

Unifying Life History Analyses for Inference of Fitness and Population Growth

Ruth G. Shaw

Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul,
Minnesota 55108

rshaw@superb.ecology.umn.edu

Charles J. Geyer

School of Statistics, University of Minnesota, Minneapolis, Minnesota 55455

charlie@stat.umn.edu

Stuart Wagenius

Institute for Plant Conservation Biology, Chicago Botanic Garden, Glencoe, Illinois 60022

swagenius@chicagobotanic.org

Helen H. Hangelbroek

Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul,
Minnesota 55108

helen_hangelbroek@hotmail.com

and

Julie R. Etterson

Biology Department, University of Minnesota-Duluth, Duluth, Minnesota 55812

jetterso@d.umn.edu

Received _____; accepted _____

Prepared with AAST_EX— Type of submission: Article

ABSTRACT

Lifetime fitness of individuals is the basis for population dynamics, and variation in fitness results in evolutionary change. Though the dual importance of individual fitness is well understood, empirical datasets of fitness records generally violate the assumptions of standard statistical approaches. This problem has plagued comprehensive study of fitness and impeded empirical unification of numerical and genetic dynamics of populations. Recently developed aster models address this problem by explicitly modeling the dependence of later expressed components of fitness (e.g. fecundity) on those expressed earlier (e.g. survival to reproductive age). Moreover, aster models allow different sampling distributions for different components of fitness, as appropriate (e.g. binomial for survival over a given interval and Poisson for fecundity). The analysis is conducted by maximum likelihood, and the resulting compound distributions for lifetime fitness closely approximate the observed data. To illustrate the breadth of aster's utility, we provide three examples: a comparison of mean fitness among genotypic groups, a phenotypic selection analysis, and estimation of finite rates of increase. Aster models offer a unified approach to addressing the wide range of questions in evolution and ecology for which life history data are gathered.

Subject headings: *Chamaecrista fasciculata*, community genetics, demography, *Echinacea angustifolia*, fitness components, *Uroleucon rudbeckiae*

26 The fitness of an individual is well understood as its contribution, in offspring,
to the next generation. Fitness has both evolutionary significance, as an individual's
28 contribution to a population's subsequent genetic composition, and ecological importance,
as an individual's numerical contribution to a population's growth. The simplicity of these
30 closely linked ideas belies serious complications that arise in empirical studies. Because
lifetime fitness comprises multiple components of fitness expressed over one to many
32 seasons or stages, its distribution is typically multimodal and highly skewed in shape
and thus corresponds to no known parametric distribution. This problem has long been
34 acknowledged (Mitchell-Olds and Shaw 1987; Stanton and Thiede 2005), yet to date there is
no single rigorously justified approach for jointly analyzing components of fitness measured
36 sequentially throughout the lifetime of an individual. This limitation severely undermines
efforts to link ecological and evolutionary inference.

38 Here we present applications of a new statistical approach, *aster*, for analyzing
life-history data with the goal of making inferences about lifetime fitness or population
40 growth. Within a single analysis, *aster* permits different fitness components to be modeled
with different statistical distributions, as appropriate. It also accounts for the dependence of
42 fitness components expressed later in the life-span on those expressed earlier, as is necessarily
the case with sequential measurements of aspects of the life-history, e.g. reproduction
44 depending on survival to reproductive age. Geyer et al. (2007) present the statistical theory
for *aster* models. Our goal here is twofold. First, we illustrate the problem by reviewing the
46 limitations of approaches that have previously been employed in empirical studies. Second,
we describe how *aster* models resolve these problems, illustrating these points with three
48 empirical examples.

1. The problem and previous efforts to address it

50 Individual fitness realized over a lifespan typically does not conform to any well known
distribution that is amenable to parametric statistical analysis. In contrast, components
52 of fitness over a given interval, such as survival to age x , reproduction at that age, and
the number of young produced by a reproductive individual of that age, generally conform
54 much more closely to simple parametric distributions. For this reason, components of
fitness are sometimes analyzed separately to circumvent the distributional problem of
56 lifetime fitness. For example, in a study of genetic variation within a population of *Salvia*
lyrata in its response to conspecific density, Shaw (1986) provided separate analyses of
58 two components of fitness, survival over two time intervals and size of the survivors, as a
proxy for future reproductive capacity in this perennial plant. This approach considers size,
60 or in other cases fecundity, conditional on survival. It has the appeal that the statistical
assumptions underlying the analyses tend to be satisfied, but it offers no way to combine the
62 analyses to yield inferences about overall fitness. Separate analyses of fitness components
cannot substitute for an analysis of overall fitness, particularly considering the possibility
64 of tradeoffs between components.

A common approach to analyzing fitness as survival and reproduction jointly is to use
66 fecundity as the index of fitness and retain fecundity values of zero for individuals that
died prior to reproduction. When observations are available for replicate individuals, a
68 variant of this approach is to use as the measure of fitness the product of the proportion
surviving and the mean fecundity of survivors (e.g. Belaoussoff and Shore 1995; Galloway
70 and Etterson 2007). In either case, the resultant distributions typically have at least two
modes (one at zero) and are highly skewed, such that no data transformation yields a
72 distribution that is suitable for parametric statistical analysis. Authors frequently remark
on the awkwardness of these distributions in their studies (e.g. Etterson 2004), but rarely

74 publish fitness distributions. Antonovics and Ellstrand (1984), however, presented the
distribution of lifetime reproductive output (their Fig. 2) in their experimental studies of
76 frequency-dependent selection in the perennial grass, *Anthoxanthum odoratum*, noting its
extreme skewness. Finding no transformation that yielded a normal distribution suitable
78 for analysis of variance, they assessed the robustness of their inferences by applying three
distinct analyses (categorical analysis of discrete fecundity classes, ANOVA of means, and
80 nonparametric analysis). In this study, results of the three analyses were largely consistent,
but, in general, results are likely to differ.

82 Others have noted the importance of complete accounting of life-history in inferring
fitness or population growth rate, as well as evaluation of its sampling variation, and
84 have presented methods to accomplish this. Caswell (1989) and Morris and Doak (2002)
explain how to obtain population projection matrices from life-history records and, from
86 them, to estimate population growth rate. They also describe methods for evaluating its
sampling variation and acknowledge statistically problematic aspects of these methods
88 (their Chapters 8 and 7, respectively). Lenski and Service (1982) considered the complete
life-history record of individuals as the unit of observation and used jackknife resampling to
90 estimate population growth rate and its sampling variance. Recent efforts to evaluate the
nature of selection have likewise taken a comprehensive demographic approach. McGraw
92 and Caswell (1996) considered individual life-histories but chose the maximum eigenvalue
of an individual's Leslie matrix (λ) as its fitness measure. They regressed λ on the fitness
94 components, age at reproduction and lifetime reproductive output to estimate selection on
them. Van Tienderen (2000) advocated an alternative approach for studying phenotypic
96 selection. This approach involves evaluating the relationships between each component
of fitness and the phenotypic traits of interest via separate multiple regression analyses
98 (Lande and Arnold 1983) to obtain the selection gradients in different episodes of selection.
These selection gradients are then weighted by the elasticities (Caswell 1989) of each

100 component of fitness obtained from analysis of the appropriate population projection
matrix. Beyond these approaches linking demography and fitness, methods targeting the
102 problem of “zero-inflated” data (i.e. many observations of zero distorting a distribution)
have been proposed (Cheng et al. 2000; Dagne 2004). Each method has liabilities, however.
104 For example, elasticities do not take into account sampling variation in the life-history,
and violations of distributional assumptions remain a problem (McGraw and Caswell 1996;
106 Coulson et al. 2003). Moreover, none of these methods generalize readily for inference in
the wide range of contexts that life-history data can, in principle, address.

