COMMENTARY

Issues of terminology, gradient dynamics and the ease of sympatric speciation in Adaptive Dynamics

D. WAXMAN* & S. GAVRILETS†

*Centre for the Study of Evolution, School of Life Sciences, University of Sussex, Brighton, Sussex, UK
†Department of Ecology and Evolutionary Biology, and Department of Mathematics, University of Tennessee, Knoxville, TN, USA

General comments

It has been plainly stated that Adaptive Dynamics is ‘not a scientific theory, but a mathematical framework’ (Metz, 2005); it is a ‘method but not a model’ (Kisdi & Gyllenberg, 2005). From a completely pragmatic point of view, there is, then, only one question that evolutionary biologists need to ask: Is Adaptive Dynamics a useful way of approaching problems? The 14 Commentaries present a variety of different views, indicating that there is no consensus in the answer to this question. Indeed, we view the previous pages, consisting of the Target Review and the 14 Commentaries, as a part of a much larger public discussion on Adaptive Dynamics, and we thank the other authors for their contributions to the discussion.

We feel it necessary to acknowledge that the wording of our comment in Question 6 of our Target Review (Waxman & Gavrilets, 2005) concerning ‘hidden limitations and unconscious or implicit assumptions’ was not optimal. It was not our intention to imply that limitations and assumptions of Adaptive Dynamics are hidden or unconsciously made by practitioners of Adaptive Dynamics. Rather, the statement was intended for biologists who use Adaptive Dynamics, or who wish to understand the biological implications of its mathematical results, and may not be aware of the underlying limitations and assumptions. These limitations and assumptions are not, in our opinion, sufficiently emphasised in the literature and, equally importantly, the consequences of their violation are not made clear. The Commentary of Geritz & Gyllenberg (2005) makes it admirably clear what these limitations are but not the consequences of their violation. The Commentaries of Abrams (2005), Barton & Polechová (2005), Kokko (2005), Spencer & Feldman (2005), Butlin & Tregenza (2005), and Van Dooren (2005) provide additional discussion of the issue.

We also feel it necessary to clarify our wording about the treatment of genetic drift in Adaptive Dynamics. In the second paragraph of Question 4.3, we said an implicit assumption of Adaptive Dynamics is that all beneficial mutations will always initially increase in frequency. However in the first paragraph, we noted that in the work of Dieckmann & Law (1996), fixation of new alleles occurs at a rate proportional to the probability of fixation of a mutant. Since the probability of fixation of all mutants is not unity, the two paragraphs are clearly inconsistent. Our reading of the Adaptive Dynamics literature, however, leads us to believe that it is common to make the assumption that beneficial mutations always initially increase in frequency (see Meszéna, 2005). We are also of the opinion that the treatment of genetic drift, when included in Adaptive Dynamics, has not been particularly realistic. As an example, the probability of invasion of a deleterious mutation is invariably taken as zero in Adaptive Dynamics (Geritz & Gyllenberg, 2005; Meszéna, 2005). However, if a mutation is deleterious, but only very slightly so, and the population is finite, then such near neutral mutants can invade and become fixed in the population (Kimura, 1957).

At the end of our Review, we stated our earnest and legitimate wish to see more details of numerical simulations, but this was somehow interpreted, in some Commentaries, as an attempt to denigrate Adaptive Dynamics, because of the lack of analytical results, or to attach specific significance to simulations in Adaptive Dynamics. This was not our intention.

Although we appreciate all of the Commentaries made, we will not answer them directly. Rather we discuss three controversial points raised in more than one Commentary. A separate response is, however, also necessary for a number of peripheral points raised by Dieckmann & Doebeli (2005). Our response to these points is contained in Appendix B.

Terminology

Terminology is very important, since correct and precise descriptions allow communication, and science, to progress efficiently. Adaptive Dynamics is an approximation of a complex dynamical system, and as such, we were originally of the view that to helpfully communicate ideas about the subject, terminology should be adopted that is readily recognisable and compatible with that of dynamical systems. However, we find the arguments of Geritz & Gyllenberg (2005), Metz (2005) and Dieckmann & Doebeli (2005) against the terms *equilibrium*, *stationary point* and *saddle point*, to be persuasive and we appreciate their efforts to justify the terminology of Adaptive Dynamics.

