Title:
Adaptive Evolution of Four Microcephaly Genes and the Evolution of Brain Size in Anthropoid Primates

Authors:
Stephen H Montgomery (shm37@cam.ac.uk)¹
Isabella Capellini (Isabella.Capellini@durham.ac.uk)²
Chris Venditti (c.d.venditti@reading.ac.uk)³
Robert A Barton (r.a.barton@durham.ac.uk)²
Nicholas I Mundy (nim21@cam.ac.uk)¹§

Affiliations:
¹ Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, UK
² Evolutionary Anthropology Research Group, Department of Anthropology, Durham University
³ School of Biological Sciences, University of Reading
§ Corresponding author: Department of Zoology, University of Cambridge, Downing Street, Cambridge, CB2 3EJ, UK; Email: nim21@cam.ac.uk

Key words:
ASPM, MCPH1, CDK5RAP2, CENPJ, Brain, Neurogenesis, Primates

Running head:
Evolution of Microcephaly Genes in Anthropoids
Abstract

The anatomical basis and adaptive function of the expansion in primate brain size have long been studied; however we are only beginning to understand the genetic basis of these evolutionary changes. Genes linked to human primary microcephaly have received much attention as they have accelerated evolutionary rates along lineages leading to humans. However these studies focus narrowly on apes and the link between microcephaly gene evolution and brain evolution is disputed. We analyzed the molecular evolution of four genes associated with microcephaly (ASPM, CDK5RAP2, CENPJ, MCPH1) across 21 species representing all major clades of anthropoid primates. Contrary to prevailing assumptions positive selection was not limited to or intensified along the lineage leading to humans. In fact we show that all four loci were subject to positive selection across the anthropoid primate phylogeny. We developed clearly defined hypotheses to explicitly test if selection on these loci was associated with the evolution of brain size. We found positive relationships between both CDK5RAP2 and ASPM and neonatal brain mass and somewhat weaker relationships between these genes and adult brain size. In contrast, there is no evidence linking CENPJ and MCPH1 to brain size evolution. The stronger association of ASPM and CDK5RAP2 evolution with neonatal brain size than with adult brain size is consistent with these loci having a direct effect on prenatal neuronal proliferation. These results suggest that primate brain size may have at least a partially conserved genetic basis. Our results contradict a previous study which linked adaptive evolution of ASPM to changes in relative cortex size; however our analysis indicates that this conclusion is not robust. Our finding that the coding regions of two widely expressed loci have experienced pervasive positive selection in relation to a complex, quantitative developmental phenotype provides a notable
counterexample to the commonly asserted hypothesis that cis-regulatory regions play a dominant role in phenotypic evolution.

**Introduction**

The expansion of the brain, and in particular the neocortex, is a major hallmark of primate evolution (Jerison, 1973; Martin, 1990). After correcting for allometric scaling with body mass, primates have larger brains than most other mammals (Martin 1990; Barton 2006a) and both absolute and relative brain size have increased along multiple, independent primate lineages (Montgomery *et al.*, 2010). The adaptive significance and anatomical basis of the diversity of primate brains has long been studied using comparative methods (for review see Falk & Gibson, 2001; Finlay *et al.*, 2001; Barton, 2006a), but the investigation of the genetic basis of primate brain expansion has only begun relatively recently and is currently a topic of intense interest.

The convergent evolution of increased brain size in different lineages provides an opportunity to study whether the independent evolution of complex traits involves convergence at the molecular level (Arendt & Reznick, 2007) and may provide insights into lineage specific evolution, for example on the human lineage. Both scans of brain-expressed genes in published primate genomes (Dorus *et al.*, 2004a; Shi *et al.*, 2006; Yu *et al.*, 2006; Wang *et al.*, 2007) and studies of candidate genes (e.g. Enard *et al.*, 2002; Burki & Kaessmann, 2004; Wang *et al*. 2005) have mostly focused on identifying changes along the lineage leading to humans and have largely ignored convergent increase in brain size in multiple primate lineages.

One group of genes of particular interest in relation to the evolution of gross brain size is the microcephaly genes. Autosomal recessive primary microcephaly is a congenital disorder characterised by reduced growth of the cerebral cortex in the absence of environmental, metabolic or cytogenetic aetiologies (Bond & Woods 2006; Cox *et al.*, 2006). In humans it is
inherited as a recessive Mendelian trait involving at least eight loci, of which five have now been identified at the molecular level: \textit{ASPM, MCPH1, CDK5RAP2, CENPJ} (Jackson \textit{et al.}, 1998; Bond \textit{et al.}, 2002; Bond \textit{et al.}, 2005; Thornton & Woods, 2009) and the more recently identified \textit{STIL} (Kumar \textit{et al.}, 2009).

The five genes are expressed in the foetal brain during neurogenesis (Bond \textit{et al.}, 2002, 2005; Jackson \textit{et al.}, 2002; Kouprina \textit{et al.}, 2005; Kumar \textit{et al.}, 2009). \textit{ASPM, CDK5RAP2} and \textit{CENPJ} all have roles in centrosome or microtubule formation (Bond \textit{et al.}, 2006; Cox \textit{et al.}, 2006; Fish \textit{et al.}, 2006; Buchman \textit{et al.}, 2010), and can affect neurogenic mitosis by influencing the spindle pole and astral microtubule network (Fish \textit{et al.}, 2006; Fong \textit{et al.}, 2008; Cormier \textit{et al.}, 2009; Buchman \textit{et al.}, 2010). \textit{MCPH1} functions in the DNA damage response pathway and apoptosis (Wood \textit{et al.}, 2007; Rickmyre \textit{et al.}, 2007; Wood \textit{et al.}, 2008) and may also affect the timing of cell cycle progression (Brunk \textit{et al.}, 2007). Both apoptosis and cell cycle length are known to have significant roles in brain development (Roth & D’sa 2001; Calegari & Huttner 2003). The function of \textit{STIL} is less well studied, but it localises to the centrosome and has some functional similarities to \textit{ASPM} (Kumar \textit{et al.}, 2009; Thornton & Woods, 2009).

The main hypotheses for how the number of neurons could increase during brain expansion (Rakic, 1988; Caviness \textit{et al.}, 1995; Rakic, 1995; Kriegstein \textit{et al.}, 2006) rely on switches between symmetric and asymmetric cell divisions, via changes in spindle pole orientation, at a particular stage of neurogenesis. The functions of microcephaly genes are therefore consistent with the developmental mechanisms proposed to have facilitated brain expansion (Cox \textit{et al.}, 2006; Kriegstein \textit{et al.}, 2006; Götz & Huttner, 2005). Notably, the phenotypes exhibited by individuals with microcephaly show that these loci affect cortical surface area, not thickness, consistent with a role in regulating the size of the neural progenitor.
pool (Desir et al., 2008). Interestingly, recent studies in humans identified single nucleotide polymorphisms (SNPs) in ASPM, CDK5RAP2 and MCPH1 associated with total brain size or cortical surface area (Wang et al., 2008; Rimol et al., 2010) but not cortical thickness, an observation that again is consistent with a role in controlling the size of the neural progenitor pool (Rimol et al., 2010; Montgomery & Mundy, 2010).