108 **2. Inference of individual fitness with aster**

We present aster models (Geyer et al. 2007) for rigorous statistical analysis of
110 life-history records as a general approach for addressing diverse questions in evolution
and ecology. As noted above, two standard properties of life-history data are central to
112 the statistical challenges that aster addresses. First, the expression of an individual’s
life-history at one stage depends on its life-history status at earlier stages. For example,
114 observation of an individual’s fecundity at one stage is contingent on its survival to that
stage. Second, no single parametric distribution is generally suitable for modeling various
116 components of fitness, e.g. survival and fecundity. Aster analysis models fitness components
observed through a sequence of intervals bounded by the times at which individuals are
118 observed. The intervals could be days or years, and need not all be the same duration.
The components of fitness are modeled jointly over successive intervals by explicitly taking
120 into account the inherent dependence of each stage on previous stages, e.g. that only
survivors reproduce. We represent the life-history and, in particular, the dependence of one
122 component of the life-history on another, graphically as in Fig. 1.

EDITOR: PLACE FIGURE 1 HERE.

124 The aster approach models the joint distribution of a set of variables (fitness
components). We say an arrow in the graphical model points from a variable to its *successor*
126 or, going backwards along the arrow, from a variable to its *predecessor* (Geyer et al. 2007
used “parent” and “child” instead of “predecessor” and “successor” but this is confusing
128 in biology). The theory underlying the aster approach requires modeling the conditional
distribution of each variable given its predecessor variable as an exponential family of
130 distributions (Barndorff-Nielsen 1978; Geyer et al. 2007) with the predecessor variable
playing the role of sample size. This requirement retains considerable flexibility, because
132 many distributions are exponential families, including Bernoulli, Poisson, geometric, normal,
and negative binomial.

134 If the predecessor is zero then so is the successor. If the predecessor has the value
 $n > 0$, then the successor is the sum of n independent and identically distributed variables
136 having the named distribution. For example, the binary outcome of an individual’s survival
over a given interval is modeled as a Bernoulli variable. Likewise, given that an individual
138 survived to this point, whether or not it reproduced is considered Bernoulli. Fecundity,
given that it reproduced in this interval, may be modeled according to a zero-truncated
140 Poisson distribution (i.e. a Poisson random variable conditioned on being greater than 0).

Aster analysis yields estimates of unconditional (lifetime) fitness that account for
142 the expression of all the fitness components. The modeling of each single component of
fitness with an appropriate probability distribution leads to a sampling distribution for the
144 joint expression of the fitness components as lifetime fitness that approximates the actual
distribution, as we show in Example 2 below.

Modeling the joint distribution of fitness components establishes a proper foundation
for sound analysis, but another key idea is needed. Restriction of the choice of
conditional distributions for fitness components to exponential families results in a joint

distribution of the components that is a multivariate exponential family. Its canonical parameterization, described by equation (5) in Geyer et al. (2007), is called the *unconditional* parameterization of the aster model. Let X_i denote the variables (fitness components) and φ_i the corresponding canonical parameters of a model. When overall fitness is considered a linear combination $\sum_i a_i X_i$, where a_i are known constants, then its unconditional expectation is directly controlled by a single parameter β_k if the regression part of the model has the form

$$\varphi_i = a_i \beta_k + \text{other terms not containing } \beta_k$$

146 by equation (22) in Geyer et al. (2007); increasing β_k increases the unconditional expectation of $\sum_i a_i X_i$ other betas being held fixed. Thus confidence intervals and hypothesis tests for
148 β_k address overall fitness directly. For this reason, the unconditional parameterization is used for the example in Geyer et al. (2007) and Examples 1 and 2 below. This situation
150 most often arises when the linear combination is a simple sum (so the a_i are zero or one), e.g. some of the X_i are counts of offspring in one year and $\sum_i a_i X_i$ is the total number of
152 offspring observed in all years.

The unconditional parameterization is somewhat counterintuitive because terms in the
154 regression model that nominally refer to a single component of fitness (affect its φ_i only) directly influence the unconditional expectation of overall fitness by affecting not only its
156 distribution but also the distributions of all components before it in the graphical model (its predecessor, predecessor of predecessor, etc.) This makes it somewhat difficult, but not
158 impossible (see our Example 1), to see the role played by a single component of fitness. This is, however, an unavoidable consequence of being able to address overall fitness.

160 The analysis employs the principle of maximum-likelihood, developed by Fisher (1922) and now widely applied as a rigorous, general approach to any statistical problem (Kendall
162 and Stuart 1977). Software for conducting the analysis, is a contributed package “aster”

in the R statistical language (R Development Core Team 2006) and is freely available
164 (<http://www.r-project.org>).

We demonstrate the value and versatility of the aster approach with three examples.
166 In the first, we apply aster to compare mean fitness among groups. Specifically, we
quantify effects of inbreeding on fitness of *Echinacea angustifolia*, a long-lived composite
168 plant. In our second example, we reanalyze data of Etterson (2004) to evaluate phenotypic
selection on the annual legume, *Chamaecrista fasciculata*. In the last example, we illustrate
170 inference of population growth rate via aster. We consider a small dataset that Lenski and
Service (1982) used to demonstrate their nonparametric method for inferring population
172 growth rate from a set of individual life-histories of the aphid, *Uroleucon rudbeckiae*.
The datasets for our examples are in the aster package for R. Complete analyses for our
174 examples are given in a technical report (Shaw, et al. 2007) available at the aster website
<http://www.stat.umn.edu/geyer/aster> and are reproducible by anyone who has R (see
176 Chapter 1 of the technical report).

3. Example 1: Comparison of fitness among groups

178 In this example, we illustrate use of aster models to compare mean fitnesses of defined
groups, here, genotypic classes. Specifically, we investigate how parental relatedness
180 affects progeny fitness in the perennial plant, *Echinacea angustifolia* (narrow-leaved purple
coneflower), a common species in the N. American tallgrass prairie and Great Plains. The
182 plant is self-incompatible, and Wagenius (2000) detected no deviation from random mating
in large populations. However, following the abrupt conversion of land to agriculture and
184 urbanization that started about a century ago, the once extensive populations now persist
in small patches of remnant prairie. In this context of fragmented habitat, we expect
186 that matings between close relatives in the same remnant, and perhaps also long distance

matings, have become more common.

188 To evaluate the effects of different mating regimes on the fitness of resulting progeny,
formal crosses were conducted in the field to produce progeny of matings between plants
190 a) from different remnants, b) chosen at random from the same remnant, and c) known
to share the maternal parent. The resulting seeds were germinated, and the plants were
192 grown in a growth chamber for three months, after which they were transplanted into an
experimental field plot. In this example, we focus on pre-reproductive components of fitness,
194 survival and plant size. Survival of each seedling was assessed in the growth chamber on
three dates. The seedlings were then transplanted into an experimental field plot, and their
196 survival was monitored annually 2001–2005. The number of rosettes (basal leaf clusters,
1–7) per plant was also counted annually 2003–2005. Rosette count reflects plant vigor and
198 is likely related to eventual reproductive output.

Mortality of many plants as seedlings and juveniles resulted in a distribution of rosette
200 count in 2005 having many structural zeros. We modeled survival through each of eight
observation intervals as Bernoulli, conditional on surviving through the preceding stage;
202 we modeled rosette count in each of three field seasons, given survival to that season, as
zero-truncated Poisson (Fig. 1A). To account for spatial and temporal heterogeneity, we
204 also included in the models the factors a) year of crossing (1999 or 2000), b) planting tray
during the period in the growth chamber, c) spatial location (row and position within row)
206 in the field.

