae* ancestors exhibited high frequencies of kaempferol based chemotypes. Finally, *Gallicanae* groups – Gallica, Centifolia, Moss Centifolia and Portland – were defined by high frequencies of 3-arabinosides metabolism.

Synthetic analysis

Unweighted analysis

The phenetic similarities between the 15 biological groups were organized by means of dendrograms. Euclidean distances were computed between the factorial coordinates of the 15 races after Principal Component Analysis performed on the raw data matrix. UPGMA cluster analysis revealed a major partition in 3 groups (Fig. 3A). According to inertia analysis, this partition was mainly determined by the variables of the subset A. Moreover, it is noteworthy that the dendrogram obtained from the ACT-STATIS compromise typology of races remained nearly identical, meaning that high RV coefficients between the tables of subset A concealed the information of the other descriptors, overestimating the similarities between races and justifying the group by group analysis (Fig. 3B).

The present study extended the field of application of ACT-STATIS to data sets showing various patterns of correlation. As predicted by the theory (Lavit *et al.*, 1994), the global compromise was mainly influenced by the most correlated data sets



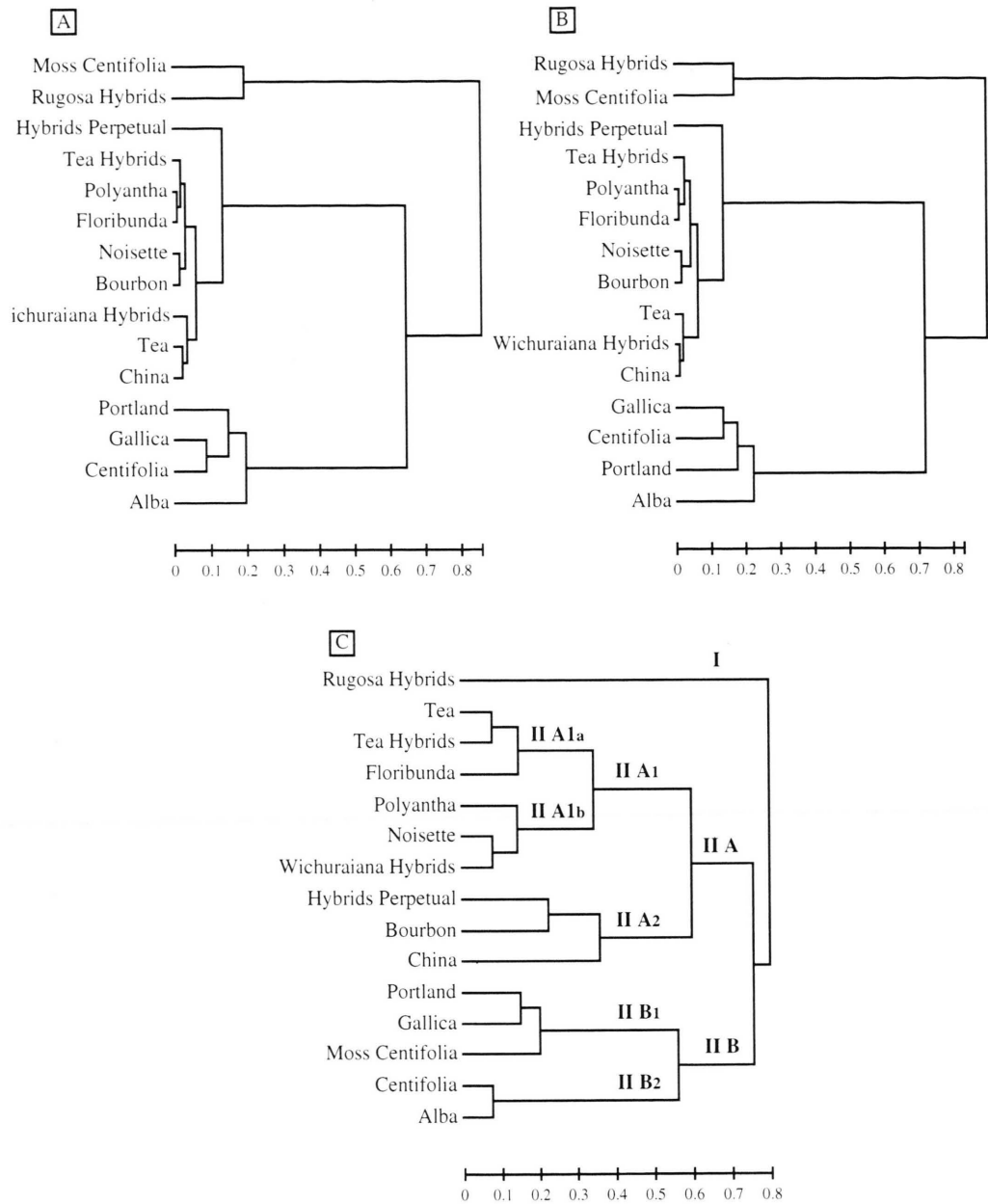


Fig. 3. A-UPGMA dendrogram of races obtained after PCA performed on the combined data sets. B-UPGMA dendrogram obtained after STATIS compromise performed on the combined data sets. C-UPGMA dendrogram of races obtained after identification and clustering of correlated data sets.

(group A), others being forgotten by this approach. A two-steps implementation protocol, where quantification of the correlation between

matrices was followed by the computation of compromises for each subset of correlated tables, was able to reveal the main underlying typologies.

Weighted analysis

In order to take into account the information provided by every subset of variables, a weighted comparison analysis was carried out. Distances were computed between the factorial coordinates resulting from the subset by subset data analysis. Each subset of factorial coordinates was weighted in order to moderate the influence of size effect: total inertia of each subset was computed after variables centring and each variable received a ponderation equal to $(\text{total inertia})^{-1}$. UPGMA cluster