Correspondence: Prof. D. Waxman, Centre for the Study of Evolution, School of Life Sciences, University of Sussex, Brighton BN1 9QG, Sussex, UK.
Tel.: +44-1-273-678559; e-mail: D.Waxman@sussex.ac.uk
Canonical equation

Consider now the ‘canonical equation’ of Adaptive Dynamics (Dieckmann & Law, 1996), which approximately describes the rate of change of a trait, due to adaptation. Mathematically similar equations have been well established in the population genetics literature, however, Dieckmann & Doebeli (2005), Geritz & Gyllenberg (2005) and Metz (2005) are of the opinion that there are some fundamental or subtle differences between the ‘canonical equation’ and analogous equations of population genetics.

For the reader’s benefit, let us write down two equations from population genetics, followed by the ‘canonical equation,’ and give a brief description of these equations. Using consistent notation, the three equations read

\[
\Delta z = G \partial \ln W(z) / \partial z, \quad \text{Lande},
\]

\[
\Delta z = G \partial \ln W(z, \bar{z}) / \partial z|_{z=\bar{z}}, \quad \text{Iwasa et al.},
\]

\[
\Delta z = \frac{1}{2} G_m \partial \ln W(z, \bar{z}) / \partial z|_{z=\bar{z}}, \quad \text{Dieckmann & Law}.
\]

The equation labelled Lande follows eqn 7 of Lande (1976). Here \( \Delta z \) is the change in the mean value of a continuously varying trait \( z \), that occurs over one generation, \( \bar{W}(\bar{z}) \) is the average fitness of a (diploid) population – written as a function of the mean trait value \( \bar{z} \), and \( G \) is the genetic variance of the trait. In deriving this equation, Lande assumed the distribution of \( z \) values, in the population, is normal and that the genetic variance, \( G \), is maintained at a constant level by a balance between mutation, selection and drift. Lande’s derivation did not allow for selection to be frequency dependent. Later, however, Iwasa et al. (1991) proposed an alternative derivation that was applicable to frequency-dependent selection, when individual fitness \( W(z, \bar{z}) \) is a function of both individual phenotype \( z \) and the mean trait value \( \bar{z} \) of the population (cf. Kirkpatrick & Roussel, 2005). Under the assumption that selection is weak Iwasa et al. (1991; eqn A9) obtained the equation above, labelled Iwasa et al. (the notation \( |_{z=\bar{z}} \) means the derivative is evaluated with \( z \) set equal to \( \bar{z} \)). Similar equations were derived by a number of other authors (see Abrams et al., 1993) and the Commentary by Abrams (2005) for a discussion. Consider now the Adaptive Dynamics approach. Let \( W(z, \bar{z}) \) be the fitness of a rare mutant with trait \( z \) in a population in which the resident trait value is \( \bar{z} \). Note that because the invader is assumed to be rare, the average trait value coincides with that of the resident. Then assuming, for simplicity, discrete generations (i.e. using a difference equation rather than a differential equation), the change of \( z \) over a generation is given by the equation labelled above as Dieckmann & Law and this is the ‘canonical equation’ (Dieckmann & Law, 1996, eqn 4.12). In this equation \( G_m = N \mu \sigma^2 \) is the genetic variance produced by mutations, each generation, in a monomorphic population of \( N \) individuals, with mutation rate \( \mu \) per individual and a variance of mutational effects of \( \sigma^2 \). Note that within the Adaptive Dynamics approximation, \( G_m \) is the only component of genetic variation available. The factor \( 1/2 \) is present in the Dieckmann & Law equation because individuals are assumed haploid; in the equations of the other authors, this factor is obscured by an additional factor of 2, because these equations describe diploid populations. According to Dieckmann & Doebeli’s (2005) Commentary, the equations used in quantitative genetics and Adaptive Dynamics ‘fundamentally differ’ and these authors colourfully dismiss any statements about similarities. While each reader will come to their own conclusions about the biological underpinnings and relatedness of the three equations, in our opinion, all have very close mathematical and biological connections.

