Received: 16 February 2019 | Revised: 7 May 2019 | Accepted: 21 May 2019

DOI: 10.1111/gbb.12582

ORIGINAL ARTICLE

Genes, Brain and Behavior

Evolution of *ASPM* coding variation in apes and associations with brain structure in chimpanzees

Sheel V. Singh[1] | Nicky Staes[1,2,3] | Elaine E. Guevara[1] | Steven J. Schapiro[4] | John J. Ely[5] | William D. Hopkins[4] | Chet C. Sherwood[1] | Brenda J. Bradley[1]

[1]Center for the Advanced Study of Human Paleobiology, Department of Anthropology, The George Washington University, Washington, District of Columbia

[2]Behavioural Ecology & Ecophysiology Group, Department of Biology, University of Antwerp, Antwerp, Belgium

[3]Centre for Research and Conservation, Royal Zoological Society of Antwerp, Antwerp, Belgium

[4]Michael E. Keeling Center for Comparative Medicine and Research, University of Texas MD Anderson Cancer Center, Bastrop, Texas

[5]MAEBIOS, Alamogordo, New Mexico

Correspondence
Sheel Singh, Center for the Advanced Study of Human Paleobiology, The George Washington University, 800 22 St. NW, Suite 6000, Washington, DC 20052.
Email: sheel274@gwmail.gwu.edu

Funding information
George Washington University Luther Rice Foundation; Office of Multidisciplinary Activities, Grant/Award Number: 1542848; NSF INSPIRE, Grant/Award Number: SMA-1542848; National Chimpanzee Brain Resource, Grant/Award Number: NIH NS092988; James S. McDonnell Foundation, Grant/Award Number: 220020293

Abstract

Studying genetic mechanisms underlying primate brain morphology can provide insight into the evolution of human brain structure and cognition. In humans, loss-of-function mutations in the gene coding for ASPM (Abnormal Spindle Microtubule Assembly) have been associated with primary microcephaly, which is defined by a significantly reduced brain volume, intellectual disability and delayed development. However, less is known about the effects of common *ASPM* variation in humans and other primates. In this study, we characterized the degree of coding variation at *ASPM* in a large sample of chimpanzees (N = 241), and examined potential associations between genotype and various measures of brain morphology. We identified and genotyped five non-synonymous polymorphisms in exons 3 (V588G), 18 (Q2772K, K2796E, C2811Y) and 27 (I3427V). Using T1-weighted magnetic resonance imaging of brains, we measured total brain volume, cerebral gray and white matter volume, cerebral ventricular volume, and cortical surface area in the same chimpanzees. We found a potential association between *ASPM* V588G genotype and cerebral ventricular volume but not with the other measures. Additionally, we found that chimpanzee, bonobo, and human lineages each independently show a signature of accelerated *ASPM* protein evolution. Overall, our results suggest the potential effects of *ASPM* variation on cerebral cortical development, and emphasize the need for further functional studies. These results are the first evidence suggesting *ASPM* variation might play a role in shaping natural variation in brain structure in nonhuman primates.

KEYWORDS
apes, brain size, genetics, neurobiology, ventricular volume

1 | INTRODUCTION

The expansion of the cerebral cortex in relation to body size is a remarkable feature that distinguishes humans from other primates.[1,2] Studying genetic mechanisms underlying variation in primate brain morphology can provide insight into the evolution of human brain size, neural reorganization and cognition.[3-6] To elucidate the proximate mechanisms underlying human brain expansion, numerous studies have investigated evolutionary changes of specific genes underlying the size and structure of the human cerebral cortex.[7-11]

Of particular interest are genes that are associated with primary microcephaly (Microcephaly Primary Hereditary or MCPH), a congenital disorder that involves a reduction in brain size that is more than two standard deviations below the mean.[12] In humans, MCPH is

strongly correlated with mild-to-moderate intellectual disability, but no further neurological deficits or malformations are present.[13-15]

ASPM (Abnormal Spindle Microtubule Assembly) is a protein-coding gene located on chromosome 1q31.3 in humans that is associated with MCPH5, one of the most frequently occurring forms of MCPH.[16,17] Yet, the extent to which *ASPM* affects natural variation in brain morphology remains contentious. Magnetic Resonance Imaging (MRI) data have shown that loss of function non-synonymous genetic mutations and single nucleotide variants (SNVs) in the protein-coding region of *ASPM* are responsible for reduced brain volumes in humans[18,19]; however, other studies have also shown the opposite.[15,20,21] Individuals with *ASPM* mutations have shown impairment of various cognitive and behavioral abilities,[14,18,22] though the direct impact of such deficits has ranged from mild to severe. In addition, adaptive evolution of *ASPM* has been associated with variation in IQ in certain human studies.[23]

Studies that measure neuroanatomical abnormalities associated with *ASPM* loss-of-function mutations find that the degree of reduction differs among various brain regions. For example, while volume and surface area of many cortical regions are impacted in microcephaly, hippocampal volume and surrounding medial temporal lobe structures remain relatively preserved.[18] The differential impact of *ASPM* on brain volumetric measures may be due to the gene's role in the regulation of neuronal migration. In fetuses, *ASPM* is preferentially expressed during cerebral cortical neurogenesis.[14,24] In mice, expression of the *ASPM* homolog is reduced postnatally, suggesting that the gene is expressed in neural progenitor cells which solely produce neurons, rather than proliferative pools that generate both neurons and glia.[14,25] In addition, studies in mice show that *ASPM* is also important for maintaining a cleavage plane orientation, which promotes symmetric proliferative divisions of apical progenitors in the brain.[26,27] This role suggests that downregulation of the gene during neurogenesis promotes asymmetric divisions, therefore limiting the expansion of the neocortex in particular.[26] These findings support the role of *ASPM* in embryonic cortical neural progenitor proliferation and suggest that ASPM protein changes can contribute to variation in the development of the cerebral cortical morphology of the brain.

