LETTER

Positive selection in *ASPM* is correlated with cerebral cortex evolution across primates but not with whole brain size

Submission to Letters section of *Molecular Biology & Evolution*

*Farhan Ali*†‡ & *Rudolf Meier*†

*Department of Psychology National University of Singapore, Singapore*

†Department of Biological Sciences, National University of Singapore, Singapore

‡To whom correspondence should be addressed: Department of Biological Sciences, National University of Singapore, S2 #02-01, 14 Science Drive 4, Singapore 117543.

Tel: +65 6516 2968. Fax: +65 6779 2486. Email: farhan.ali@alumni.nus.edu.sg

Key words: *ASPM*, cerebral cortex evolution, positive selection, primates, PAML, maximum likelihood

Title length: 117 characters
Abstract Length: 161 words
Total Length of Text: 13, 344 characters with spaces
Page Requirements: 3 journal pages for main text; 0.5 for figures and tables
Number of References: 31
Abstract

The rapid increase of brain size is a key event in human evolution. *ASPM* (abnormal spindle-like microcephaly associated) is discussed as a major candidate gene for explaining the exceptionally large brain in humans but *ASPM*’s role remains controversial. Here we use codon-specific models and a comparative approach to test this candidate gene that was initially identified in *Homo*-chimp comparisons. We demonstrate that accelerated evolution of *ASPM* (ω = 4.7) at 16 amino acid sites occurred in nine primate lineages with major changes in relative cerebral cortex size. However, *ASPM*’s evolution is not correlated with major changes in relative whole brain or cerebellum sizes. Our results suggest that a single candidate gene such as *ASPM* can influence a specific component of the brain across large clades through changes in a few amino acid sites. We furthermore illustrate the power of using continuous phenotypic variability across primates to rigorously test candidate genes that have been implicated in the evolution of key human traits.
Mutations in *ASPM* are responsible for a severely reduced brain size with no other significant abnormality (primary microcephaly) in a clinical sample of humans (Bond et al. 2002). Comparative study of sequence evolution limited mostly to humans and apes revealed that this gene has an accelerated rate of evolution in the *Homo* lineage (Zhang 2003; Evans et al. 2004), with *ASPM* possibly affecting brain size through controlling the spindle assembly during neural cell division (Fish et al. 2006). However, *ASPM*’s role as a candidate gene for brain size has recently been challenged based on gene expression studies (Kouprina et al. 2005), strong homology to genes not associated with the brain (Ponting 2006), and a lack of correlation between *ASPM* haplotypes and normal human brain size variability (Rushton, Vernon, and Bonns 2006; Woods et al. 2006; Dobson-Stone et al. 2007; Thimpson et al. 2007). One avenue for addressing such controversies surrounding candidate genes is through employing the comparative method (Goodman, Grossman, and Wildman 2005) by testing whether sequence evolution of candidate genes is correlated with quantitative phenotypic changes across a large clade. Such tests can now rely on recently developed techniques in evolutionary genetics that allow for detecting positive selection in specific codons as opposed to whole genes (Yang and Nielsen 2002; Zhang, Nielsen, and Yang 2005). Here we use these approaches to test whether changes in brain size found across primates are correlated with molecular evolution of *ASPM*.