Sympatric speciation

One especially interesting and controversial topic is the justification of the claims often made within the Adaptive Dynamics literature, that sympatric speciation occurs easily. Earlier we have identified a number of problems with numerical models these claims are based upon (see e.g. Gavrilets & Waxman, 2002; Gavrilets, 2004). In our Review we paid special attention to three underlying assumptions of a paper by Dieckmann & Doebeli (1999) which has become the standard reference in support of this claim. These assumptions are (1) the initial genetic variation has the maximum possible level, with all allele frequencies being initially set at \( 1/2 \), (2) adopting a mutation rate that has an unrealistically high value (at least two orders of magnitude higher than existing estimates), and (3) completely neglecting costs of mate choice (being choosy) in spite of very strong choosiness. Using our analysis of existing models we concluded that ‘introducing costs of choosiness and starting the population at a realistically low level of genetic variation (appropriate to mutation-selection balance) with realistic values of the mutation rate will almost definitely prevent sympatric speciation in the Dieckmann–Doebeli model or in similar models.’ We challenged the authors of these models to falsify our hypothesis.

We appreciate that Doebeli & Dieckmann (2005) took concrete steps towards answering this challenge. However, we are not satisfied with their approach and do not see how they come to believe that they have refuted our objections. Firstly, these authors decided to relax only one assumption at a time rather than all three of them simultaneously. Second, they decided to show a single simulation run for each parameter combination which apparently they view as typical. One should realize that the debate is not whether sympatric speciation may happen ‘in principle’ but whether it happens ‘easily’ and relatively quickly under biologically realistic conditions.
Given the stochastic nature of the simulations, a single run is not sufficient to see the general picture. Third, in the models of this type, the waiting time to speciation is very sensitive to parameters (see e.g. Gavrilets (2004)) and, by extension, to the way certain components of the numerical model are implemented. Unfortunately Doebeli & Dieckmann (2005) do not provide enough details of their implementation of the costs of being choosy, and this prevents us from fully understanding the degree to which their conclusions are justified.

To remove these shortcomings one of us (SG) together with Michael Vose (University of Tennessee) have implemented the Dieckmann & Doebeli (1999) numerical model following their recipe as closely as possible given the concise nature of their original publication. In our simulations, the initial genetic variation was absent for the mating and marker characters but for the ecological character the allele frequencies were set at 1/2 (because for the parameter values used, these allele frequencies evolve to this point anyway). The costs of mate choice were introduced via the reduction of the birth rate of choosy individuals. The parameter C controlling these costs is the average number of encounters with potential mates during the life-time. This parameter is analogous to parameter C in Bolnick (2004) and appears to be analogous to parameter N in Doebeli & Dieckmann (2005). The implicit assumption is that if a female does not find an acceptable mate during a specific time interval, she remains unmated. Very large C means little or no costs of choosiness whereas small values of C mean costs are important. Some additional technical details are given in Appendix A.

We limited the duration of simulations to 100 000 time units (corresponding, approximately, to 100 000 generations). We think that assuming the constancy of environmental conditions and the ability of a relatively small isolated population with several hundred individuals to survive even during this interval is problematic. In any case, if sympatric speciation by competition requires hundreds of thousands of generations this puts this process in the same category as sympatric speciation by random genetic drift and mutation via accumulation of mutually incompatible alleles (see e.g. Wu, 1985; Higgs & Derrida, 1992; Gavrilets, 2004, Chapter 9), which however requires less strict conditions.

Our numerical results on the median waiting time to speciation, \( T_{\text{in}} \), are shown in Table 1. The time unit is defined as the average time until a newborn individual gives birth (assuming no costs of choice). Doebeli and Dieckmann’s ‘direct assortative mating’ refers to the case when both the death rate and mating choice are based on the same trait. This is a ‘magic trait’ model in the terminology of Gavrilets (2004). Doebeli and Dieckmann’s ‘indirect assortative mating’ refers to the case when the death rate is based on one trait and mating choice is based on a different trait which is expressed in both sexes. This is a ‘similarity-based nonrandom mating’ in the terminology of Gavrilets (2004).

