Published in final edited form as:

Am J Phys Anthropol. 2014 April ; 153(4): 605–616. doi:10.1002/ajpa.22460.

The evolutionary history of *SLC6A4* and the role of plasticity in *Macaca*

Milena R. Shattuck¹, Jessica Satkoski-Trask², Amos Deinard³, Raul Y. Tito⁴, David G. Smith^{2,5}, and Ripan S. Malhi^{1,6}

¹Department of Anthropology, University of Illinois, Urbana, IL 61801

²Department of Anthropology, University of California, Davis, CA 95616

³Department of Anthropology, University of Minnesota, Minneapolis, MN 55455

⁴Department of Anthropology, University of Oklahoma, Norman, OK 73019

⁵California National Primate Research Center, Davis, CA 95616

⁶Department of Anthropology, Animal Biology, and Institute for Genomic Biology, University of Illinois, Urbana, IL 61801

Abstract

Serotonin has been repeatedly indicated as a biological marker of behavior. In particular, the serotonin transporter gene, *SLC6A4*, has been the focus of a large body of research. Interestingly, both rhesus macaques (*Macaca mulatta*) and humans have independently evolved a number of shared polymorphisms for this gene, which is indicative of parallel evolution between the two species. However, little is known about the evolution of this gene, particularly within macaques. While there are several hypotheses as to the adaptive values of various polymorphisms, few authors have gone beyond theoretical discussion. Here, we examined the genetic variation of *SLC6A4* within and between several species of macaques and investigate whether selection has played a significant role in its evolutionary history. In addition, we assayed the promoter region polymorphism, 5-HTTLPR, which is known to play a significant role in regulating both serotonin turnover and behavior.

In examining the distribution of the 5-HTTLPR polymorphism, we identified significant differences between Indian and Chinese populations of *M. mulatta*; furthermore, we discovered its presence in *M. cyclopis*, which has not been described before. In regard to the evolutionary history of *SLC6A4*, we found little evidence for selection and conclude that *SLC6A4* largely evolved through neutral processes, possibly due to its potential role in regulating behavioral plasticity. However, we also found very low levels of linkage between the coding regions and 5-HTTLPR. Because we limited evolutionary analyses to the coding regions, it is possible that the promoter region shows a distinct evolutionary history from *SLC6A4*.

Keywords

serotonin; natural selection; genes; behavior

In the search for the biological bases of behavior, one hormone, serotonin, has repeatedly made its way into much scientific research. For a variety of behaviors or psychological

conditions, such as aggression (Kruesi et al., 1990) alcoholism (Virkkunen et al., 1994), anxiety (Lesch et al., 1996), novelty seeking (Heck et al., 2009), depression (Caspi et al., 2003), impulse control (Linnoila et al., 1993; Mehlman et al., 1994; Westergaard et al., 2003b), or antisocialism (Flory et al., 2007), indices of serotonin levels and serotonin turnover are consistently among the best biological markers for predicting behavioral patterns. Consequently, the serotonin system is critical to the study of behavior, providing a potential proximate explanation to interspecific differences in behavior as well as a possible mechanism through which selection and behavioral evolution can occur. Thus, an examination of the evolution of the serotonin system, and the genetic variation underlying it, is warranted.

