
T. A. Dickinson; J. B. Phipps


Stable URL:
http://links.jstor.org/sici?sici=0363-6445%28198507%2F09%2910%3A3%3C322%3ASICL%28M%3E2.0.CO%3B2-R

*Systematic Botany* is currently published by American Society of Plant Taxonomists.

Your use of the JSTOR archive indicates your acceptance of JSTOR’s Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR’s Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/aspt.html.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR
ABSTRACT. Multivariate studies of flower, fruit, and leaf variation in *Crataegus crus-galli* L. s. l. have largely confirmed a priori notions of the structure of this species complex in Ontario. Four morphotypes are recognized and these are assigned to taxa published by earlier workers. Intensive sampling of local stands of the two most common morphotypes indicates a high degree of between-site heterogeneity and within-site homogeneity. It appears that the partitioning of variability observed in *C. crus-galli* s. 1. may be explained by features of dispersal, habitat preference, and the occurrence of uniparental reproduction. These results suggest possible explanations for the very narrow circumscriptions and limited distributions of many North American *Crataegus* taxa described at the beginning of the twentieth century.

Complex patterns of morphological variation were attributed to the North American species of *Crataegus* by Camp (1942) as part of an explanation for the very large number of binomials published for the eastern half of the continent (Palmer 1925). However, apart from univariate studies of variation in *Crataegus* sects. *Pruinosae* and *Crus-galli* by H. W. Rickett (Rickett 1936, 1937) and recent investigation of sect. *Pruinosae* (Sinnott and Phipps 1983) there has been no detailed description of these patterns.

We report here the results of a detailed multivariate morphometric study of a widespread eastern North American section of the genus, sect. *Crus-galli* Loud., as it occurs in Ontario. Section *Crus-galli* was selected for study because: 1) it exhibits a high degree of morphological differentiation in Ontario (Phipps and Muniyamma 1980) as well as elsewhere (Rickett 1937); 2) it is a polyploid complex (Muniyamma and Phipps 1979b); 3) it is highly distinctive and not readily confused with any other taxon; and 4) it is both frequent and abundant in much of southern Ontario (fig. 1; Phipps and Muniyamma 1980, maps 2 and 3). Until now studies of the relationship between phenotypic variation and breeding system in *Crataegus* have been restricted to species that appear to consist exclusively of sexual diploids (Bradshaw 1971, 1975; Love and Feigen 1978). Since the discontinuous morphological variation and polyploidy of sect. *Crus-galli* could be explained by the occurrence of gametophytic apomixis (Camp and Gilly 1943; Harlan and deWet 1975), this group provides an opportunity to document patterns of phenotypic variation probably associated with uniparental reproduction. We believe that such documentation is essential for an improved understanding of the taxonomic structure of North American *Crataegus*, and of the relationship between this structure and processes such as apomixis.

We report elsewhere on parallel studies that confirm the potential for uniparental reproduction of the material studied here (Dickinson 1983; Dickinson and Phipps unpubl. data).

The northern limit of sect. *Crus-galli* in Ontario coincides approximately with that of the deciduous forest region (beech-maple region, Braun 1950; Soper and Heimburger 1982). In southern Ontario it is represented by *Crataegus crus-galli* L. s. l., a polymorphic complex of at least four distinct morphotypes (Phipps and Muniyamma 1980). These morphotypes are designated A, B, C, and D for now, to leave open the question of name and rank until reference can be made to the results of the present study. The species complex is distinguished from other Ontario hawthorns by its shiny, leathery, virtually unlobed rhombic to obtrul-
late or obovate dark green leaves and medium to long (3–7 cm) thorns. As elaborated below, the morphotypes are conspicuously differentiated from each other by constant and readily recognizable combinations of flower, fruit, and foliage characteristics. Morphotypes A and B are common throughout all or most of the Ontario range of sect. *Crus-galli* (fig. 1), and frequently form extensive stands. In Ontario, *Crataegus* stands tend to occur on floodplains and poorly-managed agricultural land, especially on clay and clay loam soils (Phipps and Muniyamma 1980). Morphotypes C and D are each presently known only from a small number of isolated occurrences in the province (sites 4, 6, 14, figs. 1–2).

The objectives of the work reported here were 1) to investigate all of the main morphotypes of *C. crus-galli* s. 1. that occur in Ontario; 2) to determine the group structure of the sample obtained by objective methods; 3) to describe the apportionment of morphological variability among and within the groups in the sample; 4) to determine how the observed group structure results from descriptor variation and covariation; and 5) to compare the observed group structure with a priori perceptions. Meeting these objectives will help to provide insights into possible origins of the taxonomic complexities of North American *Crataegus*.

Since some kind of comparative basis is required to evaluate variability in *crus-galli* stands, the related species *Crataegus punctata* Jacq. was sampled as well. For purposes of comparison with *C. crus-galli* s. 1. the important feature of *C. punctata* is its virtually continuous morphological variation. *Crataegus punctata* also appears to be uniformly diploid in Ontario (Muniyamma and Phipps 1979b; Dickinson 1983). Furthermore, within the genus *Crataegus*, sect. *Punctatae* Loud. appears to be phenetically the most similar to sect. *Crus-galli* (Palmer 1946) and there are a number of putative hybrids between the two sections (Palmer 1946; Phipps and Muniyamma 1980). The distribution of *C. punctata* is broadly comparable to that of *C. crus-galli* s. 1., although it does not extend as far south, and it occurs further north (Palmer 1963; compare maps 2 and 4 in Phipps and Muniyamma 1980).