We sequenced the two large exons of *ASPM* (exons 3 and 18; 70% of the transcribed ASPM protein) for 23 primate species to complement existing *ASPM* data for 11 species from GenBank (supplementary table S1; Supplementary Material online). We chose these exons because they contain most of the mutations that cause human primary microcephaly (Bond et al. 2002), have elevated rates of gene-average \( \omega \) in humans (Zhang 2003; Evans et al. 2004), and encompass the main functional sites of the protein (Ponting and Jackson 2005). We then identified those primate lineages that had major changes in relative whole brain, cerebral cortex, and cerebellum size. Due to
the lack of phenotypic data for many primate species, only a subset of species could be analyzed in these analyses (whole brain: 28 spp.; cerebral cortex: 15 spp.; cerebellum: 15 spp.). We find that nine primate lineages have major changes in relative cerebral cortex size, the main phenotype of interest (fig. 1). These clades deviated by one or more standard deviation from the mean of change as revealed by squared change parsimony (Maddison 1991). This cut-off is biologically meaningful, because it can translate up to 300% change in absolute or >10% change in relative cerebral cortex volumes. We used these lineages as “foreground branches” in the branch-site Model A (Yang and Nielsen 2002; Zhang, Nielsen, and Yang 2005) while all other branches were treated as background branches. Branch-site model A uses the maximum likelihood approach of model-fitting to detect codon-specific positive selection by allowing \( \omega \) [the ratio of non-synonymous changes per non-synonymous sites \((d_N)\) to synonymous changes per synonymous sites \((d_S)\)] to vary across codon-positions as well as across foreground and background lineages. We find that a model of positive selection using foreground branches with major changes in relative cerebral cortex size explained the data significantly better than the null model of no positive selection \((2\Delta\ell = 14.67, P < 0.001\) as tested by a chi-square test with df = 1; table 1). The branch-site model detected a high level of positive selection for 16 codons \((\omega = 4.70)\). Specifically, 77% of 27 changes in positively-selected sites occurred in the foreground branches although these branches contributed only 29% of all branches and only 42% of all codon changes across the tree (two-tailed exact binomial tests using expected proportions of 29% and 42%, both \(Ps < .0001\)).

We used several approaches to rigorously test the robustness of our results. First, in order to rule out an undue influence of the high \( \omega \) previously found in the \textit{Homo} lineage, we excluded \textit{Homo sapiens} from the foreground but our results remained highly significant \((2\Delta\ell = 13.78, P < 0.001)\). We similarly carried out eight additional analyses where we excluded each of the remaining eight foreground branches. All \(2\Delta\ell\)
remained significant, indicating that no single branch was driving the results (supplementary table S2). Second, we randomly selected nine branches among the background branches in fig. 1 (Model B). The model did not explain the data significantly better than the null model of no positive selection ($P$s > .05). Third, we tested for the specificity of the evolutionary correlate of *ASPM* by correlating the gene’s evolution with relative whole brain size as well as the size of the cerebellum, a major sub-cortical brain component not known to have *ASPM* expression. We again examined a model of positive selection in *ASPM* whereby foreground branches had major changes in either of these structures (1 or more s.d.). These models did not explain the data significantly better than the null model (all $P$s > .05; table 1); i.e., positive selection in *ASPM* is only significantly correlated with cerebral cortex size, but not with relative whole brain or cerebellum sizes.

Our result provides strong evidence that the single gene *ASPM* is associated with major changes in relative cerebral cortex size across primates. It thus questions the validity of recent reviews that implicated *ASPM* in the brain size expansion of humans only (Ponting and Jackson 2005; Woods, Bond and Enard et al. 2005). Particularly striking is the result that only major changes of cerebral cortex size and not major changes in whole brain or cerebellum size are associated with positive selection in *ASPM*. This is consistent with an expression report indicating that *ASPM*’s expression is limited to the cerebral cortex of the brain (Bond et al. 2002). Our findings stand in contrast to recent null findings correlating *ASPM* genotypes with human brain size variation. Those studies used the relatively imprecise phenotypic trait of whole brain instead of cerebral cortex size (Rushton, Vernon, and Bonns 2006; Woods et al. 2006; Thimpson et al. 2007). Although previous studies have shown that parts of the brain scale strongly with one another and especially with whole brain (e.g., Finlay and
Darlington 1995), evidence here suggests that different brain parts still have their own evolutionary and functional differentiation with unique genetic bases (Barton 1999). Our current study thus addresses the nature of brain evolution and we recommend that future neuroimaging and genetic studies should be more specific by examining associations between *ASPM* variants and cerebral cortex size.