The general paradigm describing the association between serotonin and behavior is that serotonin helps to improve moods and to control aggressive, anxious, and impulsive behaviors. Thus, any genetic variation that modifies the degree of serotonin signaling might be expected to contribute to behavioral variation. Perhaps the most heavily studied gene in behavioral genetics is the serotonin transporter gene (*SLC6A4*, also referred to as *5-HTT* or *SERT*). In particular, there are two known polymorphic loci in this gene that correlate with behavior in humans. The first and most widely studied variation occurs in a regulatory region that lies approximately 1 kilobase (kb) upstream of *SLC6A4* (the serotonin transporter linked polymorphic region, 5-HTTLPR; Fig. 1). This polymorphic region is marked by an insertion/deletion (indel) and is typically described as having two variants, dubbed the “long” and “short” alleles (Heils et al., 1996; Lesch et al., 1996), although additional functional variants are likely (Nakamura et al., 2000). The short allele decreases transcriptional activity and serotonin turnover relative to the long allele (Lesch et al., 1996; Heils et al., 1997; Greenberg et al., 1999; Williams et al., 2001; Manuck et al., 2004; Smith et al., 2004), and seems to increase susceptibility to a wide variety of psychological conditions (for review see Hariri and Holmes, 2006; Serretti et al., 2006; Canli and Lesch, 2007; Caspi et al., 2010; Shattuck, 2011). The second polymorphism, STin2, is a variable number tandem repeat (VNTR) found in the second intron of *SLC6A4* (although alternative numbering schemes place it in the third intron: see Inoue-Murayama et al., 2008; Paredes et al., 2012). Like 5-HTTLPR, the STin2 VNTR affects transcriptional activity (Friskerstrand et al., 1999; MacKenzie and Quinn, 1999; Paredes et al., 2012) and is connected with behavior (Battersby et al., 1996; Collier et al., 1996; Ogilvie et al., 1996).

Intriguingly, macaques show a number of behavioral and genetic similarities with humans. In particular, an analogous polymorphism to 5-HTTLPR exists in rhesus macaques (*Macaca mulatta*) (Lesch et al., 1997; Rogers et al., 2006), a frequently used animal model in studies of medicine. Although independently evolved (Lesch et al., 1997), the macaque polymorphism appears to regulate serotonin functioning (Bennett et al., 2002) and behavior (e.g., Bennett et al., 1998; Trefilov et al., 2000; Champoux et al., 2002; Barr et al., 2003; Bethea et al., 2004; Brent et al., 2013) in a manner similar to humans. Furthermore, although the STin2 VNTR polymorphism has not been identified in macaques, additional polymorphisms in *SLC6A4* and other serotonin related genes have been identified in rhesus macaques that are similar to those found in humans (e.g., Newman et al., 2005; Vallender et al., 2008; Lindell et al., 2012). This suggests that parallel evolution of the serotonin system is occurring between the two species (Vallender et al., 2008). Indeed, serotonin is connected with several behaviors characteristic of invasive species, such as dispersal and exploratory behavior (Ramboz et al., 1998; Trefilov et al., 2000; Krackow and König, 2008) and rhesus macaques and humans are the two most widely distributed primate species. However, despite its potential to provide a greater understanding of behavioral evolution, as well as its impact on biomedical studies, the evolutionary history of *SLC6A4* in macaques is understudied and not well understood.

Most attempts at understanding the evolution of the serotonin system in primates have focused on the costs and benefits associated with differences in serotonin turnover. The observation that multiple polymorphisms affecting serotonin functioning have evolved and are maintained in both humans and macaques suggest that they are advantageous in some way – that is, that there is positive selection acting on *SLC6A4* and its regulatory regions. However, most of the research carried out to date seems to contradict this prediction. For example, in macaques lower serotonin levels have been shown to negatively affect rank (Higley et al., 1996b; Higley and Linnoila, 1997), survival (Higley et al., 1996a; Westergaard et al., 2003a; Howell et al., 2007), and reproductive success (Mehlman et al., 1997; Gerald et al., 2002; Cleveland et al., 2004; Hoffman et al., 2007). Therefore, any polymorphism that tends to decrease or disrupt serotonin turnover, such as the known variations in *SLC6A4*, would be expected to be selected against; and yet, such polymorphisms have not only evolved multiple times (Lesch et al., 1997; Vallender et al., 2008) but are maintain in high frequencies in both humans (e.g., Lesch et al., 1996) and macaques (e.g., Trefilov et al., 2000; Wendland et al., 2006).