Previous studies of the molecular evolution of the first four microcephaly loci to be identified supported the hypothesis that they have been subject to positive selection (Zhang, 2003; Evans et al., 2004a; Evans et al., 2004b; Wang & Su, 2004; Kouprina et al., 2004; Evans et al., 2006) but provided no direct evidence that the loci were involved in brain evolution as brain size was not incorporated into their analyses and did not include a diverse phylogenetic sample of species. A recent study analyzed ASPM evolution in relation to brain size in primates and concluded that branches with high relative telencephalon volume (reported as cerebral cortex) were associated with positive selection on ASPM (Ali & Meier, 2008).

For a comprehensive understanding of the role of genes in primate brain evolution, broad comparisons across the primate phylogeny incorporating relevant phenotypes are needed (Carroll, 2003; Goodman et al., 2005; Barton, 2006b; Pollen & Hoffman, 2008; Vallender, 2008). An important issue is which aspects of brain phenotype are most salient. Measures of brain size corrected for body size (i.e. relative brain size) are frequently used in studies investigating brain evolution as these take into account the strong correlation between brain and body mass (Barton, 2006a). However, given the implied functions of the four microcephaly genes in regulating the proliferation and survival of neurons, absolute brain mass may be a more relevant phenotypic measure as in primates it increases linearly with the total number of neurons (Herculano-Houzel et al., 2007). In agreement with quantitative genetic analysis of brain and
body size allometry (Lande, 1979), it has recently been shown that primate brain and body size
differ in their evolutionary trajectories (Montgomery et al., 2010) suggesting that these two traits
must be developmentally and genetically decoupled to some extent despite their closely
correlated evolution. Crucially, because primate neocortical neurogenesis is largely restricted to
prenatal development (Rakic, 1988; Rakic, 2002; Bhardwaj et al., 2006) and microcephaly is
primarily a disorder of foetal brain growth (Cox et al., 2006), microcephaly gene evolution
should be more closely related to neonatal brain size than to adult brain size. Postnatal brain
growth is largely driven by gliagenesis (Low & Cheng, 2006), axon growth (Sauvageot & Stiles,
2002) and myelination (Sowell et al., 2001), rather than by production of new neurons. There
are only two known sites in the primate brain, which are small and non-cortical, in which
substantial postnatal neurogenesis occurs (Jabes et al. 2010). Indeed, apoptosis eliminates large
numbers of neurons (Buss et al., 2006). Variation in these and other non-neurogenic processes
will reduce the relationship between brain size and neuron number as development progresses,
weakening any association with the molecular evolution of genes under selection in relation to
pre-natal neurogenesis. Indeed patterns of postnatal brain growth vary considerably across
primates (Leigh, 2004). Finally, if there is an association with adult brain size, given the specific
effect of microcephaly on the development of the cerebral cortex and their functions in cortical
neurogenesis (Cox et al., 2006; Thornton & Woods, 2009) we might also predict a stronger
association with adult neocortex size than adult whole brain size. Unfortunately there are
insufficient comparative data available on volumes of neonatal brain regions to test this
hypothesis.

Alternative hypotheses to explain the high evolutionary rates of microcephaly genes also
exist. The four loci are widely expressed throughout the body and *ASPM, CDK5RAP2* and
*CENPJ* are particularly highly expressed in the testis (Bond *et al.*, 2005; Kouprina *et al.*, 2005), where many genes have been shown to be under sexual selection in primates (Dorus *et al.*, 2004b; Clark & Swanson, 2005; Ramm *et al.*, 2008). However, the precise function of these genes in testes development and function is still unknown. For *ASPM*, a possible ciliary function led to the suggestion of a role in sperm flagellar movement which may affect sperm locomotion and hence be targeted by sexual selection (Ponting, 2006). If the microcephaly genes do have important roles in the testes or sperm their high rates of evolution may be associated with levels of sexual selection and be unrelated to changes in brain size. Hence explicit tests are required before the molecular evolution of microcephaly loci can be linked to brain evolution.

Here we investigate the molecular evolution of *ASPM, CDK5RAP2, CENPJ* and *MCPH1* in relation to brain size in anthropoid primates. First, we test whether these loci are under positive selection across anthropoids and whether or not different anthropoid clades have experienced different selective regimes. Second, we explore the association between the rate of molecular evolution of microcephaly genes and measures of brain size, predicting a positive and stronger association between these genes and absolute neonatal brain size than adult brain size, a stronger associations with absolute brain size as compared to relative brain size and potentially a stronger relationship with neocortex size than whole brain size. Finally, we investigate the relationship between microcephaly genes and relative testes size, a commonly used phenotypic correlate for sperm competition and sexual selection (Harcourt *et al.*, 1995; Ramm & Stockley, 2010), to test the hypothesis that these loci may have been under sexual selection. We find that whereas all four loci have been targets of selection throughout primate evolution, *ASPM* and *CDK5RAP2* but not *MCPH1* and *CENPJ* show positive associations with absolute neonatal brain
but not with any measure of relative brain size or relative testis size, suggesting a role in the evolution of total neuronal number.

Materials and methods

Phenotypic data

Data for body, brain mass and volumes of specific brain regions were obtained from previously published data (Bauchot & Stephan, 1969; Stephan et al., 1981; Zilles & Rehkemper, 1988). leading to a dataset of 37 primate genera including 14 catarrhines, 12 platyrrhines, one tarsier and 10 strepsirrhines (Table S1). Data on neonatal brain size (22 taxa) were obtained from Capellini et al., (in press). Data on testis mass for 30 genera were taken from Harcourt et al. (1995).

To test the hypothesis that the evolution of the microcephaly genes should be more strongly associated with absolute brain mass than measures of relative brain size we calculated relative brain mass by performing a phylogenetically controlled regression analysis between log(brain mass) and log(body mass), and for neocortex, log(neocortex volume) was regressed against log(rest of brain volume) separately (following Barton 1998). For testis, log(testis mass) was regressed against log(body mass). These analyses were performed using a phylogenetically controlled regression using phylogenetic generalised least squares models (PGLS) in BayesTraits (Pagel et al., 2004) with Maximum Likelihood and 1,000 runs for each analysis. Residual values from the regression line were calculated for each taxon and these were used as values of relative brain size and relative testis mass in all subsequent analyses. All phenotypic data are provided in Table S1 and additional phenotypic analyses are presented in the Supplementary Information.

