Testing hypotheses without considering predictions

Dear Sir,

In their recent essay, Poole and Penny[1] defend the view that the ability to engulf bacteria and other food particles—phagocytosis—must have evolved in the eukaryotic lineage as an absolute prerequisite for the origin of mitochondria. That hypothesis, which one could call ''phagotrophy first'', is now almost 40 years old.[2–4] By comparison, the scientific practice of evaluating hypotheses by the measure of how well their predictions fare is only about 70 years old.[5] It turns out that all of the predictions that the phagotrophy first idea ever generated have failed.[6,7] Now Poole and Penny[1] argue that new criteria, instead of predictions, should be used to evaluate the relative merits of hypotheses concerning eukaryote origin. A brief response would seem in order.

Prokaryotes are not vertebrates

As their first main point, Poole and Penny suggest that the preferred null hypothesis for eukaryote origin should be gradual character evolution along long-stem lineages, with extinction of all intermediate lineages except the one that acquired mitochondria.[1] They offer the fossil record of vertebrates as evidence in favour of their case. They suggest that vertebrate phylogeny should serve a role model for our approach to understanding microbial evolution, the prokaryote-to-eukaryote transition in particular. If we had no information at all about microbial genome evolution, one could probably agree that a purely tree-like process similar to the one that we know from the vertebrate fossil record would be the preferred null hypothesis for starters. And if we had positive evidence to indicate that microbial evolution entails exactly the same mechanisms of natural variation as found in vertebrates, with no major additional mechanisms, one could also probably agree with them.

But evidence from a few hundred prokaryotic genomes has shown that the basic mechanics of natural variation in prokaryotes entail the regular reasssortment of genes across higher taxonomic boundaries via lateral gene transfer (LGT).[8–13] For example, the genomes of two sequenced strains of *Escherichia coli*, K12 and O157, differ by about 75,000 nucleotide substitutions, but also by about 30% of their genomes (~2Mb) consisting of genes lacking in the other strain owing to differentially acquired and lost DNA.[14] By contrast, the thought that two individuals of the same vertebrate species might differ by 30% of differentially acquired DNA is absurd. Among prokaryotes, it is not uncommon to find that three individuals of the same species share less than 40% of their genes,[15] or archaebacterial genomes bearing 30% laterally acquired eubacterial genes,[16] or vice versa.[17] Poole and Penny[1] (p. 80) state that they do not accept evidence from patchy gene distributions as evidence for LGT, citing the *Thermotoga* example.[17] Presumably they prefer differential loss to explain such patterns (the only alternative to LGT for explaining patchy gene distributions), but that reasoning leads inexorably to burgeoning ancestral genome sizes, or the ''genome of Eden'' as Ford Doolittle and colleagues[18] have called it. Using the genome of Eden constraint on ancestral genome size distributions, it can be shown by gene-tree-independent means that most, it not all, prokaryotic gene families have experienced at least one LGT event—at the bare minimum—during microbial evolution.[19] It is safe to say that similar LGT rates among vertebrates are unlikely to ever be found. Because considerable mechanistic differences do exist between microbial evolution and vertebrate evolution, it seems unwise to suggest that the latter should serve as a mechanistic model for the former.

Deep phylogeny is not easy

Poole and Penny's second main point is that models for the origin of mitochondria evoking a prokaryotic host should, in their opinion, reflect an origin of eukaryotic informational genes from within the archaebacteria, rather than as a sister lineage to them. Notwithstanding the circumstance that various analyses do indeed recover such a result,[6] there is the looming issue of how accurate deep phylogenetic trees are. Penny et al.[20] for example, offered strong evidence to suggest that ''it should be difficult or impossible to recover trees accurately after divergences of more than about 300–400 million years. [. . .] Under these standard models, there is no justification for expecting correct results for ancient divergences'' and most cogniscenti would probably agree. But even if we assume that molecular phylogenetics works flawlessly for ancient divergence, the mass lineage extinction corollary among archaebacterial stem lineages that Poole and Penny[1] evoke to save their phagotrophy first model (their Fig. 2A) saves the prokaryote-host models that they disfavour equally: one only has to assume that many intermediate (archaebacterial) lineages went extinct. We would prefer to avoid models that come pre-equipped with dead-end branches, favouring the view that deep phylogenetics for individual genes is a tenuous undertaking,[6] for exactly the reasons that Penny et al.[20] have suggested.

