

Supporting Information Methods and Results S1. Data treatment, ordination and measures of disparity of Stebbins (1951) dataset.

Supporting Information Figure S1: Frequency of the 16 character states recorded by Stebbins (1951) for ten angiosperm subgroups. Colors highlight character states whose frequency exceeds 85 %.

Methods S1

Data treatment

We used the character states of eight binary characters that were recorded for 288 angiosperm families by Stebbins (1951). A character state combination is thus a binary vector of length eight. Polymorphic families ($n=90$) were recorded several times for each character state combination. We assigned each family to one of ten major angiosperm taxa (some of them non-monophyletic) according to the current classification of APG III (2009) and Soltis et al. (2011): basal angiosperms, monocots, basal eudicots, basal superrosids, fabids, malvids, Caryophyllales, basal superasterids, lamiids, and campanulids. We are aware that the taxonomic composition and therefore also the family descriptions of some angiosperm families such as Hamamelidaceae, Scrophulariaceae, or Theaceae have changed considerably since Stebbins' paper was published. Some character state combinations of Stebbins dataset might thus be differently represented in the currently delimited families. We chose angiosperm subgroups to be large enough so that most of these changes do not affect the total morphospace occupied by each category. The full dataset with all recorded character state combinations (taken from Stebbins 1951) and the subgroup assignments are given in Supplementary Table 1. The number of families per subgroup is given in Figure 4d.

Ordination

In order to visualize the area of the space occupied by each major taxonomic group, we performed an NMDS ordination with the function `metaMDS` from the *vegan* package (Oksanen *et al.*, 2013) in R vs. 3.0.0. (R Core Team, 2013). This method uses an iterative process to maximize the correspondence between pairwise distances calculated from the data, and the Euclidean distances between dots in a two or three-dimensional representation (Rabinovitz, 1975). The quality of the fit is estimated by the stress value, ranging from 0 (best solution) to 1 (worst solution). We used the distance index $D = 1 - S$ where S is the simple matching index for symmetric binary characters (Gower & Legendre, 1986), so that for two given character combinations, D is the proportion of characters that are under different states in these combinations. Finally, we tested for morphological differences among taxonomic groups using a non-parametric multivariate analysis of variance (npMANOVA; Anderson, 2001) with the function `adonis` from the *vegan* package in R, using the same dissimilarity index (D) and 10,000 repetitions. Post hoc tests were performed in the same way with a npMANOVA and a Bonferroni correction.

Measure of disparity

We computed three disparity indices for each of the ten angiosperm subgroups, following the method of Ciampaglio *et al.* (2001): the sum of variances, as the sum of the three variances associated to the scores of each NMDS ordination axis; the range, as the maximum simple matching distance between two families; and the averaged pairwise distances, calculated with the simple-matching distance index. As subgroup size varies considerably (from 15 families for basal superrosids to 84 for fabids, figure 4d), we added an analysis of rarefaction for the range (variance and average pairwise distance are relatively insensitive to sample), for which we strictly followed the method of Foote (1992). For each angiosperm subgroup, and for each possible sampling size (from 3 to the group size), we calculated the mean and 90% confidence intervals of the disparity indices obtained from 200 random samplings without replacement. The resulting curves allowed us to compare the angiosperm subgroups despite their size differences.

Results S1

Posthoc test results for the npMANOVA:

```
> posthoc: upper part: r2, lower part: p
      basal angiosperms basal eudicots basal superasterids basal superrosids campanulids
basal angiosperms      NA          0.0250          0.1170          0.0160          0.3910
basal eudicots         0.3895          NA          0.1610          0.0820          0.4290
basal superasterids    0.0002          0.0001          NA          0.0590          0.1080
basal superrosids      0.5956          0.0451          0.0092          NA          0.3420
campanulids           0.0001          0.0001          0.0004          0.0001          NA
Caryophyllales        0.0025          0.0001          0.0009          0.0038          0.0001
fabids                0.0336          0.0055          0.0002          0.5888          0.0001
lamiids               0.0001          0.0001          0.0001          0.0001          0.0001
malvids               0.0015          0.0001          0.0001          0.2683          0.0001
monocots              0.0001          0.0001          0.0514          0.1058          0.0018

      Caryophyllales fabids lamiids malvids monocots
basal angiosperms    0.1130 0.0320 0.3010 0.0630 0.100
basal eudicots       0.2060 0.0480 0.2800 0.1080 0.121
basal superasterids  0.0720 0.0720 0.1090 0.0790 0.020
basal superrosids    0.1190 0.0060 0.2350 0.0170 0.024
campanulids          0.3160 0.2140 0.1460 0.2500 0.068
Caryophyllales      NA 0.0790 0.2380 0.1320 0.091
fabids              0.0001 NA 0.1920 0.0230 0.070
lamiids             0.0001 0.0001 NA 0.2180 0.102
malvids             0.0001 0.0237 0.0001 NA 0.053
monocots            0.0001 0.0001 0.0001 0.0002 NA

> p #with Bonferroni correction
[1] 0.001111111

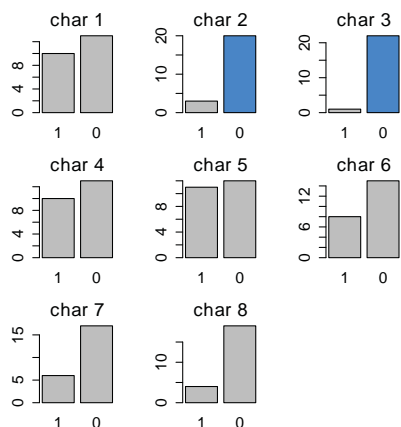
> posthoc2: significance of p after the Bonferroni correction
      basal angiosperms basal eudicots basal superasterids basal superrosids campanulids
basal angiosperms      <NA>          <NA>          <NA>          <NA>          <NA>
basal eudicots         ns          <NA>          <NA>          <NA>          <NA>
basal superasterids    *          *          <NA>          <NA>          <NA>
basal superrosids      ns          ns          ns          <NA>          <NA>
campanulids           *          *          *          *          <NA>
Caryophyllales        ns          *          *          ns          *
fabids                ns          ns          *          ns          *
lamiids               *          *          *          *          *
malvids               ns          *          *          ns          *
monocots              *          *          ns          ns          ns

      Caryophyllales fabids lamiids malvids monocots
basal angiosperms    <NA> <NA> <NA> <NA> NA
basal eudicots       <NA> <NA> <NA> <NA> NA
basal superasterids  <NA> <NA> <NA> <NA> NA
basal superrosids    <NA> <NA> <NA> <NA> NA
campanulids          <NA> <NA> <NA> <NA> NA
Caryophyllales      <NA> <NA> <NA> <NA> NA
fabids              * <NA> <NA> <NA> NA
lamiids             * * <NA> <NA> NA
malvids             * ns * <NA> NA
monocots            * * * * NA
```

