Ten questions for evolutionary studies of disease vulnerability

Randolph M. Nesse

Departments of Psychiatry and Psychology, Institute for Social Research, The University of Michigan, Ann Arbor, MI, USA

Evolutionary Applications ISSN 1752-4571

© 2011 Blackwell Publishing Ltd
of methods. Methods that fail are discarded. Those that work are kept, improved, and applied to new problems. Along the way, philosophers of science offer perspective and suggestions.

This article does not attempt to offer a shortcut for that process, and it certainly does not propose a general solution to the challenges of testing hypotheses about adaptation. Instead, it reviews strategies that have proved effective or ineffective for formulating and assessing hypotheses about traits that leave bodies vulnerable to disease. Ten questions arise routinely in the process of assessing such hypotheses. Considering each question systematically can help to minimize errors. A preliminary version of these questions has saved many students from failure and at least one professor from despair (Nesse 2007). They are like a pilot’s checklist for projects in evolutionary medicine preparing for takeoff.

Attempts to offer general advice about scientific methods are easy to ridicule. When Peter Medawar addressed, ‘What is “The Scientific Method,”’? he began by noting: ‘If the purpose of scientific methodology is to prescribe or expound a system of inquiry or even a code of practice for scientific behavior, then scientists seem to be able to get on very well without it… Of what other branch of learning can it be said that it gives its proficients no advantage; that it need not be taught or, if taught, need not be learned?’ (Medawar 1969, 8) However, he continues, ‘Of course, the fact that scientists do not consciously practice a formal methodology is very poor evidence that no such methodology exists.’ He then offers a sophisticated perspective on the gulf between the enterprise of formulating hypotheses and that of testing them, concluding that ‘Imaginativeness and a critical temper are both necessary at all times, but neither is sufficient’ (p 58). Because both are rarely well developed in the same person, science progresses by imaginative scientists proposing a variety of ideas that are winnowed by others with ‘a more critical cast of mind.’ The creative process is ineffable, so advice on scientific methods inevitably emphasizes errors.

Medawar’s observation is particularly germane for studies in this area of Darwinian medicine. Traits that leave bodies vulnerable to disease spur the imagination. Each one is a mystery. Hypothesis formation begins by imagining how existing traits could be ‘redesigned’, so they are less likely to cause disease. Some individuals are resistant to HIV infection, some have no appendix, and some fight off tuberculosis easily. Why not all of us? Why didn’t natural selection shape bodies less vulnerable to diseases? Such questions inspire students to wonderfully creative explanations, many of which are, unfortunately, preposterous. Subtle aspects of evolutionary theory are rarely the problem. Far more often, mistakes arise from difficulty framing the question or from elementary errors.

Table 1 lists four tasks common to all science, and ten specific questions that arise in the process of considering hypotheses about traits that leave the body vulnerable to disease. Each task offers opportunities for errors somewhat specific to this area of work.

The first task is to specify the object of explanation. This seems simple enough, but students overwhelmingly begin by looking for adaptive functions of diseases or rare alleles. This is usually a mistake; most diseases and rare alleles are not shaped by selection. An appropriate object of explanation is more often a shared trait that leaves almost all members of a species or a subgroup vulnerable to a disease.

Specifying the kind of explanation also seems simple, but pitfalls loom. The first arises from failure to understand the difference between a proximate and an evolutionary explanation, and the need for both. Another arises from failure to distinguish evolutionary explanations based on phylogeny from those based on the functional significance of a trait.

Giving every possible hypothesis full and fair consideration is a challenge in all science. This can be especially difficult for evolutionary explanations because some are intuitively attractive, others are hard to recognize, and multiple answers can be correct.

Finally, testing evolutionary hypotheses about vulnerability can be challenging. The most appropriate methods differ, in under-appreciated and fundamental ways, from those used to test proximate hypotheses.

A list of generic tasks of science offers little guidance on its own, but it can help to organize the challenges and questions that arise when framing and assessing specific evolutionary hypotheses about vulnerability to disease. Ten questions arise routinely in the course of such an inquiry. Addressing them systematically increases the chances of successful takeoff and landing.

Challenges associated with specifying the object of explanation (Q1–3)

Upon first recognizing that every trait needs an evolutionary explanation, the temptation to try to explain diseases directly is nearly irresistible. Why did natural selection shape cancer? What are the benefits of aging? What is the function of the appendix? How do genes that cause schizophrenia give a selective advantage? Such seemingly sensible questions are framed incorrectly (Williams and Nesse 1991). Most diseases are not shaped directly by natural selection or other evolutionary forces, so they are not, themselves, appropriate objects for evolutionary explanation.

Aspects of the body that make it vulnerable to a disease are an entirely different matter. They require an
Ten questions for evolutionary studies of disease vulnerability.