analysis produced a strikingly more informative synthetic dendrogram (Fig. 3C). The major split between I and (II A – II B) reflected the originality of Rugosa Hybrids (section *Cinnamomeae*) due to their characteristic patterns of variation of group A (prickles and pilosity) and flavonol chemotype. II A and II B were mainly explained by subset A, subset B (shape of the flower) and flavonol chemotype and have a taxonomic meaning, since II A could be associated with the *Chinenses* and *Synstylae* sections, whereas II B could be associated with the *Gallicanae*. The major partition reflected the underlying botanical organization in 4 sections and suggested that *Synstylae* and *Chinenses* are probably closer to each other than to *Gallicanae*.

The II A₁ and II A₂ dichotomy essentially layed on differences of flavonol chemotypes and flower number; it is noteworthy that this division separated the races with *Gallicanae* ancestors (II A₂) and the races with *Synstylae* ancestors (II A₁). On the contrary, the scission of II B, leading to II B₁ and II B₂, appeared to be artefactual, since the aggregation of Centifolia and Alba (II B₂) could not be explained by homogeneous patterns of variation. Finally, II A_{1a} and II A_{1b} could be related to selection trends privileging original colours (orange, yellow) and rolled corollas within the Tea, Tea Hybrids and Floribunda. Intermediate clusters reflected hybridization events, whereas lower level clusters showed selection trends in modern races.

The synthetic hierarchy of races was coherent with the botanical and historical knowledge. The obtained dendrogram illustrated the unequal influences of hybridization, mutation and artificial selection on some phenotypic characters, and was able to classify the factors of diversity by order

of decreasing discriminatory power in cultivated roses: genealogical structure, hybridization and artificial selection.

The similarities between the 15 races as seen by the synthetic phenotypic taxonomy are coherent with the typology obtained after RAPD genomic fingerprinting (Martin *et al.*, submitted). As an example, Portland and Gallica were tightly associated in both studies, giving good support to questioning the supposed *Gallicanae* × *Chinenses* origin of Portland roses. This coherency with a totally different approach strongly supports the use of ACT-STATIS to build a synthetic phenogram.