In spite of significant stochastic variability, one can clearly see a number of patterns. Firstly, having a realistic mutation rate, and assuming no initial genetic variation, greatly increases the waiting time to speciation in the ‘direct assortative mating’ model even when the costs of choosiness are absent. For example, with \( \mu = 10^{-5} \), \( T_{\text{in}} \) decreases from 130 to 14 530. Second, decreasing the mutation rates to realistic values dramatically increases the waiting time to speciation in both models for all values of C. Third, costs of being choosy prevent speciation in the ‘indirect assortative mating’ model. Even in the magic trait model (Doebeli and Dieckmann’s ‘direct assortative mating’ model), which is probably the most conducive model for sympatric speciation (e.g. Kondrashov, 1983a, b; Rice, 1984; Gavrilets, 2004, Chapter 10), these costs can significantly delay or even prevent speciation. Decreasing the strength of nonrandom mating (Bolnick, 2004) or increasing the number of loci (Dieckmann & Doebeli, 1999) will further delay speciation.

In response to our criticism of the initial conditions used, Doebeli & Dieckmann (2005) suggest that allele frequencies at 1/2 is a natural choice for neutral loci and that setting allele frequencies at 1/2 in the mating trait is reasonable because this results in random mating of ‘average’ individuals. We disagree. First, the level of genetic variation in a neutral locus (or trait) maintained by mutation in a finite population will be much smaller than the maximum possible level they use. As a result, the population may simply lack genetic variation necessary to initiate speciation. Second, the fact that the average value of the mating trait corresponds to random mating is largely irrelevant because almost all individuals in the initial population will deviate from such a spurious ‘average type’ by exhibiting strong positive or negative assortative mating. Excluding humans, we are not aware

<table>
<thead>
<tr>
<th>( C ) (average # encounters)</th>
<th>Direct assortative mating ( \mu = 10^{-3} )</th>
<th>( 10^{-4} )</th>
<th>( 10^{-5} )</th>
<th>( \mu = 10^{-3} )</th>
<th>( 10^{-4} )</th>
<th>( 10^{-5} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doebeli &amp; Dieckmann</td>
<td>128</td>
<td>120</td>
<td>130</td>
<td>15 300</td>
<td>66 860</td>
<td>2/20</td>
</tr>
<tr>
<td>( \infty )</td>
<td>327</td>
<td>1833</td>
<td>14 530</td>
<td>14 160</td>
<td>77 360</td>
<td>1/20</td>
</tr>
<tr>
<td>50</td>
<td>354</td>
<td>1555</td>
<td>7790</td>
<td>0/20</td>
<td>0/20</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1233</td>
<td>9800</td>
<td>39 975</td>
<td>( 1/20 )</td>
<td>( 4/20 )</td>
<td>( 0/20 )</td>
</tr>
</tbody>
</table>

The data shown are based on 20 runs for each parameter combination. When \( T_{\text{in}} > 10^5 \) the number of times speciation was observed in 20 runs is shown. Simulations were not done for parameter values corresponding to empty cells because it is clear that speciation will most likely not occur within the time-span used.
of any biological populations where such a mixing of mating types has been observed.

In response to our criticism of the high mutation rate used in their simulations, Doebeli & Dieckmann (2005) suggest that their ‘locus should not be understood as coding for a single protein, but more generally as describing independent stretches of DNA’ that ‘might be very much longer than a single locus, and hence the mutation rate per such stretch might be quite high.’ However, we are not aware of any data suggesting that continuously varying traits are controlled by a number of mutually independent clusters of very many tightly linked genes. In any case, such a model will be incompatible with their assumption of two alleles per ‘locus’ whereas the dynamics of multiallele models are expected to be quite different (because in a multiallelic case, each allele will experience much weaker selection). The argument of Doebeli & Dieckmann (2005) and Meszéna (2005) that one can increase mutation rate \( \mu \) to compensate for small population size \( N \) that one is forced to use in numerical simulations to have them run fast enough is not convincing either. It is only true for neutral alleles that the dynamics of diversification is largely controlled by \( N \) and \( \mu \) through their product \( N\mu \). With the inclusion of selection, the dynamics with low \( N \) and high \( \mu \) will, in general, be quite different from that with high \( N \) and low \( \mu \).