**MATERIALS AND METHODS**

**Sampling.** The operational taxonomic unit (OTU) and unit of sampling chosen was an individual hawthorn tree. Sites 1 through 5 (figs.
Crataegus crus-galli sensu lato

MORPHOTYPE A

☐ T1 (S1; N=10)
☐ T4 (S2; N=24)
◊ non-random sample, S2, S9, S11; N=7

MORPHOTYPE B

■ T2 (S1; N=17)
▲ T3 (S3; N=16)
● T5 (S2; N=7)
★ T6 (S4; N=19)
◊ non-random sample, S2, S6-S10, S12; N=16

MORPHOTYPE C

B S6, N=11

MORPHOTYPE D

★ T8 (S4; N=7)

INTERMEDIATE OTUS

★ S4, N=2

Crataegus punctata

P T7 (S5; N=18)
◊ non-random sample, S2, S3; N=3

Crataegus sect. Macracanthae

M S5, S13; N=3

1-2) were selected for intensive study because of their abundance of C. crus-galli s. l. (morphotypes A and B) or C. punctata. Sites 4 and 6 were chosen for morphotypes C and D, respectively. The remaining sites were selected because they represented important components of the Ontario distribution of sect. Crus-galli (sites 7, 8, and 9, fig. 1) or because they enabled inclusion of OTUs cited either as vouchers for cytological observations (sites 10 and 12; Muniyamma and Phipps 1979b) or as exemplars of specimens of particular taxa (sites 11 and 13; Phipps and Muniyamma 1980).

At sites 1 through 5 sampling frames were established using topographic maps and air photos. Within each a sample of hawthorns (≥2.0 m high) belonging to sect. Crus-galli (sites 2-4) or C. punctata (site 5) was chosen randomly and then permanently marked and vouchered. The random samples at sites 1 through 5 consist of seven different site-taxon combinations. Each of these is referred to as a topodeme sample (T1-T7, fig. 2), employing the -deme terminology of Gilmour and Gregor (1939; Gilmour and Heslop-Harrison 1954). We use the term topodeme to refer collectively to the individuals of a single taxon (morphotypes A or B, or C. punctata) at a given site. No other relationships among the individuals in a topodeme is implied. The seven OTUs of morphotype D at site 4 make up the entire topodeme of this taxon (T8, fig. 2; Dickinson and Phipps 1984). However, only the two OTUs of morphotype D included in the random sample at this site are considered in some of the detailed analyses described below.

At sites 6 through 13, and to a limited extent at sites 2 through 5 (fig. 2), additional OTUs were chosen preferentially. Apart from the concern to sample OTUs for which cytological or other data were already available, a major factor in the selection of OTUs, especially at sites more distant from London, Ontario, was the availability of flowers at the time of collection. In addition to the morphotypes of C. crus-galli s. l. and C. punctata, the total sample studied also included two individuals of C. macracantha Loud. and one of C. succulenta Schrad. ex Link, both species of Crataegus sect. Macracanthae Loud. These OTUs were included because of their apparent similarity to C. dispersma Ashe, a taxon of putative hybrid origin related to C. crus-galli s. l.; it was not possible to include C. dispersma in the study reported here. One of the two OTUs of C. macracantha, Stewart 2506, was cited as an exemplar of C. ?dispersma by Phipps and Muniyamma (1980). Further details of the sampling procedures and study sites are given in Dickinson (1983). Voucher specimens of all OTUs are deposited in the herbarium of the Department of Plant Sciences, University of Western Ontario (UWO).

Data collection. Descriptors known to describe variation within and among Crataegus...
taxa (table 1; Palmer 1925; Phipps and Muniyamma 1980) were chosen. Data for each of the 160 Crataegus OTUs collected at sites 1 through 13 (figs. 1–2) were obtained from 1) spring and fall herbarium vouchers, 2) a collection of flowers stored in alcohol, and 3) a record of fruit shape and size (fruit prints; Kruschke 1955). Ten of the flower and fruit descriptors listed with their abbreviations in table 1 were scored from these collections, as follows: PUB2, from the spring herbarium voucher, without replication; STYL, STAM, TCAL, and PUB1 from a total of 20 liquid-preserved flowers; PROJ from the undehisced anthers of 10 unopened flowers; WFL and LCAL from 10 open flowers; and LFR and WFR from the prints of 20 fruits. ANTH was scored in the field at the time of anthesis, without replication.

Leaf collections from 6–8 short shoots also were made for 2–7–10 OTUs randomly drawn from each topodeme represented in the random samples at sites 1–5 (fig. 2). In this way a subsample of 60 OTUs was available for which six leaf descriptors were scored in addition to the eleven flower and fruit descriptors (table 1). In order to control for variation due to leaf heteroblasty (Dickinson and Phipps 1984) the position of each leaf along the short shoot was recorded. Only data from terminal leaves were used in the studies reported here.

Replicated observations were summarized for each OTU by the descriptor mean and standard deviation, and by a descriptor-state frequency distribution (table 1). The vectors of descriptor means plus the scores for ANTH and PUB2 for each OTU made up the basic 11 x 160 and 6 x 60 data matrices used in most of the analyses described below.

Descriptor commensurability was obtained either by standardization to zero mean and unit standard deviation or by transformation to a 0–1.0 interval (condensation, Crovello 1968; ranging, Gower 1971). Data for descriptors PUB2 and ANTH were missing for three and ten randomly sampled OTUs, respectively. The corresponding topodeme sample mean scores for these descriptors were substituted for the missing value code (Sokal and Sneath 1963; Green 1979).

### Table 1. Descriptors of flowers, fruits, and short shoot leaves, showing states scored for each one. Maxima and minima observed for the flower and fruit descriptors in the 160 OTU sample are in italic. For the leaf descriptors, the maximum and minimum OTU means used in ranging are given in parentheses.