<table>
<thead>
<tr>
<th>Task 1: Define the object of explanation precisely</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1. Is the object of explanation a uniform trait in the species, or is the goal to explain variations in a trait among groups or individuals?</td>
</tr>
<tr>
<td>Q2. Has the object of explanation been influenced by evolution?</td>
</tr>
<tr>
<td>Q3. What kind of trait is the object of explanation?</td>
</tr>
<tr>
<td>a. A fixed human trait, such as the narrow birth canal</td>
</tr>
<tr>
<td>b. A facultative trait, such as the capacity for sweating</td>
</tr>
<tr>
<td>c. Human genes, in the most general sense of the term</td>
</tr>
<tr>
<td>d. Pathogen traits, such as the level of virulence</td>
</tr>
<tr>
<td>e. Pathogen genes, such as those that confer antibiotic resistance</td>
</tr>
<tr>
<td>f. Somatic cell lines, such as those in tumors or the immune system</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task 2: Specify the kind of explanation sought</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q4. Is the goal to explain the evolution of the trait, or its proximate mechanisms?</td>
</tr>
<tr>
<td>Q5. Is the goal to explain the trait’s phylogeny, or the evolutionary forces that shaped it?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task 3: List and consider all viable hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q6. Are all viable hypotheses considered and given fair consideration, or are some hypotheses neglected, while others receive favored treatment?</td>
</tr>
<tr>
<td>Q7. Could different vulnerabilities cause the disease in different individuals or subgroups?</td>
</tr>
<tr>
<td>Q8. What categories of explanation are under consideration?</td>
</tr>
<tr>
<td>a. Mismatch of bodies with environments they did not evolve in</td>
</tr>
<tr>
<td>b. Co-evolution with pathogens that evolve faster than hosts can</td>
</tr>
<tr>
<td>c. Constraints on selection, such as time required, genetic drift, and mutation</td>
</tr>
<tr>
<td>d. Trade-offs, especially costs associated with apparently superior alternative possible traits</td>
</tr>
<tr>
<td>e. Reproductive success at the expense of health</td>
</tr>
<tr>
<td>f. Defenses such as fever and pain that cause harm and suffering, but were shaped by selection because they offer protection in certain situations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Task 4: Describe the methods used to test the hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q10. What methods are used to test the hypotheses?</td>
</tr>
<tr>
<td>a. Consistency with evolutionary theory</td>
</tr>
<tr>
<td>b. Modeling using quantitative methods</td>
</tr>
<tr>
<td>c. Comparative methods</td>
</tr>
<tr>
<td>i. Comparisons among species</td>
</tr>
<tr>
<td>ii. Comparisons among subgroups of a species</td>
</tr>
<tr>
<td>iii. Comparisons among individuals who vary in a trait</td>
</tr>
<tr>
<td>d. Experimental methods</td>
</tr>
<tr>
<td>i. Extermination or disruption (e.g. studies that knock-out genes or block fever)</td>
</tr>
<tr>
<td>ii. Augmentation (e.g. administration of extra testosterone)</td>
</tr>
<tr>
<td>iii. Examining regulation of a facultative trait to see if it behaves as predicted</td>
</tr>
<tr>
<td>iv. Observing evolutionary changes in the lab or the field</td>
</tr>
<tr>
<td>e. Examining the details of fit between observed form and a postulated function</td>
</tr>
</tbody>
</table>

Evolutionary explanation. Why don’t defenses against cancer work better? What is the evolutionary explanation for senescence? Why does the appendix persist? Why do genes that predispose to schizophrenia persist? These are good questions. Shifting the focus from diseases to traits that make organisms vulnerable to disease requires a major change in perspective.

For many in medicine, an even more fundamental shift is necessary. Most medical research focuses on how bodily mechanisms work, and how differences among individuals account for why some become ill. An evolutionary approach focuses on how bodily mechanisms evolved, and how traits shared by most members of the species leave them vulnerable to a disease. The difference becomes clear only if the object of explanation and kind of explanation sought are described explicitly.

The first challenge is to specify whether the object of explanation is a trait universal in the species or variation in a trait (Q1). Attempts to explain universal traits, such as the narrow birth canal, wisdom teeth, and synthesis of bilirubin, proceed very differently from attempts to explain variations in traits, such as skin color, ability to digest lactose, or vulnerability to malaria.

The second challenge is to ensure that the trait is something likely to have evolved. This question (Q2) helps to avoid the confusion that arises from trying to posit unitary evolutionary explanations for things like war or liking miso soup. Selection shaped the behavioral traits that make war possible, but not war itself. Addressing this question also helps to minimize attempts to explain diseases directly; cancer, atherosclerosis, and obesity are not universal traits.

Question 3 encourages precise description of the trait in question. Traits require different approaches depending on which of six categories they best fit.

Fixed human traits

The first and most obvious category is for universal traits such as the birth canal, the heart’s mitral valve, the appendix, the adrenal glands, and the curve of the human spine. These offer relatively definable targets for evolutionary explanation.

Facultative adaptations

Other universal traits are facultative adaptations such as the capacities for skin tanning, adaptation to high-altitude, callus formation, cough, and fever. Natural selection shaped these protective responses in conjunction with regulation mechanisms that express them when they are likely to be useful. The time scale of such responses can range from instantaneous to a lifetime. Tissue damage
arouses pain immediately. Cold causes shivering in minutes. Callus formation and skin tanning take weeks.

Some responses have a longer horizon. When early experiences shape the long-term developmental trajectory of an individual, this is usually described as developmental plasticity (Stearns 1989; West-Eberhard 2003). One classic example remains inadequately documented; babies whose spend their earliest months in hot climates have more sweat glands as adults (Kawahata and Sakamoto 1951). A similar finding in voles is well established; those born at the end of summer have thicker coats (Lee et al. 1987). A review of such environmental influences during human development emphasizes the important distinctions among developmental disruptions (such as from folate deficiency), immediately adaptive responses (such as premature birth to escape an infected uterus), and predictive adaptive responses, such as those that use in uterine cues to adjust behavior and metabolism to varying environments (Gluckman et al. 2005a,b).