With PGLS the phylogeny is converted into a variance-covariance matrix, where the diagonal of
the matrix gives information on the path length from root to tips (the ‘variance’) and the off-diagonal values of the matrix provide information on the shared evolutionary history of any pair of species, that is the time from the root to the last common ancestry (the ‘covariance’) (Pagel, 1997, 1999; Freckleton et al. 2002; Capellini et al, in press b). With PGLS regression, the variance-covariance matrix is included into the error term of the regression model, and the resulting estimated regression parameters (i.e. slopes and intercepts) are ‘phylogenetically controlled’ (Pagel, 1997, 1999; Freckleton et al. 2002; Capellini et al, in press b).

**Phylogeny**

We used a genus level composite phylogeny of primates using published trees. The topology is taken from Goodman et al. (2005) for haplorhine primates and Horvath et al. (2008) for strepsirhines. Proportional branch lengths were obtained from recent studies of primate divergence dates (Opazo et al., 2006; Page & Goodman, 2001; Poux & Douzery, 2004; Purvis, 1995) scaled to agree with dates of divergence for the deeper primate nodes estimated by Steiper & Young (2006). The tree obtained therefore has branch length information in time and is ultrametric (Figures S1).

**Laboratory methods**

Genomic DNA samples had previously been extracted from tissue samples using Qiagen kits. Sequence data from previous studies and primate genomes were collected from the online databases GenBank and Ensembl. From these sequences primers were designed using Primer3Plus (Untergasser et al., 2007). We sequenced exons which had previously been shown to have accelerated rates of evolution or contained a large proportion of the coding sequence. For
ASPM we sequenced exons 3 and 18, totalling 6235bp (60% of the coding region). For MCPH1 three exons were sequenced: 8, 11 and 13, totalling 1556bp (62% of the coding region). Exons 2 and 7 were sequenced for CENPJ, totalling 1556bp (52% of the coding region). We sequenced 7 out of 38 exons of CDK5RAP2: exons 12, 20, 21, 24, 25, 32 and 33, (total 2120bp; 37% of the coding sequence). PCR reactions and sequencing on both strands were performed using standard protocols (see Supplementary information & Table S2 for further details and primers).

Sequences were edited in SEQMAN v. 5.05 (DNASTAR Inc.) and aligned and checked in CLUSTALW in MEGA 4.0 (Tamura et al., 2007). Exons of each locus were concatenated and subsequently analysed together; alignments are available on request. Sequences were obtained for 5 apes, 5 Old World monkeys and 10 New World monkeys representing all major clades of anthropoid primates, shown in Figure 1. Where phenotypic data were not available for the species sequenced we used closely related congeneric species for which data were available. Newly sequenced data have been submitted to Genbank (See Table S3 for accession numbers). We used the Strepsirhines Microcebus murinus and Otolemur garnetti; Ensembl IDs are shown in Table S3.

Molecular evolution

A common measure used to infer selection pressures acting on coding regions of genes is the ratio of rates of non-synonymous to synonymous fixed base changes. Estimation of $dN/dS$ ratios ($\omega$) was carried out using a codon-based maximum likelihood method (PAML version 4; Yang, 2007). Several analyses were performed to test the hypothesis that the four loci have experienced positive selection across primates, in particular, in relation to brain size evolution. Nested models are compared using the likelihood ratio test (LRT) statistic (-2[LogLikelihood$_1$ –
LogLikelihood]) to critical values of the Chi square distribution and degrees of freedom as the difference in the number of parameters estimated by each model.

**Site and branch models**

To detect positive selection across primates we implemented the site models. These allow the \( \omega \) to vary among sites but not across lineages (Nielsen & Yang, 1998; Yang *et al.*, 2000). Model M1a (NearlyNeutral) allows sites to fall into two categories with \( \omega < 1 \) (purifying selection) and \( \omega = 1 \) (neutral evolution), whilst model M2a (PositiveSelection) allows sites to fall into three categories with \( \omega < 1 \), \( \omega = 1 \) and \( \omega > 1 \) (positive selection) (Yang *et al.*, 2005).

In addition we used the branch models to test whether the \( dN/dS \) of lineages leading to humans was significantly higher than non-human lineages and whether \( dN/dS \) significantly differed between Apes, Old World Monkeys and New World Monkeys.

**Root to tip \( dN/dS \) and gene-phenotype associations**

The branch models were used to estimate the average \( dN/dS \) ratio from the ancestral anthropoid to each terminal species tip. These values were then set as species data and used in a PGLS regression with measures of brain size in BayesTraits, as explained above (Pagel, 1999; Pagel *et al.*, 2004; Organ *et al.*, 2006). Previous analyses which have tested for correlations between phenotypes and \( dN/dS \) ratios using similar methods have typically used the \( dN/dS \) of the terminal branch (e.g. Dorus *et al.*, 2004b; Nadeau *et al.*, 2007). However whilst a species’ phenotype reflects the whole phenotypic evolution, the \( dN/dS \) of the terminals branch does not reflect the whole genotypic evolution. The root-to-tip \( dN/dS \) is more inclusive of the evolutionary history of a locus and is a property of the species tip, rather than the terminal branch, and is
therefore more suitable for regressions against phenotypic data from extant species. In addition, by analysing the rate of evolution since the last common ancestor of the species in our dataset all branches are the same length and therefore not subject to temporal effects on $dN/dS$ (Wolf et al., 2009). One assumption of regression analysis is that the residuals of the model are normally distributed. As the residuals of the regression using $dN/dS$ ratios were not normally distributed, we used $\log_{10}(dN/dS)$ to improve normality. Residuals of regression analysis with log-transformed $dN/dS$ did not violate assumptions of normality and constant variance.

First we examine the relationship between microcephaly gene molecular evolution and the evolution of absolute and relative neonatal brain size. As we specifically hypothesise a positive association between brain size and the selection pressure on these loci the significance of the regression coefficient was determined using a one-tailed t-test. A significant negative association would suggest an increase in purifying selection has acted on a locus as brain mass increased and, whilst interesting, would suggest that the locus does not contribute to the genetic basis of change in that phenotype and could not explain why the locus has evolved adaptively. Hence we meet both recently suggested requirements for justifying the use of one-tailed tests; we explain why we hypothesise an association in a particular direction and why the opposite pattern can be treated the same as a non-significant trend in the expected direction (Ruxton & Neuhäuser, 2010).

As the size of our dataset for this analysis is limited by the availability of neonatal brain size data, and the imperfect overlap between phenotype and gene sequence datasets, it is highly likely that the small sample size will result in low power to reject the null hypothesis. To minimise the chances of Type I errors, we restrict our analyses to a small number of critical tests, and we determine the specificity of relationships for microcephaly genes by testing for
associations with other genes having no known role in neurogenesis. In particular we test for associations with alternative phenotypes and test for associations between the evolution of genes with no known role in neurogenesis with brain size. In addition a Jack-knife approach was taken to test the robustness of the associations found and to identify any outliers which have a dominant effect on the slope of the regression.