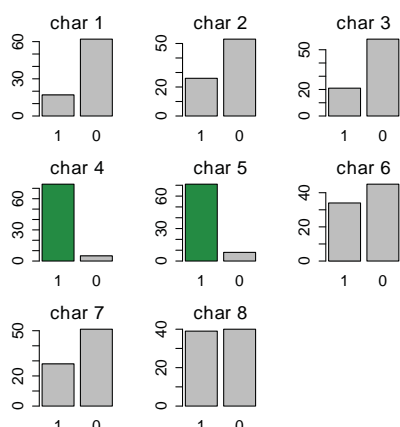
References:

- Anderson MJ. 2001.** A new method for non-parametric multivariate analysis of variance. *Austral Ecology* **26**: 32–46.
- Ciampaglio CN, Kemp M, McShea DW. 2001.** Detecting changes in morphospace occupation patterns in the fossil record: characterization and analysis of measures of disparity. *Paleobiology* **27**: 695–715.
- Foote M. 1992.** Rarefaction analysis of morphological and taxonomic diversity. *Paleobiology* **18**: 1-16.
- Gower JC, Legendre P. 1986.** Metric and euclidean properties of dissimilarity coefficients. *Journal of Classification* **3**:5-48.
- Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2013.** vegan: Community Ecology Package. R package version 2.0-7. [WWW document] URL <http://CRAN.R-project.org/package=vegan>. [accessed 1 January 2014].
- Rabinovitz GB. 1986.** An introduction to nonmetric multidimensional scaling. *American Journal of Political Science* **19**: 343-390.
- R Core Team 2013.** R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. [WWW document] URL <http://www.R-project.org/>. [accessed 1 January 2014].
- Soltis DE, Smith SA, Cellinese N, Wurdack KJ, Tank DC, Brockington SF, Refulio-Rodriguez NF, Walker JB, Moore MJ, Carlswald BS et al. 2011.** Angiosperm phylogeny: 17 genes, 640 taxa. *American Journal of Botany* **98**: 704-730.
- Stebbins GL. 1951.** Natural selection and the differentiation of angiosperm families. *Evolution* **5**: 299-324.
- Stevens PF. 2001 onwards.** *Angiosperm Phylogeny Website. Version 12, July 2012.* [WWW document] URL <http://www.mobot.org/MOBOT/research/APweb/welcome.html>. [accessed 1 January 2014].