In addition to evaluating the effects of mating treatments on overall fitness, we
208 developed models to test for differences in the timing and duration of the mating treatment
effects on fitness. At the earliest stages, in the benign conditions of the growth chamber,
210 effects of the mating treatments may be negligible. Alternatively, it may be that the
effects of mating treatment at the earliest stages largely account for their overall effects on

212 fitness. These scenarios differ in their implications concerning the inbreeding load expected
in standing populations (Husband and Schemske 1996). We developed four aster models,
214 named “chamber,” “field,” “sub,” and “super.” Each was a joint aster analysis of all 11
bouts of selection (survival over eight intervals, rosette count at three times). The “field”
216 model includes explicit mating treatment effects only on the final rosette count (variable
r05 in Figure 1A), but because of the unconditional parameterization of aster models
218 (section 2, above) these effects propagate back to earlier stages. The “chamber” model
includes explicit mating treatment effects only on the final survival before transplanting
220 (variable lds3 in Figure 1A), but, again, these effects propagate through all preceding
bouts of survival. The “sub” model is the greatest common submodel of “chamber” and
222 “field,” and the “super” model is their least common supermodel (i. e. “sub” includes no
effects of mating treatment on any aspect of fitness, whereas “super” includes effects of
224 mating treatment on both survival up to transplanting and on final rosette count).

The aster analysis revealed clear differences among the mating treatments in overall
226 progeny fitness through the end of the available set of records, (model “field” compared to
“sub”, ($P = 1.1 \times 10^{-5}$). The unconditional expected rosette count for each cross type is
228 the best estimate for the expected rosette count in 2005 for every seed that germinated in
2001. The fitness disadvantage of progeny resulting from sib-mating relative to the other
230 treatments is a 35%–42% reduction in rosette count (Fig. 2).

EDITOR: PLACE FIGURE 2 HERE.

232 Because of the aforementioned propagation of effects back to earlier stages, the effects of
mating treatment in the “field” model directly subsume overall fitness expressed over the
234 course of the experiment. Though this analysis suffices for inferring the overall effects
of mating treatment on fitness, we investigated further the timing and duration of these
236 effects using the additional models described above. The comparison of the “sub” and

“chamber” models shows that survival before transplanting differs among mating treatments
238 ($P = 0.012$). However, the comparison of the “chamber” and “field” models with the
“super” model shows that “super” fits no better than “field” ($P = 0.34$) but does fit better
240 than “chamber” ($P = 3.1 \times 10^{-4}$). Hence the “field” model fully accounts for differences
in expressed fitness. The terms of the “chamber” model that quantify the effect of mating
242 treatment on survival up to transplanting are not needed to fit the data, because the
aforementioned back propagation of effects subsumes the effects of mating treatment in the
244 growth chamber. This does not mean there are no effects of mating treatment on fitness
before transplanting. The comparison of “sub” and “chamber” confirms they exist, and
246 Fig. 2 clearly shows them. The fitness disadvantage of progeny resulting from sib-mating
relative to the other treatments is clear in the 7%–10% reduced survival up to the time of
248 transplanting but the overall fitness disadvantage of inbreds is considerably greater (Fig. 2).

4. Example 2: Phenotypic selection analysis

250 Lande and Arnold (1983) proposed multiple regression of fitness on a set of quantitative
traits as a method for quantifying natural selection directly on each trait. In practice,
252 these analyses have generally employed measures of components of fitness as the response
variable, rather than overall fitness (see e.g. examples in Lande and Arnold 1983). As a
254 result, the estimated selection gradients, the partial regression coefficients, reflect selection
on a trait through a single ‘episode of selection’, rather than selection over multiple episodes
256 or, ideally, over a cohort’s lifespan. Focusing on this limitation, Arnold and Wade (1984a)
considered partitioning the overall selection gradient into parts attributable to distinct
258 episodes of selection, and Arnold and Wade (1984b) illustrated the approach with examples.
Wade and Kalisz (1989) modified this approach to allow for change in phenotypic variance
260 among selection episodes. Whereas these developments were intended to account for

the multiple stages of selection, they do not directly account for the dependence of later
262 components of fitness on ones expressed earlier, an issue that also applies to van Tienderen
(2000).

264 Apart from the problem of dependence among fitness components, Mitchell-Olds and
Shaw (1987), among others, have noted that statistical testing of the selection gradients
266 is compromised, in many cases, by the failure of the analysis to satisfy the assumption of
normality of the fitness measure, given the predictors. This concern applies to McGraw and
268 Caswell’s (1996) approach to phenotypic selection analysis, which integrates observations
from the full life-history. To address this problem for the case of dichotomous fitness
270 outcomes, such as survival, Janzen and Stern (1998) recommended the use of logistic
regression for testing selection on traits and showed how the estimates resulting from
272 logistic regression could be transformed to obtain selection gradients. Schluter (1988) and
Schluter and Nychka (1994) suggested estimating fitness functions as a cubic spline to allow
274 for general form, but this method requires a parametric error distribution, whether normal,
binomial, or Poisson.

276 Aster explicitly models the dependence of components of fitness on those expressed
earlier. Moreover, unconditional aster analysis estimates the relationship between overall
278 fitness and the traits directly in a single, unified analysis. It thus serves as a basis for
statistically valid inference about phenotypic selection, unlike other methods whose required
280 assumptions are often seriously violated. We illustrate this use of aster with a reanalysis
of Etterson’s (2004) study of phenotypic selection on three traits in three populations of
282 the annual legume, *Chamaecrista fasciculata*, reciprocally transplanted into three sites.
The three traits, measured when the plants were 8–9 weeks old, are leaf number (LN, log
284 transformed), leaf thickness (measured as specific leaf area, SLA, the ratio of a leaf’s area
to its dry weight, log transformed) and reproductive stage (RS, scored in 6 categories,

286 increasing values denote greater reproductive advancement). Here, for simplicity, we
consider only the data for the three populations grown in the Minnesota site.

288 *C. fasciculata* grows with a strictly annual life-history. In this experiment, fitness
was assessed as 1) survival to flowering, 2) flowering, given that the plant survived, 3) the
290 number of fruits a plant produced, and 4) the number of seeds per fruit in a sample of three
fruits, the last two contingent on the plant having flowered. Preliminary analyses revealed
292 that nearly all survivors flowered and fruited, so these fitness components were collapsed
to a single one, modeled as Bernoulli (**reprod**). Consequently, overall fitness was modeled
294 based on survival and reproduction, the number of fruits per plant, and the number of
seeds per fruit, (termed **reprod**, **fruit**, and **seed**, Fig. 1B). Preliminary analyses assessed
296 the fit of truncated Poisson and truncated negative binomial distributions to the data for
both fecundity components. On this basis, the fecundity components were modeled with
298 truncated negative binomial distributions. In addition to the traits of interest, the model
included the spatial block in which individuals were planted.

300 To illustrate phenotypic selection analysis most straightforwardly, we begin by
analyzing two of the fitness components, **reprod** and **fruit** in relation to two of the traits,
302 leaf number (LN) and leaf thickness (SLA). This model detected strong dependence of
fitness on both traits such that selection is toward more, ($P < 10^{-6}$) thinner ($P = 0.007$)
304 leaves.

Extending this analysis to assess curvature in the bivariate fitness function, we detected
306 highly significant negative curvature for both traits, suggestive of stabilizing selection,
($P < 5.7 \times 10^{-43}$). The plot of the fitness function together with the observed phenotypes
308 (Fig 3) reveals that the fitness optimum lies very near the edge of the distribution of leaf
number.