We subsequently explored the relationship with volume of adult whole brain and neocortex. This analysis was performed to test our hypothesis that genes involved in neurogenesis should be more strongly associated with neonatal than adult brain mass, and that if there is an association with adult brain size it may be stronger for neocortex size than whole brain size. Comparisons of non-nested models were performed using Akaike Information Criterion (AIC: calculated as (2 x no. parameters) + (-2 x Log[Likelihood])) to identify the best supported model, where a lower value indicates a better fitting model, and a difference between models greater than two suggests a substantial difference (Burnham & Anderson 2002).

In addition to the standard \(dN/dS\) ratios, we used multiple regressions to investigate the association between phenotype and \(dN\) while controlling for \(dS\). Here we predict a negative association between brain size and \(dS\) given known relationships between \(dS\) and life history traits such as body size (Nikolaev et al., 2007). Conversely, a locus which is a target of selection in relation to brain size may show a positive association with \(dN\). Both approaches examine variation in \(dN\) and \(dS\) relative to one another, but they make different assumptions about the nature of the underlying relationship. For example, a significant \(dN/dS\)-phenotype relationship suggests an association between phenotypic evolution and selection acting on a locus and may be obtained when both change together in a tightly correlated fashion but with one changing at a faster rate than the other (so that the ratio correlates with the absolute value of the changes),
whereas in this case a multiple regression would show no significant correlation. Hence, differences in the results obtained may be informative about the nature of the gene-phenotype correlation.

In most cases the phenotypic data is based on a small number of individuals and the degree of interspecific variation is unknown. However where interspecific variation greatly exceeds intraspecific variation, as is expected to be the case for brain size, results of comparative analyses are not biased by intraspecific variation (see Nunn & Barton, 2001). It is also likely that error introduced by sampling small numbers of individuals will lead to an underestimate of correlation coefficients between two traits (Nunn & Barton, 2001; Ives et al., 2007).

Finally, we used branch-site models to test for associations between positive selection and brain evolution, but this method did not produce informative results (see Supplementary Information).

**Results**

Our full dataset comprises sequence data from 21 species for each gene including representative species from all major anthropoid clades, and 11-17 newly sequenced species for each locus. The coverage of full coding sequence in the dataset comprises 60% for *ASPM* (from 2 exons), 37% for *CDK5RAP2* (from 7 exons) 52% for *CENPJ* (from 2 exons) and 62% for *MCPH1* (from 3 exons) (see methods and Table S3). Previous datasets for *CDK5RAP2* and *CENPJ* have included only 4 species (Evans et al., 2006). With the exception of Ali & Meier’s recent *ASPM* study (2008) which was published after the completion of our dataset, analyses of this gene have considered 3 (Zhang, 2003), 7 (Evans et al., 2004) and 8 species (Kouprina et al., 2004) whilst studies of *MCPH1* have included 7 (Evans et al., 2004) and 13 species (Wang & Su, 2004).
These studies were particularly lacking in a diverse range of New World Monkeys. Although we have sequenced only partial coding sequences, we have focused on regions of the gene which contain functionally important domains or which have previously been shown to have accelerated rates of evolution. Our dataset therefore allows us to examine how widespread selection on these regions has been across anthropoids and to explore the relationship between this selection and phenotypic evolution.

Pervasive adaptive evolution in microcephaly genes

We first examined whether a signal of positive selection is present in four microcephaly loci by performing site model tests using a codon-based maximum likelihood method (Yang, 2007; Table 1). All four genes showed evidence for positive selection across anthropoids, with estimated omega of 2.57-5.39 at 2.28–6.50% of sites across the loci. The most significant results are for ASPM and CDK5RAP2 which have 2.3% and 6.5% of sites having an omega of 5.39 and 4.42 respectively. In CDK5RAP2, a notable feature is the clustering of sites identified as being under positive selection using Bayes Empirical Bayes (p >95%) (Positions 929 - 977 and 1683 - 1690), these mostly fall within an SMC domain (Evans et al., 2006), these domains are thought play a role in the chromosome segregation, regulation and repair (Hirano, 2006), but the functional significance of these particular sites are unknown.

Previous authors have proposed that ASPM and MCPH1 have evolved at a higher rate along lineages leading from the last common ancestor of apes to modern humans than on other lineages (Evans et al., 2004a & b). We tested this explicitly using a branch model and found that the evolutionary rate of change in lineages leading to humans does not significantly differ from that in other lineages (Table 2a), which is congruent with the fact that the proportional change in
brain mass along this lineage is also not exceptional compared to the rest of primate brain size evolution (Montgomery et al., 2010). In addition we tested whether the evolutionary rate of the four loci differed between apes, Old World Monkeys and New World Monkeys and found no significant differences (Table 2b). Finally we performed site models to test for positive selection on each of the three clades separately. Although the results are influenced by differences in the number of sequences and the time depth for each clade, we found no instance where positive selection was found in Apes and not in Old or New World Monkeys (Table S4).

**Associations between gene evolution and brain size**

To explicitly test the link between the molecular evolution of microcephaly genes and brain size we performed several tests. We first performed a phylogenetically controlled regression analysis in BayesTraits (Pagel, 1999; Pagel et al., 2004) between the root-to-tip $dN/dS$, estimated using the branch models, and neonatal brain mass.

We found a significant association between the molecular evolution of $CDK5RAP2$ and absolute neonatal brain mass ($t_{11} = 1.95$, $p = 0.039$, $R^2 = 0.255$) but no significant association between $ASPM$, $CENPJ$ or $MCPH1$ and this trait (Table 3). For $ASPM$ however, *Callithrix* represented a strong outlier (Figure 2, Table S5) and when this species was removed the association between the $dN/dS$ of *ASPM* and neonatal brain mass became significant ($t_{10} = 2.42$, $p = 0.018$, $R^2 = 0.369$). *Callithrix*, and the other Callitrichids, show high rates of evolution of $ASPM$ (Table S5) but have the smallest brain masses among the anthropoid primates (Stephan et al., 1981), however this is due to a secondary reduction in brain mass in this taxon (Ford, 1980, Montgomery et al., 2010). The significance of the regression between $ASPM$ and neonatal brain
mass was not affected by the removal of any other species. From here on, unless otherwise stated, regressions for *ASPM* were performed without the Callithrichids. Removing *Callithrix* does not reveal a significant association with neonatal brain mass for either *CENPJ* ($t_{10} = 0.81$, $p = 0.218$, $R^2 = 0.061$) or *MCPH1* ($t_{10} = -0.55$, $p = 1.000$, $R^2 = 0.029$). The high root-to-tip *MCPH1* $dN/dS$ ratio of *Pan* (1.39) appeared to be heavily influenced by a small number of synonymous substitutions on the terminal *Pan* branch (1 synonymous change, compared to 6 synonymous changes on the *Homo* branch), and we note previous studies have found a much lower $dN/dS$ ratio on the terminal *Pan* lineage using the full coding sequence (Evans *et al.*, 2004a; Wang & Su, 2004). We therefore repeated the regressions excluding the *Pan* data point but still found no significant result (data not shown). No locus showed any association with relative neonatal brain size (Table 3).