This has led to a variety of hypotheses to explain the presence of genetic variation in the serotonin system that as a whole appears to be negative. Many of these hypotheses center on the idea of balancing selection. For example, Trefilov (2000) suggested that heterozygous advantage may explain the presence of both the long and short alleles of 5-HTTLPR in macaques. Other authors have suggested that while lower serotonin can be negative in some contexts, it may be positive in others, causing selective pressure to act in opposing ways across the population or species. For example, Howell and colleagues (2007) showed that, in rhesus macaques, individuals with lower levels of serotonin functioning were more likely to die at an early age due to higher rates of aggressive interactions. However, those that survived to adulthood were more likely to achieve a high rank. Thus, balancing selection can occur through disparities in selective pressures in different age groups (Gerald and Higley, 2002; Howell et al., 2007), between sexes (Westergaard et al., 2003b), in different social settings (Gerald and Higley, 2002; Lindell et al., 2012), and in different habitats (Suomi, 2006; Chakraborty et al., 2010; Brent et al., 2013).

Alternatively, others have argued that the same polymorphisms in *SLC6A4* that increase the risk of behavioral disorders in bad environments also increase the ability to *thrive* in good environments (Belsky et al., 2009; Caspi et al., 2010; Homberg and Lesch, 2011). As such, these genes are best viewed as contributing to plasticity, defined here as the degree to which a genotype expresses phenotypic variation across environments, or the norm of reaction (see Supporting Information Fig. S2); that is, certain polymorphisms will confer an increased sensitivity to environmental cues, both good and bad. For example, while the short allele of *SLC6A4* seems to increase risk of depression when combined with major life stressors (Caspi et al., 2003; Eley et al., 2004; Brummett et al., 2008), in the *absence* of stress the short allele increases *resistance* to depression (Eley et al., 2004; Brummett et al., 2008; reviewed in Belsky et al., 2009), thus showing a high response to environmental input. In contrast, the risk of depression for those carrying the long allele is far more consistent across various environments. This perspective introduces several additional hypotheses about the evolution of *SLC6A4*. It has been argued that increased plasticity played a key role in human evolution (Potts, 1996) and an adaptive feature of widespread organisms that occupy a variety of habitats (see Thompson, 1991), including macaques (Chakraborty et al., 2010). If plasticity itself were under selection, then positive selection for genetic variants of *SLC6A4* that increase plasticity would explain both their presence at high frequency and their independent evolution in humans and macaques.

In contrast, if the target of selection is a specific behavioral pattern, such as increased aggression, rather than the capacity to be more or less plastic, then there are least two

different outcomes. First, when the norms of reaction of different genotypes are non-parallel (that is, they show a gene-by-environment interaction) and cross over (Supporting Information Fig. S2), different genotypes may be selected for in different environments, again leading to balancing selection (Gillespie and Turelli 1989).¹ Second, it is possible that the diminished correlation between genotype and phenotype created by increased plasticity would make it difficult to detect selection at a genetic level, and the gene might be expected to evolve according to neutral evolutionary forces, particularly if the environments are not consistently distinct or if most environments occur where the norms of reaction cross (Wright, 1931; Gupta and Lewontin, 1982; Sultan, 1987; Via, 1987).

Thus, there are several potential evolutionary scenarios for serotonin related genes. However, despite widespread interest in the potential selective advantages and disadvantages for variants of *SLC6A4*, most studies have been limited to theoretical discussions, and therefore remain speculative. In this study, we use an empirical approach to understanding the evolutionary history of the serotonin system by exploring the genetic variation of *SLC6A4* within and between several species of macaques. In addition, we apply several tests of neutrality to determine whether selection – positive, purifying, or balancing – played a significant role in shaping genetic variation. Using a similar molecular evolutionary approach, Claw and colleagues (2010) examined *SLC6A4* in humans and found evidence of positive selection. If the serotonin system of rhesus macaques is evolving in parallel with humans, then we would expect to see similar evidence of positive selection for this species as well.