310 EDITOR: PLACE FIGURE 3 HERE.

Thus, for this trait, selection against both extremes of the standing variation in the trait
312 (i.e. stabilizing selection) is not observed. The aster analysis fits the data well, as reflected
by the scatter plots of Pearson residuals which show very little trend and only three
314 extreme outliers for fruit number (Fig. 4A). This evidence that we have modeled the
data appropriately reinforces our confidence in use of aster models to estimate phenotypic
316 selection gradients.

Comparison of the bivariate fitness function inferred via aster with that obtained via
318 ordinary least squares (OLS), which has become standard (Lande and Arnold 1983), reveals
that they differ in the sign of the curvature for leaf number, LN. In contrast to the negative
320 curvature for LN, which the aster analysis strongly indicates, OLS estimates positive
curvature in this direction, suggestive of disruptive selection, such that the bivariate surface
322 is a saddle; each analysis attaches highly significant P values to this term of the quadratic
($P < 5 \times 10^{-16}$). However, the homoscedasticity and normality assumptions required for
324 OLS regression to give meaningful P -values are seriously violated (Fig 4). Such violations
of assumptions for an OLS regression analysis are expected, given that 8% of plants have
326 fitness of zero and that the distributions of numbers of fruits per plant is heavily skewed.
Violation of these assumptions distorts the surface inferred to represent selection. Because
328 the assumptions for the aster model appear satisfied, we trust the aster model P -values and
estimated fitness surface and not those from OLS regression.

330 EDITOR: PLACE FIGURE 4 HERE.

We extend the above phenotypic selection analysis to include the additional fitness
332 component, **seed**, and also the additional phenotypic predictor, reproductive stage (RS).
This analysis detected dependence of both **fruit** and **seed** on LN ($P < 10^{-15}$) and
334 dependence of **seed** on RS ($P < 10^{-6}$). In addition, the dependence of **fruit** on SLA
was marginally significant ($P = 0.054$). In each case, the relationship between fitness

336 component and trait also takes into account pre-reproductive mortality. Thus, this analysis
quantifies the relationship between the two components of fecundity and the traits; it does
338 not, however, satisfy the goal of evaluating the relationship between overall fitness and the
traits. In the following, we explain the challenge of analyzing data in the structure given,
340 as well as its resolution.

Though impractical in this case, as in many others, one could imagine having counts
342 of total numbers of seeds produced by each plant as well the number of fruits. Data in this
structure would be amenable to phenotypic selection analysis by the standard aster method
344 illustrated above, extending the graph in Fig. 1B to include total seed count as a successor
node of fruit count. Here, where seed counts are available for a set number (3) of fruits,
346 **seed** and **fruit** depend jointly on **reprod** (Fig. 1B). In this case, it might seem natural
to use the product of fruit count and number of seeds per fruit in an aster analysis. This
348 would not be valid, because this product is not distributed according to an exponential
family. Thus, the structure of this aster model precluded inference of overall selection via a
350 single aster analysis.

To account jointly for both aspects of fecundity, we conducted a parametric bootstrap
352 (Efron and Tibshirani 1993, Section 6.5) to infer and test the relationship between overall
fitness and the traits. This entailed simulating 1000 datasets from the aster model estimated
354 from the original data. Specifically, values for the fitness components, **fruit** and **seed**,
were drawn from the sampling distributions used in the aster analysis, with parameters
356 inferred from it. For each individual in the simulated datasets, we calculated the overall
absolute fitness (W) of each individual as the product of its simulated values for **seed**
358 and **fruit**. From this, individual relative fitness (w) was obtained by dividing by the
average, over the whole dataset, of absolute fitness. Estimates of the average selection
360 gradients were obtained for each dataset via multiple regression of w on the three traits

via OLS as described by Lande and Arnold (1983). Finally, each selection gradient for the
362 actual data was taken as the mean of the selection gradient estimates over all the datasets,
and its sampling distribution was approximated by the bootstrap distribution, which was
364 approximately normal (Fig 5 shows the distribution for LN; others were similar). The
estimates obtained in this way for the linear fitness function were qualitatively similar
366 to those produced by OLS multiple regression of w on the three traits. However, the
magnitudes of the estimates differed; for LN, the OLS estimate of the selection gradient
368 exceeded that from the parametric bootstrap based on the aster model by over 25%.
Note that our simulations used the aster model distribution rather than the usual OLS
370 assumptions of homoscedastic normal errors and are thus statistically valid; we used OLS
only for estimation of the bivariate linear regression for each bootstrap dataset.

372 EDITOR: PLACE FIGURE 5 HERE.

As noted, seeds were counted for a fixed number of fruits, three per individual.
374 Subsampling in this way is a common practice in studies of animals (e.g. Howard 1979), and
of plants. It is this feature of the data that dictated modeling the dependence of seed count
376 and fruit number jointly on reproductive status (Fig. 1). This, in turn, obviated the direct
aster analysis of the dependence of overall fitness on the traits and entailed application of
378 the parametric bootstrap. As a practical alternative, the seed count for each individual
could be obtained for a subset of fruits, where the number of fruits on which to count seeds
380 is determined at random, e.g. each fruit is counted with probability p independently of all
other fruits, so the number of fruits counted for an individual with n total fruits has a
382 binomial(n, p) distribution. Data in this design would be amenable to direct analysis of
selection via aster by incorporating the sampling of fruits into the graphical model.

384 In this section, we have illustrated how aster can analyze complete life-history records
to yield phenotypic selection analyses that do not suffer typical violations of assumptions

386 of the OLS regression of fitness on phenotypes. Even for analysis of an annual life-history,
commonly considered relatively straightforward, aster greatly improves the adherence of the
388 analysis to statistical assumptions and, accordingly, the validity of the inferences. We have
also shown that, even when the available data preclude modeling total reproductive output
390 as sequentially dependent on all earlier expressed fitness components, aster estimates the
parameters of a fitness model that can be used with a parametric bootstrap to yield a
392 statistically sound phenotypic selection analysis.

5. Example 3: Estimation of population growth rate, ϕ

394 Lenski and Service (1982) recognized the need for a valid statistical approach to
inferring rates of population growth (ϕ) from life-history records. In particular, they
396 emphasized the importance of accounting for individual variation in survivorship and
fecundity when inferring ϕ via the stable age equation: (their eqn 1). This expression
398 weights earlier produced offspring more heavily in their contribution to population growth
to an extent dependent on the population growth rate itself (Fisher 1930). Lenski and
400 Service (1982) presented a nonparametric approach that resamples records of individual
life histories according to the jackknife procedure. Using the properties of the jackknife,
402 they showed how to obtain estimates and sampling variances of ϕ . They illustrated the
approach with a small dataset sampled from the aphid, *Uroleucon rudbeckiae*. The survival
404 and fecundity in each of fourteen age intervals were recorded for 18 individuals in a cohort
(see Fig. 1), and these data served as the basis for estimating ϕ and its sampling variance
406 for this cohort (data printed in Lenski and Service 1982).

In using the aster approach to analyze these data, we modeled the binomial parameter
408 governing survival probability, $\text{logit}(\sigma_x)$, as a quadratic function of age, x , and found not
only that survivorship declined significantly with age ($P = 0.001$) but also that there is

410 significant deviation ($P = 0.028$) from a linear decline in $\text{logit}(\sigma_x)$. Expected fecundities,
412 β_x , modeled according to a Poisson distribution, were estimated for each age class, x , given
survival to that age, as the stable age equation requires.