We next explored whether the association with brain mass can be observed for adult phenotypes by performing regressions of root-to-tip $dN/dS$ with absolute and relative neocortex and whole brain size. We found no significant association with any measure of absolute brain size for any locus (Table S6). To confirm this result we performed the same regressions but using only the species used in the neonatal regressions. Using this reduced dataset we subsequently found a significant association between both absolute neocortex and whole brain mass with *ASPM* (whole brain: $t_{10} = 2.732$, $p = 0.010$, $R^2 = 0.427$; neocortex $t_{10} = 2.980$, $p = 0.007$, $R^2 = 0.470$) and *CDK5RAP2* (whole brain: $t_{11} = 2.131$, $p = 0.028$, $R^2 = 0.292$; neocortex $t_{11} = 2.00$, $p = 0.035$, $R^2 = 0.267$) but not *MCPH1* or *CENPJ* (not shown). To identify the cause of this discrepancy we performed a reverse jack-knife, repeating the regressions with the full dataset of adult phenotypes and removing each species, one at a time (Table S7). We found that for both *ASPM* and *CDK5RAP2* the removal of *Papio* results in a significant association with
absolute neocortex and whole brain mass (ASPM: whole brain: $t_{15} = 2.063$, $p = 0.028$, $R^2 = 0.221$: neocortex: $t_{14} = 2.128$, $p = 0.0257$, $R^2 = 0.244$; CDK5RAP2: whole brain: $t_{18} = 2.418$, $p = 0.013$, $R^2 = 0.245$, neocortex: $t_{17} = 2.238$, $p = 0.019$, $R^2 = 0.228$), whereas removing any other species does not have this affect (Table S5), suggesting *Papio* is an outlier to a general trend.

We subsequently added the Callitrichids back into the regressions with *ASPM* and find that the addition of any one Callitrichid or all three together substantially reduces or negates the association (Addition of *Saguinus* alone: $t_{16} = 0.612$, $p = 0.274$, *Leontopithecus* alone: $t_{16} = 1.842$, $p = 0.042$, *Callithrix* alone: $t_{16} = 0.612$, $p = 0.265$, All three: $t_{18} = 0.302$, $p = 0.2383$). In contrast, after removal of all three Callitrichids the association with *CDK5RAP2* is still significant ($t_{15} = 0.612$, $p = 0.042$) and removal or any single species (removal of *Saguinus*: $t_{17} = 2.010$, $p = 0.031$, *Leontopithecus*: $t_{17} = 2.359$, $p = 0.015$, *Callithrix*: $t_{17} = 2.010$, $p = 0.031$) or any pair of Callitrichids (retention of *Saguinus*: $t_{16} = 1.886$, $p = 0.039$, *Leontopithecus*: $t_{16} = 1.988$, $p = 0.032$, *Callithrix*: $t_{16} = 2.460$, $p = 0.012$) does not significantly affect the association. This confirms Callitrichids are outliers to the general trend for *ASPM* but not for *CDK5RAP2*. It is notable that *Leontopithecus* has a much smaller effect on the *ASPM* result than *Saguinus* and *Callithrix* suggesting a complex pattern of selection on this locus, more data will be required to tease out the pattern within Callitrichids.

In addition we repeated the analysis using relative measures of adult brain and neocortex sizes. For *ASPM* there was a significant association with relative adult brain size on both the full dataset ($t_{19} = 1.77$, $p = 0.046$, $R^2 = 0.142$; without Callitrichids $t_{15} = 2.241$, $p = 0.020$, $R^2 = 0.239$) and the reduced dataset used for regressions with neonatal data ($t_{10} = 2.59$, $p = 0.013$, $R^2 = 0.402$). However in both cases the result was dependent on the human data point and when this was removed the association is no longer significant (adult: $t_{18} = 0.574$, $p = 0.287$, $R^2 = 0.018$;
without Callitrichids: \( t_{15} = 0.923, p = 0.185, R^2 = 0.054; \) neonate \( t_{18} = 0.880, p = 0.200, R^2 = 0.079 \). No other significant relationship was detected (Table S6).

We next used AIC to determine the best supported model for the associations between \textit{ASPM} and \textit{CDK5RAP2} and neonatal and adult brain mass using only the species for which data are available for both phenotypes. For both loci the regression with neonatal brain mass has a substantially lower AIC than whole brain mass (\textit{ASPM}: neonate AIC = 2.44, adult AIC = 7.39; \textit{CDK5RAP2}: neonate AIC = 12.19, adult AIC = 14.67) suggesting a closer association with neonatal brain mass. Comparing adult neocortex and whole adult brain size we found no substantial difference (\textit{ASPM}: neocortex AIC = 7.43, whole AIC = 7.40; \textit{CDK5RAP2}: neocortex AIC = 16.29, whole AIC = 14.67). Hence both the results are for \textit{ASPM} and \textit{CDK5RAP2} are highly consistent.

To further explore the relationship between brain evolution and the molecular evolution of the microcephaly genes we performed multiple regressions with neonatal brain size and root-to-tip \( dN \) and \( dS \) (log transformed) as independent variables. Significant negative partial regression coefficients were found, as predicted, for the three cytoskeletal genes and \( dS \) (1-tailed \textit{ASPM}: \( t_8 = -2.958, p = 0.007; \) \textit{CDK5RAP2}: \( t_9 = -1.919, p = 0.042; \) \textit{CENPJ}: \( t_9 = -1.859, p = 0.046 \) but interestingly not for \textit{MCPH1} (\( t_9 = -1.294, p = 1.000 \)). Neither \textit{CENPJ} nor \textit{MCPH1}, which show no association between brain size and \( dN/dS \), show a significant association with \( dN \) (\textit{CENPJ}: \( t_9 = 0.049, p = 0.481; \) \textit{MCPH1}: \( t_9 = -2.322, p = 1.000 \)). For \textit{ASPM} we do see an association between \( dN \) and neonatal brain size (\textit{ASPM}: \( t_8 = 2.032, p = 0.035 \)) but we do not find a significant association for \textit{CDK5RAP2} (\( t_9 = 1.127, p = 0.142 \)). This suggests the association between \textit{ASPM} and \( dN/dS \) may be driven predominantly by an accelerated \( dN \) whilst the association for \textit{CDK5RAP2} may have a more complex basis.
As a final analysis we used the branch-site test to detect positive selection on branches along which brain size is estimated to have expanded greatly (see Ali & Meier, 2008). However, as discussed in the Supplementary Information, we found this method was not informative as it is possible to get a positive result by selecting branches at random. This is likely to be due to positive selection acting across the phylogeny, as demonstrated by our site model tests, and strongly questions the strength of the methodology used and results obtained by Ali & Meier (2008), namely that episodes of adaptive evolution of *ASPM* are specifically associated with expansion of the relative size of the telencephalon.

