ASPM and mammalian brain evolution: A case study in the difficulty in making macroevolutionary inferences about gene-phenotype associations

Stephen H. Montgomery[1], Nicholas I. Mundy[2] & Robert A. Barton[3]

[1] Dept. of Genetics, Evolution & Environment, University College London
[2] Dept. of Zoology, University of Cambridge
[3] Dept. of Anthropology, University of Durham

Identifying the genetic basis of adaptive phenotypes can be a significant step towards understanding how that phenotype evolved. With the increased availability of interspecific molecular data one approach to uncover such genes has been to search for signatures of adaptive evolution at the molecular level. Many analyses have adopted a candidate gene approach, focusing on genes with important developmental roles. One such candidate gene is *ASPM*, which is involved in neurogenesis and associated with major neurological disorders [1]. The molecular evolution of *ASPM* has been investigated for a decade (Table S1), under the hypothesis that it contributes to primate brain evolution. A recent study by Xu et al. [2] extends the taxonomic scope by demonstrating that *ASPM* evolved adaptively in cetaceans. However, descriptive studies of patterns of selection are now being supplanted by those that explicitly test for gene-phenotype associations. Using such an approach we find that Xu et al.'s conclusion that *ASPM* is linked to increases in cetacean EQ, a measure of relative brain size, is not supported. We highlight developments in the analysis of molecular data and phylogenetic methods that are capable of resolving major issues in functional gene-phenotype co-evolution

One approach to making gene-phenotype associations is to test for shifts in selection pressure acting on a gene in taxa that display the phenotype of interest. This frequently involves comparing estimates of *dN/dS*, a measure of the strength of selection acting on a protein coding gene, using a range of tests implemented in software such as PAML (Table S2) [3]. The results of these tests can be influenced by the nature of the data and, in particular, require sufficient evolutionary variation to make reliable estimates. Data with few substitutions or from a restricted number of taxa can lead to spurious results. These effects are evident in Xu *et al*.'s analysis.

First, they suggest that a high proportion of branches in the cetacean phylogeny have an elevated *dN/dS*, which they interpret as evidence of increased positive selection but do not perform explicit tests of this hypothesis. Further analysis (Supplementary Information) suggests that none of these is significantly greater than one, the threshold for rejecting neutral evolution. The apparent elevation in dN/dS is likely influenced by the low number of substitutions on short branches. This problem is particularly strong for cetaceans, which have low substitution rates [4]. Second, it is suggested that positive selection is limited to mammalian orders with high EQs. However, this result is likely to be due to a sampling bias, and inclusion of further taxa provides evidence for positive selection across mammals (Supplementary Information). Identifying robust shifts in selection pressure clearly requires both adequate and even sampling, and sufficient numbers of substitutions.

A related method involves testing for shifts in the selection acting on a gene and changes in the associated phenotype along a subset of branches in a phylogeny. This method is particularly useful when applied to novel, or discrete traits, but has also been applied to continuously variable, quantitative traits. This can lead to two problems; first, identifying the branches which show high rates of phenotypic evolution, and second, applying models of molecular evolution which assume episodic positive selection in the presence of pervasive positive selection. A previous study on *ASPM* suggested an association between episodic positive selection and branches showing major increases in cortical volume in primates, identified using parsimony based ancestral state reconstructions [5]. However, closer analysis revealed this result was not robust, as positive selection was not episodic but pervasive, and the identification of key branches was not supported by alternative methods [6]. Xu *et al*. suggest an association between high rates of evolution and major increases in cetacean relative brain size but do not explicitly test for phenotypic shifts. Instead, they rely on previous assumptions about cetacean evolution to highlight key branches. Recent comparative analyses unfortunately suggest these assumptions are not valid [7]. Furthermore, their results demonstrate positive selection was again pervasive, and not limited to a subset of branches. Hence, although this approach may be valid for some phenotypes care is needed on both the phenotypic and molecular side of the analyses. Methods are available that explicitly identify phenotypic rate shifts [8] and, combined with tests for episodic vs. pervasive positive selection, robust tests for gene-phenotype association can be performed in some situations.

Comment [SM1]: does this unfortunately soften the blow? I see Nick's point about criticizing them for not knowing something they couldn't have known, but its hard to find (and also to fit in) another reference as the cetacean literature seems just to repeat the same assumptions over and over again without really testing them. But the main point is that Xu et al didn't test them either when they could/should have done. Not sure whats best here…

Comment [NM2]: It will seem unfair to criticize them solely on the basis of an unpublished study, so important to include another reference here.

If positive selection acting on a locus was pervasive and the phenotype did not evolve in a punctuated manner, a potentially more relevant approach is to test for correlated rates of gene and phenotypic evolution across the whole phylogeny. Several methods have now been proposed to perform such analyses [6, 9-10], and a handful of studies have found evidence for macroevolutionary gene-phenotype associations. For example, one method that has been applied to *ASPM* is to test for a significant regression between the selection pressure acting on a gene during the descent of each species (measured by root-to-tip *dN/dS*) and alternative phenotypes along branches of the phylogeny[6]. Using this approach selection on *ASPM* has been linked to absolute brain mass, and in particular neonatal brain mass, in anthropoid primates [6]. This result is supported by a significant association being found in two largely independent datasets representing both increases and decreases in brain mass [6,11], and is consistent with the hypothesis that selection on *ASPM* may contribute to the evolution of neurogenic output.