METHODS

Subjects

Genotyping of 5-HTTLPR was carried out using aliquots of previously extracted DNA of 70 *M. mulatta*, 12 *M. fascicularis*, 11 *M. fuscata*, six *M. nemestrina*, six *M. sylvanus* and one individual each from *M. assamensis*, *M. cyclopis*, *M. nigra*, and *M. silenus* (Supporting Information Tables S1 and S2). These same samples were used to obtain sequence data for evolutionary analyses, although for *M. mulatta* only a subset of the samples were analyzed (N=27; see below). Samples in this study have been used in previous studies (Smith and McDonough, 2005; Smith et al., 2007; Satkoski et al., 2008), and detailed information about sample origins can be found there. Individuals were of mixed sex, and nonrelatives from various geographic origins were selected based on breeding records or SNP/STR analyses (Satkoski et al., 2008). Samples that indicated possible hybridization were avoided, including between Chinese and Indian populations of *M. mulatta* (see below).

Most analyses focused on those species for which multiple samples were available: *M. mulatta*, *M. fascicularis*, *M. fuscata*, and *M. nemestrina*. (Only one *M. sylvanus* sample was successfully sequenced in all regions examined; Table S2.) The remaining samples were included in analyses requiring a phylogenetic framework. Where relevant, *M. sylvanus* was used as an outgroup (Supporting Information Fig. S1), being the sister species to the Asian macaques and the first to diverge in phylogenetic analyses (Tosi et al., 2003; Vos, 2006; Li et al., 2009).

¹Note, however, that this is slightly different from the other balancing selection hypotheses we described here. In those cases, the genotype-phenotype relationship remains stable, but the adaptive value of the phenotype varies depending on context. For example, an allele that causes anxious or fearful behavior may be negative in a low-risk environment, but beneficial when it produces that same behavior in a high-risk environment. In the case of plasticity and gene-by-environment interactions, the adaptive value of a phenotype remains the same, but the genotype-phenotype relationship varies across environments. Thus, anxiety might be universally negative, but the allele that produces anxiety differs across environments.

The *M. mulatta* samples come from two distinct areas within Asia: China and India (Satkoski et al., 2008). Previous studies indicate a significant amount of differentiation between these two populations of *M. mulatta* (Melnick et al., 1993; Morales and Melnick, 1998; Tosi et al., 2003; Smith, 2005; Hernandez et al., 2007; Satkoski et al., 2008). Because the presence of population structure can have a significant influence on several of the tests employed here, an analysis of molecular variation (AMOVA) (Weir and Cockerham, 1984; Excoffier et al., 1992; Weir, 1996) was carried out to estimate the amount of population differentiation seen in this gene using the program Arlequin (Excoffier et al., 2005). If the AMOVA was significant, we applied separate analyses to each of the *M. mulatta* populations.

Sequencing

All sequences generated for *SLC6A4* are available through GenBank (accession numbers KF623295-KF623524). The total size of *SLC6A4* is approximately 70,000 base pairs (bp), with 14 exons and 1,890 bp of coding regions (based on human reference sequence, Entrez gene ID 6532). To increase cost efficiency, we limited amplification and sequencing to three regions dispersed across the gene: 1) the start of the coding region, which includes Exon 2 and its flanking regions (982 bp); 2) Exons 3–7, their introns, and flanking regions (3,291 bp); and 3) Exon 14, the 3'UTR, and flanking regions (761 bp) (Fig. 1). For ease of discussion, hereafter these areas shall be referred to as the 5' region, the mid-region, and the 3' region, respectively. In total, we sequenced 5,034 bp of the gene, including 1,157 bp of coding sequences. Because of the large emphasis on 5-HTTLPR in the literature, we also genotyped all individuals of all species for this polymorphism. Finally, five additional noncoding regions were amplified and sequenced for use in the HKA test (Hudson et al., 1987; see below). We designed these noncoding regions to be at least 20,000 bp from the nearest coding region and spread across multiple chromosomes (Satkoski-Trask et al., 2011).