Predictive adaptive responses are an important research area in evolutionary medicine (Gluckman et al. 2005a,b). Responses to cues of two kinds have been studied extensively. Mothers exposed to severe stress give birth to offspring with increased responses to stress (Meaney et al. 2007). Mothers exposed to limited nutrition give birth to offspring prone to obesity, hypertension, diabetes, and atherosclerosis if they grow up in nutritionally rich environments (Gluckman et al. 2005a,b; Barker 2007). In each case, the change is in the direction that would conceivably increase the fitness of an offspring living in environmental conditions similar to those experienced by the mother. Work is ongoing to determine whether these mechanisms are indeed facultative adaptations. In both cases, explanation of the epigenetic mechanisms that mediate the effects will likely provide useful guidance, as will data on the prevalence and fitness effects of nutrition availability during pregnancy (Rickard et al. 2010).

**Human genes**

Human genes are obvious targets for evolutionary explanation, and population genetics offers reliable methods (Childs 1999; Lewontin et al. 2000). Attention usually focuses on variations, but fixed or very common genes deserve attention also. For instance, most vertebrates have the enzyme uricase, but the hominid line lost the responsible gene in the Miocene, leaving us with high levels of uric acid and vulnerability to gout (Varela-Echavarria et al. 1988; Johnson et al. 2010; Kebaugh and Thomas 2010). Is this because of the antioxidant properties of uric acid, its effect on blood pressure, some other benefit, costs associated with the gene, or just chance? The active reabsorption of uric acid in the kidney suggests some adaptive function, but a definitive answer remains elusive (Álvarez-Lario and Macarrn-Vicente 2010). The answer is important, because it may help to explain the relationship between modern diets, obesity and atherosclerosis (Johnson et al. 2010). A proximate view that assumes uric acid is merely a metabolic byproduct neglects the important possibility that high uric acid levels may also offer fitness advantages.

Genes that give a reproductive advantage can go to fixation even if they harm health. Men would, on average, live 7 years longer if their metabolism and behavior were like that of women. Why aren’t they? Competitive ability increases fitness more for males than females in a polygynous species, while ability to repair tissues increases fitness relatively more for females. Data from multiple species, and from diverse human groups across the past century, support this hypothesis (Kruger and Nesse 2006).

Could genes that make human mood regulation systems vulnerable to mania have become fixed because they give selective advantages, separate from mania, that more than counter-balance the disadvantages of illness experienced by some people? This purely hypothetical example illustrates the possibility that alleles can improve fitness at the expense of health. Of course, the focus is usually on genetic variation and on why individuals with a variant might have advantages as well as disadvantages; this has been proposed for manic depressive illness (Wilson 1998).

The textbook exemplar is the sickle cell hemoglobin allele. Individuals with two copies get sickle cell disease. Individuals homozygous for ordinary hemoglobin are vulnerable to malaria. Heterozygote individuals have decreased vulnerability to malaria and limited symptoms of sickle cell disease; their relative fitness advantage explains the persistence of the sickle cell allele where malaria is prevalent (Livingstone 1960; Wiesenfeld 1967; Piel et al. 2010). This explanation is solid, but it is by no means a generalizable exemplar for evolutionary medicine. The variation is in a single base pair, it is of recent origin, and almost all other documented examples of balancing selection have also been shaped by malaria (Evans and Wellens 2002).

This has not inhibited attempts to propose heterozygote advantage as an explanation for other traits. As the most common fatal disease caused by a recessive allele, carried by 4% of European Americans, cystic fibrosis is a good candidate. While 70% of cases arise from the Delta F508 allele, 1400 other causal mutations have been identified (Collins 1992). Heterozygote mice are protected from dehydration caused by cholera toxin (Bertranpetit and Calafell 1996), and their intestinal cells are resistant to penetration by Salmonella typhi (Pier et al. 1998). Does this explain the high prevalence and diversity of the cystic fibrosis allele? It does not fit well with epidemiologic data.
showing higher frequencies in northern climates, where death from diarrhea may be less common, but this is weak counter-evidence. Attempts to understand the prevalence of cystic fibrosis illustrate the human tendency to think of alleles as either normal or abnormal, as if they were components in a designed machine. Mutation-selection balance, and mutation vulnerability of the allele, may account for most of the prevalence of the cystic fibrosis, with frequencies perhaps increased because of selective advantages in certain circumstances.

Recessive alleles causing Tay Sachs and other diseases of sphingolipid metabolism in Askanazi Jews have also been attractive targets (Zlotogora and Bach 2003), but definitive conclusions have been hard to find. Analysis of new genetic evidence shows positive selection in Askanazi Jews for ability to metabolize alcohol and lactose, but the frequency of sphingolipid diseases seems better accounted for by bottleneck effects (Bray et al. 2010). Another review considers the general difficulties of reaching firm evolutionary conclusions about the evolutionary significance of recessive alleles and the excessive attention associated with hypotheses about heterozygote advantage (Valles 2010). This review also documents the prevalence of confusion arising from failure to specify the exact object of explanation, and it provides a useful taxonomy of evolutionary explanations for alleles that cause disease (see Table 2).

Such sophisticated analysis is needed and welcome; however, many mistakes are much more fundamental. For instance, untenable hypotheses based on naïve group selection are published remarkably often. Hypotheses about Mendelian defects such as color blindness (Yokoyama and Takenaka 2005) are especially likely to rely on group benefits that are inconsistent with modern evolutionary theory.