Of endosymbionts and mitochondria

Poole and Penny's third main point[1] builds upon the many known examples of prokaryotic and eukaryotic endosymbionts that live within eukaryotic cells. From that they conclude that phagocytosis is a conditio sine qua non for the origin of mitochondria. In this sense, their argument is identical to that of Cavalier-Smith,[21] and both arguments suffer from the

BioEssays 29:500–503, © 2007 Wiley Periodicals, Inc.

same flaw, as a brief consideration reveals. The interesting examples of endosymbionts living within eukaryotic cells that Poole and Penny recite can be expanded ad infinitum. The proteobacterial endosymbionts of insects such as *Wigglesworthia*[22] and *Wolbachia*,[23] the methanogenic endosymbionts of anaerobic ciliates,[24] the purple endosymbionts of the ciliate *Strombidium*,[25] the sulfur-metabolizing symbionts of clam gills,[26] the chemosynthetic endosymbiont consortia of gutless tubeworms,[27] endosymbionts that live within the endoplasmic reticulum of diatoms,[28] the cyanobacterial endosymbionts of sponges,[29] and endosymbionts with genomes smaller than some plastid genomes[30] constitute a few additional examples. Any literature search with the query ''endosymbiotic bacteria'' will return hundreds more published examples, as will Buchner's 1953 book.[31] And let us furthermore grant the existence of similarly abundant intracellular endosymbioses since the origin of phagocytosis over geological time.

The point is this: although it is unquestionably true that phagocytosis promotes the establishment of intracellular endosymbionts, the issue is not how common endosymbionts within phagotrophs are. The issue is the origin of mitochondria. Against the backdrop of all eukaryote individuals harbouring intracellular endosymbioses over time since the origin of phagocytosis (a very large number of endosymbionts), we can vividly contrast the observation that mitochondria arose from proteobacteria only once—one single time—in all of Earth's history among the cells that we know (the only ones that we have to explain). Phagocytosis promotes the establishment of intracellular symbioses, but it has no influence whatsoever on the rate at which mitochondria arose from intracellular endosymbionts. Like others,[32] Poole and Penny[1] fail to distinguish between endosymbionts and organelles; the main difference is a protein import apparatus,[33,34] which is present in organelles but lacking in endosymbionts.

Two examples of prokaryote hosts harboring prokaryotic endosymbionts[35,36] show that phagocytosis is not an absolute prerequisite for the establishment of intracellular endosymbiosis. Furthermore, phagocytosis cannot even be linked in a causal manner to the origin of mitochondria since, if phagocytosis were in some way rate-limiting for the origin of organelles, then organelles descended from bacteria would be as common and as diverse as endosymbionts themselves are. But the observation is that mitochondria only arose once in all of evolution.[6,7,37] Phagocytosis thus influences the rate at which endosymbionts arise in evolution, but it has no bearing upon the rate at which mitochondria arise in evolution, because the latter rate (with a frequencey of one event per 4 billion years) is not correlated (nor can it be correlated, because of its singularity) with the former.

Thus, a central conclusion from their paper that ''all data point to a mechanism of cell engulfment being a prerequisite for the origin of mitochondria'' is merely Cavalier-Smith's 2002 opinion[21] disguised as the result of a logical endeavour. The flaw in their phagocytosis argument is that they assume a particular process (phagocytosis) to be the 'known cause' to an event (the singular origin of mitochondria), while in fact it is the 'known cause' to different events (the myriad origin of endosymbionts among phagocytosing cells) and is related to their explanandum (the origin of mitochondria) solely via their assumption. By analogy, if we try to infer the mechanism of a break-in bank robbery by observing how people normally withdraw money from their accounts, we will inevitably (and erroneously) conclude that the thief simply walked in the front door during office hours and made a legal withdrawl. Because there are examples of endosymbionts within prokaryotic hosts,[35,36] but no examples of cells that became phagocytotic without the evolutionary participation of mitochondria,[6,7] the logical faux pas of describing past processes in terms of unobserved (imaginary) mechanisms, as opposed to envoking known quantities, as Occam's razor would precribe, sits in Poole and Penny's[1] corner, not in the prokaryotic host corner.

If a premise is untrue...

Poole and Penny's final conclusion is that theories for the origin of mitochondria that entail a prokaryotic host rather than a eukaryotic (phagocytosing) host should be avoided,[1] because ''no archaea have been shown to carry bacterial endosymbionts'' (p. 80). That argument entails an evident flaw: they assume a priori that the alternative hypothesis that they purport to be testing is untrue. If the host that acquired the mitochondrion was an archaebacterium, as some of us are suggesting, then eukaryotes are therefore such examples of archaea that host endosymbionts, and then there is no shortage of examples at all. We could state that ''No mammals have been shown to have wings'' but the truth of that statement hinges upon the truth of the premise that bats are not mammals. One cannot draw upon the assumed truth of a premise as evidence in favour of its strength.