**Controls for specificity of gene-phenotype associations**

To exclude the possibility that the gene-phenotype correlations reported above are coincidental, we investigated eight loci (Table S8), with no known function in neurogenesis, for which data were already available for a reasonably large number of species (n = 10-20) across the anthropoid phylogeny. This control set includes both genes which have previously been shown to have experienced positive selection across anthropoids and genes which appear to be under purifying selection (Table S9a). We tested for an association between the root-to-tip $dN/dS$ of these loci and absolute adult and neonatal brain mass in the same way as described above. No locus was found to have a significant association with either phenotype (Table S9b), and neither did the removal of any one species result in a significant change in the regression slope and an positive association with brain size (data not shown). This suggests that the significant results presented above are unlikely to be Type I errors.

Given the strong allometric relationship between brain and body mass we also checked whether molecular evolution of *ASPM* and *CDK5RAP2* could be more strongly associated with
adult body mass than brain mass, due for example to differences in effective population size
(Nikolaev et al., 2007) or general scaling effects, using the neonate dataset to allow comparisons
using AIC. For ASPM there was no significant association with adult body mass with ($t_{13} = -0.368, p = 1.000, R^2 = 0.012$) or without inclusion of Callithrix ($t_{12} = 0.846, p = 0.210, R^2 = 0.067$). For CDK5RAP2 there is a marginally significant association with body mass ($t_{13} = 1.823, p = 0.048, R^2 = 0.232$) but the AIC score for an association with absolute neonatal brain mass is
substantially lower (Body mass AIC = 20.515; Neonatal brain mass AIC = 12.19) showing that
the association with brain mass is much better supported, and that the weak association with
body mass is likely to be due to correlated evolution between brain and body mass.

We next tested whether the adaptive evolution of the four loci could be better explained
by sexual selection in relation to their expression in the testes. We performed phylogenetically
controlled regressions between root-to-tip $dN/dS$ and relative testes size, a commonly used
correlate of sexual selection (see methods). No significant association was observed for ASPM,
CDK5RAP2 or CENPJ. A significant association was found for MCPH1 ($t_{18} = 1.959, p = 0.033, R^2 = 0.176$; Table 4) but when Pan (which has an unexpectedly high $dN/dS$) is removed, this
association is no longer significant ($t_{17} = 0.530, p = 0.301, R^2 = 0.016$), suggesting more data will
be required to confirm this result.

**Discussion**

**Molecular evolution of microcephaly genes and brain evolution**

Studying the molecular basis of convergent phenotypes has enhanced our understanding of the
evolutionary genetics of adaptation and the constraints which act on phenotypic evolution
(Arendt & Reznick, 2007). Here we show that independent increases in brain mass across
anthropoids may share a common genetic basis. By sampling a substantial number of phylogenetically diverse species we have demonstrated that positive selection acted on four microcephaly loci across the anthropoid phylogeny and was not, as previously reported, restricted to lineages leading to humans. To our knowledge our study is the first to implement robust codon based models to test for positive selection acting on these loci across anthropoids. [a previous study reporting such a finding (Wang & Su, 2004) used Model 3 in PAML which is not a robust test for positive selection (Anisimova et al., 2002)]. This is a striking result as pervasive positive selection is considered rare. However, as brain size has increased multiple times independently, and is likely to have been under strong selection in all major groups of anthropoids (Montgomery et al., 2010), such widespread positive selection on genes involved in the evolution of brain size should perhaps be expected.

We explored whether this selection is relevant to gross brain size evolution using phylogenetically controlled regressions which show that the average $dN/dS$ across *ASPM* and *CDK5RAP2* is significantly related to absolute neonatal brain mass and that in both cases a relatively large proportion of the variance is explained ($R^2 = 0.369$ and 0.255 respectively). Furthermore, two key predictions of a gene’s involvement in prenatal neurogenesis were verified for *ASPM* and *CDK5RAP2*: (i) an association with absolute brain mass as this correlates closely with total neuron number (Herculano-Houzel et al., 2007) and (ii) a stronger association with neonatal than adult brain mass as cortical neurogenesis is largely restricted to prenatal development (Rakic, 1988; Rakic, 2002; Bhardwaj et al., 2006).

These results are not explained by a general association with body mass, nor can they be attributed to random or genome-wide effects, since no associations were found with two other microcephaly genes (*CENPJ* & *MCPH1*) and eight control genes. Thus, although the results are
only marginally significant, our control tests and the highly consistent pattern observed in the
significant results strongly suggest the associations found are unlikely to be Type I errors. The
results using multiple regressions suggest that while positive selection on ASPM has brought
about an increase in \( dN/dS \) mainly through an acceleration in \( dN \) relative to \( dS \), the pattern for
CDK5RAP2 may be more complex. This suggests that caution should be exercised in interpreting
\( dN/dS \) ratios, and we recommend the use of supplementary analyses such as multiple regression
to disambiguate correlations involving this measure. Together with the demonstration of positive
selection in the site analyses, these results imply that adaptive evolution on ASPM and
CDK5RAP2 has been involved in independent changes in brain size along multiple lineages
during primate evolution through a role in prenatal neurogenesis.

Although we detect a general positive association for ASPM and CDK5RAP2 there are
notable outliers, which may suggest a more complex relationship between the evolution of these
loci and brain size. For both loci Papio has a much lower \( dN/dS \) than would be predicted given
our results and the size of its the brain in this species. Assuming the association between ASPM,
CDK5RAP2 and brain evolution has a genuine, functional basis, the Papio discrepancy may
indicate brain expansion can occur independently of the evolution of these loci. However as
Papio is not represented in our neonatal dataset we cannot say whether it is an outlier due to pre
or postnatal developmental processes. The high rate of evolution of ASPM during Callitrichid
evolution is also striking. A plausible hypothesis is that this acceleration is related to phyletic
dwarfism and the reduction of brain size in this clade (Ford, 1980; Montgomery et al., 2010).
The lower rate of evolution of CDK5RAP2 in the Callitrichids raises the interesting possibility
that different selection pressures have acted on these loci during episodes of brain size reduction.
Further work is required to address these issues.
Despite demonstrating *CENPJ* and *MCPH1* have experienced pervasive positive selection during anthropoid evolution we found no significant relationship between either locus and any measure of absolute or relative brain mass. This potentially indicates an interesting dichotomy in evolutionary roles among microcephaly genes and raises the issue of whether these loci are involved in the evolution of other traits or more specific aspects of brain phenotype that were not considered here. It is important to emphasize that the phenotype(s) on which selection for *CENPJ* and *MCPH1* is acting in primates has not been established and this study provides no evidence that they are involved in the evolution of gross measures of size of the whole brain or neocortex.