Interest focuses primarily on estimating ϕ , but also on its sampling variance, as noted
414 by Alvarez-Buylla and Slatkin (1994), because of its importance in assessing whether a
population is growing or declining. We estimated ϕ directly as a nonlinear function of the
416 expectations of survivorship to age x , σ_x , and fecundity at that age, β_x . From the data
given, we estimated $\phi = 1.677$ with a standard error of 0.056. Our estimate is similar
418 to the estimate 1.688 from the method Lenski and Service (1982) recommend, and 95%
confidence intervals are also similar, ours being (1.57, 1.79) and theirs (1.52, 1.85). Whereas
420 we see only modest improvement over that of Lenski and Service (1982), the aster method
can be used in more complicated situations where the jackknife is inapplicable. Moreover,
422 the generality of the aster framework permits incorporation of age structure into fitness
comparisons, like those in Examples 1 and 2. Because aster employs parametric models, we
424 expect it also to surpass the approach of McGraw and Caswell (1996) in statistical power,
while avoiding the distributional problems and estimation bias they acknowledge.

426

6. Discussion

Both the numerical and genetic dynamics of populations depend fundamentally on
428 the individuals' contribution to the next generation, their fitness. Extensive theoretical
work (e.g. Fisher 1930; Charlesworth 1980) has formalized and extended this insight of
430 Darwin, yet statistical challenges have continued to compromise the empirical evaluation
of fitness. The aster approach (Geyer et al. 2007) addresses these challenges by explicitly
432 modeling the dependence of components of life-history on those expressed earlier. It thus
takes full advantage of available data to yield comprehensive assessments of fitness that

434 are as precise as possible. The precision of aster modeling not only offers statistical power
for tests of null hypotheses; it also promotes quantitative comparison of fitnesses. Most
436 important, as a general framework for inferring fitness from life-history, aster can address
questions that arise in diverse evolutionary and ecological contexts. The examples presented
438 here illustrate the breadth of aster’s applicability, including comparison of mean fitness
among genotypically distinct groups, inference of phenotypic selection, and estimation
440 of population growth rate. Thus, as a single, general framework for addressing diverse
questions in evolution and ecology, aster modeling offers unification of these empirical
442 efforts. Even beyond analysis of life-history data, aster modeling is appropriate for any
set of responses in which there are dependencies analogous to those characteristic of
444 life-histories. In an experimental study of foraging behavior, for example, individual
subjects may forage in a given interval or not and, given that they forage, may take varying
446 numbers of prey. In a medical context, aster could be used to expand on survival analysis
to incorporate measures of quality of life in evaluation of the relative benefits of different
448 procedures. We emphasize that aster obviates the common practice of multi-step analysis
(e.g. van Tienderen 2000), which cannot provide valid statistical tests or sound estimates
450 of sampling error. A single aster model can encompass the real complexities not only of
life-histories, but also of discrete and continuous predictors, and thus provide a full analysis
452 to yield direct inferences about fitness and population growth.

Lifetime fitness rarely, if ever, conforms to any distribution amenable to parametric
454 statistical analysis less complex than aster analysis. This pathology of fitness distributions
has plagued empirical studies of fitness. Resampling approaches are sometimes used, but
456 this is not a general solution, because valid resampling schemes are not generally available
for complex data structures. Moreover, resampling methods sacrifice statistical precision
458 relative to parametric analysis. As an alternative, transformations are often attempted, but
the prevalence of mortality before reproduction typically results in fitness distributions with

460 many individuals at zero, such that no transformation produces a well known distribution.
Even if individuals that don't survive to reproduce are excluded, the fitness of survivors
462 alone is often highly skewed and may not readily transform to a suitable form. Moreover,
even if such a transformation could be found, analyses of fitness on an alternative scale
464 can result in misinterpretation (Stanton and Thiede 2005). Aster avoids these problems by
directly modeling each distinct component of fitness with a suitable parametric distribution
466 and accounting for the dependence of each fitness component on those expressed earlier.
As a consequence, it suitably models the sampling variation and yields results on the
468 biologically natural scale of expected number of individuals produced per individual.
Even when life-history records are available for only a portion of the life-span, as in our
470 example of genotypic differences in fitness of *Echinacea angustifolia*, joint analysis via an
unconditional aster model provides comparisons based on the most comprehensive fitness
472 records at hand.

Studies of variation in fitness often focus on a single component of fitness (e.g. Arnold
474 and Lande 1983). These are less subject to distributional problems and can yield insight
into the nature of fitness variation during a particular episode of selection. However, the
476 resulting understanding of fitness and its variation is fragmentary and can be misleading
when the relationship between components of fitness, on the one hand, and traits or
478 genotypes, on the other, varies over the lifespan (Prout 1971). Whereas Arnold and Wade
(1984a) proposed an approach to evaluate phenotypic selection over multiple episodes
480 (modified by Wade and Kalisz 1989), this approach comprises separate analyses of each
episode, ignoring the dependence structure of fitness components. The sampling variance of
482 the resulting estimates of selection cannot readily be determined.

Our first example demonstrates the most basic use of unconditional aster models to
484 estimate and compare mean fitness for groups produced by different mating schemes and,

thus, differing in genetic composition. Here, individual size is considered a component of
486 fitness during the juvenile period; the typically strong positive relationship between size and
eventual fecundity justifies this here, as elsewhere. This analysis reveals highly significant
488 inbreeding depression and yields an overall estimate of at least 70%, when extrapolated
linearly to inbreeding arising from one generation of selfing. We have also used aster to
490 analyze survival and annual production of flower heads jointly for samples of remnant
populations of *E. angustifolia* grown in a common field environment. This analysis, which
492 used an aster model similar in structure to that of example 1, revealed significant differences
in fitness among remnants (Geyer et al. 2007). In the case we provide here, we further
494 demonstrate how the likelihood framework of aster permits straightforward tests of several
alternative hypotheses. We show that the significant early disadvantage in size and survival
496 of inbred plants does not adequately account for the fitness differences at the end of the
period of observation. Rather, inbreeding depression in growth and survival persists beyond
498 the first three months. Thus, in addition to providing statistically rigorous comparisons of
overall fitness among groups, aster yields insights into the sources of fitness differences.

500 Aster readily extends beyond this basic application to address further questions
in evolution and ecology, as our remaining examples show. To accomplish phenotypic
502 selection analysis, the aster model includes the traits under consideration as predictors of
cumulative fitness. Inference of quadratic and correlational selection is also straightforward.
504 Phenotypic selection analysis via aster establishes the relationship between individuals’
overall demographic-genetic contribution to the next generation and the traits they express.
506 This was the goal of a method presented by van Tienderen (2000), but because that method
does not take into account variation among individuals in demographic inferences, it cannot
508 validly represent the statistical uncertainty of inferred parameters. Moreover, because van
Tienderen’s approach involves separate analyses to estimate selection gradients for each
510 fitness component, it does not take into account the dependence relationships of the

fitness components and is subject to the usual distributional problems (e.g. Coulson et al.
512 2003).