An important question concerns the level of costs of being choosy in natural populations. Unfortunately, not much information seems to exist. Some data mostly on birds suggest that the number of males sampled by females is typically in the range of 2–20 (Jennions & Petrie, 1997). With these values of \( C \), sympatric speciation would be very difficult to achieve in the models being discussed. We note that similar effects of costs of choosiness on the possibility of sympatric speciation are observed and similar conclusions are made in other recent papers (e.g. Kirkpatrick & Nuismer, 2004; Bolnick, 2004; Gourbiere, 2004 as well as in the Commentary by Gourbiere & Mallet, 2005).

Doebeli & Dieckmann (2005) suggest that in some groups, like chimpanzees and humans, individuals evaluate many potential mates before reproduction so that parameter \( C \) can be rather large. We are not convinced this observation is relevant to potential cases of sympatric speciation that have been identified so far (e.g. Coyne & Orr, 2004). At the same time we note that more empirically-based information on the process of mate choice is indeed necessary.

One also has to keep in mind that reduced mating success, as captured by parameter \( C \), is only one of many costs of choosiness that have been identified. Other costs include reduced viability due to travel time and energy, assessment time and energy, risks of predation, risk of disease transmission, risk of injury from males, as well as reduced fertility (e.g. Pomiankowski, 1987; Jennions & Petrie, 1997). These costs will have a similar effect on reducing genetic variation in mating characters and inhibiting sympatric speciation. Costs of being choosy have been identified recently as a major factor opposing speciation whose importance is comparable to that of recombination (Gavrilets, 2004).

We conclude that although sympatric speciation is theoretically possible, the conditions are restricted. These conditions are: strong combined effects of disruptive selection and nonrandom mating, strong association of the genes controlling traits subject to selection and those underlying nonrandom mating, high levels of genetic variation, and the absence of costs on being choosy. For more discussion of Dieckmann & Doebeli’s (1999) models see Gavrilets (2005) whereas Gavrilets (2004) provides a general discussion of theoretical speciation research as well as a large number of simple models in which conditions for sympatric speciation were found analytically.

**Dedication**

We would like to end this work by dedicating it to the memory of John Maynard Smith (JMS). This seems highly appropriate, given that JMS was extremely interested in Adaptive Dynamics and, as has been pointed out several times in the above discussion, Adaptive Dynamics has very strong overlaps with his work. In his last years, JMS spent a considerable amount of time and effort thinking deeply about the issues raised by Adaptive Dynamics – and running simulations; indeed he wrote a first draft of a Commentary on our Review, before his untimely death. We miss hearing his views on the state of the discussion so far.

**Acknowledgments**

We are grateful to Ted Kawecki, Kate Lessells, John Welch and an anonymous reviewer for their helpful comments and suggestions. SG thanks Michael D. Vose for implementing a numerical model used to produce the data in Table 1.

One of us (DW) was supported by the Leverhulme Trust and the other (SG) was supported by National Institutes of Health Grant GM56693 and by National Science Foundation Grant DEB-0111613.

**Appendix A**

Here, we discuss some details of the numerical model used to generate the data presented in Table 1.

Two changes relative to the basic model in Dieckmann & Doebeli (1999) have been made. First, we used four loci for each trait rather than 5. This was done in order to be able to set an initial population to be both monomorphic and randomly mating. (Note that with an odd number of equivalent diallelic loci contributing to an
additive trait, it is impossible to have a homozygote with the trait value exactly at the middle of the range of possible trait values.) Because the time to speciation generally increases with the number of loci this change should actually decrease the waiting time to speciation. Second, the preference function used by Dieckmann and Doebeli was very asymmetric, with the variance of the preference function of individuals following positive assortative mating being 400 times smaller than that of individuals following negative assortative mating. In the absence of any justification of this assumption we decided to use a simpler symmetric preference function proposed by Bolnick (2004).