Evolutionary changes in the coding sequence of *ASPM* have also been shown to be associated with brain size evolution in nonhuman primates and other mammals.[28-31] For example, acceleration of non-synonymous substitutions in *ASPM* is significantly related to encephalization quotient (EQ) in cetaceans, suggesting that the gene may have contributed to relative brain size expansion.[31] Odontocete lineages with increased EQ show significant evidence of positive selection; however, absolute brain and body mass do not relate to *ASPM* selection rates.[31] In addition, *ASPM* variation is related to neonatal brain size in non-primate clades such as Glires and Euungulata.[32] Studies have also found signatures of positive selection in the hominoid (ie, ape) phylogeny, with higher ratios of non-synonymous to synonymous change found along the lineages leading to the human branch.[4,29,30,32] By contrast, the rates of ASPM protein evolution were significantly lower in New World monkeys, Old World monkeys, artiodactyls, carnivores, and rats.[28,31] Thus, accelerated evolution of *ASPM* may be unique to the hominoid lineage and is suggested to have been an important mechanism in the evolutionary history of human brain size.

ASPM's evolution in modern humans seems to have been driven by positive selection, causing recently emerged *ASPM* haplotypes to increase in frequency.[29,30,33] It is possible that human haplogroup variants arose in Eurasia and are still in the process of dispersion. Although *ASPM* shows strong signatures of adaptive evolution, studies examining *ASPM* variation have yielded mixed and controversial results in humans, as significant effects are largely associated with microcephaly and not common DNA variants in the general population.[18,21,23,34]

Despite the major interest in *ASPM* for its potential role in the evolution of the human brain, little research has been conducted regarding within-species inter-individual genetic variation in other primates. Therefore, the aims of this study were to identify coding variation in *ASPM* in chimpanzees (*Pan troglodytes*), one of the closest living relatives of humans, and to examine the potential association of such variation with measure of brain structure. Additionally, we sought to reexamine signatures of positive selection in *ASPM* among hominoids based on newly available data that increase depth of coverage and now includes bonobos.

2 | MATERIAL AND METHODS

2.1 | Subjects

This study used DNA samples for a total of 241 adult chimpanzees (*Pan troglodytes verus*). All subjects were members of the colonies of apes residing at the Yerkes National Primate Research Center (YNPRC) (N = 50 females and 44 males, ages ranging 6-53 years old) and the Michael E. Keeling Center for Comparative Medicine and Research of The University of Texas MD Anderson Cancer Center (KCCMR) (N = 84 females and 63 males, ages ranging from 8-51 years old). American Psychological Association and National Institute of Health guidelines for the ethical use of chimpanzees in research were adhered to during all aspects of this research.

2.2 | DNA extraction and genotyping

The reference genome Pan_tro_3.0 for chimpanzees is largely based on a single *P. t. verus* male, but includes additional lower-coverage genomic sequence data from seven other individuals.[35,36] We used these combined data for eight individuals and five previously deposited sequences from Genbank[37] (Accession numbers: NM_001008994 XM_514079, AY497016 AY365047, XM_016948364, XR_001715725, XM_016948430) to identify 24 non-synonymous single nucleotide variants (SNVs) in *ASPM*. Five of these SNVs are known to (a) cause amino acid substitutions, and (b) were polymorphic in our population based on an initial genotype screening using a panel of 10 unrelated individuals (Table 1). The other 19 non-synonymous loci were not polymorphic in our population. The five polymorphic SNVs were therefore considered promising and were further genotyped for the whole population of

TABLE 1 SNV genotype frequencies

SNV	Exon	N	Genotype (amino acid)	Number observed	Frequency
			AA (VV)	215	0.89
V588G	3	241	AC (VG)	24	0.10
			CC (GG)	2	0.01
			AA (KK)	4	0.03
Q2882K	18	165	AC (QK)	28	0.17
			CC (QQ)	133	0.80
			AA (EE)	2	0.01
K2796E	18	156	AG (KE)	14	0.09
			GG (KK)	140	0.90
			AA (YY)	10	0.07
C2811Y	18	148	AG (CY)	46	0.31
			GG (CC)	92	0.62
			AA (VV)	75	0.33
I3427V	27	228	AG (IV)	121	0.53
			GG (II)	32	0.14

241 chimpanzees. These candidates included an A/C SNV in exon 3 of *ASPM*, which causes a valine to be substituted for a glycine at amino acid position 588 (V588G). We also identified one A/C and two A/G SNVs in exon 18 of *ASPM* (Q2772K, K2796E, C2811Y) and one A/G SNV in exon 27 (I3427V). Identifications of ancestral and derived alleles were based on phylogenetics via the Ensembl genome browser.[38]

Genomic DNA was extracted from 200 μL blood samples (N = 234) using the QIAampDNA Mini Kit automated on a QiaCube (Qiagen, Hilden, Germany). For an additional seven individuals, we extracted DNA from formalin-fixed brain tissue using a QIAmp DNA FFPE Tissue Kit also automated on a QiaCube (Qiagen). Brain tissue sections were 40 μm thick with a maximum area of 250 mm^2 per section. DNA concentrations were quantified using a Nanodrop 2000 (ThermoFisher Scientific, Waltham, Massachusetts) spectrophotometer. Controls were included in all procedures.

Subsequent genotyping of the SNVs of interest (V588G, Q2772K, K2796E, C2811Y, I3427V) was performed using High-Resolution Melt Analysis (HRMA),[39,40] which detects DNA variants by characterizing the sequence-specific melting behavior of short PCR amplicons. All qPCR amplifications and melting curves were performed on a Rotor-Gene Q (Qiagen) real time platform. Primer pairs flanking the SNVs in exon 3 (forward: 5′-GGATTCTTTTGATTTCTCGCAC-3′ and reversed: 5′-CTTCGACAACAGCTTCAGTTG-3′), exon 18 (forward: 5′-CTGT GGTTATTCAGGCTGCATA-3′ and reversed: 5′-GTAGGGCAGCACA TTTCTGTG-3′), and exon 27 (5′-GAAGTAGGTCCAAAGTTGTTG-3′ and reversed: 5′-TATTCTGGTCCTTACAGGTGT-3′) were designed to target a short segment (approximately 100 bp) containing the polymorphic sites. The 25 μL PCR reaction mix contained 12.5 μL HRM MasterMix (Qiagen: (HotStarTaq plus DNA polymerase, EvaGreen dye, Q-Solution, deoxynucleotides, and $MgCL_2$), 1.75 μL primer mix (10 μM concentration of forward and reverse primers), and approximately 20 ng of genomic DNA. The PCR began with a 5-minute initial activation at 95°C, followed by 40 cycles of denaturation at 95°C for 10 seconds, and annealing/extension at 54°C for 30 seconds. Temperature was increased from 65 to 95°C at 0.1°C/2 second increments immediately after amplification in order to acquire melting pattern data for the qPCR products. Fluorescence data were plotted as a function of temperature during DNA denaturation (melting) and visualized and compared using the Rotor-Gene Q HRM software package (Qiagen).