Thus, while it remains within the realm of the (n.b.) imaginable that the host that acquired the mitochondrion was a phagotroph, Poole and Penny[1] have failed to offer any evidence in support of that view, other than statements of their opinion that it was so. Nor have they supplied a logical foundation in support of that view, nor do the new criteria that they formulate to test hypotheses for eukaryote origin improve matters for their case, on the contrary. Nonetheless, they are fully entitled to their *assumption* that their hypothesis is correct.

Conclusion

Most of us would probably agree that a theory is a set of mutually compatible hypotheses that generates testable predictions while accounting for available observations.

Observations that are at odds with the predictions or otherwise left unaccounted under the theory then have to be explained with the help of corollary assumptions, each of which adds a quantum of weight to the foundation of the theory. If that weight becomes too severe, and if a competing alternative theory that generates testable predictions can account for the same (or preferably more) observations as the old one while requiring fewer corollaries, it is to be preferred until something better comes along, and so forth. But Poole and Penny's "preferred" null hypothesis[1] fails to account directly i) for the ubiquity of mitochondria among eukaryotes, ii) for the common ancestry of mitochondria, hydrogenosomes and mitosomes, iii) for the lack of phagocytotic prokaryotes, and iv) for the circumstance that far more eukaryotic genes reflect a eubacterial ancestry than reflect an archebacterial ancestry. A competing alternative hypothesis entailing a prokaryotic host[38,39] can account directly for those observations, while also mechanistically accounting for the origin of the nucleus–cytosol compartmentation.[40]

But neither microbial genome comparisons nor congruence between predictions and observations stand in the foreground of Poole and Penny's essay.[1] On the bottom line, they argue that it is preferrable to adopt new criteria for hypotheses-testing that circumvent the currency of predictions and corollaries, so that particular *ideas* about eukaryote origin remain constant while the *criteria* that uphold them are adapted to account for new observations. Such can hardly have been their aim or intent, but it is the end result of their argumentation—and that needs to be said.