The null hypotheses for increased prevalence of a rare genetic variant in a subpopulation are founder effects and drift; however, selection effects are being confirmed. Lactase is the best-studied example. Alleles that allow lactose digestion in adulthood have dramatically different prevalence in different geographic areas; nearly absent in Asia, they are almost uniformly present in northern Europe (Simoons 1978). Inability of adults to digest lactose is the ancestral state, so alleles for lactose intolerance should not be considered defective. New genetic data show mutations for lactase persistence arising and being selected for repeatedly (Tishkoff et al. 2006). The story is not, however, as simple as strong selection increasing frequencies of lactase alleles in dairying cultures. It now appears that such cultures in Europe expanded quickly, with a selective advantage perhaps as high as 20%, so subpopulations migrated north, carrying along alleles for adult expression of lactase (Ingram et al. 2009).

Still unaddressed is the question of why lactase synthesis in adults is inhibited in most human populations. Is the selection force simply the cost of manufacturing an enzyme that is usually unnecessary? Does the presence of lactase increase vulnerability to certain pathogens? Does inability to digest milk prevent older children from competing with younger siblings? Or, does lactase inhibition arise from drift of a mutation that is neutral in most populations? These questions have been surprisingly neglected.

Adaptation to low oxygen pressure at high altitudes offers another example; fitness-enhancing haplotypes went from rare to over 80% prevalence in just 4000 years in Tibetan Highlanders (Simonson et al. 2010; Yi et al. 2010). The responsible gene, EPAS1, inhibits hypoxia-induced hemoglobin synthesis, and thereby decreases medical complications at high altitudes. In the Andes, by contrast, the same environmental challenge has shaped very different physiologic mechanisms that are similarly effective (Beall 2007). Strong selection causes changes, but it is hard to predict what they will be.

Variations at the ApoE locus also have medical relevance; those with the ApoE 4 allele have high risks of heart disease and Alzheimer’s disease. These variations are especially appropriate for evolutionary examination given marked geographic differences in frequency and near-fixation of Apo-E 4 in our primate ancestors (Finch 2010). These alleles may give advantages to meat eaters, but the exact trade-offs remain unclear (Finch and Stanford 2004).

Searches for the adaptive significance of specific alleles are most fruitful when selection forces are specific to a particular geographic area. For instance, the absence of Duffy antigen makes it harder for *Plasmodium falciparum* to enter blood cells, giving a potent fitness advantage where malaria is prevalent. However, it also seems to increase vulnerability to other infections and to predispose to more malignant prostate cancer (Shen et al. 2006). Similarly, the rare individuals who lack the CCR5 antigen are also protected against pathogens entering cells, in this case, HIV, but they are more vulnerable to other infections, including West Nile Fever (Ahuja and He 2010). Both examples illustrate trade-offs that help to prevent thinking of an allele as all good or all bad.

**Table 2. Evolutionary explanations for high frequencies of disease-causing alleles (Valles 2010, 185).**

<table>
<thead>
<tr>
<th></th>
<th>Evolutionary explanations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Elevated mutation rate</td>
</tr>
<tr>
<td>2</td>
<td>Segregation distortion</td>
</tr>
<tr>
<td>3</td>
<td>Reproductive compensation</td>
</tr>
<tr>
<td>4</td>
<td>Founder effects and genetic drift</td>
</tr>
<tr>
<td>5</td>
<td>Heterozygote advantage (overdominance)</td>
</tr>
</tbody>
</table>
Ten questions for evolutionary studies of disease

Attempts to propose natural selection as the explanation for geographic distribution of rare polygenic diseases illustrate the perils of advocating for one hypothesis without fully considering all possibilities. For instance, the increased prevalence of Type I diabetes in northern versus southern Europe has been used to support the hypothesis that genes predisposing to diabetes were selected for during the ice age because high levels of glucose protect against tissue damage from freezing (Moalem et al. 2005). However, the predisposing genes do not directly increase glucose levels; they mediate autoimmune reactions that destroy pancreatic beta cells, leading, in the absence of exogenous insulin, to early death. Even if not fatal, Type I diabetes results in extravagant caloric loss via glucose in the urine. It seems most unlikely that cold conditions in northern Europe 14 000 year ago selected for genes predisposing to Type I diabetes. Much more likely are drift and selection driven by infectious agents.

Pathogen traits and genes

Pathogen traits, such as level of virulence, are easier to study because they evolve fast enough to observe in the laboratory, or across a few months or years in natural populations (Anderson and May 1979; Ewald 1995). Likewise, the frequency of pathogen genes is a fine object of explanation because it changes in response to laboratory manipulations (Lenski 1998). They also lend themselves well to phylogenetic studies of practical use in tracking different clones (Manges et al. 2001) and sources of infection (Dombek et al. 2000), although such analyses can be problematic (Gordon 2001).

Cell lines undergoing somatic evolution

Evolution occurring in somatic cell lines poses different challenges and opportunities. These have been studied mostly in immune cells and tumors, but somatic evolution has also been recognized in neurons (Edelman 1987). Diverse applications in cancer are proving important (Greaves 2000). Tumors can be viewed as ecosystems in which cells compete for resources, with more successful cells displacing others, thus changing the genetic signature of a tumor as it evolves (Merlo et al. 2006). Apoptotic cell death after telomere shortening protects against cancer at a cost of faster aging (Newbold 2002), and apoptosis more generally is essential in development, and in coping with pathogens and stressors (LeGrand 1997).

A significant proportion of cancer arises from mutations early in development resulting in mosaicism (Frank 2010), and positive selection in antagonistic co-evolutionary processes may account for maintenance of alleles that confer cancer vulnerability (Crespi and Summers 2006).

Some genetic abnormalities in laboratory-maintained cancer cell lines are now recognized to result from the peculiar selection forces acting on cells grown in bottles. The evolutionary genetics of subpopulations of cells has been used to understand metastasis (Nguyen and Massague 2007) and to stage cancers. Evolutionary analysis of cancer cell lines has enormous clinical significance.

