THE ECOLOGY OF ANTIOXIDANTS & OXIDATIVE STRESS IN ANIMALS
How to measure oxidative stress in an ecological context: methodological and statistical issues

Peeter Hõrak[1] and Alan Cohen[2],*

[1]*Institute of Ecology and Earth Sciences, Tartu University, Vanemuise 46, 51014 Tartu, Estonia; and* [2]*Groupe de recherche PRIMUS, Centre Hospitalier Universitaire de Sherbrooke, 3001 12*e *Ave Nord, Sherbrooke, Quebec J1H 5N4, Canada*

Summary

1. Reactive oxygen and nitrogen species can damage biomolecules if these lack sufficient antioxidant protection. Maintaining and up-regulating antioxidant defenses and repair of the damaged molecules require resources that could potentially be allocated to other functions, including life-history and signal traits.

2. Identifying the physiological mechanisms causing and counteracting oxidative damage may help to understand evolution of oxidative balance systems from molecular to macroevolutionary levels. This review addresses methodological and statistical problems of measuring and interpreting biomarkers of oxidative stress or damage.

3. A major methodological problem is distinguishing between controlled and uncontrolled processes that can lead either to shifts in dynamic balance of redox potential or cause pathological damage. An ultimate solution to this problem requires establishing links between biomarkers of antioxidant defenses and oxidative damage and components of fitness.

4. Biomarkers of redox balance must correspond to strict technical criteria, most importantly to validated measurement technology. Validation criteria include intrinsic qualities such as specificity, sensitivity, assessment of measurement precision, and knowledge of confounding and modifying factors.

5. The complexity of oxidative balance systems requires that assay choice be informed by statistical analyses incorporating context at biochemical, ecological and evolutionary levels. We review proper application of statistical methods, such as principal components analysis and structural equation modelling, that should help to account for these contexts and isolate the variation of interest across multiple biomarkers simultaneously.

Key-words: antioxidants, complex systems, oxidative damage, principal component analysis, technical validation, multi-level models

Introduction

Life-history theory relies on the concept of trade-offs, where a change in one trait or function to improve fitness necessarily has negative consequences for other traits or functions such that they reduce fitness (Stearns 1992; Zera & Harshman 2001). Micro- and macroevolutionary trade-offs at population, species, or lineage levels lead to continua of life-history and physiological strategies. Within individuals (i.e. at tissue, cellular and molecular levels) these trade-offs can be observed in the controlled down-regulation of one function to give priority to another.

Examples include down-regulation of immune responsiveness during reproductive season, pregnancy and lactation (Martin, Weil & Nelson 2008) and antagonistic interactions between Th1-cell-mediated inflammatory responses and Th2-cell-mediated humoral immune responses (Mosmann & Sad 1996). Alternatively, physiological trade-offs can be revealed in somatic damage resulting from uncontrolled and/or unbalanced activation of some physiological process. These pathological effects have negative impacts on individual health and may impinge on fitness by reducing mating, reproductive success, or survival. Examples include immunopathological side-effects of the inflammatory response (e.g. sepsis) or chronic elevation of baseline corticosterone.

*Correspondence author. E-mail: alan.cohen@usherbrooke.ca

Measurement of oxidative stress in ecology **961**

As indicated by the collection of the papers in the *Functional Ecology* special feature 'The Ecology of Antioxidants & Oxidative Stress in Animals', oxidative stress (OS) is currently believed to be involved in life-history trade-offs at the levels of individuals, populations and lineages (see also Von Schantz *et al.* 1999; Costantini 2008; Dowling & Simmons 2009; Monaghan, Metcalfe & Torres 2009). Understanding the physiological mechanisms causing oxidative damage may help to understand evolution of resource allocation patterns and *vice versa*. Building off of the conceptual challenges highlighted at the end of Costantini *et al.* 2010 in the *Functional Ecology* special feature, this article will tackle methodological and statistical hurdles that confront scientists aiming to quantify OS or damage. We start by outlining the problems in distinguishing between controlled and uncontrolled (i.e. regulated and unregulated) processes which can lead either to shifts in dynamic balance of redox potential or cause pathological damage. We continue by explaining how problems with interpreting different biomarker values can be addressed by manipulation of different components of the antioxidant system and outline potential lines of research for establishing the ultimate links between markers of oxidative damage and fitness. We next review the criteria for biomarkers of OS, with a special emphasis on applicability for field studies. Finally, we review proper application of statistical methods, such as principal components analysis and structural equation modelling, that should help isolate the variation of interest across multiple biomarkers simultaneously. We include a detailed hypothetical example of how our recommendations and approach might be applied to an actual study, from design through analysis. Throughout, our background as avian ecologists may occasionally be more apparent than we would wish, but the principles we describe should be equally applicable at least across vertebrates and probably more generally. Moreover, the statistical methods we outline can be applied equally well to other aspects of physiology such as hormones and the immune system.

Distinguishing damage from shifting balance

A major task in assessing the role of OS in physiological trade-offs is to distinguish whether observed patterns are due to controlled (regulated) or uncontrolled (unregulated) processes, not least because the predictions about the directions of induced changes can be opposite. Controlled processes result in physiological changes that are regulated in ways that tend to be advantageous to the organism, given its circumstances. Uncontrolled processes occur when the regulatory system is overwhelmed or doesn't exist, and are thus often not advantageous to the organism. For instance, one might predict reduced values of biomarkers of antioxidant protection and increased oxidative damage in response to increased mitochondrial reactive oxygen species (ROS) production due to increased work load (e.g. Costantini, Dell'ariccia & Lipp 2008) or exposure to cold (e.g. Selman *et al.* 2008). On the other hand, regular exercise has a positive impact on redox balance, which has been ascribed to the up-regulation of anti-

oxidant and DNA repair enzymes (e.g. Radak *et al.* 2008); also, hypothermia can attenuate OS (Stefanutti *et al.* 2005). Similarly, preconditioning by induction of mild OS can lead to suppression of oxidative brain damage in ischaemia/reperfusion, presumably because this 'trains' the tissue to expect or withstand OS (Wang *et al.* 2007). Another example concerns high levels of plasma total antioxidant capacity (TAC, also known as TAS, total antioxidant status, or TEAC, Troloxequivalent antioxidant capacity), which have sometimes been interpreted as a sign of beneficial redox state (Hõrak *et al.* 2006; Tummeleht *et al.* 2006; Saino *et al.* 2008; Geens, Dauwe & Eens 2009). However, several experiments have demonstrated up-regulation of TAC levels in response to oxidative challenge induced by either exhaustive exercise (Vider *et al.* 2001), immune activation (Hõrak *et al.* 2007), or stress resulting from exposure to cold (Cohen, Hau & Wikelski 2008a) or restraint (Cohen, Klasing & Ricklefs 2007). Furthermore, in humans, high TAC values can even associate positively with severity of illness and mortality (Chuang *et al.* 2006). These findings suggest that high values of antioxidant capacity can reflect adaptive and compensatory responses to oxidative (or physiological) stress rather than optimal health condition. Such problems are inherent to feedback-based homeostatic systems, so caution is needed in interpreting the changes in pro-oxidant/antioxidant balance.