Conclusion

Within the phylogenetic framework, when numerous matrices are involved, each of them can be analysed separately (Sennblad *et al.*, 1998; Graham *et al.*, 1998); combination of data prior to a global analysis is also a possible strategy (Remsen and DeSalle, 1998). Whereas the first approach is adapted to the detection of heterogeneity between data sets, the second permits to benefit from the whole information. Nevertheless, separate analyses are often based on arbitrary partitions of the data (Lanyon, 1993; Baker and DeSalle, 1997). The ACT-STATIS approach united the advantages of both strategies: it permitted a combined analysis, but after benefitting from the separate analyses. In the prospects outlined by de Queiroz *et al.* (1995), developing tests dedicated to the detection of heterogeneity between data sets appeared as a priority. RV coefficient could be the decisional tool dedicated to build the clustering on an objective basis. However, testing the significance of such coefficients is still to be undertaken.

Independance of characters is a basic requirement of any taxonomical analysis, which take congruence between their distributions as reflecting a shared genealogical pattern. This assumption is violated by characters sharing a common determinism: they are actually a single character recorded more than once. RV coefficient, delineating clusters of characters with similar distribution, helps to spot out the potential sources of this analytical bias.

Finally, ACT-STATIS can help to distinguish, among the similarities between taxa, the phylogenetically significant ones from those reflecting a

convergence phenomenon, e.g. potentially to begin the separation between adaptative feature and phylogenetic burden : in our example, group A and chemotype appeared to be representative of the genealogical structure ; on the contrary, colour revealed common diversity patterns between chinese races (China, Bourbon, Hybrids Perpetual) and (Gallica, Portland) reflecting a parallel evolu-

tion due to similar artificial selection pressure on a shared metabolic basis.

Aknowledgements

Financial support by the Conseil Général du Val-de-Marne is greatly aknowledge. The authors thank Mrs Marys-Bessou, Mr Franchellin and Mr Brunel for their valuable assistance.

- Arisumi K. (1963), Studies on the flower colour in *Rosa*, with special references to the biochemical and genetic analyses and the applications of those results to the practical breeding. I. Science Bulletin of the Faculty of Agriculture, Kyushu University **20**, 131–148.
- Bandelt H. J. (1995), Combination of data in phylogenetic analysis. Pl. Syst. Evol. [Suppl.] **9**, 355–361.
- Beales P. (1985), *Roses* (Beales P., ed.), du Chêne, Paris.
- Biolley J. P., Jay M. and Viricel M. R. (1994), Flavonoid diversity and metabolism within 100 *Rosa* × *hybrida* cultivars. Phytochemistry **35**, 413–419.
- Buth G. M. and Misri K. (1984), Seed structure of some species of *Rosaceae*. J. Pl. Anat. Morph. **1**, 63–68.
- Carroll J. D. (1968), Generalization of canonical correlation to three or more sets of variables, Proc. Am. Psychol. Ass., 227–228.
- De Queiroz A., Donoghue M. J. and Kim J. (1995), Separate versus combined analysis of phylogenetic evidence, Annu. Rev. Ecol. Syst. **26**, 657–681.
- Debener T., Barttels C. and Mattiesch L. (1996), RAPD analysis of genetic variation between a group of rose cultivars and selected wild rose species. Molec. Breeding **2**, 321–327.
- Escoufier B. and Pagès J. (1994), Multiple factor Analysis (AFMULT package), Computational Statistics and Data Analysis **18**, 121–140.
- Fu J. and Murphy, R. W. (1999), Discriminating and locating character covariance : an application of permutation tail probability (PTP) analyses, Syst. Biol. **48**, 380–395.
- Graham S. W., Kohn, J. R., Morton, B. R., Eckenwalder, J. E. and Barrett S. C. H. (1998), Phylogenetic congruence and discordance among one morphological and three molecular data sets from Pontederiaceae, Syst. Biol. **47**, 545–567.
- Grossi C., Raymond O. and Jay M. (1998), Flavonoid and enzyme polymorphisms and taxonomic organization of *Rosa* sections : *Carolinae*, *Cinnamomeae*, *Pimpinellifoliae* and *Synstylae*, Biochem. Syst. Ecol. **26**, 857–871.
- Grossi C., Raymond O. and Jay M. (1999), *Rosa* taxonomy and hierarchy of markers defined by ACT STATIS, Z. Naturforsch. **54c**, 25–34.
- Hotelling, H. (1936), Relations between two sets of variates, Biometrika **28**, 321–377.
- Kettenring, J. K. (1971), Canonical analysis of several sets of variables, Biometrika **58**, 433–451.
- L'Hermier Des Plantes H. (1976), Structuration des tableaux à trois indices de la statistique. Théorie et application d'une éthode d'analyse conjointe. Thesis, Montpellier, 98 pp.
- Lavit C., Escouffier Y., Sabatier R. and Traissac P. (1994), The ACT (STATIS method), Computational statistics and data analysis **18**, 97–117.
- Lewis W. H. and Basye R. E. (1961), Analysis of nine crosses between diploid *Rosa* species. Proc. Amer. Soc. Hort. Sci. **78**, 573–579i
- Maia N. and Venard P. (1976), Cytotaxonomie du genre *Rosa* et origine des rosiers cultiés. In: Travaux sur le Rosier de Serre. INRA ed., Paris, pp. 7–20.
- Martin M., Piola F., Chessel D., Jay M. and Heizmann P. (submitted), The domestication process of the modern rose : genetic structure and allelic composition of the rose complex, Theor. App. Gen.
- Mickevich M. F. (1978), Taxonomic congruence. Syst. Zool. **27**, 143–158.
- Mikanagi N., Saito N. and Yokoi M. (1990), Flavonoid distribution in the flowers of genus *Rosa*, sections *Synstylae*, *Cinnamomeae* and *Pimpinellifoliae*. In: Flavonoids in Biology and Medicine III. Current Issues in Flavonoid Research, 89–96.
- Mikanagi Y., Yokoi M., Ueda Y. and Saito N. (1995), Flower flavonol and anthocyanin distribution in subgenus *Rosa*. Biochem. Syst. and Ecol. **23**, 183–200.
- Millan T., Osuna F., Cobos S., Torre A. M. and Cubero J. I. (1996), Using RAPDs to study phylogenetic relationships in *Rosa*. Theor. Appl. Genet. **93**, 273–277.
- Miyamoto M. M. (1985), Consensus cladograms and general classifications. Cladistics **1**, 186–189.
- Nixon K. C. and Carpenter, J. M. (1996), On simultaneous analysis. Cladistics **12**, 221–241.
- Normark B. B. and Lanteri A. A. (1998), Incongruence between morphological and mitochondrial-DNA character suggests hybrid origins of parthenogenetic weevil lineages (genus *Aramigus*). Syst. Biol. **47**, 475–494.
- Raymond O., Biolley J. P. and Jay M. (1995), Fingerprinting the selection process of ancient roses by means of floral phenolic metabolism. Biochem. Syst. Ecol. **23**, 555–565.