The above six objects of explanation are not necessarily all-inclusive; however, specifying the object of explanation carefully can help prevent mistakes. While an appropriate object of explanation is always something that can be influenced by evolution, this does not necessarily mean natural selection; mutation, drift and migration need full consideration.

Challenges associated with specifying the kind of explanation sought (Questions 4–5)

The distinction between proximate and evolutionary explanations, and the need for both, is so natural for evolutionary biologists that it is like riding a bicycle. It was emphasized by Mayr (1983) and others, but most health professionals have never heard of it. As a result, proximate explanations are sometimes proposed as alternatives to evolutionary explanations. For instance, considering the functions of fever sometimes elicits suggestions that it can be explained by the actions of cytokines. Considering the evolutionary reasons for the narrow birth canal may elicit a suggestion that it results from developmental mechanisms, as if they are an alternative to explanations based on the costs and benefits of a larger or smaller birth canal.

Tinbergen (1963) provides a more detailed foundation for evolutionary medicine by outlining the four kinds of questions that must be answered to provide a full biologic explanation for any trait: mechanism, development, function, and phylogeny. The first two are proximate explanations, the latter two, evolutionary. His profound main insight, still widely unrecognized, is that answers to all four questions are essential; they are not competing alternatives, but complementary components of a full explanation.

Confusion often arises when information about proximate mechanisms is used to test a proposal about a trait’s adaptive significance. For instance, if fever is a direct result of higher metabolic rate, this would undermine the hypothesis that it is a defense against infection. Data showing that fever results from an increase in the body’s temperature set-point helps to confirm that it is an adaptation (Kluger 1979).

Confusion sometimes arises from failure to recognize that a full evolutionary explanation has two components, the phylogeny of the trait and its adaptive significance. Phylogenetic explanations are sometimes framed as if they...
are alternatives to functional explanations. Usually, the author is trying to argue that apparently maladaptive aspects of a trait result from phylogenetic inertia (Blomberg and Garland 2002). Sometimes this refers to suboptimal traits that are similar to those in ancestral species; the appendix might be an example. Sometimes it refers to a mismatch between a slow-evolving organism and a changing environment. It can also refer to genetic, mechanical, and other constraints that slow or prevent change in a trait. Wisdom teeth offer a good example. Do they cause problems for modern humans because they were previously useful but now are costly? Or, do constraints make reducing the number of molars difficult? Or, is the explanation not a phylogenetic constraint at all, but a change in the developmental environment, with smaller jaws resulting from chewing foods softer than our ancestors ate?

Much interest has attended the ways in which phenotypes are altered by mechanisms that regulate gene expression, some of which can be transmitted for several generations. For instance, in-utero stress imprints genes in ways that increase fetal cortisol receptors (Zhang and Meaney 2010). This is an important and interesting mechanism. Its adaptive significance is a separate question. Is this mechanism an adaptation, shaped by selection to adapt individuals to varying environments? Or, is it an epiphenomenon of some other process?

Challenges associated with considering all possible explanations (Questions 6–9)

Failing to consider all possible hypotheses (Q7) is a general problem in science, one magnified in evolutionary studies because all the alternatives are rarely obvious and some may have intrinsic appeal. Questions about the adaptive significance of apparently maladaptive traits arouse special fascination. For instance, if you had thought fever, cough, vomiting, and anxiety were problems, it is a revelation to realize that they are actually adaptations. Is menstruation also an adaptation? What about vomiting associated with high pressure in the eyeball? How can we tell?

Systematic consideration of all possible hypotheses is the antidote to contentious general debates about studying adaptations. Vociferous critiques e.g. (Gould and Lewontin 1979) have left some with the impression that all attempts to understand the adaptive significance of traits are scientifically illegitimate. This global generalization is the equally incorrect flipside of attempts to find adaptive explanations for everything (Reeve and Sherman 1993; Queller 1995; Alcock 2001). Most physicians and medical researchers are unfamiliar with the history of these controversies (Ruse 2000; Segerstråle 2000) and the many attempts to resolve them (Maynard Smith and Holliday 1979; Reeve and Sherman 1993; Rose and Lauder 1996). This may be just a well; debates about adaptation in general offer little help. Progress is coming thanks to studies of specific questions that consider every possible hypothesis, one by one, with as much objectivity as humans can manage. Examples illustrate the benefits of this approach.

Dogs and cats can synthesize vitamin C, so they rarely get scurvy, but apes have mutations in the gulonolactone oxidase gene so they cannot synthesize vitamin C. The accepted evolutionary explanation has been that our ancestors had plenty of fruit in their diets, so there was no selection to maintain vitamin C synthesis pathways (Jukes and King 1975). This is plausible, but it does not consider the possibility that the mutations offered a selective advantage, perhaps by reducing reactive oxygen production, or by fine tuning stress response regulation via effects on hypoxia inducible factor 1z, or by noncoding effects of the gulonolactone oxidase pseudogene (De Tullio 2010; Johnson et al. 2010). Understanding why humans cannot synthesize vitamin C could be important for clinical recommendations.

Vomiting associated with acute angle glaucoma has been hypothesized to be a specific adaptation to reduce salt in the body, thus reducing eyeball pressure (Wood 2008). It would be remarkable if a rare disorder of adult life could shape such a specific protective response, especially one with little likelihood of actually relieving the condition (Pasca and Nesse 2008). The neglected alternative hypothesis is that vomiting is an epiphenomenon that arises from mechanisms that are more general.