Variation in antioxidant levels and biomarkers of OS or damage within individuals may reflect three types of phenomena: (i) direct regulation as a response to changing conditions; (ii) side-effects of other regulated processes (e.g. uric acid may indicate protein metabolism more than oxidative status; Hollmén *et al.* 2001); or (iii) pathological processes such as damage accrual. It is critical to know which of these is being measured in order to properly interpret results. Often, changes in markers rather than absolute levels *per se* can be informative: good long-term predictors of fitness may be derived from the responses of short-term measures to changes in conditions. An example is a positive association between stressor-induced corticosterone secretion and survival in American redstarts (Setophaga ruticilla, Angelier, Holberton & Marra 2009). Biomarkers of OS too change in response to stressors and oxidative insults, but it is not always clear whether and which of these changes are regulated (e.g. Cohen, Hau & Wikelski 2008a). Establishing links between such changes and components of fitness would be of utmost importance for understanding the function of such biomarkers. Unless such links have been established in a given species and context, caution is required in interpreting the changes in biomarkers of redox balance as these do not necessarily indicate occurrence of oxidative stress.

The ultimate link to fitness

Although it is widely believed that reactive species are involved in many diseases (Halliwell & Gutteridge 2007), the question of whether and how increased OS impinges on fitness in the wild is difficult to answer. The medical approach is straightforward: if oxidative damage contributes significantly

to disease pathology then actions that decrease it should be therapeutically beneficial (Halliwell & Whiteman 2004). Application of such an approach to ecological settings, however, immediately results in problems of measuring pathology, or more generally, components of fitness such as morbidity, mortality, mating or reproductive success. For instance, factorial experiments manipulating oxidative status and antioxidant levels and measuring response in the biomarkers of oxidative damage can successfully reveal shifts in redox balance. However, such experiments do not enable us to establish the biological impact of increased OS unless the link between OS and some fitness component is shown.

So far, the links between oxidative damage and fitness have been demonstrated in only a few ecological studies, reviewed in two papers from the *Functional Ecology* special feature (Costantini *et al.* 2010; Metcalfe & Alonso-Álvarez 2010). We therefore emphasize the urgent need for further research in this area. Studies monitoring individuals over their lifetime would be ideal and feasible in short-living species like passerine birds, lizards or voles. Although cross-sectional studies have their limitations as discussed by Nussey *et al.* (2009), they still appear valuable for obtaining information about the role and regulation of OS in long-lived animals, which may largely differ from that of short-lived ones (e.g. Cohen *et al.* 2008b). A potentially useful but as yet unexplored possibility would be measurement of biomarkers of oxidative damage in the age classes with highest mortality, i.e. juveniles such as nestling birds. In such cases, direct observations of mortality in the nest enable us to distinguish it from dispersal and no effort is required for capture and recapture. An example of how such an approach enables us to establish a link between haematological parameters and nestling mortality is provided by Nadolski *et al.* (2006). Yet another valuable source of information is provided by ecotoxicological studies in environmental pollution gradients (e.g. Isaksson *et al.* 2005; Berglund *et al.* 2007) or intoxication experiments (e.g. Kenow *et al.* 2008; Ansari *et al.* 2009), which enable us to interpret the variation in the markers of OS in association with pathological processes.

Another promising research perspective is the study of the effects of OS on ornamental traits. In species where the connection between ornament elaboration and mating success is firmly established, detection of the effect of experimentally induced OS on ornaments is an indirect proof for the impact of OS on fitness. The link between sexually selected traits and OS was first outlined in a landmark paper by Von Schantz *et al.* (1999) and the experimental evidence, concerning mainly birds and fish, is accumulating [see reviews by Dowling & Simmons (2009) and Monaghan, Metcalfe & Torres (2009) and papers from this special feature Costantini *et al.* (2010) and Metcalfe & Alonso-Álvarez (2010)]. We stress the particular value of studies which apply well-established chemical treatment protocols for induction of OS, such as administration of pro-oxidants like paraquat (Isaksson & Andersson 2008), diquat (Galván & Alonso-Alvarez 2009), or hemin (Seaman *et al.* 2008), because it warrants that resulting effects on phenotype were genuinely caused by OS rather than

immune activation or increasing metabolic costs. A disadvantage of such methods is the obvious potential harm to individuals treated: strong justification and proper ethical approval are necessary. Suppression of the systemic levels of glutathione, an important endogenous antioxidant, is promising (and technically feasible) and has been shown to increase both melanin pigmentation and plasma TAC levels (Galván & Alonso-Alvarez 2008) and lipid peroxidation (Hõrak *et al.* 2010). In these experiments, birds were treated with buthionine sulfoximine, a synthetic amino acid which specifically inhibits glutathione synthesis without other pathological side-effects. Such treatment is perhaps more humane to induce oxidative stress than administration of strong pro-oxidants like paraquat. A review of methods used to induce OS in biomedical models is provided by Knasmüller *et al.* (2008).

Antiradical effects of dietary or other treatments can also be assessed by stimulation of the oxidative burst in whole blood samples or phagocytic cells *ex vivo* with substances inducing endogenous ROS production (e.g. He *et al.* 2006; Sureda *et al.* 2007; Olsson *et al.* 2009) [but see Knasmüller *et al.* (2008) for a critique on using cell cultures]. Manipulation of dietary antioxidants can provide information about their function (Pike *et al.* 2007; Catoni, Peters & Schaefer 2008). To summarize, measuring associations between biomarkers of OS and fitness is of ultimate importance. In addition to ecological field studies, valuable information for interpreting the variation in these biomarkers can be obtained from ecotoxicological studies and experimental induction or suppression of ROS production or synthesis of antioxidants.

Criteria for biomarkers

There are a number of excellent reviews about biomarkers to estimate ROS, antioxidant defenses, oxidative damage and repair mechanisms (e.g. Dalle-Donne *et al.* 2006; Halliwell & Gutteridge 2007; Hermans *et al.* 2007; Knasmüller *et al.* 2008; Monaghan, Metcalfe & Torres 2009). Instead of reproducing this material we will concentrate on the general methodological problems relating to work with such biomarkers and outline some specific issues of concern in ecological studies.

In biomedical research it is considered fundamental to demonstrate that changes in biomarkers do reflect the later development of disease. To that end, ecological approaches to OS research may have some advantages due to access to a wider range of model organisms and experimental settings where associations between various biomarkers and fitness components can be tracked (see above). However, picking the right model organism and relevant setup is not enough, because biomarkers must also correspond to rigorous technical criteria, equally important for ecological and biomedical research. Below we briefly review the most relevant such criteria as outlined by Dalle-Donne *et al.* (2006) and Halliwell & Gutteridge (2007).

1. The biomarker should detect a major part of oxidative damage on the target molecule *in vivo*. Measurements of

peroxidizability of a substrate *in vitro* under artificially created oxidative stress do not necessarily indicate that the same substrate is oxidizing faster *in vivo*. Halliwell & Gutteridge (2007) present the analogy that if building A can resist explosion less well than building B, this does not mean that building A is falling down. However, popular and feasible spectrophotometric assays of antioxidant capacity or total antioxidant status predominantly relay on such *in vitro* techniques.