<table>
<thead>
<tr>
<th>STYL</th>
<th>Number of styles per flower (0, 1, 2, 3, 4, 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STAM</td>
<td>Number of stamens per flower (2–4, 5–7, 8–10, 11–13, 14–16, 17–19, 20–22, 23–25)</td>
</tr>
<tr>
<td>PROJ</td>
<td>Presence or absence of an apical projection of the stamen filament between the anther lobes (0, 1)</td>
</tr>
<tr>
<td>ANTH</td>
<td>Color of the undehisced anthers at time of anthesis (ivory, 0; faint pink, 1; pink, 2; red, 3)</td>
</tr>
<tr>
<td>TCAL</td>
<td>Degree of calyx lobe toothing (absent, or only one or two teeth on a single lobe, 0; isolated teeth on more than one lobe, 1; several on most lobes, 2; each lobe densely toothed, 3)</td>
</tr>
<tr>
<td>WFL</td>
<td>Flower width in mm (8 states at 0.5 mm intervals 2.5–6.5 mm)</td>
</tr>
<tr>
<td>LCAL</td>
<td>Length of longest calyx lobe in mm (14 states at 0.5 mm intervals 3.0–10.0 mm)</td>
</tr>
<tr>
<td>PUB1</td>
<td>Ovary pubescence (absent, 0; sparse, 1; dense, at least locally, 2; very dense, 3)</td>
</tr>
<tr>
<td>PUB2</td>
<td>Pedicel pubescence, scored as for PUB1</td>
</tr>
<tr>
<td>LFR</td>
<td>Fruit length in mm (10 states at 1.15 mm intervals 7.0–18.5 mm)</td>
</tr>
<tr>
<td>WFR</td>
<td>Fruit width in mm (10 states at 1.35 mm intervals 7.5–21.0 mm)</td>
</tr>
<tr>
<td>X</td>
<td>Leaf length above the widest point, in mm (12.14–26.40)</td>
</tr>
<tr>
<td>Y</td>
<td>Maximum leaf width perpendicular to X and Z, in mm (14.14–31.20)</td>
</tr>
<tr>
<td>Z</td>
<td>Leaf length below widest point, to leaf base, in mm. Z and X are colinear, and are aligned between the leaf apex and the point of attachment of the leaf to the stem (29.14–56.14)</td>
</tr>
<tr>
<td>NUMSEC</td>
<td>Number of secondary veins on one side of the midrib (5.00–9.38)</td>
</tr>
<tr>
<td>ANGSEC</td>
<td>Angle between the midrib and the fourth secondary vein from the apex, to the nearest five degrees (25.00–43.57)</td>
</tr>
<tr>
<td>TEETH</td>
<td>Number of teeth per 1.0 cm of leaf margin on one side of the leaf apex (5.25–11.38)</td>
</tr>
</tbody>
</table>

Data Analysis

The numerical analyses described below were carried out using an assortment of program packages (BMDP, Dixon and Brown 1979; SPSS, Nie et al. 1975; MINITAB, Ryan et al. 1981), published programs (Cooley and Lohnes 1971; Orlóci 1978; Podani 1980) and programs adapted or written for the purpose at hand by the first author. All of these were run on either the CYBER 170/835 or DECSYSTEM 1090 installation of the University of Western Ontario Computing Centre.

Data description. Univariate and bivariate distributions of the data were examined within the individual and pooled topodeme samples. This was done by calculating skewness and kurtosis statistics, and by means of scatterplots of OTU descriptor scores with, respectively, their normal probability scores (Ryan et al. 1981)
and with each other. The correlation structure of the flower and fruit data (111 OTU random sample) and that of the leaf data (58 OTU subsample) was examined by comparing descriptor product-moment correlation coefficients with the corresponding partial correlation coefficients (Rohlf 1977), and by means of some of the graphical methods suggested by Campbell (1981).

Multivariate variability of the individual topodeme samples (T1–T7) was assessed by testing a null hypothesis of the equality of their variance-covariance matrix determinants (Cooley and Lohnes 1971). Because this test is sensitive to departures from normality (Layard 1974) it was supplemented by a more robust statistic, the total variance (Van Valen 1978). This is calculated as the Euclidean distance in p-dimensional space between each OTU and the centroid of the group to which it belongs, as

\[ y_{Bi} = \left( \sum_{j} (x_{ij} - \bar{x}_k)^2 \right)^{1/2} \]

for the ith OTU, the kth group, and for \( j = 1, \ldots, p \) descriptors, or as the corresponding squared generalized distance

\[ y_{Mi} = (x_i - \bar{x}_k)W^{-1}(x_i - \bar{x}_k) \]

where \( W \) is the pooled within-groups covariance matrix for the entire random sample (see below). The variability of the group is then described in terms of the mean value of these distances \( \bar{y}_p, \bar{y}_s \). Statistical comparisons were carried out by means of one-way analyses of variance (program BMDP7D, Dixon and Brown 1979). A related multivariate measure of variability, Gilmartin's (1974) mean phenetic distance (MPD), was calculated for each topodeme sample as the average of the between-OTU Euclidean distances. Differences in the dispersion of topodeme samples, as well as in their location also have been illustrated by means of the plotting procedure described by Andrews (1972; Gnanadesikan 1977). In this method, ordination scores for each OTU \( (y_1, \ldots, y_s) \) for \( t \) ordination axes) are used as coefficients of the terms of a trigonometric expression so as to perturb a standing wave in a characteristic way. Thus for the ith OTU and 6 axes,

\[ f(x) = y_{1i}/\sqrt{2} + y_{2i}\sin(x) + y_{3i}\cos(x) + y_{4i}\sin(2x) + y_{5i}\cos(2x) + y_{6i}\sin(3x). \]

In figure 6 the values for \( x \) are multiplied by \( 2\pi \) so as to normalize the abscissa to a 0–1.0 range (Gnanadesikan 1977).

**Group structure.** The group structure of the entire 160 OTU sample is understood as its differentiation into polythetic groups of similar OTUs. Similarity was measured pairwise between OTUs, using three dissimilarity coefficients and data for the 11 flower and fruit descriptors. Euclidean distances and generalized distances were calculated from matrices of OTU descriptor scores standardized to zero mean and unit standard deviation. Both coefficients were calculated according to Orloci (1978); generalized distances were calculated from the total sample correlation matrix eigenvalues and component scores produced by R-algorithm principal components analysis (R-PCA). The information radius between OTUs (Jardine and Sibson 1971; Orloci 1978) was calculated directly from the observed OTU descriptor-state frequency distributions according to the formulation given by Prentice (1979). To distinguish between data-dependent and method-dependent features of the results, three sorting algorithms were used to obtain groups of similar OTUs: single-linkage, average-linkage (UPGMA), and minimum variance. These methods are described fully by Sneath and Sokal (1973) and Orloci (1978); the calculations were carried out using program NCLAS (Podyani 1980).