Aster also encompasses inference of population growth rates. Lenski and Service
514 (1982) first noted the importance of sound statistical modeling for population growth.
They recognized that the appropriate unit of observation is the individual and its complete
516 life-history. Our use of aster for inferring ϕ builds on their work by employing parametric
models for each life-history event. The resulting estimate and confidence interval for ϕ is
518 similar to that obtained by Lenski and Service’s (1982) method using the jackknife. This
application of aster differs from our other examples in using aster models to obtain estimates
520 of life-history events, conditional on previous aspects of the life-history. In our example,
expected fecundity at a given age x , which is the product $\beta_x \sigma_x$ in the notation of Lenski
522 and Service (1982), is an unconditional mean value parameter directly estimated by the
aster model. As with any life table, $\sigma_x \beta_x$ for all x implicitly determines ϕ . The important
524 point is that aster analysis provides a sound parametric basis for evaluating the variance of
 ϕ . In cases where there is direct interest in inference about particular components of fitness,
526 aster models offer the capability of obtaining estimates of expected values of a life-history
component, conditional on earlier life-history status (e.g. fecundity at a given age, given
528 survival to that age), as we show in our estimation of the finite rate of increase from the
aphid data. Conditional estimation could also be used for inferring phenotypic selection in
530 a given episode.

In the example of Lenski and Service (1982), complete records are available for each
532 individual in a single cohort, so the life table can be based on age. Life histories are often
tabulated in relation to size or stage categories instead of age, because size or stage often
534 predicts survival and fecundity better than age does (Caswell 1989). Thus, life-history can
be informative even when ages of individuals are unknown, for example, in censuses of

536 populations in nature, as in the examples of Alvarez-Buylla and Slatkin (1994). Though our
examples include only cases based on age, aster analysis applies equally well to life-history
538 data according to stage or size (Caswell 1989, chap. 2).

Because aster accommodates the complex dependencies of life-history data, as well as
540 the varied probability models that are appropriate, it yields rigorous and precise inferences
about population growth and fitness. Our examples demonstrate aster’s versatility, which
542 suits it to address the full breadth of questions that arise in relation to life-histories.

RGS and CJG cordially thank Janis Antonovics for his encouragement when we
544 began work on the basic idea about 1980 and for funding its development then, as well as
for his enthusiasm about its eventual realization. Computational challenges stymied the
546 initial efforts, and other work intervened until the richness of life-history data from recent
experiments stimulated us to revisit the idea. For very helpful suggestions for clarifying the
548 manuscript, we thank J. Antonovics, K. Mercer, M. Price and N. Waser. Examples 1 and 2
are drawn from research funded by NSF (DMS-0083468, DEB-0545072, DEB-0544970) and
550 EPA, respectively, as well as the University of Minnesota Center for Community Genetics.

REFERENCES

- 552 Alvarez-Buylla, E. R., and M. Slatkin. 1994. Finding confidence limits on population
growth rates: three real examples revised. *Ecology* 75:255–260.
- 554 Antonovics, J., and N. C. Ellstrand. 1984. Experimental studies of the evolutionary
significance of sexual reproduction. I. A test of the frequency-dependent selection
556 hypothesis. *Evolution* 38:103–115.
- Arnold, S. J., and M. J. Wade. 1984a. On the measurement of natural and sexual selection:
558 theory. *Evolution* 38:709–719.
- Arnold, S. J., and M. J. Wade. 1984b. On the measurement of natural and sexual selection:
560 applications. *Evolution* 38:720–734.
- Barndorff-Nielsen, O. E. 1978. Information and exponential families. John Wiley,
562 Chichester.
- Belaoussoff, S., and J. S. Shore. 1995. Floral correlates and fitness consequences of
564 mating-system variation in *Turnera ulmifolia*. *Evolution* 49:545–556
- Caswell, H. 1989. Matrix population models: construction, analysis, and interpretation.
566 Sinauer Associates, Sunderland, Mass.
- Charlesworth, B. 1980. Evolution in age-structured populations. Cambridge University
568 Press, Cambridge.
- Cheng, S., D. Wang, and E. P. Smith. 2000. Adjusting for mortality effects in chronic
570 toxicity testing: mixture model approach. *Environmental Toxicology and Chemistry*
19:204–209.

- 572 Coulson, T., L. E. B. Kruuk, F. Tevacchia, J. M. Pemberton, and T. H. Clutton-Brock.
2003. Estimating selection on neonatal traits in red deer using elasticity path
574 analysis. *Evolution* 57:2879–2892.
- Dagne, G. A. 2004. Hierarchical Bayesian analysis of correlated zero-inflated count data.
576 *Biometrical Journal* 46:653–663.
- Efron, B., and R. J. Tibshirani. 1993. *Introduction to the bootstrap*. Chapman and Hall,
578 London.
- Etterson, R. R. (2004) Evolutionary potential of *Chamaecrista fasciculata* in relation to
580 climate change. I. Clinal patterns of selection along an environmental gradient in the
great plains. *Evolution* 58:1446–1458.
- 582 Fisher, R. A. 1922. On the mathematical foundations of theoretical statistics. *Philosophical*
Transactions of the Royal Society of London, A 222:309–368.
- 584 Fisher, R. A. 1930. *The genetical theory of natural selection*. Clarendon Press, Oxford.
[Reprinted Dover]
- 586 Galloway, L. F., and J. R. Etterson. 2007. Inbreeding depression in an autotetraploid herb:
a three cohort field study *New Phytologist* 173:383–392.
- 588 Geyer, C. J., S. Wagenius, and R. G. Shaw. 2007. Aster models for life history analysis.
Biometrika in press.
- 590 Howard, R. D. 1979. Estimating reproductive success in natural populations. *American*
Naturalist 114:221–231.
- 592 Husband, B. C., and D. W. Schemske. 1996. Evolution of the magnitude and timing of
inbreeding depression in plants. *Evolution* 50:54–70.

- 594 Janzen, F. J., and H. S. Stern. 1998. Logistic regression for empirical studies of multivariate
selection. *Evolution* 52:1564–1571.
- 596 Kendall, M., and A. Stuart. 1977. *The advanced theory of statistics*, Vol. 1, 4th Ed.
Macmillan, New York.
- 598 Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters.
Evolution 37:1210–1226.
- 600 Lenski, R. E., and P. M. Service. 1982. The statistical analysis of population growth rates
calculated from schedules of survivorship and fecundity. *Ecology* 63:655–662.
- 602 Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. A. Field, S. J. Low-Choy,
A. J. Tyre, and H. P. Possingham. 2005. Zero tolerance ecology: improving ecological
604 inference by modelling the source of zero observations. *Ecology Letters* 8:1235–1246.
- McGraw, J. B., and H. Caswell. 1996. Estimation of individual fitness from life-history
606 data. *American Naturalist* 147:47–64.
- Mitchell-Olds, T., and R. G. Shaw. 1987. Regression analysis of natural selection: statistical
608 inference and biological interpretation. *Evolution* 41:1149–1161.
- Morris, W. F., and D. F. Doak. 2002. *Quantitative conservation biology: theory and
610 practice of population viability analysis*. Sinauer Associates, Sunderland, Mass.
- Prout T. 1971. The relation between fitness components and population prediction in
612 *Drosophila*. I: the estimation of fitness components. *Genetics* 68:127–149.
- R Development Core Team. 2006. *R: A language and environment for statistical computing*.
614 R Foundation for Statistical Computing, Vienna. <http://www.r-project.org>.
- Schluter, D. 1988. Estimating the form of natural selection on a quantitative trait.
616 *Evolution* 42:849–861.

