Response to Comment on “Evolutionary Trade-Offs, Pareto Optimality, and the Geometry of Phenotype Space”

Oren Shoval,1 Hila Shetel,1 Guy Shinar,1 Yuval Hart,1 Omer Ramote,1 Avi Mayo,1 Erez Dekel,1 Kathryn Kavanagh,2 Uri Alon1*2

Edelaar raises concerns about the way we tested our theory. Our mathematical theorem predicts that despite the high dimensionality of trait space, trade-offs between tasks lead to phenotypes in low-dimensional regions in trait space, such as lines and triangles. We address Edelaar’s questions with statistical tests that eliminate pseudoreplication concerns, finding that our predictions remain convincingly supported.

In Shoval et al. (1), we presented a theory in which trade-off between tasks can lead natural selection to phenotypes that lie on simple shapes in the space of traits (morphospace). Two and three tasks lead to best-compromise phenotypes that fall on a line and triangle, respectively. Good agreement was demonstrated in three realms: *Escherichia coli* gene expression, insect castes, and cross-species morphology.

Edelaar (2) calls for improving the statistical test we used to compare the data to a triangle, for the case of cross-species morphology, focusing on Darwin’s finches (3). The improvement suggested is controlling for possible dependence of data points due to phylogenetic relationships. In the other examples (gene expression and insect castes), this concern is not relevant because comparisons are within the same species or population.

Here, we present improved statistical tests that stringently take into account possible phylogenetic dependence between data points, finding that the predictions remain strongly supported.

(i) Darwin’s finches’ triangle remains highly significant when reducing the data to six points representing species averages over islands. In (1), we used data for each island, because we reasoned that phylogenetic effects are small: Islands were colonized more than a million years ago, whereas the traits we consider show large adaptations to island climate within years to decades (3, 4); indeed, morphology is uncorrelated with distance between islands (5).

To stringently control for possible dependence between the same species on different islands, we take Edelaar’s suggestion and study a reduced data set of six points: the five traits of each of the six species averaged over all islands. Projecting these six points onto the plane (the two first principal components (2PCs)) yields a perfect triangle, in the sense that the convex hull is equal to the minimal area triangle. Using the test of (1), the *P* value is *P* = 0.09, because 9% of randomized six-point data sets (randomizing the projected coordinates of the points on the 2PC plane) also yield a perfect triangle. With six points on a plane, there is not enough statistical power to reject the null hypothesis.

However, randomizing the positions of the points along a plane does not make use of the full power of the data. Five traits were measured, so the points lie in a five-dimensional (5D) morphospace (3). Because a triangle is on a plane, the data should fall on a plane in this 5D morphospace and, on this plane, on a triangle. To test this, we used the 5 data 107 times and calculated for each randomized set how planar it is by computing the total variance on the first 2PCs of that data set (each data set thus had its own 2PCs). We then evaluated planarity by projecting on the plane defined by that data set’s 2PCs. The measured data has 99.6% of the variance on its 2PCs. Only *P*plane = 8 × 10–7 of the randomized sets had equal or smaller variance. Among these, only *P*triangle = 0.1 had equal or better triangularity than the original data when projected on the plane. Thus, planarity + triangularity has *P* = *P*plane × *P*triangle ≈ 10–7.

We repeated this also by normalizing the 5D data by a proxy for organism size, dividing each trait by the wing or tarsus length. When randomizing in the remaining four dimensions, we find *P* = 7 × 10–7 and *P* = 4 × 10–7 for normalizing by tarsus and wing. We conclude that triangularity is significant even with a conservative test for pseudoreplication.

Finally, our theory also predicts that specialists should lie on the vertices of the triangle and generalists in the middle, as occurs in Darwin’s finches. The probability of this, compared with random shuffling of generalist/specialist label in the six-point dataset, is 1 in 20, so that the *P* value is 0.05.

### Table 1. All eight ant trait pairs in (6) show statistically significant lines or triangles. The *P* values that are shown were found using our software (www.weizmann.ac.il/mcb/UriAlon/groupParetoSW.html). The triangle in panel 1 (pronotal spine length) is occupied only along two edges, a situation that can result if the trade-off in each individual is only between two of the three tasks.

<table>
<thead>
<tr>
<th>Panel number</th>
<th>Trait that the y axis corresponds to in the panel</th>
<th>Type</th>
<th><em>P</em> value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pronotal spine length ⊥ pronotal width</td>
<td>Triangle</td>
<td>1 × 10–4</td>
</tr>
<tr>
<td>2</td>
<td>Poison sac length ⊥ pronotal width</td>
<td>Triangle</td>
<td>&lt;10–4</td>
</tr>
<tr>
<td>3</td>
<td>Postpharyngeal gland width ⊥ pronotal width</td>
<td>Line</td>
<td>2 × 10–3</td>
</tr>
<tr>
<td>4</td>
<td>Labial gland width ⊥ pronotal width</td>
<td>Line</td>
<td>4 × 10–5</td>
</tr>
<tr>
<td>5</td>
<td>Hypopharyngeal gland width ⊥ pronotal width</td>
<td>Triangle</td>
<td>3 × 10–2</td>
</tr>
<tr>
<td>6</td>
<td>Metapleural gland (bulla) length ⊥ pronotal width</td>
<td>Line</td>
<td>&lt;10–5</td>
</tr>
<tr>
<td>7</td>
<td>Mandibular gland width ⊥ pronotal width</td>
<td>Line</td>
<td>4 × 10–5</td>
</tr>
<tr>
<td>8</td>
<td>Duver’s gland width ⊥ pronotal width</td>
<td>Line</td>
<td>3 × 10–4</td>
</tr>
</tbody>
</table>
values mentioned above can be divided by 20 to test the predictions of triangularity and specialists at vertices.

Similar conservative tests can be carried out for bats and other species when high-dimensional data are available.

(iii) All eight ant caste data sets of Wilson show lines or triangles, supporting our prediction. Edelaar made two additional suggestions of statistical problems in some of the other data sets we used. The first is multiple tests of the same hypothesis in the case of ant traits. Edelaar claims (without calculation) that none of the other seven trait pairs measured by Wilson (6) shows a triangle. We note that our theory predicts triangles or lines, the latter in the case of two tasks. We find, using our online software (7), that all eight ant trait pairs in the Wilson data set show statistically significant lines or triangles; thus, all support our predictions. In fact, three cases show significant triangles (Table 1). We conclude that the theory predictions are strongly supported by all eight of the trait pairs in the Wilson data set.

(iv) The bacterial gene expression data set does not have pseudo replication. Edelaar raised the possibility that the bacterial gene-expression data set may have autocorrelations that cause pseudoreplication. This data set was generated in our laboratory (7), and we carefully constructed it not to include significant autocorrelations. For example, time between samples is much longer than green fluorescent protein maturation time.

Even if only one out of every four time points is used, statistical significance of the line is still $P < 10^{-4}$.

References
8. September 2012; accepted 19 December 2012
10.1126/science.1228921