Amplification was carried out using primers based on the *Macaca mulatta* draft assembly (Rhesus Macaque Genome Sequencing and Analysis Consortium et al., 2007) and designed using the programs Primer3 (Rozen and Skaletsky, 2000) and GeneRunner (generunner.net). Specific PCR protocols differed for each of the regions amplified and is available at the request of the first author. Amplified samples were cleaned up using the ExoSAP-IT protocol (usb.com) and sent to the W. M. Keck Center for Comparative and Functional Genomics, UIUC, for Sanger sequencing. We took multiple steps to ensure the accuracy and quality of the sequences. First, sequencing primers were designed to provide overlap so that we had multiple reads for much of the sequence data. Second, to address the possibility of allelic dropout, which occurs when PCR preferentially amplifies only one chromosome, any sample that appeared to be homozygous at all sites across a region was re-amplified, beginning with the original DNA sample, using a different set of primers. Finally, a small subset of the samples were cloned and sequenced, and checked against the original data obtained.

Once received, sequences from all three regions (5', mid, and 3') were aligned and edited manually using Sequencher (www.genecodes.com). Each heterozygote base pair was confirmed visually by identifying clear double peaks in the chromatogram. All SNPs (single nucleotide polymorphisms) and indels (insertions/deletions) were identified. Haplotypes for each of the regions were determined using the program Phase v.1 (Stephens et al., 2001; Li and Stephens, 2003; Stephens and Donnelly, 2003). Haplotypes generated by Phase were checked against the haplotypes obtained from the subset of cloned samples.

In order to genotype 5-HTTLPR, approximately 400 bp surrounding the polymorphism were amplified. We then ran the PCR product through a 2% agarose gel for approximately 2

hours at 250V, which separates DNA bands based on their size. The gels were soaked in ethidium bromide and visualized under UV transillumination. The presence of the long or short allele was determined visually; a heterozygous individual showed two distinct bands, while those homozygous for the long or short allele showed a single band at specified levels in the gel. In order to accurately ascertain the frequency of the 5-HTTLPR alleles within and across *M. mulatta* populations, genotyping was carried out on a larger sample of *M. mulatta* (N=70) than was sequenced, in addition to all other non-*M. mulatta* individuals.

To visualize the relationship between the haplotypes, a network was created using the reduced median method in Network (fluxus-engineering.com). Additionally, a minimum spanning tree was generated using Arlequin (Excoffier et al., 2005) and HapStar (Teacher and Griffiths, 2011).

Analyses

The regions sequenced are widely separated from each other. In particular, there are approximately 15 kb separating the promoter region from the 5' region, and about 18 kb separating the mid-region from the 3' region. Although analyzing a single gene, the large size of *SLC6A4* makes recombination likely, a factor which can affect certain analyses (Hudson et al., 1987; Anisimova et al., 2003; Andolfatto, 2008). To account for this possibility, we performed separate analyses on each of the three regions sequenced, as well as on the combined haplotypes (including 5-HTTLPR). To determine the level of recombination, we implemented the method of Hey and Wakeley (1997) to obtain an estimate of the parameter C where $C = 2Nc$, and c is the rate of recombination per generation per base pair and N is the effective population size (Hey and Wakeley, 1997). We did this for each of the regions sequenced, the gene as a whole, and the coding region as a whole. We also conducted an exact test of linkage disequilibrium (LD) between all pairs of polymorphic loci to determine which pairs of loci are in significant LD, implemented in Arlequin (Lewontin and Kojima, 1960; Slatkin, 1994; Slatkin and Excoffier, 1996; Excoffier et al., 2005).

General indices of molecular variation were calculated using Arlequin (Excoffier et al., 2005). These included two different estimates of theta ($\theta = 4N\mu$, where N is the effective population size and μ is the mutation rate): θ_S (Watterson, 1975) and θ_π (Tajima, 1983), which were used to estimate within-species diversity. Nucleotide diversity (averaged over all loci) (Tajima, 1983) was used to estimate within- and among-species diversity. For the 5-HTTLPR polymorphism, a chi-square test of independence was used to determine if the populations are in Hardy-Weinberg equilibrium (HWE) and to see if there is a significant difference between the Chinese and Indian populations of *M. mulatta*.