**Contributions of descriptors to group structure.** Ordination methods were used to corroborate the results of the cluster analyses and to discover the way in which individual descriptors contributed to observed group structure. R-algorithm principal components analysis (R-PCA) of the entire 160 OTU sample employed the total sample correlation matrix for the 11 flower and fruit descriptors (table 1). This ordination depicts the configuration of OTUs represented by the matrix of Euclidean distances calculated from standardized descriptors described above (Gower 1966).

The remaining analyses of the flower and fruit data concerned only the randomly sampled OTUs and employed that subset of six de-
scripts whose univariate distributions most nearly approximated normality (see below). Canonical correlation analysis (CCA) was used to study the congruence of the two data sets available for the 60 OTU subsample (six flower and fruit descriptors plus six descriptors of short shoot terminal leaves; table 1). In this analysis, as in the R-PCA described above, the sample was treated as if it represented a single homogeneous group.

To examine the extent to which taxa and topodeme samples were differentiated from each other, these were used as alternative sets of a priori groups in another series of ordinations. Depending on whether or not the dispersions of these groups were found to be equal (as described above) either a multivariate analysis of variance (and canonical variates (CVA) ordination; Gittins 1979; Pimentel 1979) or multigroup R-PCA (Campbell 1976; Pimentel 1979) was carried out. In the latter case the ordination is based on eigenanalysis of the pooled within-groups dispersion matrix W. It should be noted that these latter analyses evaluating the differentiation of the taxonomic groups and topodeme samples examined the original covariance structure of the data, since they were based on descriptors rendered commensurate by ranging rather than by standardization (Orlóci 1978).

With each of these ordination methods the contribution of descriptors to the group structure portrayed was assessed by means of their canonical weights and (or) correlations with the ordination axes. These relationships are represented graphically as vectors. Descriptor communalities indicating the proportion of descriptor variance accounted for by selected ordination axes also were calculated. In the canonical analysis described above, intraset and interset communalities were calculated. These represent the proportion of the descriptor variance for a domain (leaf, or flower and fruit) accounted for by the canonical variates of that domain or of the other domain, respectively. It is also possible to calculate for each domain the percent of its total variance explained by the canonical variates of that domain or of the other one. Thus, under columns $H_{1}$ and $H_{2}$ in table 5, respectively, we give the total percent domain variance explained by all of its canonical variates, and by all of the canonical variates of the other domain (total redundancy; Cooley and Lohnes 1971). Details of the calculations are given by Cooley and Lohnes (1971), Green (1978) and Pimentel (1979).

**Results**

Distributions of the descriptors within the individual topodeme samples were normal, except in the case of T2, which contained two large-fruited outlier individuals (fig. 6). All of the multistate descriptors, some of which were invariant or nearly so within groups, had non-normal distributions. Likewise, meristic descriptors such as STAM and TEETH were distributed bimodally over the random samples for which they were scored. The distributions of continuous descriptors (WFL, LFR, WFR, X, Y, Z) tended to be more nearly normal. In all cases the departures from normality were associated with the taxonomic heterogeneity of the sample (figs. 3–4).

Examination of bivariate distributions was restricted to the six flower and fruit descriptors whose univariate distributions were least affected by the heterogeneity of the sample, and which were not invariant in any of the topodeme samples (STYL, TCAL, WFL, LCAL, LFR, WFR), and to the six leaf descriptors. The generally linear, continuous relationships observed among these descriptors are summarized by the semi-matrices of product-moment correlation coefficients in tables 2 and 3. The heterogeneity of the samples was evident however, either in these scatterplots (X-TEETH, NUMSEC-TEETH, ANGSEC-TEETH) or in plots of the corresponding regression residuals (STYL-TCAL, STYL-LFR, WFL-LCAL). The contrast between topodeme samples T6 (morphotype A, site 4) and T7 (C. punctata, site 5) was especially pronounced in comparisons involving the descriptor STYL.

Comparison of product-moment correlations ($r_{hi}$) with the corresponding partial correlations ($r_{hi}$, estimating the correlation between descriptors $h$ and $i$ independent of the effects of all the other descriptors considered) provides a means of determining how patterns of descriptor covariation are repeated or are unique (Rohlf 1977; Orlóci 1978). The absolute values of $r_{hi}$ for STYL-LCAL, WFL-LFR, LFR-WFR, X-Z, and Y-Z are markedly greater than those of the corresponding partial correlations (tables 2–3),
indicating that the correlations between these descriptors are largely explained by their correlations with other descriptors. A similar situation was observed with STYL-STAM, WFL-PROJ and WFR-STAM correlations in the 160 OTU sample (fig. 4). Conversely, the absolute values of the correlations between TEETH and X, Y, NUMSEC, and ANGSEC are all considerably less than those of the corresponding partial correlations (table 3), suggesting that
variation in TEETH among the OTUs (between sects. Crus-galli and Punctatae) is different from that of any of the other descriptors (fig. 5).

For descriptors STYL, TCAL, WFL, LCAL, LFR, and WFR the within-topodeme sample standard deviation did not increase as the corresponding mean increased (Dickinson 1983). As a consequence, comparison of topodeme samples by means of coefficients of variation is inappropriate (Van Valen 1978). Both data sets available for the random sample indicated the inequality of the dispersions of the three taxonomic groups present. With the smaller sample and more variable descriptors the leaf data supported acceptance of the null hypothesis of equal topodeme sample dispersions, on the basis of tests of both covariance matrix determinants and average total variances (Dickinson 1983). The same null hypothesis was rejected using the determinants of the flower and fruit descriptor dispersion matrices ($P < 0.001$; for both T1–T6 and T1–T7). Heterogeneity of the topodeme sample dispersion matrices appears to have been exaggerated by the effect of particularly small variances and large correlations on the magnitude of some of the crus-galli to-
Table 2. Flower and fruit descriptor correlation coefficients, $r_{hi}$ (upper semi-matrix), and partial correlations, $r_{hi}$ (lower semi-matrix; selected descriptors, table 1; topodeme samples T1–T7, figs. 1–2).