2. The biomarker should be stable, not being lost or formed artefactually in stored samples. The problem of artefactual formation is known for colorimetric determination of MDA by thiobarbituric acid reactive substances (TBA or TBARS assay). Although being repeatedly and severely criticized for a long time (e.g. Halliwell & Gutteridge 2007), this assay is still popular, not least due to its simplicity, including for ecological research. An example of a biomarker quite sensitive to collection and assay conditions is glutathione disulfide, GSSG (e.g. Rossi *et al.* 2002); yet good repeatabilities have been obtained even in samples collected in an ecological setting (Alonso-Alvarez *et al.* 2010).

3. The biomarker must employ validated measurement technology. Validation criteria include intrinsic qualities such as specificity, sensitivity, assessment of measurement precision, and knowledge of the confounding and modifying factors. This criterion is of utmost importance when applying commercial biomedical assays or other methods elaborated in humans to different species. For instance, measurement of an important antioxidant, albumin, is not possible via colorimetric test (bromcresol green) in avian blood (Harr 2002), yet this assay is persistently used in avian studies. Some methods require modification or adjustment; it is mandatory that all such modifications of standard procedures are reported. Field samples also need to be tested for potential confounding effects of time of trapping and sampling, temperature, season, stage of reproductive cycle, sex and the state of the individual. Non-automated measurement techniques, such as the COMET assay for DNA damage, also require control for different persons performing the analyses. In the presence of consistent covariation with such confounding variables, the problem can sometimes be controlled by statistical elimination of interference (see later). The ultimate question of the usefulness of each particular assay depends on measurement precision. Assessment of the repeatability on the basis of intraclass correlation coefficients (Lessells & Boag 1987) is ideal because it accounts for measurement precision in the biological context, i.e. with respect to between-individual variation. This requires replicating an assay on multiple aliquots per individual for at least a subset of individuals. (In addition to assessment of measurement precision, repeatabilities can also be used to characterize the stability of biomarkers within individuals over time as discussed below; these two applications must not be

confused). If it is not possible to subdivide individual samples into multiple aliquots then pooling the leftover samples from many individuals and assessing the CV of repeated measurement of that mixture would constitute an alternative for presenting the repeatability. Insufficient description of assay procedures and measurement precision are widespread and can become a bad practice in ecological literature unless editors and reviewers impose strict demands on authors. Poor measurement precision constitutes an especially serious problem when inferences are made on the basis of lack of association between studied variables. Equally important is following the guidelines of reliable and unbiased sample collection and statistical treatment (Ruxton & Colegrave 2006; Biro & Dingemanse 2009) and good laboratory practice.

4. Collection of samples should interfere minimally with normal life activities of the studied organism. For instance, in the case of birds analyses should be performed on blood quantities not exceeding 2·5% of body mass [However, even smaller amounts may increase mortality of passerines in the wild (Brown & Brown 2009)]. Operating with small amounts of blood means that not all standard protocols and kits developed for humans are mechanically applicable for small animal research. However, in many cases, such kits can be still used by applying appropriate dilutions or reductions in sample volumes. To establish the minimum required sample volume, running a pilot study to establish the range of biomarker values in a given population before planning massive sampling would also be helpful. Emphasis on non-invasive methods naturally does not deny the value of terminal experiments for assessment of oxidative damage in different organs and tissues post mortem. However, such experiments too should aim at finding systemic markers of damage, i.e. those that can be determined from blood samples and correlate with local damage.

5. The biomarker must not be confounded by diet (unless the effects of diet *per se* are the focus of interest). This problem has been reported for some of the most potent markers of oxidative damage, the plasma MDA and HNE concentrations (e.g. Dalle-Donne *et al.* 2006). Another popular measure, plasma total antioxidant capacity (TAC/-TAS/TEAC), correlates very strongly with plasma uric acid levels, which could be an indication of incidental amino acid catabolism rather than regulated antioxidant protection (Cohen, Klasing & Ricklefs 2007). In laboratory settings, when animals are kept on a uniform diet and blood can be drawn after nocturnal fast, problems of dietary interference (as well as diurnal variation) can be largely ignored. On the other hand, captive diet, *ad libitum* feeding and lack of predators can systematically affect the values of some (but not all) biomarkers (Sepp, Sild & Hõrak 2010).

6. The biomarkers should preferably be consistent within individuals over time, or should be intended as markers of transient condition (e.g. response to a stressor). Different

biomarkers measure different aspects of redox balance and thus also reflect physiological disorders that vary in duration. Such variation is also best described by calculating repeatabilities (like in the case for assessment of measurement precision) but in this case the same individuals are repeatedly sampled over time. Good long-term measures that reflect steady differences between individuals show high repeatabilities over longer time periods. Measures that change in response to oxidative insults are good short-term indicators of condition. Yet other parameters fluctuate uninformatively, in which case they should not be used. In the ecological literature, the data about individual consistency of different biomarkers of OS have seldom been reported (but see Norte *et al.* 2008 about antioxidant enzymes; and Galván & Alonso-Alvarez 2009 for total GSH, TBARS and TAS). Clearly, comparison of long- and short-term repeatabilities of different biomarkers would help to compose an optimal battery of assays for testing experimental outcomes. The same information is also valuable for understanding the function and diagnostic value of these traits. Bell, Hankison & Laskowski (2009) provide a good example about addressing similar problems in measuring behaviour and Garamszegi *et al.* (2009) introduce statistical remedies for handling poor quality data and missing observations.

As noted by Halliwell & Gutteridge (2007), no currently used biomarker of oxidative damage meets all these technical criteria, but some are better than others. A combination of different markers is thus always preferable to measuring a single response variable (see also below regarding the measurement of complex systems). The challenge for both ecologists and medical biochemists is then to sort out the optimal combinations of such markers, keeping in mind the above-listed technical requirements and a core criterion of a link to pathology or fitness components. A major way to proceed in such a search is to register the response of different markers to controlled oxidative insults *in vivo* and/or antioxidant administration. Such experiments enable extracting the parameters responding most sensitively (see e.g. Rossi *et al.* 2006; Galván & Alonso-Alvarez 2009), and having thus the highest diagnostic values. For example, it would be informative to find out whether and how other measures of oxidative damage (such as blood levels of lipid peroxidation products or protein carbonyls) correlate with resistance of erythrocytes to OS *in vitro*. Emergence of such correlations likely depends on the variation of general health state of the animals studied. For instance, an extensive meta-analysis of biomedical literature indicates that only under the severe pathological conditions all the indices of OS correlated with each other (Dotan, Lichtenberg & Pinchuk 2004). Finally, there is a need for studies comparing and verifying simple assays by more sophisticated analytical methods (e.g. McGraw, Tourville & Butler 2008). For instance, such an approach is essential for clarifying the potential utility of the colorimetric TBARS assay, which has been criticized because of non-specificity and ar-

tefactual generation of TBARS (Halliwell & Gutteridge 2007). To summarize, ecological research on biomarkers of OS requires more focus on validation of measurement technology and detection of confounding and modifying factors.