The persistence of genes that cause schizophrenia poses a puzzle many have attempted to solve. Some postulate selective advantages for schizophrenia itself (Allen and Sarich 1988), or an advantage to the group (Stevens and Price 1996). Some propose genetic linkage to a strongly selected locus (Burns 2005). Some attempt to incorporate phylogenetic and proximate mechanisms into a unified explanation (Horrobin 1998). Others offer a more complex explanation, based on the origins of language and cerebral asymmetry, that includes aspects of proximate mechanisms, phylogeny, development, and functional significance in a complex mix (Crow 1997). Progress has been slow, in large part because most of these hypotheses have been considered in isolation from each other. More comprehensive approaches are finally appearing (Polimeni and Reiss 2003; Brüne 2004), just in time to take advantage of new genetic data that will likely answer the question.

An intriguing recent proposal about schizophrenia builds on Haig’s work on mechanisms that advance the competing interests of paternal versus maternal genes (Haig 1993). The weight of an offspring that maximizes...
the genetic interests is slightly higher for the father than the mother; the next child may have a different father, so calories conserved for a subsequent pregnancy give greater benefits to maternal than paternal genes. As sperm form, a process called imprinting turns off some genes that make offspring smaller. The process of egg formation turns off genes that tend to make offspring larger. When these opposing forces balance, all is well, however deficient influences from either side result in offspring substantially larger or smaller than normal. This illustrates the possibility of disease vulnerability arising from competitions that create an arms race between maternal and paternal genomes.

Genes imprinted during gamete production can also influence behaviors that differentially influence maternal and paternal fitness. This combines with evidence for variation in expression of maternally and paternally imprinted genes in different areas of the brain to suggest a creative hypothesis (Badcock and Crespi 2006). Gross imbalance of epigenetic effects that favor the interests of the mother or the father could help to explain the syndromes of schizophrenia and autism, respectively (Crespi et al. 2009). The full argument is complex, and relies on interpreting rapidly increasing new data on imprinted genes, but it illustrates the heuristic value of an evolutionary approach.

The possible adaptive significance of menopause has spurred intense research ever since it was proposed to give a selective advantage because older women may advance their genetic interests more by assisting existing children than by additional direct reproduction with its attendant risks to a mother and her existing children (Williams 1957). Theoretical models and increasingly sophisticated data from hunter-gatherer populations have been brought to bear on this hypothesis, and the related idea that surviving grandmothers increase their grandchildren’s reproductive success (and thus, their own inclusive fitness) (Rogers 1993; Hawkes 2004; Shanley et al. 2007). This line of research has implications for understanding other phenomena, including the duration of the human life-span (Hawkes 2004), even as new data challenge its original formulation (Lahdenperä et al. 2010). However, menopause may turn out to be an epiphenomenon of competition between eggs in the ovary (Reiber 2010), perhaps one that is also influenced by grandmother effects.

The appendix offers another instructive example. Is it just an atavistic remnant? Or is it an adaptation that stores helpful bacteria to repopulate the gut after it has been purged by an infection (Bollinger et al. 2007)? Comparative data bear on the question, but anatomic studies suggest that the trait is difficult even to define with exactness (Fisher 2000). Could the human appendix persist because individuals with a smaller and thinner appendix are more vulnerable to appendicitis (Nesse and Williams 1994). It is peculiar to think that appendicitis could be the selection force that maintains the appendix! I rather doubt that this will turn out to be correct, but it illustrates how the persistence of a trait can conceivably be explained by the very factors that make it disadvantageous.

Six categories of explanation

Bodies have traits that leave them vulnerable to disease for six main reasons (Q8). As listed in Table 1, they are mismatch with the modern environments, co-evolution with pathogens, trade-offs, constraints on natural selection, reproductive success at the expense of health, and protective defenses that are easily confused with diseases (Nesse and Williams 1994; Nesse 2005). Each deserves consideration.

They are not, however, mutually exclusive; multiple factors may contribute to a full explanation (Q9). This makes this line of evolutionary research very different from most proximate research. In experimental studies on mechanisms, evidence for one hypothesis tends to undermine others; DNA either is a double helix, or it is not. For evolutionary questions, multiple answers may be correct. For instance, our vulnerability to atherosclerotic disease arises substantially from exposure to novel aspects of our modern environment (Eaton et al. 1988), but is also probably a product of trade-offs between the advantages of having an immunologically responsive endothelium and the disadvantages of plaque formation (Nesse and Weder 2007). Direct infectious causes may also prove important (Ewald and Cochran 2000). Genetic differences among human subpopulations may also play an important role, especially differences in mitochondrial DNA (Wallace 2005) and ApoE (Finch and Stanford 2004). These factors interact in complex causal networks, and the relative importance of each factor may be different in different populations, or even in different individuals (Q7). Ignoring such complexities is a mistake.

When the breadth of the evolutionary medicine is acknowledged, a narrow focus can be helpful, for instance, on mismatch (Gluckman and Hanson 2006). However, some evolutionary approaches to disease emphasize one of the six factors to the exclusion of others. For instance, a book on evolutionary psychiatry emphasizes the role of novel aspects of modern social environments (Stevens and Price 1996). The role of novel environments is certainly important, but other factors are also. Others approaches emphasize the role of pathogens and co-evolutionary arms races in shaping vulnerability to disease (Ewald and Cochran 2000). There is no doubt...
that many diseases will turn out to have surprising infectious causes, but considering one factor separate from others makes interpretation difficult. Some authors interpret almost all responses as protective defenses. Others, especially geneticists, interpret most phenomena as arising from mutation and drift unless strong evidence forces consideration of other possibilities. It is best, when possible, to consider multiple explanations in concert.