We generated and compared melting temperatures and curve shape to genotype samples with up to four independent qPCR and high-resolution melt analyses per sample. This HRMA method of genotyping was validated and confirmed using an initial set of 25 individuals that were also genotyped via Sanger sequencing in both directions on an Applied Biosystems Genetic Analyzer platform (DNA Analysis Facility at Yale University). Multiple-sequence alignments of the resulting DNA sequences were performed using ClustalW[41] as implemented in Geneious (Version 6.0.6).

2.3 | MRI data

Magnetic resonance images (MRI) scans were previously collected for the same 241 adult and sub-adult chimpanzees (N = 107 males and 134 females, age range 6-53 years old). Chimpanzee MRIs were obtained from a data archive of scans collected prior to the 2015 implementation of U.S. Fish and Wildlife Service and National Institutes of Health regulations governing research with chimpanzees. These scans were made available through the National Chimpanzee Brain Resource (www.chimpanzeebrain.org). MRI scans were obtained in vivo where subjects were first immobilized by a ketamine injection (10 mg/kg) and then anesthetized with propofol (40-60 mg/[kg/h]) following standard procedures at the YNPRC and KCCMR. Subjects were subsequently transferred to the MRI facility and placed in the scanner chamber in a supine position with their head placed inside the human-head coil. Subjects were then scanned using a 1.5 Tesla scanner (N = 143) (Phillips,

Model 51, Bothell, Washington) or a 3.0 Tesla scanner (N = 66) (Siemans Inc., Malvern, Pennsylvania). The 1.5 Tesla scanner T1-weighted images were collected in the transverse plane using a gradient echo protocol (echo time = 8.5 ms, pulse repetition = 19.0 ms, number of signals averaged = 8, and a 256 × 256 matrix), and images from the 3.0 Tesla scanner were obtained using a 3D gradient echo sequence (echo time = 4.4 ms, pulse repetition = 2300 ms, number of signals averaged = 3, matrix size = 320 × 320). Scan duration ranged between 40 and 60 minutes as a function of brain size. All subjects remained anesthetized throughout the scanning process, including transportation time to and from the imaging facility (total time approximately 2 hours). After finishing the MRI procedure, subjects returned to their home facility and were held in a single cage for 6 to 12 hours in order to allow the anesthesia to wear off. They returned to their home cage immediately afterwards.

The acquired MRI data were transferred to a PC running Analyze 8.0 (Mayo Clinic, Rochester, Minnesota) software for post-image processing. MRI scans were skull-stripped and segmented into gray matter, white matter, and intracerebral ventricle space. They were subsequently aligned in the axial plane and cut into 1 mm slices using Analyze 8.0, and then imported into BrainVISA 4.0.1 (BV). The BV program used the gray/white interface generated in the sulci extraction pipeline process to calculate hemispheric gray and white matter volumes for each subject. Gray and white matter volumes included subcortical structures such as the thalamus and basal ganglia, but not the cerebellum or brainstem structures. Total cortical surface area and cerebral ventricular volume were also measured by the BV software.

2.4 | Statistical analysis

Statistical analysis was performed using statistical software program R (www.r-project.org, version 3.3.2). We tested for population adherence to Hardy-Weinberg equilibrium using the "HWAlltests" function in the Hardy-Weinberg package in R.[42] In addition, we analyzed group differences using a χ^2 test for independence.[43]

For testing *ASPM* genotype effects on brain morphology data extracted from MRI, we fitted linear mixed models using the lme4 function in R. We conducted normality testing and log transformed ventricular volume to adhere to linear model assumptions. Relatedness coefficients were used to correct for the degree of relatedness of each individual to all other individuals in the colony. Relatedness coefficients were extracted from pedigree data using the kinship2 package in R.[44] Given that we correct all our models for relatedness, all means reported in the results section are estimated marginal means. Fixed effects (sex, genotype) and covariates (age at time of MRI, relatedness) remained the same, but we included scanner as a random intercept in the linear mixed model to account for non-independence of data derived from two different scanners. Given that the CC genotype in exon 3 was rare (<1% of the population), we pooled it with the AC genotype when conducting statistical analysis. We did the same with two SNVs in exon 18 (Q2772K, K2796E), pooling the AA genotype with the AC and AG genotypes respectively. The Akaike information criterion (AIC) was used for statistical model selection for each combination of *ASPM* genotype and brain volumetric measure. Models with the lowest AIC were chosen as the best fit model. The significance of the fixed effects was tested by an F-test with a Kenward-Roger correction for the number of degrees of freedom using the add-on package, pbkrtest.

We conducted a gene linkage disequilibrium analysis using SNPStats, a web-based application that processes data using matrices with selected statistics (D, D', Pearson's *r*, and associated *P* values).[45] We evaluated the power (sensitivity) of our datasets using G*Power 3.1 software.[46] We assumed an at-best moderate effect, and followed Cohen recommended levels of effect sizes: small (w = 0.1), medium (0.3) and large (0.5).[47] We calculated the achieved power of each sample set using a medium effect size with α = 0.05.

2.5 | Function of coding variants

We assessed the potential effect of variants on protein function using SNAP2, a trained classifier that uses a neural network to infer functional consequences of known coding variants.[48] It considers various sequence and variant features to distinguish between effect and neutral variants. If the SNAP2 score exceeds 50, the effect of a variant is believed to be of importance to the native protein function. Variants with scores below −50 are considered to have a neutral effect, and scores between 50 and −50 are considered unreliable.

2.6 | Protein evolution

Altered selective pressures throughout evolution may affect patterns of genetic polymorphisms; therefore, we tested for evidence of positive selection on *ASPM*. We tested for evidence of positive selection on *ASPM* using a d_N/d_S-based framework implemented in codeml[49] within the python wrapper ete3.[50] We obtained aligned sequence data for 30 mammals, including 27 primates, from the UCSC Genome Browser (UCSC multiZ hg38 30-way primate alignment; http://genome.ucsc.edu/cgi-bin/hgTrackUi?db=hg38&g=cons30way). This dataset is more complete than previous protein evolution assessments of ASPM, as it is the first to include bonobos (*Pan paniscus*). We downloaded a species tree comprising the species in the alignment from the 10KTrees website.[51] The topology of this tree is concordant with a recent primate phylogeny estimated using genomic data.[52] We used branch[53] and branch-site[54] models to test whether accelerated evolution is observed across *ASPM* or at particular sites within *ASPM* on the human and ancestral African ape lineages, relative to the rest of the tree, as determined by likelihood ratio tests using an alpha threshold of 0.05.