The problem of measuring complex systems

The appropriate assays to measure oxidative balance cannot be determined by biochemical considerations alone – the complexity of oxidative balance systems requires that biochemical assay choice be informed by statistical analyses incorporating context. Each assay is a proxy for a biochemical process, and our hope is that by measuring these proxies, we can gain an understanding of how the system works. However, because the biochemical processes are part of a system rather than independent, the interpretation of any given measure is highly dependent on what is happening in the rest of the system. For example, it is now well known that vitamin E can become pro-oxidant rather than antioxidant in certain biochemical environments; the same is likely true for many other antioxidants, including carotenoids (Neuzil & Stocker 1994; Surai 2002). This is an example of biochemical context, just one of several types of context that are important when considering measures of oxidative balance. By 'context' here, we are not simply referring to co-variates or confounding variables, but to factors that alter the interpretation of oxidative balance markers with respect to health. For example, a high vitamin E level is likely to mean something different for health/fitness in (i) individuals with high or low carotenoid levels; (ii) measurements taken before or after an acute stressor; and (iii) species with high versus low vitamin E contents in the diet, etc.

Broadly speaking, interpretation of measures can depend on ecological context (factors such as external stressors, food source availability, season and phenotypic plasticity of the internal biochemical environment) and evolutionary context (factors such as resistance to or tolerance for oxidative damage, life history strategy and evolution of biochemical physiology in response to changes in species diet). Ecological and evolutionary context are of course mediated by biochemical context, but often through mechanisms that are still poorly understood, and often in ways that are at least partially predictable based on ecological or evolutionary variables. For example, many bird species have been observed to show marked changes in antioxidant levels in response to capture stress (Cohen, Klasing & Ricklefs 2007). Mediation of this stress response appears not to be completely attributable to changes in corticosterone level; the precise mechanism remains unknown (Cohen, Hau & Wikelski 2008a). Thus, if we want to incorporate the effects of stress level into an analysis of oxidative balance, we should include organism-level measures of stress rather than a biochemical covariate such as corticosterone.

The three types of context mentioned so far – biochemical, ecological and evolutionary – can conveniently be distinguished by how they are measured. We can consider the

effects of biochemical context as the co-variance among various biochemical assays across measurements or individuals. The ecological context can be considered the way that interpretation of biochemical measures depends on non-biochemical factors measurable at the individual level (stress level, season, recent diet, sex, etc.). Lastly, the evolutionary context is the difference across species or lineages in any of the lower-level associations.

The structure outlined so far indicates the potential for substantial complexity. The interpretation of any given biochemical assay may depend simultaneously on (i) the levels of other biochemical parameters; (ii) the immediate condition of the individual; (iii) the history and genetic factors particular to the individual; (iv) season and other environmental variables likely to affect populations; and (v) biochemical and physiological characteristics of the species. Indeed, there are many examples in the literature of all such effects (e.g. Neuzil & Stocker 1994; Tella *et al.* 2004; Costantini & Dell'omo 2006; Costantini 2008; Cohen, McGraw & Robinson 2009). The problem we face in trying to measure oxidative balance is that we want measures that allow a clear interpretation and comparability across whatever groups we are comparing – for example, we don't want a measure that shows oxidative stress during breeding season but indicates nothing of importance at other times. The less generally applicable our measures, the harder it is to interpret our results both within the context of the given study and externally.

In many cases, it may not be possible to compare the same assays across disparate groups. For example, β-carotene levels tend to be much higher in songbirds than in many other avian lineages (Cohen & McGraw 2009). Even though β-carotene can be measured easily in birds, and furthermore may be important for health in many or all species of various lineages, it is not possible to have a uniform interpretation of a given level across species. Similarly, it appears that naked mole-rats have evolved special adaptations for tolerating extremely high levels of oxidative damage (and living extraordinarily long lifespans in spite of sustaining substantial damage) (Blazej *et al.* 2006). A comparison of oxidative damage across species would thus need to incorporate a consideration of tolerance.

From a measurement perspective, the amount of work that needs to be done to validate any given assay can be problematic. Even if we are not trying to make interspecific comparisons, the interpretation of an assay within a study species could be different than for published accounts in other species, but validation itself is too labour-intensive to replicate for all species. For example, if we did not know that naked mole-rats have an apparently high tolerance for oxidative damage, we could easily misinterpret a study showing differences in production of antioxidant enzymes across age-classes. Often, there will be no good solution to this problem. However, the combination of careful thought and application of statistical tools can often be used to improve our ability to characterize complex systems such as oxidative balance. In the section below we outline several such approaches. It will usually be the case that the question of how to measure a

system is nearly identical to the question of how to understand it; luckily, this means that methodological research will also be substantive research.

Statistical methods for measuring and understanding oxidative balance systems

A GENERAL FRAMEWORK

Most or all studies of oxidative balance in ecological context to date (including our own) have had at least some shortcomings in their characterization of oxidative balance systems; this was inevitable given the state of knowledge at the time, and it is only with hindsight that we are able to suggest new approaches. Often, a single measure is used without a complete understanding of what this measure means for oxidative balance more broadly (e.g. Hõrak *et al.* 2006). In other situations, multiple measures are used, but results can be conflicting and there is no clear understanding of how the measures relate to each other (e.g. Cohen *et al.* 2008b). Often the measures are considered completely independently, and attempts to understand interrelationships have generally been only partially successful (e.g. Tummeleht *et al.* 2006; Cohen & McGraw 2009). At an ecological level, few studies consider time-of-day, diet, or seasonality (but see Cohen, McGraw & Robinson 2009; Filho *et al.* 2001). To our knowledge, no study has successfully accounted for all of the above problems. These problems mean that there is rarely if ever a single clear interpretation for results. Many researchers are aware of the problems and acknowledge limitations in their studies, but there has been no clear way to address them. This section attempts to provide some initial guidance, and will undoubtedly be improved upon by others in the future. If the discussion seems abstract at points, readers may wish to refer to the appropriate sections of Appendix S1, which provides a hypothetical example of how these principles might be applied.

As noted above, there are several different levels at which context matters for our interpretation of biochemical assays of oxidative balance. From a statistical perspective, we recommend an approach that allows both assessment of whether context can be adequately incorporated into models and, if so, proper methodology for doing so. At the biochemical level, this is achieved by variable simplification methods (such as principal components analysis, described later) which identify the best set of measures to use to identify underlying processes. Once this set of variables is identified, it should be checked that the interpretation of the variables is stable across the variation to be represented in the study (ecological or evolutionary); if so, multi-level models can be applied to parse out the hierarchical effects. We provide a worked hypothetical example in Appendix S1 to illustrate.