To introduce the costs of being choosy we followed the approach of Dieckmann & Doebeli (2003) by setting the birth rate of an individual at \( b_i = (1 + c/P_i)^{-1} \), where \( P_i = \sum_{j \neq i} \psi_{ij} \) has the meaning of the effective number of mating partners for individual \( i \), and \( \psi_{ij} \) is the strength of preference of individual \( i \) for individual \( j \) (0 ≤ \( \psi_{ij} \) ≤ 1) as defined in Bolnick (2004). The coefficient \( c \) controls the costs of being choosy. Small \( c \) implies that birth rate \( b_i \) weakly depends on \( P_i \) and thus costs are low, whereas large \( c \) means that \( b_i \) strongly depends on \( P_i \) and thus costs are high. Another way to interpret \( c \) is based on the fact that \( C = N/c \), the ratio of the population size \( N \) to \( c \), can be interpreted as the average number of potential mates (both acceptable and unacceptable) encountered during a life-span. This means that \( C \) is analogous to the maximum number of males a female is allowed to sample as specified in the model of Gavrilet & Boake (1998) which was used by Matessi et al. (2001) and Bolnick (2004).

The parameter values were the same as given in the Methods section of Dieckmann & Doebeli (1999) except the mutation rates which are specified in Table 1. The date shown in Table 1 are based on 20 runs for each parameter combination.

Appendix B

As stated above, this Appendix addresses a number of peripheral issues raised by the Commentary of Dieckmann & Doebeli (2005) that have not been addressed in the main text. For the purposes of this Appendix, we shall refer to the relevant Commentary simply as D&D.

1. Origin and maintenance of genetic variation: D&D declare ‘fallacious’ our statement in Question 4.2 that polymorphism cannot be maintained when the singular point is locally stable. We note that our statement refers to the classical model of Christiansen & Loeschcke (1980). As we showed in Question 2.9, in this model AD predicts that genetic variation cannot be maintained if \( B < 0 \) whereas the exact results of Christiansen & Loeschcke (1980) show that even under this condition the maintenance of variation is possible.

2. Size of mutational steps: In answer to Question 2.1, we characterised AD as being based on ‘very small changes in existing phenotypic values.’ D&D stated this to not be accurate, because of the inclusion of pairwise invasibility plots as a tool of analysis in Adaptive Dynamics. The use, or not, of pairwise invasibility plots is irrelevant to this issue, since it is taken as a basic assumption of AD that there are ‘small mutational steps’. This assumption is clearly stated in most AD papers including the Commentaries by Geritz & Gyllenberg (2005), Metz (2005) and Meszéna (2005). Given the disagreements between practitioners of Adaptive Dynamics about the size of mutations, there is surely room for clarifying where Adaptive Dynamics stands on this. We emphasise that this is an important issue, given the empirical and theoretical results concerning the size of mutations contributing to adaptation (Orr & Coyne, 1992; Barton & Polechová, 2005; Tregenza & Butlin, 2005).

3. Range of trait values: D&D declare ‘clearly false’ our statement that mutant phenotype can take any value from \( \infty \) to \( -\infty \). Our statement was too broad; we should have added the qualification that in ‘most cases’ this is assumed. We recognise, of course, that some traits, such as mating or dispersal probabilities, are naturally constrained (lying between zero and one). Alternatively, if one explicitly introduces restrictions on the range of possible trait values then clearly the limits can be different from \( -\infty \) and \( \infty \). In this context, we note that recent work indicates that limiting the range of traits may create artefacts that significantly change conclusions about the ease of sympatric speciation (Polechová and Barton; in press).

4. Dependence of mutations: D&D declare ‘erroneous’ the statement that ‘it is... assumed that the distribution of the deviation of the mutant from the parental type is independent of the parental phenotype.’ They omitted the qualification that was present at the beginning of the quoted sentence ‘In most cases...’

5. Definition of invasion fitness: D&D declare ‘incorrect’ our definition of invasion fitness and our use of the term ‘frequency’ rather than ‘density’ in our discussion of mutant invasion. We note that our goal was to explain AD in standard population genetic terms rather than in terms used by AD, and in this context our terminology is standard: Diekmann et al. (2003) also use mutant frequencies instead of densities to derive invasion fitness.

6. Referencing: D&D reject our recommendation to be more open to referencing relevant previous and current literature on the basis that we ‘refer to only one instance to back up’ the lack/deficit of citations. It is readily apparent that the paragraph in question (in Question 6) gives three different areas, where this deficit of citations occurs, rather than one example, as claimed. More examples of this deficit are given in the rest of our Review.

References