In order to determine the role of selection on the genetic variation of *SLC6A4*, we applied several types of selection tests that use different aspects of the sequence data. Since several of these tests are sensitive to non-selective evolutionary forces such as demographics and recombination, a single significant result, without confirmation from other tests, should be treated with caution. However, if significant results are obtained from multiple tests, evidence of selection is much more robust. For all selection tests requiring a measure of interspecific diversity, *M. sylvanus* was used as the outgroup.

First, the HKA test (Hudson et al., 1987) examines the ratio of genetic variation within and between species across several loci, to determine if a gene of interest shows a distinct pattern from other loci. Specifically, a relative increase in between-species variation for *SLC6A4* would indicate positive selection, while a relative increase in within-species variation would indicate balancing selection. For this test, we used the gene of interest,

SLC6A4, and the five additional, presumably neutral loci. This was carried out using the program HKA, provided by J. Hey (<http://genfaculty.rutgers.edu/hey/software>).

We also examined the ratio of nonsynonymous to synonymous substitutions (ω) (Kimura, 1977) by using a z-test to determine whether ω was significantly different from one (indicative of neutral evolution). Specifically, $\omega > 1$ indicates positive selection, while $\omega < 1$ indicates purifying selection. The ω was calculated using the Nei-Gojobori method (Nei and Gojobori, 1986) and variance was estimated using the bootstrap method (Nei and Kumar, 2000). These analyses were carried out using the program MEGA v.4 (Tamura et al., 2007). In addition, we used McDonald-Kreitman (MK) test (McDonald and Kreitman, 1991) to look at ω within and between species using DNAsp (Rozas et al., 2003), which should remain constant for genes evolving neutrally.

Maximum likelihood methods were carried out using the codeml program in PAML (Phylogenetic Analysis using Maximum Likelihood) (Yang, 2007). PAML assesses models of evolution in a phylogenetic framework and determines a maximum likelihood value based on the given data. Models that incorporate selection are compared to null models using a likelihood ratio test (LRT) to determine which model best fits the data. We included all nine macaque species using a phylogeny based on previous molecular work (Deinard and Smith, 2001; Tosi et al., 2003; Vos, 2006; Li et al., 2009; Fig. S1). PAML uses only one sequence per species, so haplotypes were determined based on fixed differences between species and we did not perform separate analyses for Indian and Chinese macaques. Furthermore, since the codeml program employed here only uses information from coding regions, we pooled the data from the three regions sequenced and did not do separate analyses for each region. PAML was used to determine if 1) positive selection has occurred on a specific lineage, as indicated by an elevated ω (Yang, 1998), and 2) positive selection had occurred on specific sites within the gene (Nielsen and Yang, 1998; Yang et al., 2000; Swanson et al., 2003).

Finally, both Tajima's D (Tajima, 1989) and Fay and Wu's H (Fay and Wu, 2000) were calculated to detect skews in the frequency spectrum not expected under neutrality. Both tests are calculated by comparing two different estimates of theta. Tajima's D uses the estimates θ_S and θ_π . An excess of rare alleles will cause $\theta_S > \theta_\pi$ (leading to a significantly negative Tajima's D) and is consistent with positive selection. An excess of intermediate-frequency alleles will cause $\theta_S < \theta_\pi$ (leading to a significantly positive Tajima's D) and is consistent with balancing selection. For Fay and Wu's H, the two parameters that are compared are θ_π and θ_H . The parameter θ_H is most heavily affected by derived genetic variants (as determined by an outgroup) at high frequency. As such, an excess of high frequency, derived alleles will produce a significantly negative H-value ($\theta_\pi < \theta_H$), indicating positive selection. Significance for Fay and Wu's H was determined by running coalescent simulations using an estimate of the recombination rate and theta (θ_π) as parameters (Fay and Wu, 2000). DNAsp (Rozas et al., 2003) was used to determine the value and significance of Fay and Wu's H. The value and significance of Tajima's D was determined using Arlequin.