Explicit hypothesis testing is challenging but clearly favourable when arguing for a gene-phenotype association at a macro-evolutionary level where comparative functional tests may not be forthcoming. Careful planning is required to ensure maximum statistical power in such analyses, for example by targeting the collection of genetic data according to the availability of phenotypic data when the latter is a restrictive commodity. This is clearly an issue with brain volume data. The overlap between Xu *et al.*'s genetic data and cetacean brain size data is incomplete, nevertheless one can still test hypotheses while acknowledging this caveat. When the available data are used to test for a macroevolutionary association between selection on *ASPM* and either EQ or absolute brain size, no significant association is found (EQ: $t_9 = 0.445$, $p = 0.667$; brain mass: $t_9 = -0.741$, $p = 0.478$) (Supplementary Information). We therefore find no support for an association between *ASPM* and cetacean brain size either based on the patterns of positive selection within cetaceans or across mammals, or through explicit hypothesis testing.

This absence of evidence does not of course rule out the possibility that *ASPM* does indeed play some role in cetacean brain evolution. Xu *et al*. clearly demonstrate that *ASPM* evolved adaptively in cetaceans, and patterns of evolution in primates are suggestive of a link between *ASPM* and brain mass raising the possibility that *ASPM* has a conserved role in the mammalian brain evolution. Explicit tests using

comparative methods, combined with functional data, are necessary to assess this hypothesis.

The methodology for such tests is in its infancy and further developments are required. In addition to poor overlap between genetic and phenotypic datasets one can envisage several other limitations. For example, if selection is restricted to a subset of sites or domains the signal of a gene-phenotype association could be lost when using gene-wide *dN/dS* values. Should we then perform association tests on functional domains, or is a sliding-window analysis across a locus desirable? If phenotypic reversals are common the signal could again be lost as *dN/dS* may increase during both increases and decreases of a phenotypic trait [6], is it possible to account for such effects? For polygenic traits how do we detect real associations with genes that are only targeted by selection intermittently? Beyond candidate genes do we have sufficient power to perform genome-wide scans for macroevolutionary phenotypic associations? And beyond protein coding genes, what tests can be applied to promoter regions or levels of gene expression? The development of new methods may begin to offer answers to these questions [9-10, 12].

Xu *et al.*'s study of the evolution of *ASPM* in cetaceans is a welcome addition to a field frequently mired by a narrow focus on the singular case of human brain evolution. Furthermore, it raises important questions about the genetic basis of complex and convergent phenotypes. However, the issues discussed above limit the conclusions derived regarding the phenotypic relevance of selection on *ASPM* in cetaceans. These problems are frequently found in similar studies and we highlight them here only because they need to be addressed if we are to move beyond the descriptive phase of comparative adaptive genetics to one capable of applying powerful statistical tests to gene-phenotype associations.

Acknowledgments

We are grateful to Guang Yang for providing the cetacean *ASPM* data, to three anonymous reviewers and Hans Heesterbeek for constructive criticism.

References

1. Ponting, C. & Jackson, A.P. 2005. Evolution of primary microcephaly genes and the enlargement of primate brains. *Curr. Opin. Genet. Dev.* 15 (3), 241-248.
2. Xu, S., Chen, Y., Cheng, Y., Yang, D., Zhou, X., Xu, J., Zhou, K. & Yang., G. 2012 Positive selection at the ASPM gene coincides with brain size enlargements in cetaceans. *Proc. Roy. Soc. B* 279 (1746), 4433-4440.
3. Yang, Z. 2007. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol*. 24 (8), 1586-1591.
4. McGowen, M.R., Grossman, L.I. & Wildman, D.E. 2012 Dolphin genome provides evidence for adaptive evolution of nervous system genes and a molecular rate slowdown. *Proc. Roy. Soc. B* 279 (1743), 3643-3651.
5. Ali, F. & Meier, R. 2008. Positive selection in *ASPM* is correlated with cerebral cortex evolution across primates but not with whole-brain size. *Mol. Biol. Evol*. 25 (11), 2247-2250.
6. Montgomery, S.H., Capellini, I., Venditti, C., Barton, R.A. & N.I Mundy. 2011 Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. *Mol. Biol. Evol.* 28 (1), 625-638.
7. Montgomery, S.H., Geisler, J.H., McGowen, M.R., Fox, C., Marino, L. & Gatesy, J. 2013. The evolutionary history of cetacean brain and body size. *Evolution, in press.*
8. Venditti, C., Meade, A. & Pagel, M. 2011. Multiple routes to mammalian diversity. *Nature* 479, 393-396.
9. Lartillot, N. & Poujol, R. 2011. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. *Mol. Biol. Evol.* 28 (1), 729-744.
10. O'Connor, T.D. & Mundy, N.I. 2009. Genotype-phenotype associations: substitution models to detect evolutionary associations between phenotypic variables and genotypic evolutionary rate. *Bioinformatics* 25 (12), i94-i100.
11. Montgomery, S.H. & N.I Mundy. 2012 Evolution of ASPM is associated with both increases and decreases in brain size in primates. *Evolution* 66 (3), 927-932.
12. Dunn, C.D., Luo, X. & Wu, Z. 2013. Phylogenetic analysis of gene expression. *Integr. Comp. Biol.* Online early, doi: 10.1093/icb/ict068.