- Schluter, D., and D. Nychka. 1994. Exploring fitness surfaces. *American Naturalist*
618 143:597–616.
- Shaw, R. G. 1986. Response to density in a wild population of the perennial herb *Salvia*
620 *lyrata*: variation among families. *Evolution* 40:492–505.
- Shaw, R. G., C. J. Geyer, S. Wagenius, H. H. Hangelbroek, and J. R. Etterson. 2007.
622 Supporting data analysis for “Unifying life history analysis for inference of fitness
and population growth”. Technical Report No. 658, School of Statistics, University
624 of Minnesota.
- Stanton, M. L., and D. A. Thiede. 2005. Statistical convenience vs. biological insight:
626 consequences of data transformation for the analysis of fitness variation in
heterogeneous environments. *New Phytologist* 166:319–337.
- 628 van Tienderen, P. H. 2000. Elasticities and the link between demographic and evolutionary
dynamics. *Ecology* 81:666–679.
- 630 Wade, M. J. and S. Kalisz. 1989. The additive partitioning of selection gradients. *Evolution*
43:1567–1569.
- 632 Wagenius, S. 2000. Performance of a prairie mating system in fragmented habitat:
self-incompatibility and limited pollen dispersal in *Echinacea angustifolia*. Ph.D.
634 dissertation. University of Minnesota.

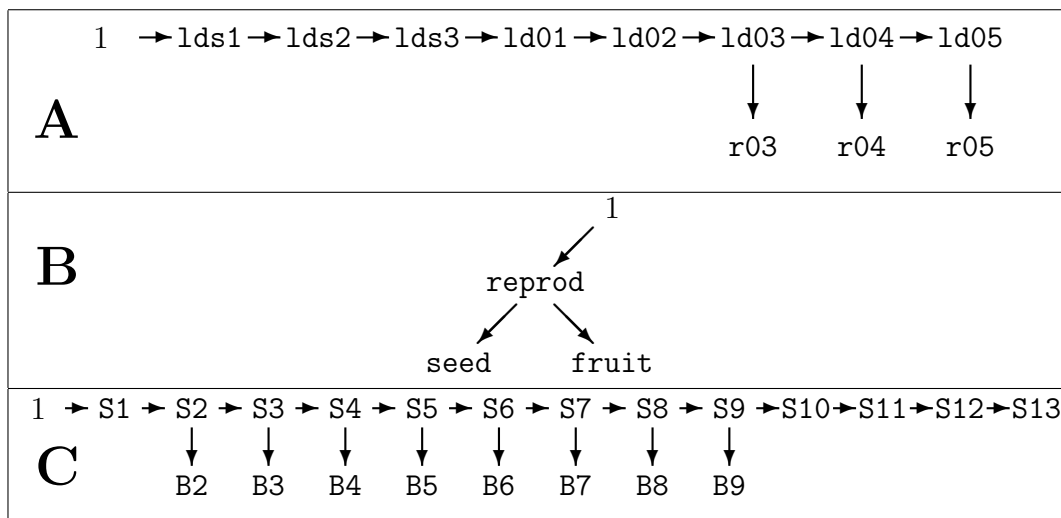


Fig. 1.— Graphical models used to analyze data for the three empirical examples presented. In each case, each individual present at the outset is indicated by the constant variable 1, the root node. Arrows lead from one life history component to another that immediately depends on it (from predecessor node to successor node of the graph). Successor variables are conditionally independent given the predecessor variable. If a predecessor variable is nonzero, then a particular conditional distribution of the successor variable is assumed. If a predecessor variable is zero for a given individual, for example due to mortality, then its successor variables are also zero. A: Example 1: *Echinacea angustifolia*, a perennial plant. Fitness comprises juvenile survival at three times up to transplanting into the field (lds_i) and subsequent survival through five years (ld0_i), as well as the plant’s number of rosettes (r0_i) in three years. The survival variables are modeled as (conditionally) Bernoulli (zero indicates mortality, one indicates survival), and r0_i is (conditionally) zero-truncated Poisson (i.e. a Poisson random variable conditioned on being greater than 0). B: Example 2: *Chamaecrista fasciculata*, an annual plant. Success or failure of reproduction (here, including survival to reproduce) is given by **reprod**, modeled as Bernoulli (zero indicates no seeds, one indicates nonzero seeds). Fecundity, given that a plant reproduces, is given by its number of fruits (**fruit**), modeled as two-truncated negative binomial and number of seeds per fruit in a sample of three fruits (**seed**), modeled as zero-truncated negative binomial. C: Example 3: *Uroleucon rudbeckiae*, an aphid. S_i is (conditionally) Bernoulli and B_i is (conditionally) zero-truncated Poisson (indicating count of offspring).

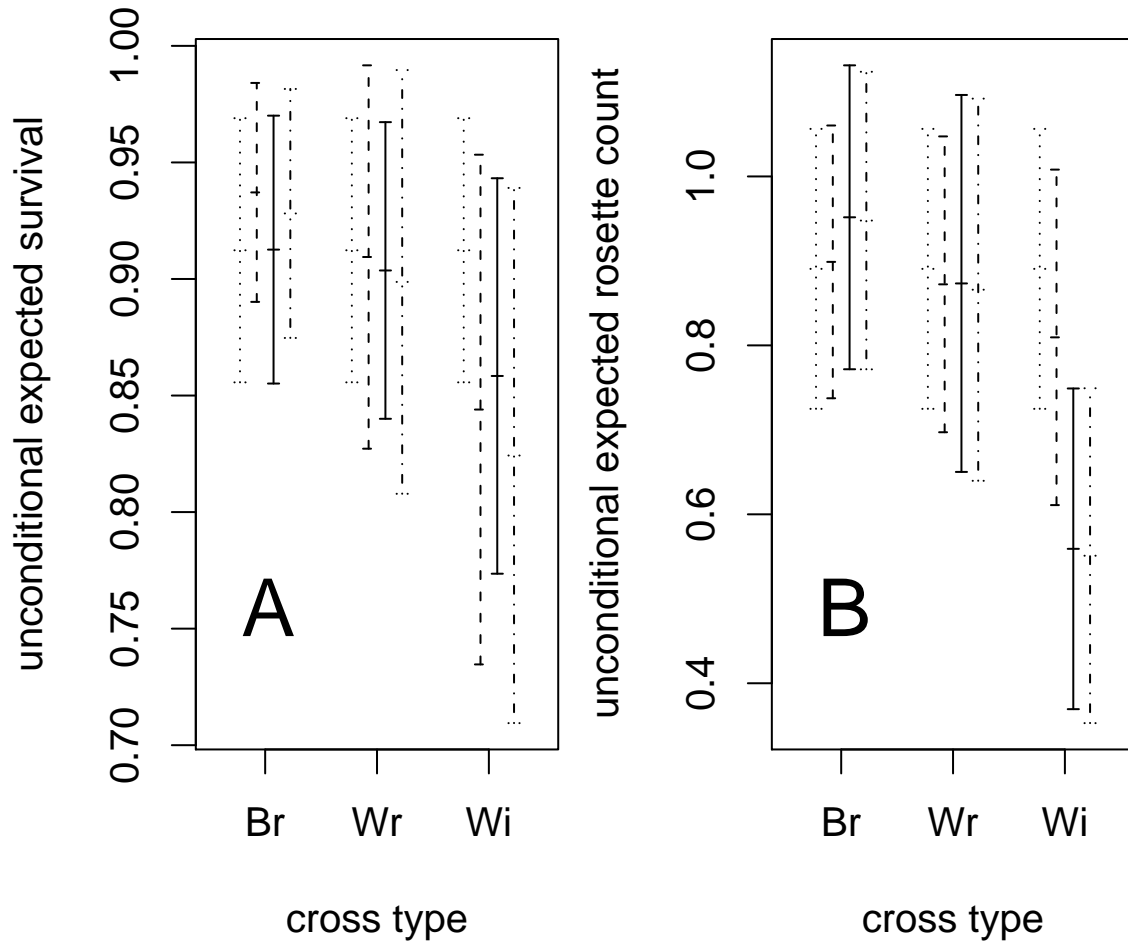


Fig. 2.— Predicted values and 95% confidence intervals for the unconditional mean value parameter for (A) survival up to transplanting and (B) rosette count in the last year recorded (i.e. overall fitness over the study period) for a “typical” individual for each cross type. The experimentally imposed crossing treatments are Br, between remnant populations; Wr, within remnant populations; and Wi, inbred within remnants (i.e. Between sibs). Lines indicate model: dotted, “sub” model; dashed, “chamber” model; solid, “field” model; dot-dash, “super” model.