Previous large-scale analyses have extended our understanding of brain and cognition genes at the genomic level, showing that most brain and cognition genes are under strong negative selection (Shi, Bakewell, and Zhang 2006; Yu et al. 2006; Wang et al. 2007). These studies, however, used a different technique for detecting positive selection (gene-average ω). The same technique applied to our *ASPM* data identifies only four primate branches with gene-average ω > 1.0 and only two clades with major changes in cerebral cortex size have ω > 1.0 (supplementary fig. S1). Overall, these results suggest that future research should use codon-specific models to investigate positive selection in additional genes associated with brain and cognition phenotypes. Moreover, these models are more biologically relevant given that active sites in proteins comprise only a few codons. Finally, our study documents the power of utilizing the phenotypic variability across primates for testing candidate genes that are initially identified in clinical *Homo sapiens* samples and/or in *Homo-Pan* comparisons. Such an approach has been recently applied to various taxa including primates (Kelly and Swanson 2008), but has yet to be expanded to brain and cognition genes for which the relevant phenotypes may be continuous in nature.

**Methods**
We obtained primate tissues and DNA from the Singapore Zoological Gardens and the Coriell Institute, USA. Whole genomic DNA was extracted using the phenol-chloroform method. We used standard PCR recipes and newly-designed primers to target exons 3 and 18 of the gene \textit{ASPM}. New \textit{ASPM} data for 23 primate species were generated in this study by sequencing the amplified fragments in both directions. Eleven published primate sequences from GenBank were added (supplementary table S1) to yield a total alignment of about 7kb for 34 species. The dataset is over 80\% complete in terms of sequence data for all sites. To correlate brain size evolution with changes in \textit{ASPM}, we used literature data on body mass, whole brain, cerebral cortex, and cerebellum sizes (Stephan, Frahm, and Baron 1981; Harvey and Clutton-Brock 1985; Zilles and Rekhamper 1988; Smith and Jungers 1997). Whole brain was corrected for body size using the encephalization quotient (Jerison 1961). As for cerebral cortex and cerebellum volumes, we followed previous authors by taking the ratio of the respective brain part to the rest of the brain and log-transforming the data (Kudo and Dunbar 2001). The primate tree was reconstructed using a supermatrix comprising 71 genes (15 mitochondrial and 56 nuclear genes). The sequence data were either newly sequenced or obtained from Genbank, yielding a dataset of 98036 aligned bps. Bayesian analyses (Ronquist and Huelsenbeck 2003) utilized the GTR + \gamma + invariant sites model as recommended by Modeltest (Nylander 2004). The analysis yielded a phylogeny with all clades having a posterior probability of 1.00. The phenotypic measures were then mapped onto a rooted tree using squared change parsimony (Maddison 1991).

We applied the maximum likelihood branch-site Model A to detect correlations between positive selection in \textit{ASPM} and phenotypic evolution (Yang and Nielsen 2002; Zhang, Nielsen, and Yang 2005). In Model A, the branches on the tree are classified
into foreground and background branches. We classified as foreground branches those that experienced major changes in the evolution of relative cerebral cortex, whole brain, and cerebellum sizes as inferred by the squared change parsimony method. Phenotypic changes of 1 or more standard deviations (s.d.) away from the mean of change were used as the criterion for determining branches with major changes in relative cerebral cortex, cerebellum and whole brain sizes. The branch-site models were tested against the recommended null hypothesis of no positive selection in any of the foreground or background branches. The likelihood ratio test was used whereby 2 times the change in log-likelihood scores ($2\Delta l$) of the more complex model vs. the null hypothesis is computed and compared against a chi-distribution (d.f. = 1) (Zhang, Nielsen, and Yang 2005). A statistical control, Model B, was also used to establish whether the observed results are different if random branches were used as foreground branches. Specifically, nine of the background branches for relative cerebral cortex dataset were randomly selected and set as foreground branches. Model B was applied five times in order to obtain an average $2\Delta l$ across different sets of random branches. In order to avoid type 1 errors, we followed recommendations (Anisimova and Yang 2007) to use a correction for multiple-testing by setting a more conservative of $\alpha = .01$ in all models. For molecular evolution modelling, we used the software HyPhy (Pond, Frost, and Muse 2005) that utilizes the codon models of Codeml (Yang 1997). All models except Model B were applied twice to ensure convergence of log-likelihood scores and in order to avoid results based on local optima.
Acknowledgements