MEASURING A SUITE OF BIOCHEMICAL VARIABLES

It is by now well established that interesting variation in oxidative balance often cannot be sufficiently represented by a single variable (Costantini *et al.* 2007; Cohen, McGraw &

Robinson 2009). Levels of different antioxidants may not correlate, and levels of damage may vary across tissues or target molecules (e.g. Lopez-Torres *et al.* 1993). However, straight measurement of an extensive suite of variables also has its drawbacks. Many of the measurements may be largely redundant even if some are not, and examining many graphs of many assays is much more cumbersome than examining one, especially if the many graphs show differing trends and their relative interpretations are not well understood. Our goal then should be to understand (i) what the most relevant measures are and (ii) how to represent them as succinctly as possible. The preferred tools for this are principal components analysis (PCA) and factor analysis (FA), statistical methods for understanding the relationships among a suite of variables. Detailed mathematical description of these methods can be found elsewhere (e.g. Johnson & Wichern 2007); a brief introduction and suggestions for implementation are provided in Appendix S2.

From an oxidative balance perspective, there are several key things we can hope to gain from PCA or FA. First, we can identify the dimensionality of the system. Each orthogonal axis that explains substantial variation can be interpreted as representing an independent aspect of the system, and the number of such axes is thus the dimensionality. Second, the loadings of those axes can be used to understand the nature of each dimension of the system. Third, the axes generated can then be used as new variables that better represent the underlying processes than any of the original biomarkers themselves. A careful PCA or FA analysis can thus provide the basis for both understanding and measuring a complex system. However, the results can depend to some extent on which variables are included in the analysis, and in some cases the results may not be stable [for example, as sample size increases (Cohen & McGraw 2009)]. Care must be exercised to avoid spurious results in these cases.

GENERALIZABILITY OF MEASURES

Once a set of measures has been established by variable simplification (or just using the raw measures), it is important to establish how general the results are. The results may apply only at a particular level of analysis (e.g. individual or species), or for the particular subset of individuals or species used in the analysis. In Appendix S2 we provide the outline of a method that can be used to test the generalizability of summary measures generated from PCA or FA. No such approach can be readily applied to raw variables. An analysis of generalizability can aid other researchers in interpreting the results even if the measures are not intended for use outside the study in question.

Often a set of axes generated by PCA or FA will not be generalizable across the contexts of interest for a given study. This may be problematic from the perspective of the original research question, and serious consideration must be given to whether that question can truly be asked and answered, given the data. However, the lack of generalizability and the particulars of the differences can themselves be a substantive result.

For instance, in an analysis of micronutrient antioxidant levels in 78 bird species, the full set of birds and one of four phylogenetic subgroups show an axis with strong positive loadings for all four carotenoids; however, in the other three subgroups, no such axis is apparent, and there is instead an axis with strong positive loadings for vitamin E, zeaxanthin, and β-cryptoxanthin. Since diet varies greatly within each of the four groups, this result suggests that physiology has evolved to use different combinations of lipid-soluble micro-molecular antioxidants in different lineages, or to use the same molecules in different ways (Cohen, McGraw & Robinson 2009).

In many situations, there may be ways to account for different axis configurations across subgroups and still answer the original research questions. This will be particularly true when the axes are generally stable except with respect to a single type of variation, such as sex or season. In this situation, the data should be stratified into the relevant groups, and, if possible, separate axes should be generated for each group, explaining the variation relevant to it. In this case, the study may actually benefit from the axis instability because the original question can be addressed and substantive information on differences between the groups can be explored.

WHEN MEASURES CAN BE GENERALIZED: MULTI-LEVEL MODELS

In cases where a set of summary variables is found to be generalizable across the groups being studied, they can for the most part be treated as any other variables. As noted above, it will often be the case that the relevant summary variables will differ across levels of analysis, such as within versus among species. However, in some cases the same axes may be applicable across hierarchical levels; in such cases, multi-level modelling will often be a preferred strategy for analysis (Rabe-Hesketh & Skrondal 2008).

Multi-level models are a subset of random-effects or mixed-effects models in which there is a clear nested structure to the data. A familiar example to many ecologists will be analysis of offspring within broods. Multi-level models are often used by ecologists to control for, say, brood effects when the outcome of interest is on the nestlings, and this is an appropriate use of the models (e.g. Costantini *et al.* 2006). However, there is a larger potential for use of these models in other contexts and applications that has perhaps been underappreciated.

Beyond simply controlling for variation at one hierarchical level in the analysis of another, these models allow the partitioning of variance across levels. For example, in a supplemental feeding trial examining nutrient uptake in nestling birds through blood and diet measures, it may be of particular interest not just to control for brood but also to examine its effects. A finding that there is substantial variation at the individual level within broods would be of interest – it could imply sibling competition, genetic variation, or factors such as hatch order affecting nutrient absorption. Multi-level models

could also be used in this case to ask whether there is more variation within some broods than others, and if so, what co-variates are associated with higher intra-brood variation. Broods where parents are struggling to provide enough food would be expected to have higher within-brood variance.

At the opposite end of the spectrum, such models are also appropriate for looking at variation within versus across species (Additional layers of nesting could be built in for higher taxonomic groups, as an alternative to traditional phylogenetic control methods.) It is again critical to make sure that the oxidative balance measures used are appropriate at all levels of the analysis. If so, this approach can allow characterization of each level separately, a comparison across levels, and a breakdown of the variance. Co-variates can also be incorporated at each level of the analysis – allowing, for example, a model that simultaneously controls for habitat type as a species-level variable and age as an individual-level variable. For example, a toxicology study might look for differences in oxidative stress in several related species across habitats varying in level of contamination. Such a study could simultaneously nest individuals within species and individuals with habitats (a 'crossed' model), and additionally include habitat-level variables such as vegetation type and individual-level variables such as sex.

Multi-level models are typically employed as a type of regression model, modelling the effect of A on B. However, they can also be extended to compare the correlation structure across groups (Cohen, McGraw & Robinson 2009). For example, we may hypothesize that free radical levels and superoxide dismutase levels correlate differently in different species, but we may not have a clear hypothesis as to which is the cause and which is the effect. In this case, we can use standard normal transformations of the variables to put them all on the same scale (mean = 0, SD = 1). Then the results for a model analysing the effect of A on B become equivalent to one analysing the effect of B on A. The model can be built to estimate a parameter for how much this effect differs across species, and whether the parameter is significantly different from zero.

A final application of multi-level models is to create empirical Bayes estimates (You & Rao 2000). These are statistical adjustments to actual measured values based on incorporating the other variation observed in the data set to correct for measurement error. For example, if we measure oxidative damage levels in a number of different species with varying sample sizes for each species, we might be inclined to doubt a particularly high value found in a species with particularly small sample size. Rather than discard this observation, we could adjust it downward to reflect the fact that the original measurement is highly uncertain and, based on what we see from our other observations, that the observed value seems highly unlikely. A multi-level model always includes an assumption about the distribution of the random parameters. For example, in this case we might assume that the species-specific average values are normally distributed. The specific distribution assumed may or may not be correct (and may need to be adjusted based on the

data, with caution), but the observed species means are not truly independent – they are a sample drawn from the actual distribution of all species means. This lack of independence means that we can infer something about one species mean from what we know about the others. Empirical Bayes estimates are the results generated from the formal method of balancing observed variance within a species with the expected distribution of all means, arriving at a most likely set of values for all species.