The appropriate initial tests differ depending on what kind of explanation is proposed. Testing a proposal about mismatch begins with a search for evidence that the disorder is more common in modern environments and that it varies in predicted ways with environmental characteristics. Proposals about co-evolution begin by demonstrating an infectious cause, or the cost of defenses against infection. Explanations based on constraints begin by demonstrating mutation rates, path-dependence, or other constraints that limit optimality. Hypotheses about trade-offs consider variations in the trait, observed or hypothetical, and the costs and benefits of deviations from the observed mean. Proposals about reproductive success at the expense of health begin by looking for a reproductive advantage associated with a trait that also causes harm. Finally, proposals about defenses are assessed by demonstrating that they are, indeed, protective and that the mechanisms that regulate their expression respond appropriately to cues indicating the presence of the relevant danger. These methods are only the start of hypothesis testing, but it is worth recognizing that the initial approach differs depending on what kind of explanation is under consideration.

Challenges arising from choosing methods for testing hypotheses (Q10)

Testing evolutionary hypotheses about disease uses the same basic principles as the rest of evolutionary biology, but some of these methods are unfamiliar to many in medicine. A brief review of several is no substitute for full descriptions, but it can alert readers to the range of possibilities.

The first method is easy and powerful; hypotheses must be consistent with modern evolutionary theory (Q10a). A remarkable number of proposed explanations do not qualify. For instance, a noted researcher recently replied to an audience member’s question by saying aging was necessary so the species can evolve faster. Another used the same explanation for why DNA is vulnerable to mutations that cause cancer. Errors arising from such naive notions about group selection remain common in medicine. These errors have nothing to do with sophisticated debates about levels of selection (Dugatkin and Reeve 1994; Keller 1999).

Hypotheses must also be consistent with plausible speeds of selection. For instance, genes for nearsightedness are sometimes thought to be prevalent now because eyeglasses have eliminated their effects on fitness for the past few hundred years. However, even if nearsightedness were selectively neutral, drift would not bring it to a prevalence of over 30% in just a few generations. Epidemiologic data confirm the conclusion; high rates of nearsightedness emerge in a society within one generation after children are first subjected to early schooling (Norn 1997). Relaxed selection is not a viable explanation for myopia.

Hypotheses based on implausibly slow selection can also be ruled out. Some researchers assume that human genomes have not changed in the past 10 000 years; this is contradicted by new evidence for substantial changes in human subpopulations. The spread in the last 5000 years of genes that express lactase in adults is a good example (Tishkoff et al. 2006), as is evidence for selection acting in the past 4000 years on alleles that allow adaptation to high altitudes (Beall 2007). Changes in social structures over the past few thousand years have also given rise to potent selection forces (Cochran and Harpending 2009), as confirmed by evidence that resistance to tuberculosis is greater in populations that have usually lived in urban settings (Barnes et al. 2010).

Mathematical models offer another method unfamiliar to many in medicine. Quantitative descriptions of proposed evolutionary processes have supported and challenged proposed explanations for phenomena such as menopause (Rogers 1993) and senescence (Hamilton 1966; Kirkwood and Rose 1991) and extraordinary behaviors (Grafen 1991) (Q10b). Mathematical models have been especially valuable in infectious disease, where they often correct flawed intuitions (Anderson and May 1979; Bergstrom et al. 2004).

The comparative method is the gold standard for testing evolutionary hypotheses (Harvey and Pagel 1991) (Q10c). Comparisons of similarities and differences among traits can help to establish phylogenetic relationships among species. A related application of the comparative method can address the adaptive significance of traits. The hypothesis that the white coat of arctic hares and polar bears is an adaptation can be tested by examining closely related species that do not live in the Arctic. The hypothesis that higher uric acid levels are selected to as confirmed by evidence that resistance to tuberculosis is greater in populations that have usually lived in urban settings (Barnes et al. 2010).
selection forces or because of close phylogenetic relationships.

Our preoccupation with the woes of a single somewhat peculiar hominid species limits applications of the comparative method in medicine, especially when the trait is only in humans, such as absence of the enzyme uricase. However, the comparative method can test hypotheses about the origins of genetic differences between human subgroups by mapping environmental variations to allele frequencies. Studies of lactase (Tishkoff et al. 2006) and sickle cell disease (Piel et al. 2010) provide exemplars.

Experimental methods (Q10d) can offer powerful tests for evolutionary hypotheses, but the underlying logic is somewhat different from that in proximate science. Instead of varying one factor to observe downstream changes that illuminate a mechanism, experiments to test hypotheses about function tend to disrupt or augment a trait to look for resulting malfunctions and trade-offs. There is nothing new about the method. Physiologists have long extirpated organs to investigate their functions. More recently, gene knockout studies address evolutionary questions about the functional significance of specific genes.

Experimental disruptions are especially effective methods for studying facultative adaptations. For instance, the utility of fever can be demonstrated by the slower recovery times of individuals who take an antipyretic during an infection (Doran et al. 1989; Stephenson 1993; Kluger et al. 1998). However, lack of obvious detrimental effects from blocking a defense does not prove it is just an epiphenomenon – the body has redundant protective mechanisms. The role of temperature can be distinguished from other aspects of inflammation by studying cold-blooded animals. Infected lizards crawl closer to a heat source and increase their body temperature. Those prevented from doing so are more likely to die, thus demonstrating the value of increased body temperature (Kluger et al. 1998). Fever can also be studied by artificial augmentation (Q10e ii). The effects of sauna baths on infection offer an example, and Wagner-Jauregg won the 1927 Nobel Prize for showing that fever induced by malaria slowed the progression of otherwise-fatal syphilis (Brown 2000).