References

1. Poole AM, Penny D. 2007. Evaluating hypotheses for the origin of eukaryotes. BioEssays 29:74–84.
2. de Duve C. 1969. Evolution of the peroxisome. Annals NY Acad Sci 168:369–381.
3. Stanier Y. 1970. Some aspects of the biology of cells and their possible evolutionary significance. Symp Soc Gen Microbiol 20:1–38.
4. Cavalier-Smith T. 1975. The origin of nuclei and of eukaryotic cells. Nature 256:463–468.
5. Allen JF. 2001. Bioinformatics and discovery: Induction beckons again. BioEssays 23:104–107.
6. Embley TM, Martin W. 2006. Eukaryotic evolution, changes and challenges. Nature 440:623–630.
7. van der Giezen M, Tovar J, Clark CG. 2005. Mitochondrion-derived organelles in protists and fungi. Int Rev Cytol 244:175–225.
8. Doolittle WF. 1999. Phylogenetic classification and the universal tree. Science 284:2124–2128.
9. Eisen J. 2000. Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. Curr Opin Genet Dev 10:606–611.
10. Ochman H, Lawrence JG, Groisman EA. 2000. Lateral gene transfer and the nature of bacterial innovation. Nature 405:299–304.
11. Koonin EV, Makarova KS, Aravind L. 2001. Horizontal gene transfer in prokaryotes: Quantification and classification. Annu Rev Microbiol 55:709–742.
12. Canchaya C, Fournous G, Chibani-Chennoufi S, Dillmann ML, Brussow H. 2003. Phage as agents of lateral gene transfer. Curr Opin Microbiol 6:417–424.
13. Doolittle WF, Bapteste E. 2007. Pattern pluralism and the tree of life hypothesis. Proc Natl Acad Sci USA (In press).
14. Perna NT, Plunkett G, Burland V, Mau B, Glasner JD, et al. 2001. Genome sequence of enterohaemorrhagic Escherichia coli O157:H7. Nature 409:529–533.
15. Welch RA, Burland V, Plunkett G III, Redford P, Roesch P, et al. 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic Escherichia coli. Proc Natl Acad Sci USA 99:17020–17024.
16. Deppenmeier U, Johann A, Hartsch T, Merkl R, Schmitz RA, et al. 2002. The genome of Methanosarcina mazei: Evidence for lateral gene transfer between bacteria and archaea. J Mol Microbiol Biotechnol 4:453–461.
17. Nelson KE, Clayton RA, Gill SR, Gwinn ML, Dodson RJ, et al. 1999. Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. Nature 399:323–329.
18. Doolittle WF, Boucher Y, Nesbo CL, Douady CJ, Andersson JO, et al. 2003. How big is the iceberg of which organellar genes in nuclear genomes are but the tip? Philos Trans R Soc Lond B 358:39–57.
19. Dagan T, Martin W. 2007. Ancestral genome sizes specify the minimum rate of lateral gene transfer during prokaryote evolution. Proc Natl Acad Sci USA 104:870–875.
20. Penny D, McComish BJ, Charleston MA, Hendy MD. 2001. Mathematical elegance with biochemical realism: The covarion model of molecular evolution. J Mol Evol 53:711–723.
21. Cavalier-Smith T. 2002. The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol 52:297–354.
22. Zientz E, Dandekar T, Gross R. 2004. Metabolic interdependence of obligate intracellular bacteria and their insect hosts. Microbiol Mol Biol Rev 68:745–770.
23. Wu M, Sun L, Vamathevan J, Riegler M, Deboy R, et al. 2004. The genome sequence and evolution of the reproductive parasite *Wolbachia pipientis w* Mel: A streamlined α-Proteobacterium massively infected with mobile genetic elements. PLoS Biology 2:327–341.
24. Embley TM, Finlay BJ. 1993. Systematic and morphological diversity of endosymbiotic methanogens in anaerobic ciliates. Antonie Van Leeuwenhoek 64:261–271.
25. Fenchel T, Bernard C. 1993. A purple protist. Nature 362:300.
26. Duplessis MR, Ziebis W, Gros O, Caro A, Robidart J, et al. 2004. Respiration strategies utilized by the gill endosymbiont from the host lucinid *Codakia orbicularis* (Bivalvia: Lucinidae). Appl Environ Microbiol 70:4144–4150.
27. Woyke T, Teeling H, Ivanova NN, Huntemann M, Richter M, et al. 2006. Symbiosis insights through metagenomic analysis of a microbial consortium. Nature 443:950–955.
28. Schmid AMM. 2003. Endobacteria in the diatom *Pinnularia* (Bacillariophyceae). I. "Scattered ct-nucleoids" explained: DAPI-DNA complexes stem from exoplastidial bacteria boring into the chloroplasts. J Phycol 39:122–138.
29. Thacker RW. 2005. Impacts of shading on sponge-Cyanobacteria symbioses: A comparison between host-specific and generalist associations. Integr Comp Biol 45:369–376.
30. Nakabachi A, Yamashita A, Toh H, Ishikawa H, Dunbar HE, et al. 2006. The 160-kilobase genome of the bacterial endosymbiont *Carsonella*. Science 314:267.
31. Buchner P. 1953. Endosymbiose der Tiere mit pflanzlichen Mikroorganismen. Birkhäuser, Basel. 771 pp.
32. Yoon HS, Reyes-Prieto A, Melkonian M, Bhattacharya D. 2006. Minimal plastid evolution in the *Paulinella* endosymbiont. Curr Biol 16:R670–R672.
33. Cavalier-Smith T, Lee JJ. 1985. Protozoa as hosts for endosymbioses and the conversion of symbionts into organelles. J Protozool 32:376–379.
34. Theissen U, Martin W. 2006. The difference between endosymbionts and organelles. Curr Biol 16:R1016–R1017.
35. von Dohlen CD, Kohler S, Alsop ST, McManus WR. 2001. Mealybug β-proteobacterial endosymbionts contain γ-proteobacterial symbionts. Nature 412:433–436.
36. Wujek DE. 1979. Intracellular bacteria in the blue-green-alga *Pleurocapsa minor*. Trans Am Micros Soc 98:143–145.

37. Dolezal P, Likic V, Tachezy J, Lithgow T. 2006. Evolution of the molecular machines for protein import into mitochondria. Science 313:314–318.
38. Martin W, Müller M. 1998. The hydrogen hypothesis for the first eukaryote. Nature 392:37–41.
39. Dagan T, Martin W. 2006. The tree of one percent. Genome Biol 7:118 [7 pages].
40. Martin W, Koonin EV. 2006. Introns and the origin of nucleus-cytosol compartmentation. Nature 440:41–45.

Tal Dagan and William Martin
Institute of Botany, University of Düsseldorf,
40225 Düsseldorf, Germany
E-mail: w.martin@uni-duesseldorf.de

DOI 10.1002/bies.20566
Published online in Wiley InterScience (www.interscience.wiley.com).