3 | RESULTS

We identified 24 SNVs across *ASPM* in chimpanzees that resulted in non-synonymous substitutions, from which we identified five loci that resulted in amino acid changes and were polymorphic in our sampled population (Figure 1, Table 1). The A/C SNVs in exons 3 and 18 cause

valine to glycine and glutamine to lysine substitutions, respectively (V588G and Q2772K; Figure 1). Genotyping of these SNVs revealed that the derived C allele of V588G is present at a lower frequency in the population (0.048), while the derived C allele of Q2772K is present at a higher frequency in the population (0.887). The A/G SNVs in exon 18 cause lysine to glutamic acid and cysteine to tyrosine substitutions, respectively (K2796E and C2811Y). Both derived G alleles are present at higher frequencies in the population (0.936, 0.803). Lastly, genotyping of the A/G SNV in exon 27, which causes an isoleucine to valine substitution (I3427V), revealed that the derived allele (G) is present at a lower frequency in the population (0.401).

Three of the five SNVs were in Hardy-Weinberg equilibrium (V588G: X^2 = 1.39, df = 1, P = .239; K2796: X^2 = 1.89, df = 1, P = .169; I3427V: X^2 = 2.03, df = 1, P = .153). Although two loci were not in Hardy-Weinberg equilibrium, excess homozygotes were not present, indicating that allelic dropout is not likely to be a problem for our genotyping (Table 1). Additionally, a χ^2 test for all loci showed that allele frequencies between colonies were not significantly different (V588G: P = .064; Q2772K: P = .862; K2796: P = .917; C2811Y: P = .651; I3427V: P = .673). Population structure is unlikely to be a confounding issue as only a small portion of dyads in our sample set were first order relatives. (Yerkes: approximately 5% of dyads; Bastrop: approximately 3% of dyads).

SNAP2 prediction of the functional consequences of V588G, Q2772K, K2796E, C2811Y and I3427V substitutions resulted in SNAP2 effect scores of 43 (expected accuracy 73%), 75 (expected accuracy 85%), 13 (expected accuracy 59%), −17 (expected accuracy 57%), and −13 (expected accuracy 57%), respectively (Figure 2). In addition, gene linkage analysis showed high levels of linkage between several SNVs in *ASPM* (Supporting Information Figure S1).

The *ASPM* V588G genotype was significantly associated with total cerebral ventricular volume (F[1236] = 0.04), with AC individuals having greater ventricular size than individuals with an AA genotype (P = .037, B = 0.07) (Figure 3). No significant associations were found between other *ASPM* genotypes and total cortical surface area or measures for the volume of whole brain, white matter and gray matter. For a detailed description of all statistical results, see Appendix S1. For all five datasets, sample sizes were sufficient (at power ranging: 0.87-0.99, α = 0.05, medium effect size) to detect null results.

The branch model that best fit the dataset of *ASPM* evolution across hominoid phylogeny was one in which d_N/d_S ratios varied across all branches on the tree (P = .0001; significant after Bonferroni correction for multiple testing). While a model in which the human lineage has an elevated d_N/d_S relative to the rest of the tree (background d_N/d_S = 0.36, human branch = 0.87, P = .036) was better than the null model of a single d_N/d_S across the tree, this model was worse than the model in which d_N/d_S varied across lineages (P = .0003; significant after correction for multiple testing). Under the model of varying d_N/d_S, the human, chimpanzee, bonobo and ancestral great ape branches showed relatively high d_N/d_S ratios of >0.8, though no hominoid lineage exceeded 1.0 (Figure 4). None of the branch-site models were significantly better than the null model.