<table>
<thead>
<tr>
<th></th>
<th>STYL</th>
<th>TCAL</th>
<th>WFL</th>
<th>LCAL</th>
<th>LFR</th>
<th>WFR</th>
</tr>
</thead>
<tbody>
<tr>
<td>STYL</td>
<td>0.599</td>
<td>0.676</td>
<td>0.337</td>
<td>0.311</td>
<td>0.669</td>
<td></td>
</tr>
<tr>
<td>TCAL</td>
<td>0.575</td>
<td>0.145</td>
<td>0.407</td>
<td>−0.129</td>
<td>0.147</td>
<td></td>
</tr>
<tr>
<td>WFL</td>
<td>0.406</td>
<td>−0.242</td>
<td>0.226</td>
<td>0.532</td>
<td>0.753</td>
<td></td>
</tr>
<tr>
<td>LCAL</td>
<td>0.036</td>
<td>0.269</td>
<td>0.147</td>
<td>−0.032</td>
<td>0.142</td>
<td></td>
</tr>
<tr>
<td>LFR</td>
<td>−0.153</td>
<td>−0.143</td>
<td>0.091</td>
<td>−0.065</td>
<td>—</td>
<td></td>
</tr>
<tr>
<td>WFR</td>
<td>0.395</td>
<td>−0.054</td>
<td>0.334</td>
<td>−0.029</td>
<td>0.587</td>
<td></td>
</tr>
</tbody>
</table>

Cluster analyses of the entire 160 OTU sample (fig. 3) consistently demonstrated both the homogeneity of the individual topodeme samples and their marked differentiation from one another, regardless of resemblance function and sorting algorithm (Dickinson 1983). They also supported differentiation of the major a priori taxonomic groupings of OTUs (morphotypes of C. crus-galli s. l.; sects. Crus-galli, Punctatae, and Macracanthae; fig. 3). The degree of differentiation observed depended to a considerable extent on the sorting algorithm used. Group structure appeared weakest in phenograms produced by space-contracting single-linkage clustering, as well as in most of those based on generalized distances. Conversely, the space-dilating effect of minimum variance clustering produced distinct groups (topodemes, morphotypes) even in the phenogram based on generalized distances. The minimum variance phenogram based on Euclidean distances calculated from standardized descriptors split the sample into two groups based on stamen number: C. punctata plus morphotypes A and D (15–25 stamens per flower), and morphotypes B and C (5–15 stamens per flower) (compare fig. 7e in Podani and Dickinson 1984). Ordination (R-PCA) of the entire 160 OTU sample (fig. 4; table 4) demonstrated the way in which the groups of OTUs formed by the cluster analyses (fig. 3) result from their similar scores on a number of highly correlated descriptors.

Table 3. Leaf descriptor correlation coefficients, $r_{hi}$ (upper semi-matrix), and partial correlations, $r_{hi}$ (lower semi-matrix) for the 60 OTU subsample from sites 1–5 (figs. 1–2).

<table>
<thead>
<tr>
<th></th>
<th>X</th>
<th>Y</th>
<th>Z</th>
<th>NUMSEC</th>
<th>ANGSEC</th>
<th>TEETH</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>0.562</td>
<td>0.491</td>
<td>−0.065</td>
<td>−0.050</td>
<td>−0.348</td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>0.500</td>
<td>0.616</td>
<td>0.155</td>
<td>−0.073</td>
<td>0.244</td>
<td></td>
</tr>
<tr>
<td>Z</td>
<td>0.279</td>
<td>0.426</td>
<td>0.164</td>
<td>−0.210</td>
<td>−0.552</td>
<td>0.212</td>
</tr>
<tr>
<td>NUMSEC</td>
<td>−0.048</td>
<td>0.057</td>
<td>0.164</td>
<td>−0.210</td>
<td>−0.552</td>
<td>0.212</td>
</tr>
<tr>
<td>ANGSEC</td>
<td>0.045</td>
<td>0.232</td>
<td>−0.559</td>
<td>0.695</td>
<td>0.672</td>
<td>0.134</td>
</tr>
<tr>
<td>TEETH</td>
<td>−0.581</td>
<td>0.436</td>
<td>0.036</td>
<td>0.302</td>
<td>−0.341</td>
<td>—</td>
</tr>
</tbody>
</table>
FIG. 5. Disposition of the 60 OTU subsample of randomly sampled OTUs in the leaf (U₁, U₂) and flower and fruit (V₁, V₂) domains obtained from canonical correlation analysis. Vectors represent the contribution of each descriptor to the distribution of OTUs in terms of the corresponding canonical weights. Symbols as described in figure 2; see table 5 for additional details.

that of the flower-fruit domain than vice versa (total redundancies, table 5).

As described already, testing the differentiation of the random sample into its constituent taxa and topodeme samples demonstrated the homogeneity of their dispersions only with respect to the leaf descriptors, and then only for the topodeme samples. CVA of the leaf data showed that the six descriptors contributed almost equally to the statistically significant (P < 0.01) separation of the topodeme samples (Dickinson 1983). Similar results with respect to taxa are reported elsewhere (Dickinson and Phipps 1984).

Using six ranged flower and fruit descriptors, R-PCA of the pooled within-group covariance matrix W for the seven topodeme samples yielded four significant principal components. Of these, the first two accounted for only 62% of the total sample variance. Display of all six dimensions of the analysis, by means of Andrews plots, demonstrated both the distinctness of morphotypes and topodeme samples and the contrast in variability between T1-T6 and T7 (fig. 6). Multi-group R-PCA of the leaf data likewise indicated the distinctness of morphotypes A, B and D (Dickinson and Phipps 1984).