STRUCTURAL EQUATIONS MODELS

Another statistical approach to understanding complex systems is through path analysis or structural equations models (the terms refer to similar techniques; we refer to them here as SEM, Loehlin 2003). Developed in part for looking at structures of genetic pathways, and also widely used by economists and sociologists, SEM allows analysis of complex causal networks. In the simplest possible example, knowing the correlation between two variables, A and B, tells us nothing about causality. However, if we add a third variable, C, we can start to test hypotheses about causality. Imagine that B is tightly correlated with A and C, but A is only weakly correlated with C. In this case, A causes B causes C is a reasonable hypothesis; so is C causes B causes A. A causes C causes B is not, nor is C causes A causes B, and so forth. SEM builds on this principle to evaluate the fit between various hypotheses of network structure and the data set at hand.

At the simplest level, SEM may help when there are clear hypotheses of how some components of oxidative balance systems affect others. For example, if we measure free radical production, antioxidant enzyme levels, and oxidative damage, we could compare support for a hypothesis A (that increased free radical production results in oxidative damage and upregulation of antioxidant enzymes which in turn mitigate oxidative damage) to support for a hypothesis B (that upregulation of enzymes is dependent on the damage rather than the free radicals themselves).

Many such relationships have been worked out by molecular biologists and biochemists, but usually only at short time scales. It is far less clear how, for example, a short-term increase in free radical production affects longer-term levels of antioxidants or damage. SEM can be used with longitudinal studies to examine the effects of levels at each time point on levels at the subsequent time point, potentially a powerful approach to understanding how oxidative systems function over time and in different environments. There are also other potential applications of SEM at the comparative level – for example, to disentangle how various interacting selection pressures such as diet, physiological optimization, and trade-offs might result in patterns of physiological evolution.

Conclusions

Much previous work by ecologists on oxidative balance has underestimated the methodological and statistical

challenges associated with proper quantification of the relevant aspects of physiology; however, we have learned a lot through trial and error, and are now in a position to address these complexities through use of different assays, careful study design, and application of appropriate statistical methods. All of these follow naturally from an understanding of the difficulties involved; here, we have outlined the major conceptual, methodological, and statistical issues and some potential solutions. Particularly, we would stress the importance of careful interpretation of variation in the biomarkers of OS in order to distinguish between well-controlled physiological adjustment of redox balance and accrual of damage. This distinction is difficult to make without a proper understanding of how different biomarkers relate to pathology and components of fitness. There is thus urgent need for both ecological field studies and controlled lab experiments to establish such links and to validate measurement technology and sampling design. When combined with proper statistical methodology, the field of oxidative ecology should then be able start answering the new substantive questions that are emerging, including:

1. What determines whether individuals or species take oxidative damage or regulate themselves to avoid it?
2. How and why do oxidative balance systems vary across species?
3. How does evolution shape the whole suite of oxidative balance physiology in concert?
4. Is there variation in the relevance of oxidative damage for fitness or health across taxa?

Acknowledgements

References

Alonso-Alvarez, C., Perez-Rodriguez, L., Garcia, J.T., Vinuela, J. & Mateo, R. (2010) Age and breeding effort as sources of individual variability in oxidative stress markers in a bird species. *Physiological and Biochemical Zoology*, **83**, 110–118.

Angelier, F.D.R., Holberton, R.L. & Marra, P.P. (2009) Does stress response predict return rate in a migratory bird species? A study of American redstarts and their non-breeding habitat. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 3545–3551.

Ansari, R.A., Kaur, M., Ahmad, F., Rahman, S., Rashid, H., Islam, F. & Raisuddin, S. (2009) Genotoxic and oxidative stress-inducing effects of deltamethrin in the erythrocytes of a freshwater biomarker fish species, *Channa punctata* Bloch. *Environmental Toxicology*, **24**, 429–436.

Bell, A.M., Hankison, S.J. & Laskowski, K.L. (2009) The repeatability of behaviour: a meta-analysis. *Animal Behaviour*, **77**, 771–783.

Berglund, Å.M.M., Sturve, J., Förlin, L. & Nyholm, N.E.I. (2007) Oxidative stress in pied flycatcher (*Ficedula hypoleuca*) nestlings from metal contaminated environments in northern Sweden. *Environmental Research*, **105**, 330–339.

Biro, P.A. & Dingemanse, N.J. (2009) Sampling bias resulting from animal personality. *Trends in Ecology & Evolution*, **24**, 66–67.

Blazej, A., Timothy, P.O.C., Wenbo, Q., Eric, M.D., Anson, P., Asish, R.C., Holly Van, R. & Rochelle, B. (2006) High oxidative damage levels in the longest-living rodent, the naked mole-rat. *Aging Cell*, **5**, 463–471.

Brown, M.B. & Brown, C.R. (2009) Blood sampling reduces annual survival in cliff swallows (*Petrochelidon Pyrrhonota*). *The Auk*, **126**, 853–861.

Catoni, C., Peters, A. & Schaefer, M.H. (2008) Life history trade-offs are influenced by the diversity, availability and interactions of dietary antioxidants. *Animal Behaviour*, **76**, 1107–1119.

Chuang, C.C., Shiesh, S.C., Chi, C.H., Tu, Y.F., Hor, L.I., Shieh, C.C. & Chen, M.F. (2006) Serum total antioxidant capacity reflects severity of illness in patients with severe sepsis. *Critical Care*, **10**, 1–7.

Cohen, A., Hau, M. & Wikelski, M. (2008a) Stress, metabolism, and antioxidants in two wild passerine bird species. *Physiological and Biochemical Zoology*, **81**, 463–472.

Cohen, A.A., Klasing, K.C. & Ricklefs, R.E. (2007) Measuring circulating antioxidants in wild birds. *Comparative Biochemistry and Physiology B*, **147**, 110–121.

Cohen, A.A. & McGraw, K.J. (2009) No simple measures for antioxidant status in birds: complexity in inter- and intraspecific correlations among circulating antioxidant types. *Functional Ecology*, **23**, 310–320.

Cohen, A.A., McGraw, K.J. & Robinson, W.D. (2009) Serum antioxidant levels in wild birds vary in relation to diet, season, life history strategy, and species. *Oecologia*, **161**, 673–683.

Cohen, A.A., McGraw, K.J., Wiersma, P., Williams, J.B., Robinson, W.D., Robinson, T.R., Brawn, J.D. & Ricklefs, R.E. (2008b) Interspecific associations between circulating antioxidant levels and life-history variation in birds. *American Naturalist*, **172**, 178–193.

Costantini, D. (2008) Oxidative stress in ecology and evolution: lessons from avian studies. *Ecology Letters*, **11**, 1238–1251.

Costantini, D., Dell'ariccia, G. & Lipp, H.-P. (2008) Long flights and age affect oxidative status of homing pigeons (*Columba livia*). *Journal of Experimental Biology*, **211**, 377–381.

Costantini, D. & Dell'omo, G. (2006) Environmental and genetic components of oxidative stress in wild kestrel nestlings (*Falco tinnunculus*). *Journal of Comparative Physiology B*, **176**, 575–579.

Costantini, D., Casagrande, S., De Filippis, S., Brambilla, G., Fanfani, A., Tagliavini, J. & Dell'omo, G. (2006) Correlates of oxidative stress in wild kestrel nestlings (*Falco tinnunculus*). *Journal of Comparative Physiology B*, **176**, 329–337.