**Integrating comparative genetics and neurobiology**

Our results are consistent with models of how neuron number might evolve. A single additional round of proliferative, symmetric divisions of neuroepithelial cells in the ventricular zone would double the number of neurons in the cortex (Rakic, 1988,1995; Cavines et al., 1995). Neuroepithelial cells have apical-basal polarity and the switch from proliferative, symmetric to neurogenic, asymmetric divisions is controlled by the orientation of the spindle pole during mitotic division (Chenn & McConnell, 1995; Götz & Huttner, 2005). An alternative, but not mutually exclusive, model places greater emphasis on prolonged intermediate-progenitor cell division in the sub-ventricular zone, which may also occur by changes in spindle pole orientation (Kriegstein et al., 2006). In addition, as brain size expands, neural progenitors become increasingly elongated (Smart et al., 2002; Fish et al. 2008) and selection may be acting on cytoskeletal genes in response to the need to maintain the precision of spindle orientation during mitotic division of these highly elongated cells in larger brained species (Zhang 2003; Kouprina
et al., 2004; Fish et al., 2008). In this way selection on these loci may be in response to the evolution of larger brains rather than causing the change in brain size. Whilst it is also possible to envisage scenarios where the change in spindle orientation itself leads to the production of additional neurons, whether the role of these loci in brain size evolution is causative or responsive is yet to be determined. Both scenarios are consistent with the results presented here.

Notably, there is strong agreement between the results of studies investigating the expression of these loci (Bond et al., 2002; Jackson et al., 2002; Bond et al., 2005; Kouprina et al., 2005), their function (Bonds & Woods, 2006; Cox et al., 2006; Fish et al., 2006; Brunk et al., 2007; Wood et al., 2007; Buchman et al., 2010), models of brain evolution (Rakic, 1988; Kriegstein et al., 2006) and their molecular evolution which together implicate ASPM and CDK5RAP2 as having significant roles in the evolution of neuron number and brain size.

**Implications for evolutionary genetics of adaptation**

The inferred role of ASPM and CDK5RAP2 in the evolution of primate brain mass have implications for our understanding of the evolutionary genetics of adaptation. First, they provide evidence that a complex, polygenic quantitative phenotype evolved by convergence at the molecular level (Cresko et al. 2004; Mundy, 2005; Arendt & Reznick, 2007). Second, our results go against a commonly asserted hypothesis that evolution of form occurs primarily through changes in cis-regulatory sequences (King & Wilson, 1975; Carroll, 2005; Wray, 20007; Carroll 2008).

Changes in cis-regulatory sequences are proposed to be more important for phenotypic evolution as their modular nature limits pleiotropic effects (Carroll, 2005; but see Hoekstra & Coyne, 2007; Lynch & Wagner, 2008; Stern & Orgogozo, 2008). In fact the evolution of brain
mass has been singled out as an example where regulatory evolution is likely to be the predominant evolutionary mechanism: “the evolution of complex traits such as brain size… must have a highly polymorphic and largely regulatory basis” (Carroll, 2005; emphasis ours). It is therefore interesting that the microcephaly genes are expressed throughout the body (Bond et al., 2005; Kouprina et al., 2005) but are subject to positive selection. If ASPM and CDK5RAP2 are in fact involved in brain evolution how are pleiotropic effects avoided?

The first point to note is that the pathology of primary microcephaly itself shows that pleiotropic effects of microcephaly gene disruption can be limited to the brain (Bond & Woods 2006). Both ASPM and CDK5RAP2 are alternatively spliced (Kouprina et al., 2005; Buchman et al., 2010), which may reduce pleiotropic effects (Hughes 2006; Hoesktra & Coyne, 2007; Lynch & Wagner, 2008). Alternatively the evolution of ASPM may not affect non-neural cells either due to the elongated cell morphology of neuroepithelial cells (Fish et al., 2006) or cell-dependent recruitment factors (Van der Voet et al., 2009). These explanations are not mutually exclusive and provide plausible mechanisms to reduce the pleiotropic effects of protein evolution (Hoekstra & Coyne, 2007; Lynch & Wagner, 2008). More broadly, research is showing a diversity of mechanisms for the possible genetic basis of aspects of primate brain evolution, including coding sequence evolution (e.g. this study; Wang et al., 2005; Vallender & Lahn, 2006; Uddin et al., 2008) gene duplication (Burki & Kaessman, 2004; Marques-Bonet et al., 2009), non-coding RNA evolution (Zhang et al., 2008) and changes in regulatory sequences and gene expression (Khaitovich et al., 2005; Rockman et al., 2005; Gilad et al., 2006; Khaitovich et al., 2006; Prabhakar et al., 2006; Haygood et al., 2007; Somel et al., 2009). Indeed, the mosaic nature of brain structure evolution in mammals indicates a complex genetic basis to changes in brain size (Barton & Harvey 2000).
**Conclusion**

We have presented evidence implicating *ASPM* and *CDK5RAP2* in the evolution of brain size across anthropoid primates, a result which is consistent with an effect of the two loci on neurogenesis via mitotic spindle orientation. Despite showing that *CENPJ* and *MCPH1* have been subject to positive selection we find no evidence to link these loci to the evolution of gross brain size, and the mechanism of selection acting on these loci is therefore unresolved. This study demonstrates the importance of including phenotypic data and a phylogenetically broad range of species when attempting to associate the evolution of genes with brain size evolution (Carroll, 2003; Goodman *et al*., 2005; Barton, 2006b; Pollen & Hoffmann, 2008; Vallender, 2008). This point is especially pertinent to the literature on human genetic evolution where claims are often based on differences between humans and chimpanzees or a small number of non-human primates. The results also clearly highlight the importance of including measures of neonatal brain size in studies of primate brain evolution. Finally our results suggest a conserved genetic basis for brain evolution in primates, providing an important example where genetic basis of a complex developmental phenotype has involved coding sequence evolution.

**Abbreviations:**

AIC – Akaike Information Criteria  
ML – Maximum Likelihood  
NWM – New World Monkeys  
OWM – Old World Monkeys  
PGLS - Phylogenetic Generalised Least Squares
SNP – Single Nucleotide Polymorphism

**Acknowledgments:**

We are grateful Andrew Kitchner and Drew Bain (National Museums Scotland), Mike Bruford (Zoological Society London) and Leona Chemnick (Center for Reproduction of Endangered Species, San Diego Zoo) for providing tissue samples, and Barry Keverne, Chris Jiggins, the Mundy & Jiggins Labs (University of Cambridge) and Adrian Friday helpful discussions, Jennifer Fish for correspondence on the function of *ASPM* and Mark Pagel for helpful comments on the phylogenetic analyses. SHM & NIM thank BBSRC, the Leverhulme Trusts and Murray Edwards College for financial support. IC & RAB thank BBSRC/NERC for financial support (grant number BB/E014593/1). CV thanks the Leverhulme Trust for financial support (grant number: ECF/2009/0029).