DISCUSSION

The analyses reported here are the first to describe the pattern of morphological variation exhibited within and among individual sites by a *Crataegus* species complex. Results of these analyses mostly agree with the a priori classification of the 160 OTU sample according to section (*Crus-galli, Punctatae, Macracanthae*) and morphotype (A–D). The latter result is not surprising, in view of the use made of some of the descriptors employed earlier in distinguishing these groups, or of ones highly correlated with those descriptors.

Unlike the corresponding Euclidean distances, generalized distances among OTUs are calculated so as to be independent of the pattern of descriptor correlations (Orlić 1978; Atchley et al. 1982). Accordingly, for a given sorting algorithm the differences observed between phenograms based on these distances are an indication of the extent to which the observed group structure is due merely to sampling a large number of highly correlated descriptors. In the cluster analyses taxa and topodeme samples largely remained distinct regardless of the resemblance coefficient employed (fig. 3; Dickinson 1983; Podani and
Dickinson 1984). According to Atchley et al. (1982), such a result suggests that differences among groups are not likely to be due only to the action of a small number of pleiotropic genes.

Compared to the flower and fruit data the leaf descriptors are much more variable within taxa and within topodeme samples (Dickinson 1983; Dickinson and Phipps 1984). Therefore, cluster analyses based on these descriptors demonstrate much weaker group structure (Dickinson 1983; Podani and Dickinson 1984). Even so, canonical analyses showed that variation in the leaf descriptors is highly correlated with taxonomic group and, to a lesser extent, topodeme sample (fig. 5; Dickinson and Phipps 1984). The results concerning the differentiation of the topodeme samples in the present study are very important, for they suggest possible explanations for the origin of many of the taxonomic difficulties encountered at present with Crataegus in North America.

It is clear that the topodeme sample of C. punctata (T7) is distinguished from those of C. crus-galli s. l. (T1-T6) not only by the features used in identification of these species (table 1; fig. 4) but also by the extent of its variability (fig. 6). The topodeme samples of morphotypes A and B (T1-T6) appear to be homogeneous in this latter respect (figs. 4-5; Dickinson 1983). The low levels of variability in the qualitative characteristics of flowers and inflorescences not considered in figure 6 (e.g., PROJ, PUB1, PUB2, ANTH) result in apparently extreme contrasts between topodemes, almost as great as those used in the past to distinguish species of Crataegus (Palmer 1925).

Similar patterns of local differentiation are apparent in the C. pruinosa (Wendl.) K. Koch complex (Rickett 1936; Sinnott and Phipps 1983), a group that in Ontario also appears to consist of triploid and tetraploid individuals (Muniyamma and Phipps 1979b), and in which apomixis has been demonstrated cytologically (Muniyamma and Phipps 1979a). Evidence of the morphological differentiation of local stands in other plant groups comparable to that presented here is not widely available. An exception may be the pattern of variation reported for tetraploid and putatively apomictic Bidens connata Muhl. in northwestern Ontario (Crowe and Parker 1981). Otherwise, the differentiation observed here in Ontario C. crus-galli s. l. corresponds in a general way to that described in other taxonomically complex groups in which apomixis and polyploidy have been demonstrated (e.g., Crepis, Babcock and Stebbins 1938; Ranunculus auricomus, Marklund and Rousi 1961).

The apportionment of phenotypic variation into relatively large between-topodeme and small within-topodeme components in morphotypes A and B may thus be linked to the occurrence of apomixis, polyploidy and self-fertility. Cytological and experimental evidence for these traits in C. crus-galli s. l., but not C. punctata, is reported elsewhere (Dickinson 1983; Dickinson and Phipps unpubl. data).

Such features of the breeding system of C. crus-galli s. l. also may interact with characteristics of reproductive behavior in the genus as a whole as well as with other traits of the species complex. An example of this is the way in which conditions favorable for seedling establishment by Crataegus individuals tend to be very localized both in time and space (erosion surfaces, abandoned agricultural land; Valek 1980). Another would be the variation in the evenness with which seeds of Crataegus individuals are distributed within and among suitable habitats, perhaps as the effects of non-random patterns of frugivore movement are superimposed on those of year-to-year variation in fruit and seed set.
Table 5. Canonical correlation analysis of the relationship between the six flower and fruit descriptors indicated and six leaf descriptors (table 1) for the 60 OTU subsample. Parameters shown include the correlations between the descriptors and the first two canonical variates of each domain (U₁, U₂; V₁, V₂; fig. 5). Intraset (H₁) and interset (H₂) descriptor communalities are given for the first two canonical variates. See text for further details.

(a) Leaf domain:

<table>
<thead>
<tr>
<th></th>
<th>U₁</th>
<th>U₂</th>
<th>H₁</th>
<th>V₁</th>
<th>V₂</th>
<th>H₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>0.169</td>
<td>0.766</td>
<td>0.615</td>
<td>0.153</td>
<td>0.541</td>
<td>0.317</td>
</tr>
<tr>
<td>Y</td>
<td>0.399</td>
<td>0.594</td>
<td>0.512</td>
<td>0.361</td>
<td>0.419</td>
<td>0.306</td>
</tr>
<tr>
<td>Z</td>
<td>0.555</td>
<td>0.178</td>
<td>0.340</td>
<td>0.502</td>
<td>0.125</td>
<td>0.268</td>
</tr>
<tr>
<td>NUMSEC</td>
<td>-0.526</td>
<td>-0.186</td>
<td>0.311</td>
<td>-0.477</td>
<td>-0.131</td>
<td>0.244</td>
</tr>
<tr>
<td>ANGSEC</td>
<td>-0.796</td>
<td>0.124</td>
<td>0.649</td>
<td>-0.720</td>
<td>0.088</td>
<td>0.527</td>
</tr>
<tr>
<td>TEETH</td>
<td>0.627</td>
<td>-0.519</td>
<td>0.662</td>
<td>0.568</td>
<td>-0.367</td>
<td>0.457</td>
</tr>
<tr>
<td>% Variance</td>
<td>29.9</td>
<td>21.5</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Redundancy</td>
<td>24.6</td>
<td>10.7</td>
<td>40.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(b) Flower-fruit domain:

<table>
<thead>
<tr>
<th></th>
<th>V₁</th>
<th>V₂</th>
<th>H₁</th>
<th>U₁</th>
<th>U₂</th>
<th>H₂</th>
</tr>
</thead>
<tbody>
<tr>
<td>STYL</td>
<td>0.964</td>
<td>-0.042</td>
<td>0.931</td>
<td>0.873</td>
<td>-0.030</td>
<td>0.763</td>
</tr>
<tr>
<td>TCAL</td>
<td>0.518</td>
<td>-0.655</td>
<td>0.697</td>
<td>0.469</td>
<td>-0.462</td>
<td>0.434</td>
</tr>
<tr>
<td>WFL</td>
<td>0.734</td>
<td>0.528</td>
<td>0.818</td>
<td>0.665</td>
<td>0.373</td>
<td>0.581</td>
</tr>
<tr>
<td>LCAL</td>
<td>0.406</td>
<td>-0.199</td>
<td>0.204</td>
<td>0.367</td>
<td>-0.141</td>
<td>0.155</td>
</tr>
<tr>
<td>LFR</td>
<td>0.265</td>
<td>0.849</td>
<td>0.791</td>
<td>0.240</td>
<td>0.600</td>
<td>0.417</td>
</tr>
<tr>
<td>WFR</td>
<td>0.634</td>
<td>0.565</td>
<td>0.721</td>
<td>0.573</td>
<td>0.400</td>
<td>0.489</td>
</tr>
<tr>
<td>% Variance</td>
<td>39.6</td>
<td>29.8</td>
<td>100.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Redundancy</td>
<td>32.5</td>
<td>14.9</td>
<td>50.3</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Features of *crus-galli* individuals that may contribute further to the topodeme structure observed here are their large stature (see below) and abundant flowering (10-24 flowers per inflorescence; Dickinson 1985; Dickinson and Phipps unpubl. data). Comparison might be made not only with *C. punctata* (equally large; 10-14 flowers per inflorescence) but also with *Crataegus* sect. *Rotundifoliae* Eggles., which also appears to be apomictic and polyploid (Muniyamma and Phipps 1979b; Smith pers. comm.) but does not exhibit pronounced differentiation of topodesmes within component taxa (Smith pers. comm.). Individuals of sect. *Rotundifoliae* are frequently smaller than those of *C. crus-galli* s. l. or *C. punctata* and flower more sparsely (an average of 9 flowers per inflorescence; Smith pers. comm.). The relative uniformity of a *crus-galli* topodeme may be supposed to result from the arrival of large numbers of predominantly sibling, predominantly unparental individuals at a suitable site over the relatively short period of its availability.

Differences among topodesmes could thus arise from the potential uniqueness of the reproductive events by which each one is established. This uniqueness, in turn, is a function of the pollinations predominating in the production of a given crop of seed. Topodesmes of the kind described here for *C. crus-galli* s. l. easily could have contributed to a confusion of hierarchic levels as early workers confounded variation among local stands consisting only of related individuals with variation among species. In this regard both Camp (1942) and Palmer (1946) noted the large number of *Crataegus* species that even several decades after being described were still known only from their type localities.

Taxonomists with limited field experience of *Crataegus*, when working through herbarium material, may not detect the local discontinuities so apparent in a detailed study. Instead, their observations on large samples of specimens tend to support notions of continuous variation in nearly all *Crataegus* taxa. Such workers have tended to dismiss the narrow view of other workers as incompetent and un-
tainly the result exclusively of outcrossing events (Dickinson 1983; Dickinson and Phipps unpubl. data) suggests that nomenclatural recognition of individual topodemes would be completely unwarranted. Instead, we recognize only the differentiation of morphotypes A–D. A fifth taxon, *C. ?dispersus* Ashe, is also included in the key given below even though it was not possible to include authentic material of this entity in the numerical studies described above. We do this because we believe that at least a very small number of individuals referable to this taxon occur in Ontario. The approach taken in providing names for these entities is the same as that described by Phipps and Muniyamma (1980). In view of the large number of names already created we have employed the earliest available name that has been applied to Ontario material. Similarly, all five entities are accorded specific rank at this time. There are two principal differences between the treatment of the group given here and the earlier one. First, larger-fruited individuals of morphotype A are no longer segregated as *C. livoniana* Sarg., as this distinction has proved unreliable. Second, we believe that Eggleston (1908) correctly recognized that what we have called morphotype C (treated earlier as *C. sp. cf. C. bushii* Sarg.; Phipps and Muniyamma 1980; Dickinson 1983) is actually *C. prunifolia* (Poir.) Pers. We do not, however, treat this entity as Eggleston (1908) did, as a variety of *C. crus-galli* s. str.

In the key that follows, shape descriptors are those suggested by the Systematics Association Committee for Descriptive Biological Terminology (1962). Leaf venation is described by means of the terminology proposed by Hickey (1973). In addition to the morphological features used in the key, relative flowering times also distinguish the morphotypes of *C. crus-galli* s. l. at sites where more than one occur. In general, *Crataegus crus-galli* s. l. flowers later than the other morphotypes. Since the advance of spring varies both geographically (Webber and Hoffman 1967) and from year to year in southern Ontario, calendar dates are less useful for describing and predicting flowering times than are temperature sums (daily mean temperatures in excess of 5°C, accumulated from 1 April each year; Phipps and Muniyamma 1980; Smith et al. 1980; Dickinson 1983, 1985; Dickinson and Phipps unpubl. data). Representa-
tive temperature sum ranges for sites 1 through 6 (fig. 1; taxa as in fig. 2) based on London, Ontario, meteorological data are: C. prunifolia, 200–270 degree-days; C. ?grandis, 240–260 degree-days; C. punctata, 250–300 degree-days; C. fontanesiana, 280–340 degree-days; and C. crus-galli s. str., 340–400 degree-days. These ranges are reached in mid-May to early June, differences among them usually corresponding to two to six or more days at a given site.