We thank the Singapore Zoological Gardens for contributing primate tissues and T.B. Penney, T.J. Devoogd, two anonymous reviewers, and the Associate Editor for comments on an earlier draft. This work was partially funded by Sigma Xi (F.A.).

Literature cited

Anisimova M, Yang Z. 2007. Multiple hypothesis-testing to detect lineages under positive selection that affects only a few sites. Mol Biol Evol 24:1219-1228.


Nylander JAA. 2004. MrModeltest v2.2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.


Rushton P, Vernon PA, Bonns TA. 2006. No evidence that polymorphisms of brain regulator genes Microcephalin and ASPM are associated with general mental ability, head circumference, or altruism. Biol Lett. 3.
Table 1

Molecular evolution of ASPM and its association with major changes in brain phenotypes

<table>
<thead>
<tr>
<th>Model Tested</th>
<th>$2\Delta l$</th>
<th>P-Value</th>
<th>Positively-selected sites in foreground branches</th>
<th>ω Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>i) Relative cerebral cortex size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Major changes</td>
<td>14.67</td>
<td>&lt;0.001</td>
<td>4.70</td>
<td>16</td>
</tr>
<tr>
<td>- Major increases only</td>
<td>20.75</td>
<td>&lt;0.00001</td>
<td>29.23</td>
<td>10</td>
</tr>
<tr>
<td>- Major decreases only</td>
<td>11.52</td>
<td>&lt;0.001</td>
<td>5.08</td>
<td>10</td>
</tr>
<tr>
<td>ii) Relative whole brain size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Major changes</td>
<td>0.10</td>
<td>&gt; .05</td>
<td>1.0</td>
<td>0</td>
</tr>
<tr>
<td>- Major increases only</td>
<td>0.11</td>
<td>&gt; .05</td>
<td>1.0</td>
<td>0</td>
</tr>
<tr>
<td>- Major decreases only</td>
<td>0.15</td>
<td>&gt; .05</td>
<td>1.0</td>
<td>0</td>
</tr>
<tr>
<td>iii) Relative cerebellum size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Major changes</td>
<td>0.1</td>
<td>&gt; .05</td>
<td>1.0</td>
<td>0</td>
</tr>
<tr>
<td>- Major increases only</td>
<td>0.02</td>
<td>&gt; .05</td>
<td>1.0</td>
<td>0</td>
</tr>
<tr>
<td>- Major decreases only</td>
<td>0.06</td>
<td>&gt; .05</td>
<td>1.0</td>
<td>0</td>
</tr>
<tr>
<td>Model B</td>
<td>&lt;0.01</td>
<td>&gt; .05</td>
<td>1.0</td>
<td>0</td>
</tr>
</tbody>
</table>

Note. – See text for explanation of Models. $2\Delta l$ is twice the difference in log-likelihood scores between Models A & B (positive selection) and the null hypothesis (no positive selection). Result for Model B is the average from 5 independent runs.

Fig. 1. – Evolution of relative cerebral cortex size in primates. Relative cerebral cortex sizes (log-transformed) inferred using squared-change parsimony are mapped above the branches whereas the change ($\Delta$: descendant value minus ancestral value) is mapped below. Dotted lines are foreground branches (labelled a-d and terminals) that have major increases (plus signs) or decreases (minus signs). The remaining branches are background branches (solid lines) with small changes in relative cerebral cortex size.
Figure 1