Costantini, D., Coluzza, C., Fanfani, A. & Dell'omo, G. (2007) Effects of carotenoid supplementation on colour expression, oxidative stress and body mass in rehabilitated captive adult kestrels (*Falco tinnunculus*). *Journal of Comparative Physiology B*, **177**, 723–731.

Costantini, D., Rowe, M., Butler, M.W. & McGraw, K.J. (2010) From molecules to living systems: historical and contemporary issues in oxidative stress and antioxidant ecology. *Functional Ecology*, **24**, 950–959.

Dalle-Donne, I., Rossi, R., Colombo, R., Giustarini, D. & Milzani, A. (2006) Biomarkers of oxidative damage in human disease. *Clinical Chemistry*, **52**, 601–623.

Dotan, Y., Lichtenberg, D. & Pinchuk, I. (2004) Lipid peroxidation cannot be used as a universal criterion of oxidative stress. *Progress in Lipid Research*, **43**, 200–227.

Dowling, D.K. & Simmons, L.W. (2009) Reactive oxygen species as universal constraints in life-history evolution. *Proceedings of the Royal Society B-Biological Sciences*, **276**, 1737–1745.

Filho, D.W., Torres, M.A., Tribess, T.B., Pendrosa, R.C. & Soares, C.H.L. (2001) Influence of season and pollution on the antioxidant defenses of the cichlid fish acará (*Geophagus brasiliensis*). *Brazilian Journal of Medical and Biological Research*, **34**, 719–726.

Galván, I. & Alonso-Alvarez, C. (2008) An intracellular antioxidant determines the expression of a melanin-based signal in a bird. *PLoS One*, **3**, e3335.

Galván, I. & Alonso-Alvarez, C. (2009) The expression of melanin-based plumage is separately modulated by exogenous oxidative stress and a melanocortin. *Proceedings of the Royal Society B: Biological Sciences*, **276**, 3089–3097.

Garamszegi, L.Z., Calhim, S., Dochtermann, N., Hegyi, G., Hurd, P.L., Jorgensen, C., Kutsukake, N., Lajeunesse, M.J., Pollard, K.A., Schielzeth, H., Symonds, M.R.E. & Nakagawa, S. (2009) Changing philosophies and tools for statistical inferences in behavioral ecology. *Behavioral Ecology*, **20**, 1363–1375.

Geens, A., Dauwe, T. & Eens, M. (2009) Does anthropogenic metal pollution affect carotenoid colouration, antioxidative capacity and physiological condition of great tits (*Parus major*)? *Comparative Biochemistry and Physiology C*, **150**, 155–163.

Halliwell, B. & Gutteridge, J.M.C. (2007) *Free Radicals in Biology and Medicine*. Oxford University Press, Oxford.

Halliwell, B. & Whiteman, M. (2004) Measuring reactive species and oxidative damage in vivo and in cell culture: how should you do it and what do the results mean? *British Journal of Pharmacology*, **142**, 231–255.

Harr, K.E. (2002) Clinical chemistry of companion avian species: a review. *Veterinary Clinical Pathology*, **31**, 140–151.

He, H., Genovese, K.J., Nisbet, D.J. & Kogut, M.H. (2006) Profile of Toll-like receptor expressions and induction of nitric oxide synthesis by Toll-like receptor agonists in chicken monocytes. *Molecular Immunology*, **43**, 783–789.

Hermans, N., Cos, P., Maes, L., De Bruyne, T., Vanden Berghe, D., Vlietinck, A.J. & Pieters, L. (2007) Challenges and pitfalls in antioxidant research. *Current Medicinal Chemistry*, **14**, 417–430.

Hollmén, T., Franson, J.C., Hario, M., Sankari, S., Kilpi, M. & Lindstrom, K. (2001) Use of serum biochemistry to evaluate nutritional status and health of incubating common eiders (*Somateria mollissima*) in Finland. *Physiological and Biochemical Zoology*, **74**, 333–342.

Hõrak, P., Zilmer, M., Saks, L., Ots, I., Karu, U. & Zilmer, K. (2006) Antioxidant protection, carotenoids, and the costs of immune challenge in greenfinches. *Journal of Experimental Biology*, **209**, 4329–4338.

Hõrak, P., Saks, L., Zilmer, M., Karu, U. & Zilmer, K. (2007) Do dietary antioxidants alleviate the cost of immune activation? An experiment with greenfinches. *American Naturalist*, **170**, 625–635.

Hõrak, P., Sild, E., Soomets, U., Sepp, T. & Kilk, K. (2010) Oxidative stress and information content of black and yellow plumage coloration: an experiment with greenfinches. *Journal of Experimental Biology*, **213**, 2225–2233.

Isaksson, C. & Andersson, S. (2008) Oxidative stress does not influence carotenoid mobilization and plumage pigmentation. *Proceedings of the Royal Society B-Biological Sciences*, **275**, 309–314.

Isaksson, C., Örnborg, J., Stephensen, E. & Andersson, S. (2005) Plasma glutathione and carotenoid coloration as potential biomarkers of environmental stress in great tits. *EcoHealth*, **2**, 138–146.

Johnson, R.A. & Wichern, D.W. (2007) *Applied Multivariate Statistical Analysis*. Pearson Education, Upper Saddle River, NJ.

Kenow, K.P., Hoffman, D.J., Hines, R.K., Meyer, M.W., Bickham, J.W., Matson, C.W., Stebbins, K.R., Montagna, P. & Elfessi, A. (2008) Effects of methylmercury exposure on glutathione metabolism, oxidative stress, and chromosomal damage in captive-reared common loon (*Gavia immer*) chicks. *Environmental Pollution*, **156**, 732–738.

Knasmüller, S., Nersesyan, A., Mišik, M., Gerner, C., Mikulits, W., Ehrlich, V., Hoelzl, C., Szakmary, A. & Wagner, K.H. (2008) Use of conventional and -omics based methods for health claims of dietary antioxidants: a critical overview. *British Journal of Nutrition*, **99**, ES3–ES52.

Lessells, C.M. & Boag, P.T. (1987) Unrepeatable repeatabilities: a common mistake. *Auk*, **104**, 116–121.

Loehlin, J.C. (2003) *Latent Variable Analysis: An Introduction to Factor, Path, and Structural Equation Analysis*. Lawrence Erlbaum, Mahwah, NJ.

Lopez-Torres, M., Perez-Campo, R., Cadenas, S., Rojas, C. & Barja, G. (1993) A comparative study of free radicals in vertebrates-II. Non-enzymatic antioxidants and oxidative stress. *Comparative Biochemistry and Physiology B*, **105**, 757–763.

Martin, L.B., Weil, Z.M. & Nelson, R.J. (2008) Seasonal changes in vertebrate immune activity: mediation by physiological trade-offs. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **363**, 321–339.

McGraw, K., Tourville, E. & Butler, M. (2008) A quantitative comparison of the commonly used methods for extracting carotenoids from avian plasma. *Behavioral Ecology and Sociobiology*, **62**, 1991–2002.

