wider use. However, in ecological studies repeatability is often assessed for another reason – to quantify within-individual variation (see Boake 1989). Falconer & Mackay (1996) distinguish between ‘spatial’ and ‘temporal’ repeatability in this sense, reflecting the two forms of ‘special environmental’ variation. Thus, an analysis of variance of repeated measurements of the same individuals (taken either simultaneously from different points on the same individual, or with an appropriate time interval between sampling, respectively) is used to partition out the within-individual and among-individual variances. The ratio of the among-individual variance to the total variance provides the repeatability estimate.

In one sense, this level of repeatability analysis may seem of secondary importance compared with the obvious utility of repeatability as a statistical tool for estimating measurement error. For example, where the main aim is to compare mean PHA responses between various experimental and control groups (as in many ecotoxicological studies where experimental groups may be subjected to a particular toxin), an investigator may not be primarily concerned with establishing how consistently different individuals respond to the assay. However, in contrast, the main aim of behavioural ecological studies is usually to quantify population level variation in PHA-induced swelling, using the magnitude of the swelling as a putative correlate of ‘immunocompetence’ and hence individual quality or fitness (e.g. in relation to female choice for ‘immunocompetent’ males; Johnsen et al. 2000). In such studies it would be misleading to interpret PHA responsiveness as an index of ‘immunocompetence’ if (unquantified) within-individual variation was high enough to undermine the value of single measurements. Temporal repeatability is particularly relevant in this context, because one is usually interested in the potential of selection to shape immune function as an ecological character. Clearly, since PHA has antigenic – as well as mitogenic – properties, one cannot ignore the potential for acquired responses to confound in vivo estimates of temporal repeatability. However, this difficulty does not in itself detract from the need to demonstrate that the object parameter of the assay (T-cell proliferation following exposure to a mitogen) behaves consistently enough within individuals to undergo selection at an ecological level. In this context it should be readily apparent that spatial and/or temporal repeatabilities are of fundamental importance. Even where the main aim of a study is to compare mean PHA responses between different experimental and control groups, spatial and/or temporal repeatability still provides an important base-line from which to interpret results. Without an estimate of these parameters, one would have to assume that within-individual variation was low enough so as not to interfere with the detection of statistically significant differences. Otherwise, the likelihood of Type II statistical errors may be increased considerably (although the detection of significant differences among means would over-ride this concern).

In summary we strongly disagree with the recommendation that negative controls should be abandoned because they show consistently small effects compared to the treatment. Moreover, we recommend that spatial and temporal repeatabilities (rather than just the repeatability of the investigator’s measurement technique) for the PHA response are measured in order to set a limit to the biological significance that can be attached to the results from the assay.

References


M. T. Siva-Jothy
Department of Animal & Plant Sciences,
University of Sheffield,
Sheffield, UK
J. J. Ryder
Department of Biology,
National University of Ireland Maynooth,
Maynooth, Co. Kildare,
Ireland

Measurement repeatability and the use of controls in PHA assays: a reply to Siva-Jothy & Ryder

The validity of any scientific technique must never be taken for granted. In our recent paper (Smit, Bortolotti & Tella 1999), we questioned whether a well-established measure of T-cell proliferation in birds

could be improved methodologically. We thus welcome Siva-Jothy & Ryder’s (2001) commentary on our paper questioning whether such a previously established protocol should be changed. Their main concerns are the abandonment of the negative control, and the nature of repeatability of measurements. We feel that clarification of the terms ‘control’ and ‘repeatability’ is required, and that what we have proposed in simplifying the phytohaemagglutinin (PHA) assay is both valid and valuable for future research.

There is no question that an experiment without a control is not science. The so-called control application of injecting phosphate buffered saline (PBS) into the opposite wing is in no way part of the true requisite control for the objective of ecological or toxicological studies. In Smits et al. (1999) we showed seven of our various studies to have a treatment (PCBs, mine tailings, etc.) and hence by inference there were both treatment and control groups (e.g. no exposure to PCBs or mine tailings). The same PHA test was applied to both groups of animals. This is how we understand the scientific method. The terminology ‘control’ for PBS injection was originally devised to demonstrate that the T-cell proliferation was a response to PHA itself but not to the vehicle (PBS) (Goto et al. 1978). As such PBS injection was a valuable and true control in the initial study whose objective was to validate the PHA test. The PBS injection does not in any way therefore allow one to ‘interpret the resultant data in the context of the biological predictions that underpin the manipulation’.

Siva-Jothy & Ryder claim that without a negative control within individual birds ‘how can those practitioners (let alone reviewers and editors) have an objective assessment of (a) the practitioner’s technical ability to conduct the assay and (b) the nature of the response to the manipulation?’ Clearly, Smits et al. (1999) demonstrate that in determining the ‘nature of the response to manipulation’ it is not necessary to prove over and over again that PBS is not responsible in provoking skin swelling at the PHA injection site. Hence the continued use of the word ‘control’ in reference to the PBS injection shows a lack of understanding of how the PHA test is used as a tool in ecological studies.