**References:**


Figure legends

Figure 1: Phylogeny of anthropoid primates with an indication of (a) Absolute brain mass – the area of the square shows the mass of the brain as a percentage of the human brain and (b) Relative brain mass (on body mass), taken from Montgomery et al., 2010. * The second species of Colobinae varied between loci (see supplementary information).

Figure 2: Phylogenetically controlled regressions between root-to-tip dN/dS and absolute neonatal brain mass: for a) ASPM, b) CDK5RAP2, c) CENPJ, d) MCPH1. Data points are raw species values, the phylogenetically controlled regression line was estimated in BayesTraits and superimposed on top of species data. For ASPM two lines are displayed, the dashed line shows the regression when the outlier, Callithrix (labelled C) is included, the solid line shows the regression when it is excluded.
Table 1: Site models detecting positively selected sites for anthropoid primates*  

<table>
<thead>
<tr>
<th>Gene</th>
<th>N</th>
<th>lnL(null/M1a)</th>
<th>lnL(positive selection/M2a)</th>
<th>LRT statistic</th>
<th>p-value</th>
<th>Proportion of sites ω &gt;1</th>
<th>ω</th>
<th>Positively selected sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASPM</td>
<td>21</td>
<td>-20522.7</td>
<td>-20487.0</td>
<td>71.4</td>
<td>&lt;0.001</td>
<td>0.0228</td>
<td>5.394</td>
<td>195, 1407, 1437, 1558, 2016, 2185</td>
</tr>
<tr>
<td>CDK5RAP2</td>
<td>21</td>
<td>-7858.2</td>
<td>-7818.7</td>
<td>79.0</td>
<td>&lt;0.001</td>
<td>0.0650</td>
<td>4.418</td>
<td>853, 881, 929, 930, 948, 964, 977, 1581, 1683, 1687, 1688, 1690</td>
</tr>
<tr>
<td>CENPJ</td>
<td>21</td>
<td>-7695.8</td>
<td>-7684.1</td>
<td>23.4</td>
<td>&lt;0.001</td>
<td>0.0255</td>
<td>3.897</td>
<td>527, 813</td>
</tr>
<tr>
<td>MCPH1</td>
<td>21</td>
<td>-6375.3</td>
<td>-6369.0</td>
<td>12.6</td>
<td>0.002</td>
<td>0.0558</td>
<td>2.574</td>
<td>-</td>
</tr>
</tbody>
</table>

* Positively selected sites identified using Bayes Empirical Bayes are shown in the right hand column, numbered according to the full human coding sequence. Only sites with p>95% are presented; sites in bold have p>99%.
Table 2: Branch and Site model tests on separate anthropoid clades

<table>
<thead>
<tr>
<th>Gene</th>
<th>Model 2</th>
<th>Model 0</th>
<th>LRS</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dN/dS lineages leading to Homo</td>
<td>dN/dS 'non-human' lineages</td>
<td>lnL(M2)</td>
<td>lnL(M0)</td>
</tr>
<tr>
<td>ASPM</td>
<td>0.489</td>
<td>0.466</td>
<td>-21351.430</td>
<td>-21351.451</td>
</tr>
<tr>
<td>CDK5RAP2</td>
<td>0.729</td>
<td>0.676</td>
<td>-8718.412</td>
<td>-8718.440</td>
</tr>
<tr>
<td>CENPJ</td>
<td>0.623</td>
<td>0.503</td>
<td>-8664.800</td>
<td>-8664.986</td>
</tr>
<tr>
<td>MCPH1</td>
<td>0.793</td>
<td>0.579</td>
<td>-7301.153</td>
<td>-7301.505</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Model 2</th>
<th>Model 0</th>
<th>LRS</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dN/dS Apes</td>
<td>dN/dS OWM</td>
<td>dN/dS NWM</td>
<td>lnL(M2)</td>
</tr>
<tr>
<td>ASPM</td>
<td>0.473</td>
<td>0.415</td>
<td>0.479</td>
<td>-21350.890</td>
</tr>
<tr>
<td>CDK5RAP2</td>
<td>0.781</td>
<td>0.627</td>
<td>0.687</td>
<td>-8717.931</td>
</tr>
<tr>
<td>CENPJ</td>
<td>0.582</td>
<td>0.549</td>
<td>0.522</td>
<td>-8663.445</td>
</tr>
<tr>
<td>MCPH1</td>
<td>0.606</td>
<td>0.724</td>
<td>0.538</td>
<td>-7300.305</td>
</tr>
</tbody>
</table>
Table 3: Phylogenetically controlled regression analysis between root-to-tip dN/dS and brain size in anthropoid primates

<table>
<thead>
<tr>
<th>Gene</th>
<th>n</th>
<th>t-statistic</th>
<th>p-value</th>
<th>R²</th>
<th>t-statistic</th>
<th>p-value</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASPM</td>
<td>13</td>
<td>0.040</td>
<td>0.484</td>
<td>0.000</td>
<td>0.797</td>
<td>0.221</td>
<td>0.055</td>
</tr>
<tr>
<td>ASPM (without Callichinds)</td>
<td>12</td>
<td>2.420</td>
<td><strong>0.018</strong></td>
<td>0.369</td>
<td>0.967</td>
<td>0.178</td>
<td>0.086</td>
</tr>
<tr>
<td>CDK5RAP2</td>
<td>13</td>
<td>1.955</td>
<td><strong>0.039</strong></td>
<td>0.255</td>
<td>-0.500</td>
<td>1.000</td>
<td>0.022</td>
</tr>
<tr>
<td>CENPJ</td>
<td>13</td>
<td>0.631</td>
<td>0.270</td>
<td>0.035</td>
<td>-1.204</td>
<td>1.000</td>
<td>0.116</td>
</tr>
<tr>
<td>MCPH1</td>
<td>13</td>
<td>-0.697</td>
<td>1.000</td>
<td>0.042</td>
<td>-1.126</td>
<td>1.000</td>
<td>0.103</td>
</tr>
</tbody>
</table>

Table 4: Phylogenetically controlled regression analysis between root-to-tip dN/dS and relative testis mass

<table>
<thead>
<tr>
<th>Gene</th>
<th>n</th>
<th>t-statistic</th>
<th>p-value</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASPM</td>
<td>20</td>
<td>-1.396</td>
<td>1.000</td>
<td>0.097</td>
</tr>
<tr>
<td>CDK5RAP2</td>
<td>20</td>
<td>-0.621</td>
<td>1.000</td>
<td>0.021</td>
</tr>
<tr>
<td>CENPJ</td>
<td>20</td>
<td>0.824</td>
<td>0.210</td>
<td>0.036</td>
</tr>
<tr>
<td>MCPH1</td>
<td>20</td>
<td>1.96</td>
<td><strong>0.032</strong></td>
<td>0.176</td>
</tr>
</tbody>
</table>
Supplementary material:

Phenotypic data, accession numbers, additional analyses and results provided in a single doc file
Figure 2