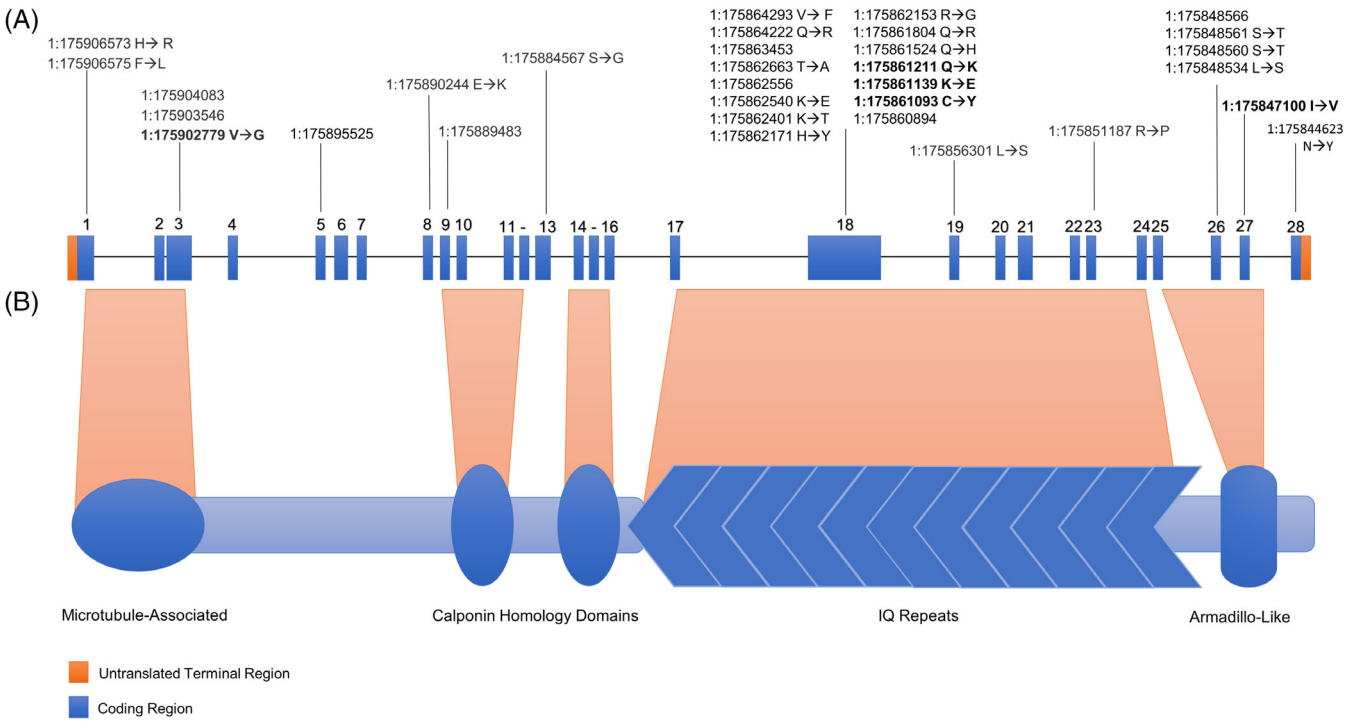

FIGURE 1 *ASPM* gene, protein, and sites of SNVs of interest. A, The exon/intron structure of *ASPM*. Positions of SNVs of interest are indicated in bold. B, The known and predicted domains of the ASPM protein: a microtubule binding domain, two calponin homology (CH) domains, a region of 81 IQ/calmodulin binding domains, and an armadillo-like C terminal domain. Modified from Nicholas et al 2009[64]

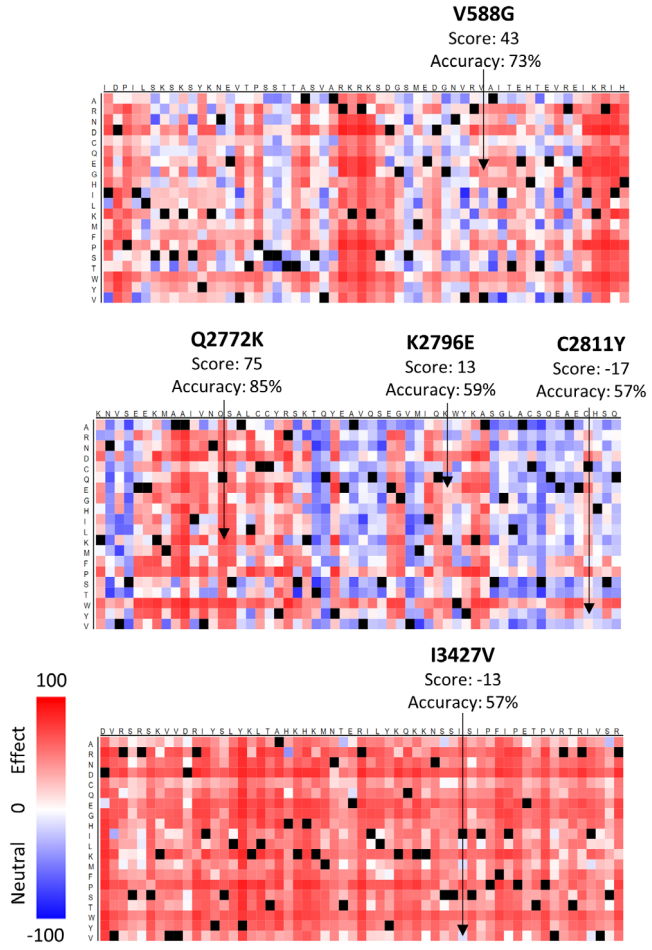

FIGURE 2 Heatmaps showing predicted functional consequences of V588G, Q2772K, K2796E, C2811Y, and I3427V. Stronger predicted effects are redder, and stronger predicted neutralities are bluer

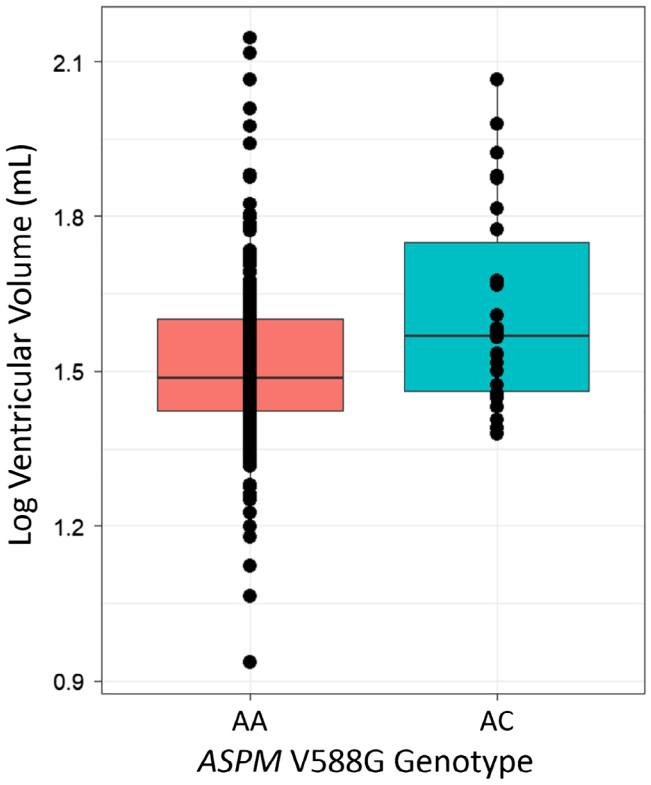

FIGURE 3 Log total ventricular volume for each *ASPM* genotype. The band inside the box represents the median, edges of the box represent lower and upper quartiles and whiskers represent minimums and maximums of the data. Outliers are indicated by individual points. Raw untransformed data show a similar distribution

4 | DISCUSSION

To date, research on *ASPM* variation has mostly focused on humans.[15,20,21] Our study aimed to characterize within-species variation in the coding region of *ASPM* in a nonhuman primate, the chimpanzee, using a relatively large sample size. We identified 24 single nucleotide polymorphisms across the full gene in chimpanzees, of which five were polymorphic in our subject population. We found one coding variant (V588G) to be potentially associated with a brain structural measure: cerebral ventricular volume.