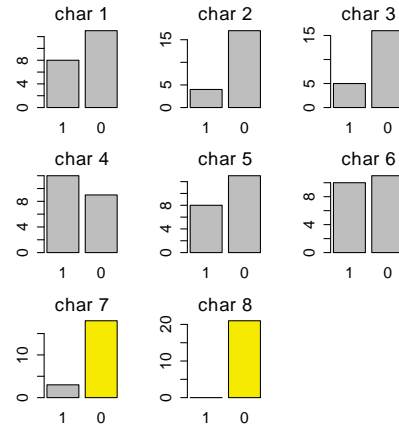
basal angiosperms



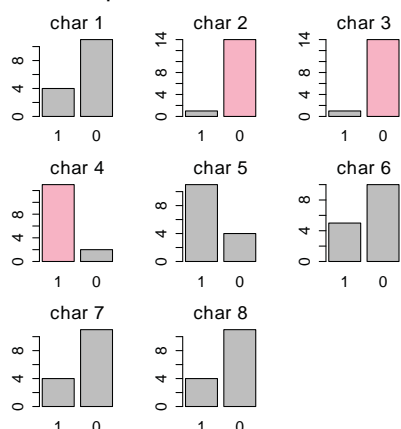
monocots



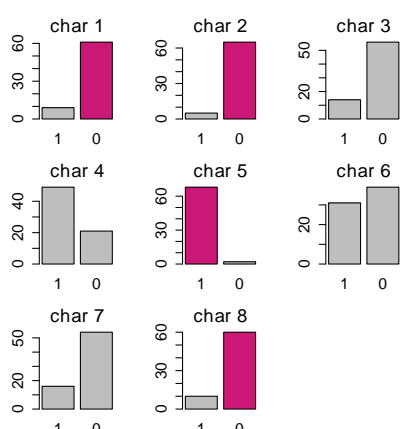
basal eudicots



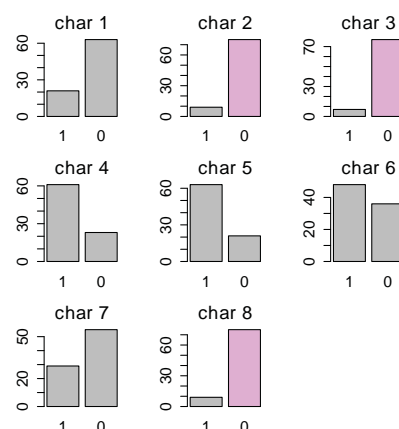
basal superrosids



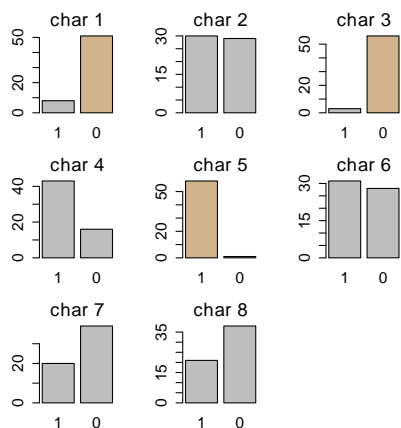
malvids



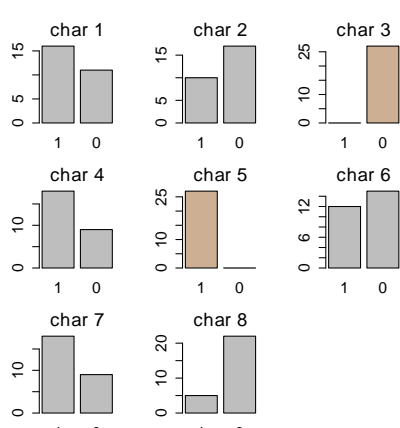
fabids



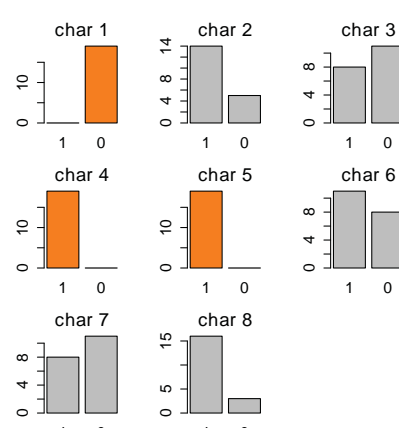
basal superasterids



Caryophyllales



campanulids



lamiids

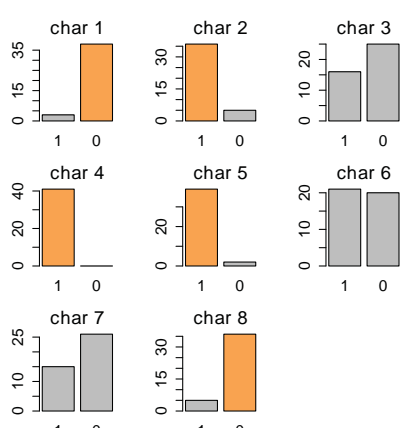


Table 1.

| Character | State "0" | State "1" |
|-----------|--|--|
| 1 | Calyx and corolla both present (dichlamydeous) | Corolla or entire perianth absent (mono- or achlamydeous) |
| 2 | Petals separate (polypetalous) | Petals (sepals in monochlamydeous forms) united (sympetalous) |
| 3 | Calyx and corolla regular (actinomorphic) | Perianth irregular (zygomorphic) |
| 4 | Stamens more numerous than perianth members (polystemonous) | Stamens the same number as or fewer than perianth members (oligostemonous) |
| 5 | Carpels separate (apocarpous) | Carpels united (syncarpous) |
| 6 | Seeds more than one per carpel (many) | Seeds not more than one per carpel (few) |
| 7 | Placentation axial | Placentation parietal, basal or free central |
| 8 | Carpels not united with the receptacle or perianth (hypogyny or perigyny, i.e. superior ovary) | Carpels united with receptacle, perianth, or both (epigyny, i.e. inferior ovary) |