In contrast to Siva-Jothy & Ryder, we fail to understand how doing both a PHA and PBS injection in any way corrects for technical problems experienced by the practitioner. Simply, two wrongs do not make a right. As Smits et al. (1999) demonstrate quite clearly, incorporating the PBS measurement only increases error in estimation of the PHA response. The negative control is also not relevant to Siva-Jothy & Ryder’s concern that people with minimal training can apply this test. Our emphasis on minimal training was made because the ability of a person to inject and measure a wing is independent of their understanding of the immune system, just as their ability to measure a wing is independent of a knowledge of aerodynamics. The only tool available for objective assessment of ‘the practitioner’s technical ability’ rests in an evaluation of the repeatability of measurements, and thus we turn to Siva-Jothy & Ryder’s second concern.

The term repeatability, as Siva-Jothy & Ryder correctly point out, can have two applications: within-individual variation as may exist in space and time, and variation attributable to measurement error. Smits et al. (1999), and most conscientious researchers, use the latter to validate their technical expertise. We reviewed 35 ecological papers published recently (1997–2001) where the PHA assay was used, and found PHA responses to have repeatabilities (r) averaging 0.94 (range: 0.70–0.99). Therefore, this objective assessment of ability shows measurement error (1–r) is typically only 6% and so the technique can, and has, been applied with an exceptional degree of reliability.

The nature of repeatability in space and time, to understand within-individual variation, was emphasized by Siva-Jothy & Ryder as being important. There are certainly questions where such information is of interest; however, they rest more within the realm of immunology, i.e. where immune function itself is the focus of the study. The PHA test can be used in ecology as a tool for answering questions of a different focus. For example, if we want to study the relationship between prelaying mass and clutch size in a bird, it would not be necessary to document mass of all body parts or mass of the bird in the future. Mass variation is a dynamic process as is immune response. While physiologists may wish to study how mass can vary among organ systems, across seasons and so on, it does not diminish the use of mass as a variable in a simpler form in cross-sectional studies typical of ecology or toxicology.

We could envision how temporal variation in immune function could provide valuable information regarding long-term constancy of immune response, which could be linked to a genetic component. Genetic and environmental components to the T-cell mediated immune response have been well evaluated using cross-fostering experiments (Brinkhof et al. 1999; Christie et al. 2000; Tella et al. 2000a). Alternatively, temporal repeatabilities offer a way of estimating maximum heritability, which in turn fail to disentangle genetic from environmental factors (Falconer & Mackay 1996). Moreover, attempts to quantify such variation have at least three potentially serious problems.

First, there are many confounding variables associated with longitudinal research. Many studies are conducted on nestlings, which show PHA responses different from adults (Martin et al. 2001; J. E. Smits & G. R. Bortolotti, unpublished data). Therefore, ontogenetic effects complicate interpretation of
temporal variation. Other variables that would have to be controlled for include seasonal and or age-related patterns in the production of immune-modulating hormones, body condition, and a variety of stressors including temperature, food supply, diet, parasites and social conditions (see, e.g. Dohms & Metz 1991; Lloyd 1995; Duffy et al. 2000; Alonso-Alvarez & Tella 2001; Tella et al. 2001). There is also the potential for additional artefacts caused by differential stress/habituation caused by repeated handling of the same birds given the immunosuppressive nature of corticosterone. To account for all of these would be a formidable task even in captive birds.

Second, the use of temporally separated tests on the same individuals must consider the artefacts created by injections or other modification of the skin surface (e.g. feather plucking at the injection site). Insertion of the needle plus a space-occupying fluid into the thin and delicate skin of a bird causes some disruption to the subepithelial and hypodermal tissue, and resulting stimulation of fibroblasts leads to changes in the matrix of the tissue and hence one's ability subsequently to measure the PHA response without bias.

Third, the immune system has a memory. The whole idea of using a plant enzyme is that it will be a novel mitogen. Repeated use of the same compound thus invalidates the independence of test points given PHA is no longer a novel antigen and one must account for adaptive aspects of subsequent immunological responses. Responses to second injections of PHA have been larger than the first (Johnsen et al. 2000; Alonso-Alvarez & Tella 2001). While Siva-Jothy & Ryder recognize this as a problem, they believe it to be of secondary importance to the need to demonstrate that individuals respond in a consistent manner. We believe their rationale is flawed.

Siva-Jothy & Ryder propose temporal stability of the PHA response as being necessary for studies in behavioural ecology, but not ecotoxicology. We see no scientific justification for their position related to discipline. Furthermore, we disagree that stability is relevant to 'the potential of selection to shape immune function as an ecological character' or how this relates to setting a 'limit to the biological significance' of the results. It appears that the authors do not recognize that natural selection can operate on traits subject to temporal variation. To give an example of mass dynamics again, it is well documented that selection acts on fledging body mass. Heavy fledglings survive the best; information on daily, seasonal or annual fluctuations in mass after that period is irrelevant to that fact. It would not be a surprise in fact if heavy nestlings remained heavy throughout life, just as immune response measured at a single point in time has been correlated (Saino et al. Saino et al. 1997; González et al. 1999; Soler et al. 1999) or implicated (Tella et al. 2000b) with survival in the future.

Through these arguments, we continue to support our modification of the PHA skin test as a valid research tool in broadly based studies in ecology and ecotoxicology. The simplified method has many advantages both in the application of the test and the interpretation of the results. This test has a valuable role as one of a suite of applications, as it is recognized as measuring one early aspect of the immune response potential.

References


---