Within-species comparison of V588G in chimpanzees revealed that the derived allele (C allele) is present at a very low frequency (0.048). The polymorphism leading to the V588G substitution is located on the third exon of *ASPM* which is one of the two exons that contain most of the coding region in the gene.[25] Exon 3 of *ASPM* is located in the microtubule binding domain of the ASPM protein (Figure 1), which causes ASPM to be transferred to the spindle poles during mitosis in a microtubule-dependent manner.[55] V588G is also strongly linked to two other SNVs on exons 18 and 27 (Figure S1). Although it is unclear which pathway V588G could affect, due to its location in the gene and its potential association with ventricular volume in this study, we suggest that it could have an impact on cerebral development.

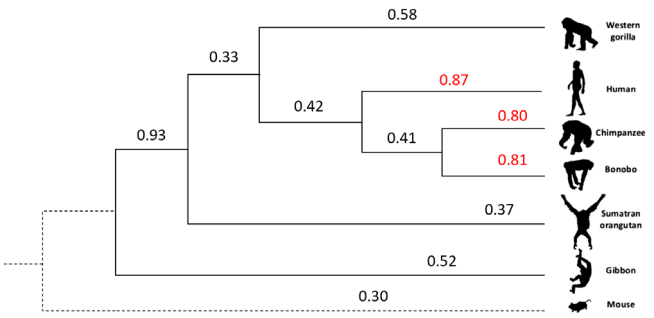

FIGURE 4 ASPM d_N/d_S ratios across apes. Higher ratios indicate higher rates of functional non-synonymous amino acid changes vs neutral synonymous changes

Our findings indicate that *ASPM* variation is potentially associated with cerebral ventricular volume in chimpanzees, but not with any of the other brain structure measures. Ventricles are a critical site of neuronal proliferation in early development.[56-58] Furthermore, the cerebrospinal fluid which circulates through the ventricles throughout life carries proteins that play important roles in central nervous

system development and maintenance,[57] like Sonic Hedgehog protein and Insulin-like Growth Factor 2.[56,57] Thus, variation in ventricular volume may affect the circulation of growth factors that could potentially influence the regulation of cerebral cortical development. Alternatively, because *ASPM* has a significant effect on neural progenitor cycling along the ventricles in fetal life,[26,59] the association shown in our study may be a result of how brain size is patterned by *ASPM* during neurogenesis in early development. It has been shown that *ASPM* plays a role in regulating the affinity of ventricular radial glial cells (VRGs) for the ventricular surface.[59]

Previous studies have shown that certain *ASPM* variants can have a significant impact on volume and surface area of cerebral cortical regions,[14,18,25,59] while other research suggests that common variants of *ASPM* do not contribute to human brain size variability.[15] In our study, we found a potential association with ventricular volume, but not other measures of brain volume and cortical surface area in chimpanzees. In any case, *ASPM* probably plays a subtle role in the development of brain structure but associations with any single variant should be interpreted with caution.

The abundance of *ASPM* mutations in human patients with microcephaly suggests that the gene plays a significant role in the regulation of brain size; however, variation in the gene has not always shown direct impact on brain circumference, volume, and intelligence in non-pathological populations.[15,23] It is possible that *ASPM* interacts with other genes to affect brain volume, and thus associations depend on genetic background.[60] Furthermore, selective pressure on *ASPM* may be associated with other aspects of neuronal function that do not lead to overt changes in brain structure, or might have a pleiotropic effect in other areas of the body, as *ASPM* is also expressed outside of the brain.[61] For example, *ASPM* is involved in sperm flagellar function, which may contribute to the rapid evolution of the gene as other proteins associated with male reproduction have also evolved rapidly.[5] Alternatively, it is possible that the SNVs studied here do not have a sufficient functional or structural impact on the protein, despite the amino acid change. However, one of the five SNVs we studied (Q2772K) had functional prediction scores that suggest the variants affect protein function.

It has been suggested that the accelerated evolution of *ASPM* may play a role in the expansion of the cerebral cortex in humans.[28] Studies show increased variation in *ASPM* across hominoids, as well as signatures of positive selection.[29,62] Unlike previous studies, we included bonobos when assessing protein evolution in *ASPM*. We ran branch-site and branch models and found high d_N/d_S ratios on the human, chimpanzee, bonobo and ancestral great ape lineages, suggesting that they have high rates of potentially functional (non-synonymous) amino acid changes compared to neutral (synonymous) changes.

Although the d_N/d_S ratios on all branches are less than one, the view that only a d_N/d_S greater than one could reflect positive selection has been recognized as overly conservative. Most proteins evolve under strong purifying selection at most sites while a few or even a single non-synonymous mutation can change protein structure and function.[54] What is even more intriguing is that accelerated evolution of *ASPM* occurred independently on each of these branches (Figure 4). Additionally, the ratios are particularly high among humans, chimpanzees, and bonobos compared to gorillas; however, the exact phenotypic impact of the variation found in this gene that is causing these species differences in evolutionary rate remains unclear. Moreover, since ASPM variation might also play a role in sperm function,[5] comparing ASPM evolutionary rates among apes might be confounded by variation in mating systems.

Overall, our results demonstrate the potential role of *ASPM* in cerebral cortical development in chimpanzees as suggested by the association with ventricular volume.[57] Further analysis of the functional and physiological mechanisms underlying *ASPM's* association with ventricular volume is needed. Although single nucleotide coding variants can have a marked impact on phenotype,[63] it is not necessarily surprising that these single SNVs, while some are likely to be functional, nevertheless are not directly associated with all measures of brain surface area and volume in chimpanzees.[15,21,23] Additional research on other mechanisms (eg, gene regulation) and other loci and pathways will help put our results in a broader context for understanding the proximate bases of primate brain variation.