Predictions about regulation mechanisms can test hypotheses about facultative responses. If selection shaped a response to deal with a certain situation, the regulation mechanism should express the response to cues that indicate the presence of that situation. Many examples are too obvious to be interesting. Foreign matter in the respiratory system arouses cough. Overheating sets off sweating. Other examples deserve more study. For instance, if vomiting is a protective response to toxins in the gut, it should be expressed whenever the net benefits are greater than the net costs. Likewise, if social anxiety is useful to protect against attack by dominant others, its intensity should be proportional to the degree of risk present.

Our knowledge about the effects of blocking protective responses has surprising gaps. This is unfortunate because so much of medicine consists of using drugs to relieve symptoms. While every doctor knows the dangers of blocking cough after surgery, much uncertainly attends decisions about treatment of everyday diarrhea and rhinorrhea. Existing studies tend to be purely empirical, without grounding in an evolutionary approach. Findings from new experiments that disrupt defensive responses will have important clinical implications.

Finally, the best test of a hypothesis may be considering the details of its form in light of its proposed function (Q10e). The correspondence is sometimes obvious; eyes were shaped for vision. Only a fool would demand studies comparing the reproductive success of blind and sighted hunter-gatherers to determine the function of the eyes. Human eyebrows are slightly more challenging, but their form is well-suited to directing sweat to away from the eyes; the sufferers of those who shave their eyebrows confirm this function. As is often the case, however, other functions are also important, in this case, social signaling.

Mapping form to function can also be treacherous, however. Some have suggested that white hair is an adaptation that accurately signals age and high status, but it seems more likely to be simply an epiphenomenon of aging melanocytes. Conversely, human female breasts have been viewed as merely for fat storage, but they also likely serve signaling functions (Low et al. 1987). Menstruation is more problematic. It is clearly costly, and many of its characteristics suggest the plausible hypothesis that it might clear pathogens from the reproductive tract (Profet 1993). However, physiologic evidence refutes the proposal; the costs of menstruation are less than those of continual maintenance of the endometrium, and menstruation is more likely to increase than decrease the pathogen load in the reproductive tract (Strassmann 1996).

Global skepticism about explanations that deduce function from form is common but unjustified. Each proposal needs to be assessed on its own merits. The form of a trait can sometimes rule out a proposed function. For instance, the proposal that vomiting is an adaptive response to acute angle glaucoma fails because there is no evidence that it actually lowers intraocular pressure (Pasca and Nesse 2008). The proposal that depression is an adaptation to focus cognition on solving a major problem is worthy but insufficient (Andrews and Thomson 2009); low mood aroused by a life problem may be a different phenomenon from serious clinical depression, many causal pathways can result in depression, rumination can sometimes be pathologic, and depressed mood can have other functions. Nonetheless, the suggestion calls useful
attention to the possible adaptive functions of rumination in certain situations, and it offers a possible, and testable, explanation for why depression arises so often in difficult life situations.

Sometimes a simple correspondence of form to function can be convincing. For instance, about 15% of human breast milk consists of indigestible complex oligosaccharides. This offers a classic evolutionary mystery. Why waste calories making something a baby cannot digest? It turns out that pathogenic bacteria cannot make use of these oligosaccharides, but bifidobacteria, major components of the normal gut flora, can. These indigestible sugars give helpful bacteria a head start (Zivkovic et al. 2010). This is not conclusive proof, but the match of form to function is strong. Further support comes from the timing; synthesis of these oligosaccharides is highest immediately after birth, then they decline. Such evidence pushes this hypothesis far ahead of the competing hypothesis that it is just an epiphenomenon, but work continues. We will know more after comparing the gut flora of babies who nurse at the breast to those fed with sterile breast milk or milk substitutes.

Conclusion

Much of the recent interest in evolutionary applications in medicine comes from attempts to understand traits that leave bodies vulnerable to disease. Such questions are especially fascinating because each is a mystery wanting a solution. It is not surprising that they inspire creative hypotheses. However, creativity and criticism need each other. As Medawar puts it: ‘The most imaginative scientists are by no means the most effective; at their worst, uncensored, they are cranks. Nor are the most critically minded. The man notorious for his dismissive criticism, strenuous in the pursuit of error, is often unproductive, as if he had scared himself out of his own wits – unless indeed his critical cast of mind was the consequence rather than the cause of his infertility’ (Medawar 1969, 58). Darwinian explanations of traits that leave us vulnerable to disease will flourish and advance medicine to the extent that they can maintain a balance between the creative and the critical.

The ten questions posed here are no substitute for knowledge and experience, just as a pilot’s preflight checklist is no substitute for flight training. Nonetheless, they may help to prevent confusion and common errors. Systematically, addressing all ten questions may help stream creativity into productive channels, thereby establishing, ‘a dialog between fact and fancy, the actual and the possible, between what could be true and what is in fact the case…[Science] begins as a story about a Possible World – a story which we invent and criticize and modify as we go along, so that it ends by being, as nearly as we can make it, a story about real life’ (Medawar 1969, 59).

Acknowledgements

Warm thanks to Peter Gluckman and Scott Carroll for encouraging me to write this article, and to Scott Carroll and two anonymous reviewers for very helpful suggestions.

Literature cited


Wiesenfeld, S. L. 1967 Sickle-cell trait in human biological and cultural evolution. development of agriculture causing
increased malaria is bound to gene-pool changes causing malaria reduction. Science 157:1134–1140.