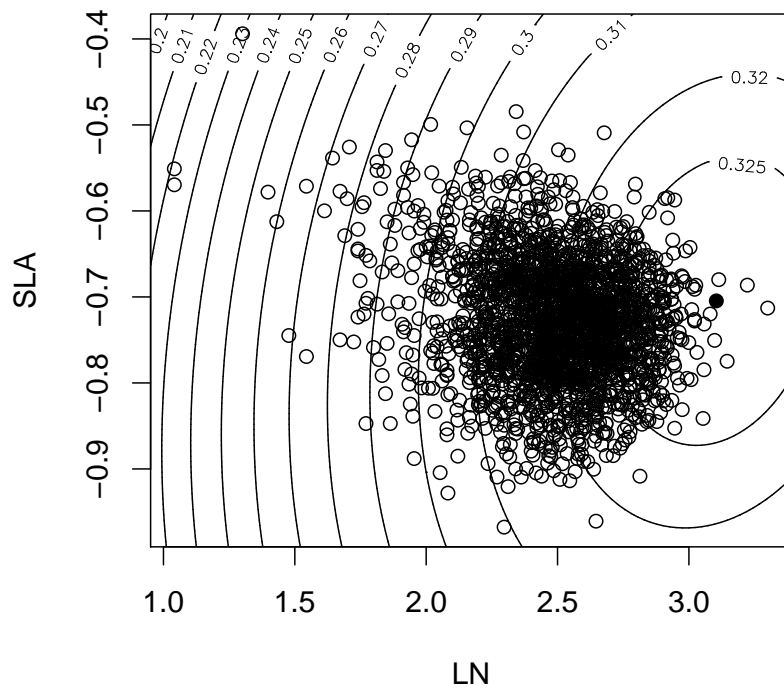


Fig. 3.— Scatterplot of SLA (specific leaf area, ln transformed) versus LN (leaf number, ln transformed) with contours of the estimated quadratic fitness function. Solid dot is the point where the estimated fitness function achieves its maximum.

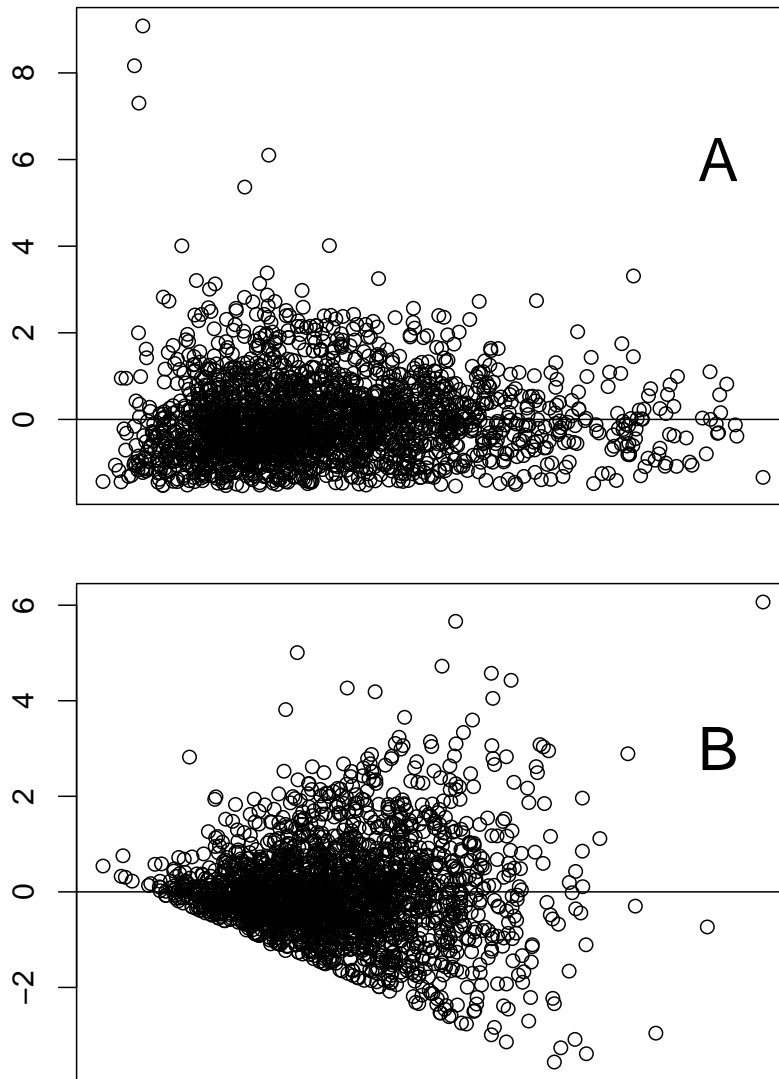


Fig. 4.— Scatter plot of residuals from phenotypic selection analysis for *Chamaecrista fasciculata*. A. Pearson residuals for fruit count conditional on survival are plotted against values fitted from the quadratic aster model including two traits, leaf number (LN) and specific leaf area (SLA). B. Residuals for fruit count are plotted against values fitted for the quadratic model including two traits, leaf number (LN) and specific leaf area (SLA) via ordinary least squares (OLS).

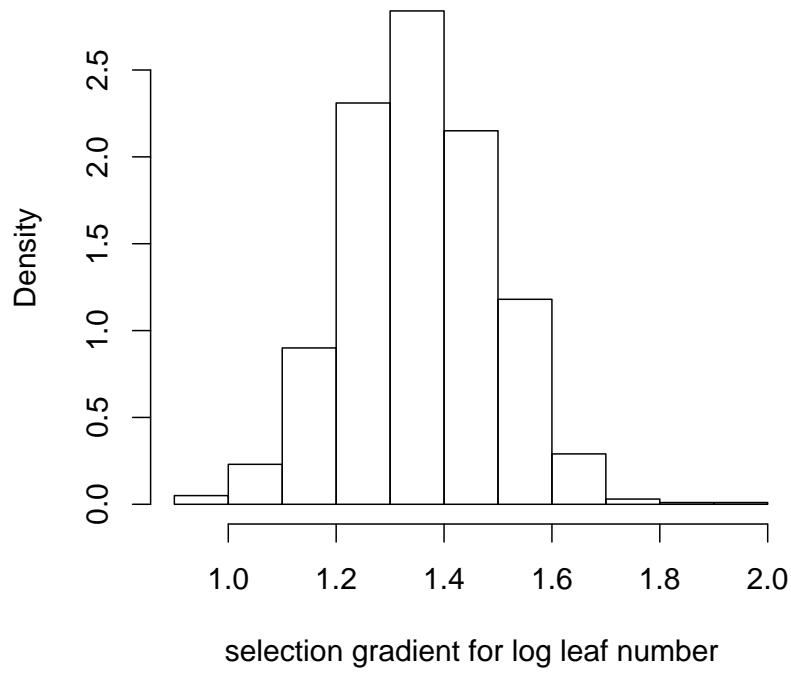


Fig. 5.— Histogram of the parametric bootstrap distribution for the selection gradient for log leaf number, LN, obtained for each bootstrap dataset as the partial regression of seed number on log leaf number, with specific leaf area (SLA) and reproductive stage (RS) also included in the model.