ACKNOWLEDGMENTS

We thank collaborators for help with providing samples, specifically: the Yerkes National Primate Research Center and The Michael E. Keeling Center for Comparative Medicine and Research. We also thank all members of the Primate Genomics Lab and the Laboratory for Evolutionary Neuroscience at the George Washington University for helpful feedback on the project. This work was partially supported by the Luther Rice Foundation at the George Washington University, National Chimpanzee Brain Resource (NIH NS092988), James S. McDonnell Foundation (220020293), and NSF INSPIRE (SMA-1542848). We thank two anonymous reviewers for helpful comments.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

ORCID

Sheel V. Singh https://orcid.org/0000-0003-0877-0950
Brenda J. Bradley https://orcid.org/0000-0002-7084-0385

REFERENCES

1. Isler K, Van Schaik CP. How humans evolved large brains: comparative evidence. *Evol Anthropol*. 2014;23(2):65-75.
2. Rilling JK. Neuroscientific approaches and applications within anthropology. *Am J Phys Anthropol*. 2008;47:2-32.
3. Enard W. The molecular basis of human brain evolution. *Curr Biol*. 2016;26(20):R1117.
4. Evans PD, Anderson JR, Vallender EJ, Choi SS, Lahn BT. Reconstructing the evolutionary history of microcephalin, a gene

controlling human brain size. *Hum Mol Genet*. 2004;13(11):1139-1145. https://doi.org/10.1093/hmg/ddh126.
5. Ponting PC. A novel domain suggests a ciliary function for ASPM, a brain size determining gene. *Bioinformatics*. 2006;22(9):1031-1035.
6. Shi L, Lin Q, Su B. Human-specific hypomethylation of CENPJ, a key brain size regulator. *Mol Biol Evol*. 2014;31(3):594-604.
7. Beck KD, Powell-Braxton L, Widmer HR, Valverde J, Hefti F. Igf1 gene disruption results in reduced brain size, CNS hypomyelination, and loss of hippocampal granule and striatal parvalbumin-containing neurons. *Neuron*. 1995;14(4):717-730.
8. Bond J, Woods CG. Cytoskeletal genes regulating brain size. *Curr Opin Cell Biol*. 2006;18(1):95-101.
9. Montgomery SH, Capellini I, Venditti C, Barton RA, Mundy NI. Adaptive evolution of four microcephaly genes and the evolution of brain size in anthropoid primates. *Mol Biol Evol*. 2011;28(1):625-638.
10. Thompson PM, Cannon TD, Narr KL, et al. Genetic influences on brain structure. *Nat Neurosci*. 2001;4:1253-1258.
11. Peper JS, Brouwer RM, Boomsma DI, Kahn René S, Hulshoff Pol HE. Genetic influences on human brain structure: a review of brain imaging studies in twins. *Hum Brain Mapp*. 2007;28(6):464-473.
12. Woods CG, Bond J, Enard W. Autosomal recessive primary microcephaly (MCPH): a review of clinical, molecular, and evolutionary findings. *Am J Hum Genet*. 2005;76(5):717-728.
13. Kaindl AM, Passemard S, Kumar P, et al. Many roads lead to primary autosomal recessive microcephaly. *Prog Neurobiol*. 2010;90(3):363-383.
14. Kumar A, Blanton S, Babu M, Markandaya M, Girimaji S. Genetic analysis of primary microcephaly in Indian families: novel ASPM mutations. *Clin Genet*. 2004;66(4):341-348.
15. Woods PR, Freimer BN, De Young AJ, et al. Normal variants of microcephalin and ASPM do not account for brain size variability. *Hum Mol Genet*. 2006;15(12):2025-2029.
16. Bond J, Roberts E, Mochida GH, et al. ASPM is a major determinant of cerebral cortical size. *Nat Genet*. 2002;32(2):316-320.
17. Roberts E, Hampshire DJ, Pattison L, et al. Autosomal recessive primary microcephaly: an analysis of locus heterogeneity and phenotypic variation. *J Med Genet*. 2002;39(10):718-721.
18. Passemard S, Verloes A, Billette de Villemeur T, et al. Abnormal spindle-like microcephaly-associated (ASPM) mutations strongly disrupt neocortical structure but spare the hippocampus and long-term memory. *Cortex*. 2016;74:158-176.
19. Rimol LM, Agartz I, Djurovic S, et al. for the Alzheimer's Disease Neuroimaging Initiative Sex-dependent association of common variants of microcephaly genes with brain structure. *Proc Natl Acad Sci USA*. 2010;107(1):384-388.
20. Dobson-Stone C, Gatt JM, Kuan SA, et al. Investigation of MCPH1 G37995C and ASPM A44871G polymorphisms and brain size in a healthy cohort. *NeuroImage*. 2007;37(2):394-400.
21. Rushton JP, Vernon PA, Bons TA. No evidence that polymorphisms of brain regulator genes microcephalin and ASPM are associated with general mental ability, head circumference or altruism. *Biol Lett*. 2006;3(2):157-160. https://doi.org/10.1098/rsbl.2006.0586.
22. Shen J, Eyaid W, Mochida GH, et al. ASPM mutations identified in patients with primary microcephaly and seizures. *J Med Genet*. 2005;42(9):725-729.
23. Mekel-Bobrov N, Posthuma D, Gilbert LS, et al. The ongoing adaptive evolution of ASPM and microcephalin is not explained by increased intelligence. *Hum Mol Genet*. 2007;16(6):600-608. doi.
24. Buchman JJ, Durak O, Tsai L-H. ASPM regulates Wnt signaling pathway activity in the developing brain. *Genes Dev*. 2011;25(18):1909-1914. https://doi.org/10.1101/gad.16830211.
25. Bond J, Scott S, Hampshire DJ, et al. Protein-truncating mutations in ASPM cause variable reduction in brain size. *Am J Hum Genet*. 2003;73(5):1170-1177.
26. Fish JL, Kosodo Y, Enard W, Pääbo S, Huttner WB. Aspm specifically maintains symmetric proliferative divisions of neuroepithelial cells. *Proc Natl Acad Sci USA*. 2006;103(27):10438-10443. http://www.pnas.org/content/103/27/10438.abstract.
27. Fish JL, Dehay C, Kennedy H, Huttner WB. Making bigger brains–the evolution of neural-progenitor-cell division. *J Cell Sci*. 2008;121(17):2783-2793.
28. Evans PD, Anderson JR, Vallender EJ, et al. Adaptive evolution of ASPM, a major determinant of cerebral cortical size in humans. *Hum Mol Genet*. 2004;13(5):489-494.
29. Kouprina N, Pavlicek A, Mochida GH, et al. Accelerated evolution of the ASPM gene controlling brain size begins prior to human brain expansion. *PLoS Biol*. 2004;2(5):e126. https://doi.org/10.1371/journal.pbio.0020126.
30. Mekel-Bobrov N, Gilbert SL, Evans PD, et al. Ongoing adaptive evolution of ASPM, a brain size determinant in *Homo sapiens*. *Science*. 2005;309(5741):1720-1722.
31. Xu S, Sun X, Niu X, et al. Genetic basis of brain size evolution in cetaceans: insights from adaptive evolution of seven primary microcephaly (MCPH) genes. *BMC Evol Biol*. 2017;17(1):206.
32. Montgomery SH, Mundy NI. Microcephaly genes evolved adaptively throughout the evolution of eutherian mammals. *BMC Evol Biol*. 2014;14(1):120.
33. Ali F, Meier R. Positive selection in ASPM is correlated with cerebral cortex evolution across primates but not with whole-brain size. *Mol Biol Evol*. 2008;25(11):2247-2250.
34. Timpson N, Heron J, Smith GD, Enard W. Comment on papers by Evans et al. and Mekel-Bobrov et al. on evidence for positive selection of MCPH1 and ASPM. *Science*. 2007;317(5841):1036.
35. The Chimpanzee Sequencing and AC. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*. 2005;437:69-87.
36. Kuderna LFK, Tomlinson C, Hillier LW, et al. A 3-way hybrid approach to generate a new high-quality chimpanzee reference genome (Pan_tro_3.0). *Gigascience*. 2017;6(11):1-6.
37. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. *Nucleic Acids Res*. 2016;44:D72-D72.
38. Kersey Julian P, Allen EJ, Allot A, et al. Ensembl genomes 2018: an integrated omics infrastructure for non-vertebrate species. *Nucleic Acids Res*. 2017;46(D1):D808-D808.
39. Smith BL, LU C-P, Alvarado Bremer JR. High-resolution melting analysis (HRMA): a highly sensitive inexpensive genotyping alternative for population studies. *Mol Ecol Resour*. 2010;10(1):193-196. https://doi.org/10.1111/j.1755-0998.2009.02726.x.
40. Frankel DC, Jacobs RL, Staes N, et al. Rapid inexpensive genotyping and barcoding of primates: multiple applications for high-resolution melt analysis in primatology and anthropology. *Am J Phys Anthropol*. 2017;162:184-185.
41. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinform*. 2002;(1):2.3.1-2.3.22. https://doi.org/10.1002/0471250953.bi0203s00.
42. Jan Graffelman M. Package "HardyWeinberg" type package title graphical tests for Hardy-Weinberg Equilibrium, 2014. http://www-eio.upc.edu/~jan. Accessed February 14, 2019.
43. McHugh ML. The chi-square test of independence. *Biochem Med*. 2013;23(2):143-149. https://doi.org/10.11613/BM.2013.018.
44. Pedigree Functions [R package kinship2 version 1.6.4]. https://cran.r-project.org/web/packages/kinship2/index.html. Accessed February 14, 2019.
45. Solé X, Guinó E, Valls J, Iniesta R, Moreno V. SNPStats: a web tool for the analysis of association studies. *Bioinformatics*. 2006;22(15):1928-1929.
46. Faul F, Erdfelder E, Lang A-G, Buchner A. GPower 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175-191.
47. Cohen J. Statistical power analysis. *Curr Dir Psychol Sci*. 1992;1(3):98-101.