**KEY TO ONTARIO FORMS OF CRATAEGUS sect. CRUS-GALLI Loud.**

C. CRUS-GALLI L. S. L.

Small trees to 6 m high, with a crown up to twice as broad. Thorns on distal branches simple, slightly curved, 3–7 cm long, 2–3 mm wide at base. Thorns on trunk and larger branches frequently compound. Leaves (of short shoots) unlobed, glabrous, dark green and conspicuously glossy adaxially, lighter green abaxially, (ob- )lanceolate to obovate to rhombic to obtrullate, the leaf base usually decurrent; margins simply to doubly serrate above, entire at base. Inflorescences cymose; flowers pentamerous; calyx lobes 5; stamens 5–20, the undehisced anthers ivory, yellow, pink or red; styles (0-)1-3(-4). Fruit red, a polypyrenous drupe, 8-15 mm long, 8-16 mm wide, the hypanthial collar and calyx lobes persistent. Pyrenes not excavated on their radial surfaces, usually the same number as the styles.

A. Stamens 5-10(-15); anthers 0.5-1.0 mm in diameter, containing pollen; styles (pyrenes) (0-1)-2(-3).

B. Calyx lobes entire to sparsely toothed; short shoot leaves narrowly obtrullate to obtrullate, the leaf base usually decurrent; margins simply to doubly serrate above, entire at base. Inflorescences cymose; flowers pentamerous; calyx lobes 5; stamens 5-20, the undehisced anthers ivory, yellow, pink or red; styles (0-)1-3(-4). Fruit red, a polypyrenous drupe, 8-15 mm long, 8-16 mm wide, the hypanthial collar and calyx lobes persistent. Pyrenes not excavated on their radial surfaces, usually the same number as the styles.

C. Inflorescence axes and flowers sparsely villous or glabrous; calyx lobes frequently spathulate, with secondary venation camptodromous; common throughout the range .......... 1. C. crus-galli s. str. (morphotype B)

C. Inflorescence axes and flowers moderately pubescent; short shoot leaves frequently spathulate, with secondary venation craspedodromous; rare (site 14) .......... 2. C. ?disperma Ashe

B. Calyx lobes densely toothed; short shoot leaves obovate to broadly obtrullate, the margins conspicuously doubly serrate above; modal stamen number 10–15, but may vary widely within individuals or a topodeme; rare (sites 6, 14) ............

D. Leaves, inflorescence axes, and flowers densely short pubescent, at least early in the growing season; calyx lobe margination variable; modal stamen number 20–22; undehisced anthers less than 0.5 mm in diameter, red, and containing no pollen; rare (site 4) ............

D. Leaves, inflorescence axes, and flowers sparsely villous or glabrous; calyx lobes entire to sparsely toothed; modal stamen number 17–19; undehisced anthers 0.5–1.0 mm in diameter, ivory, yellow or pink, and containing pollen; common in the southwestern portion of the range ............ 4. C. fontanesiana (Spach) Steud. (morphotype A)

D. Leaves, inflorescence axes, and flowers densely short pubescent, at least early in the growing season; calyx lobe margination variable; modal stamen number 20–22; undehisced anthers less than 0.5 mm in diameter, red, and containing no pollen; rare (site 4) ............

More study of sect. Crus-galli over its entire range is required to evaluate the merits of a formal classification below the species level, such as that employed by Eggleston (1908), Rickett (1936, 1937), Palmer (1963), and also Phipps and Muniyamma (1980) and Sinnott and Phipps (1983), in dealing with other Crataegus species complexes. Field and herbarium studies by Phipps suggest that the extent of variation in the crus-galli complex throughout its range is much greater than that described here for Ontario. A similarly detailed analysis of material from the U.S.A. would undoubtedly lead to continued nomenclatural recognition of much of this variation, even though up to four-fifths of the binomials attributed to the complex are expected to disappear in synonymy. The indications given here and elsewhere that many of the names published for the group (Palmer 1925) may refer merely to individual genotypes rather than populations suggest that it could prove useful to consider adopting the informal parenthetic subspecific classification proposed by Burtt (1970). Such an approach would make it possible to refer to variants, however they may have arisen, without burdening the literature with new circumscriptions and changes in rank that would be of limited biological significance.

Likewise, more study is required to determine what components within C. crus-galli s. l. are products of hybridization with other sec-
tions. In this connection we note the male sterility of C. ?grandis (morphotype D) and its resemblance to C. punctata with respect to the suite of descriptors employed here (figs. 3–4). Both C. ?disperma and C. ?grandis may represent derivatives of hybrids between sects. Crus-galli and Punctatae (Palmer 1963; Phipps and Muniyama 1980). Similarly, as pointed out by Eggleston (1908, 1923), C. prunifolia (morphotype C) resembles Crataegus sect. Macracanthae (figs. 3–4) and a number of putative Crus-galli × Macracanthae hybrids. Crataegus prunifolia differs from these taxa primarily in the absence of the cavities on the radial surfaces of the pyrenes and are characteristic of sect. Macracanthae.

Acknowledgments. This paper is based on a portion of a thesis by the first author submitted in partial fulfillment of the requirements for the Ph.D. degree at the University of Western Ontario. The study was supported in part by a Province of Ontario Graduate Scholarship to the first author, and by NSERC Operating Grant No. A-1726 to the second author. We are grateful to the owners of the study sites for allowing us to work on their properties. M. Macleod and M. Shaik provided valuable assistance in the field and in other aspects of the work. We are indebted to W. J. Crins of the Biology Department, Erindale College, Mississauga, Ontario, for bringing to our attention site 14 (fig. 1), at which C. prunifolia and possibly also C. ?disperma are found, and to P. G. Smith for sharing with us his data on Crataegus sect. Rotundifoliae.

Literature Cited