- Rehder A. (1940), In: Manual of Cultivated Trees and Shrubs. MacMillan, New York, 426–451.
- Remsen J. and DeSalle R. (1998), Character congruence of multiple data partitions and the origin of the hawaiian *Drosophilidae*. *Molec. Phylogen. Evol.* **9**, 225–235.
- Robert P. and Escoufier Y. (1976), A unifying tool for linear multivariate statistical methods : the RV coefficient. *Appl. Stat.* **25**, 257–265.
- Sennblad B., Endress M. E. and Bremer B. (1998), Morphology and molecular data in phylogenetic fraternity : the tribe Wrightiae (Apocynaceae) revisited, *Amer. J. Bot.* **85**, 1143–1158.
- Starck J. M. (1998), Non-independance of data in biological comparisons. A critical appraisal of current concepts, assumptions and solutions. *Theor. Biosci.* **117**, 109–138.
- Sytsma K. J. (1990), DNA and morphology : inference of plant phylogeny. *Tree* **5**, 104–110.
- Tehler A. (1995), Morphological data, molecular data, and total evidence in phylogenetic analysis. *Can. J. Bot.* **73**, 667–676.
- Thioulouse J., Chessel D., Doledec S. and Olivier J. M. (1997), ADE-4 : a multivariate analysis and graphical display software. *Statistics and Computing* **7**, 75–83.
- Ueda Y. and Okada Y. (1994), Discrimination of rose cultivar groups by pollen exine surface structure. *J. Hort. Sci.* **69**, 601–607.
- Wylie A. P. (1954), The history of garden roses. *J. Royal Hort. Soc.* **12**, 555–571.