48. Hecht M, Bromberg Y, Rost B. Better prediction of functional effects for sequence variants. *BMC Genom*. 2015;16(8):S1.
49. Yang ZPAML. 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol*. 2007;24(8):1586-1591.
50. Huerta-Cepas J, Serra F, Bork P. ETE 3: reconstruction, analysis, and visualization of phylogenomic data. *Mol Biol Evol*. 2016;33(6):1635-1638.
51. Arnold C, Matthews LJ, Nunn CL. The 10kTrees website: a new online resource for primate phylogeny. *Evol Anthropol Iss*. 2010;19(3):114-118.
52. Reis Dos M, Gunnell FG, Barba-Montoya J, Wilkins A, Yang Z, Yoder DA. Using phylogenomic data to explore the effects of relaxed clocks and calibration strategies on divergence time estimation: primates as a test case. *Syst Biol*. 2018;67(4):594-615.
53. Yang Z, Nielsen R. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol*. 2002;19(6):908-917.
54. Zhang J, Nielsen R, Yang Z. Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol*. 2005;22(12):2472-2479.
55. Higgins J, Midgley C, Bergh A-M, et al. Human ASPM participates in spindle organisation, spindle orientation and cytokinesis. *BMC Cell Biol*. 2010;11(1):85.
56. Huang X, Liu J, Ketova T, et al. Transventricular delivery of sonic hedgehog is essential to cerebellar ventricular zone development. *Proc Natl Acad Sci USA*. 2010;107(18):8422-8427.
57. Lehtinen MK, Zappaterra MW, Chen X, et al. The cerebrospinal fluid provides a proliferative niche for neural progenitor cells. *Neuron*. 2011;69(5):893-905.
58. Martín C, Bueno D, Alonso MI, et al. FGF2 plays a key role in embryonic cerebrospinal fluid trophic properties over chick embryo neuroepithelial stem cells. *Dev Biol*. 2006;297(2):402-416.
59. Johnson MB, Sun X, Kodani A, et al. Aspm knockout ferret reveals an evolutionary mechanism governing cerebral cortical size. *Nature*. 2018;556(7701):370-375.
60. Jayaraman D, Kodani A, Gonzalez DM, et al. Microcephaly proteins Wdr62 and Aspm define a mother centriole complex regulating centriole biogenesis, apical complex and cell fate. *Neuron*. 2016;92(4):813-828.
61. Kouprina N, Pavlicek A, Collins Keith N, et al. The microcephaly ASPM gene is expressed in proliferating tissues and encodes for a mitotic spindle protein. *Hum Mol Genet*. 2005;14(15):2155-2165.
62. Zhang J. Evolution of the human ASPM gene, a major determinant of brain size. *Genetics*. 2003;165(4):2063-2070.
63. Flanagan N, Healy E, Ray A, et al. Pleiotropic effects of the melanocortin 1 receptor (MC1R) gene on human pigmentation. *Hum Mol Genet*. 2000;9(17):2531-2537.
64. Nicholas AK, Swanson EA, Cox JJ, et al. The molecular landscape of ASPM mutations in primary microcephaly. *J Med Genet*. 2009;46(4):249-253.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Singh SV, Staes N, Guevara EE, et al. Evolution of *ASPM* coding variation in apes and associations with brain structure in chimpanzees. *Genes, Brain and Behavior*. 2019;1–9. https://doi.org/10.1111/gbb.12582