Metcalfe, N. & Alonso-Álvarez, C. (2010) Oxidative stress as a life-history constraint: the role of reactive oxygen species (ROS) in shaping phenotypes from conception to death. *Functional Ecology*, **24**, 984–996.

Monaghan, P., Metcalfe, N.B. & Torres, R. (2009) Oxidative stress as a mediator of life history trade-offs: mechanisms, measurements and interpretation. *Ecology Letters*, **12**, 75–92.

Mosmann, T.R. & Sad, S. (1996) The expanding universe of T-cell subsets: Th1, Th2 and more. *Immunology Today*, **17**, 138–145.

Nadolski, J., Skwarska, J., Kalinski, A., Banbura, M., Sniegula, R. & Banbura, J. (2006) Blood parameters as consistent predictors of nestling performance in great tits (*Parus major*) in the wild. *Comparative Biochemistry and Physiology A*, **143**, 50–54.

Neuzil, J. & Stocker, R. (1994) Free and albumin-bound bilirubin are efficient co-antioxidants for α-tocopherol, inhibiting plasma and low density lipoprotein lipid peroxidation. *Journal of Biological Chemistry*, **269**, 16712–16719.

Norte, A.C., Sheldon, B., Sousa, J.P. & Ramos, J.A. (2008) Repeatability and method-dependent variation of blood parameters in wild-caught Great Tits *Parus major*. *Acta Ornithologica*, **43**, 65–75.

Nussey, D.H., Pemberton, J.M., Pilkington, J.G. & Blount, J.D. (2009) Life history correlates of oxidative damage in a free-living mammal population. *Functional Ecology*, **23**, 809–817.

Olsson, M., Wilson, M., Isaksson, C. & Uller, T. (2009) Polymorphic ROS scavenging revealed by CCCP in a lizard. *Naturwissenschaften*, **96**, 845–849.

Pike, T.W., Blount, J.D., Lindström, J. & Metcalfe, N.B. (2007) Availability of non-carotenoid antioxidants affects the expression of a carotenoid-based sexual ornament. *Biology Letters*, **3**, 353–356.

Rabe-Hesketh, S. & Skrondal, A. (2008) *Multilevel and Longitudinal Modeling Using Stata*. Stata Press, College Station, TX.

Radak, Z., Chung, H.Y., Koltai, E., Taylor, A.W. & Goto, S. (2008) Exercise, oxidative stress and hormesis. *Ageing Research Reviews*, **7**, 34–42.

Rossi, R., Milzani, A., Dalle-Donne, I., Giustarini, D., Lusini, L., Colombo, R. & Di Simplicio, P. (2002) Blood glutathione disulfide: in vivo factor or in vitro artifact? *Clinical Chemistry*, **48**, 742–753.

Rossi, R., Dalle-Donne, I., Milzani, A. & Giustarini, D. (2006) Oxidized forms of glutathione in peripheral blood as biomarkers of oxidative stress. *Clinical Chemistry*, **52**, 1406–1414.

Ruxton, G.D. & Colegrave, N. (2006) *Experimental Design for the Life Sciences*. Oxford University Press, Oxford.

Saino, N., Bertacche, V., Bonisoli-Alquati, A., Romano, M. & Rubolini, D. (2008) Phenotypic correlates of yolk and plasma carotenoid concentration in yellow-legged gull chicks. *Physiological and Biochemical Zoology*, **81**, 211–225.

Seaman, C., Moritz, J., Falkenstein, E., Van Dyke, K., Casotti, G. & Klandorf, H. (2008) Inosine ameliorates the effects of hemin-induced oxidative stress in broilers. *Comparative Biochemistry and Physiology A*, **151**, 670–675.

Selman, C., McLaren, J.S., Collins, A.R., Duthie, G.G. & Speakman, J.R. (2008) The impact of experimentally elevated energy expenditure on oxidative stress and lifespan in the short-tailed field vole Microtus agrestis. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 1907–1916.

Sepp, T., Sild, E. & Hõrak, P. (2010) Hematological condition indices in greenfinches: effects of captivity and diurnal variation. *Physiological and Biochemical Zoology*, **83**, 276–282.

Stearns, S.C. (1992) *The Evolution of Life Histories*. Oxford University Press, Oxford.

Stefanutti, G., Pierro, A., Vinardi, S., Spitz, L. & Eaton, S. (2005) Moderate hypothermia protects against systemic oxidative stress in a rat model of intestinal ischemia and reperfusion injury. *Shock*, **24**, 159–164.

Surai, P.F. (2002) *Natural Antioxidants in Avian Nutrition and Reproduction*. Nottingham University Press, Nottingham, UK.

Sureda, A., Tauler, P., Aguilo, A., Cases, N., Llompart, I., Tur, J.A. & Pons, A. (2007) Antioxidant supplementation influences the neutrophil tocopherol associated protein expression, but not the inflammatory response to exercise. *Central European Journal of Biology*, **2**, 56–70.

Tella, J.L., Figuerola, J., Negro, J.J., Blanco, G., Rodríguez-Estrella, R., Forero, M.G., Blázquez, M.C., Green, A.J. & Hiraldo, F. (2004) Ecological, morphological and phylogenetic correlates of interspecific variation in plasma carotenoid concentration in birds. *Journal of Evolutionary Biology*, **17**, 156–164.

Tummeleht, L., Mägi, M., Kilgas, P., Mänd, R. & Hõrak, P. (2006) Antioxidant protection and plasma carotenoids of incubating great tits (*Parus major* L.) in relation to health state and breeding conditions. *Comparative Biochemistry and Physiology C*, **144**, 166–172.

Vider, J., Lehtmaa, J., Kullisaar, T., Vihalemm, T., Zilmer, K., Kairane, C., Landõr, A., Karu, T. & Zilmer, M. (2001) Acute immune response in respect to exercise-induced oxidative stress. *Pathophysiology*, **7**, 263–270.

Von Schantz, T., Bensch, S., Grahn, M., Hasselquist, D. & Wittzell, H. (1999) Good genes, oxidative stress and condition-dependent sexual signals. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **266**, 1–12.

Wang, Q., Sun, A.Y., Simonyi, A., Kalogeris, T.J., Miller, D.K., Sun, G.Y. & Korthuis, R.J. (2007) Ethanol preconditioning protects against ischemia/reperfusion-induced brain damage: role of NADPH oxidase-derived ROS. *Free Radical Biology and Medicine*, **43**, 1048–1060.

You, Y. & Rao, J.N.K. (2000) Hierarchical Bayes estimation of small area means using multilevel models. *Survey Methodology*, **26**, 173–181.

Zera, A.J. & Harshman, L.G. (2001) The physiology of life history trade-offs in animals. *Annual Review of Ecology and Systematics*, **32**, 95–126.

Supporting Information

Additional Supporting Information may be found in the online version of this article.

Appendix S1. Here we present a worked hypothetical example of how the principles outlined in this article could be applied to an actual research project, including choice and validation of assays and statistical analysis of data.

Appendix S2. Here we present a intuitive and practical summary of principal components analysis and factor analysis, as well as a method for validating the application of such techniques across different sub-groups of a population.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.