RESULTS

Sequencing, substructure, and recombination

For all three sequenced regions combined, we were able to obtain complete sequence from 20 *M. mulatta*, 7 *M. fascicularis*, 9 *M. fuscata*, 5 *M. nemestrina*, and 1 individual each of the remaining species, although for each region the number of samples successfully sequenced varied (Supporting Information Table S2).

An analysis of molecular variation (AMOVA) shows significant population differences between Chinese and Indian *M. mulatta* populations ($p < 0.0001$), with 17.9% of the total genetic variation for this species existing between groups. This is similar to results obtained for other serotonin genes examined (Shattuck, 2011). We therefore present the results for each of these populations separately, as well as the results obtained for the species as a whole.

In general, estimates of recombination are moderately high, even within regions; this is particularly true for the mid-region, where the highest estimates of C were obtained (Supporting Information Table S3). The results for the recombination analyses are reflected visually in the linkage disequilibrium (LD) plot shown in Figure 2. This plot shows a low level of significant LD between the polymorphic sites. Notably, the 5-HTTLPR polymorphism shows no LD with the 5' region, and only low LD with the other regions. This is in contrast to humans, which generally show more moderate levels of linkage between 5-HTTLPR and the coding portion of the gene (Claw et al., 2010). Despite the high level of recombination observed, most results do not differ between regions. For the remaining analyses, we report only the results for the entire gene, except where results of analyses of separate regions differ.

Molecular diversity

General indices of molecular diversity are shown in Table 1 and in Supporting Information Table S4. There are very few nonsynonymous polymorphisms; in fact, there are very few mutations within the coding regions in general. Instead, most mutations, both within and between species, occur either within the introns, at an average rate of 1 mutation per 22 bp, or in the UTR regions, at an average rate of 1 mutation per 29 bp. In contrast, mutations within the coding region only occur once per 129 bp. A minimum spanning tree showing the relationship between all haplotypes is shown in Figure 3.

We ascertained the frequency of the 5-HTTLPR alleles across *M. mulatta* populations using a much larger sample size ($N=70$). In total we genotyped 29 individuals from India, 37 individuals from China, and 4 individuals of unknown origin. All of the individuals of unknown origin are homozygous for the long (L) allele. The frequency of the alleles and genotypes for the Chinese and Indian populations are shown in Table 2. The Chinese population has a higher frequency of the short allele (S) than the Indian population, and this is significant ($p = 0.035$). Both populations are in HWE.

In addition to the *M. mulatta* samples, we also screened the other macaque species for the presence of the 5-HTTLPR polymorphism. Most species do not show any variation; however, the *M. cyclopis* individual appears to be heterozygous (i.e., it possesses both a long and a short allele). This is the first time that this polymorphism has been reported in *M. cyclopis*, although the sample awaits cloning and sequencing for confirmation.

Selection

The results for PAML are all nonsignificant. That is, models incorporating selection, either along a single lineage or within parts of the gene, do not fit the data better than null models that assumed neutral evolution (Supporting Information Table S5). This is also true for the remainder of the selection tests when applied to all three sequenced regions combined, results of which are shown in Table 3.

Separate analyses of each of the regions largely replicate these results. However, analyses of the 5' region alone do generate significantly different results for Fay and Wu's H . In *M. mulatta*, the H -value calculated is significant (Table 4). Examination of the two *M. mulatta*

subgroups show that the H-value for the Indian population is significant ($p = 0.008$), while it is not significant for the Chinese population ($p = 0.246$). A look at the haplotype network for this region shows the presence of a derived haplotype at a high frequency in Indian rhesus macaques, but rare in Chinese macaques (Fig. 4); this haplotype is most likely contributing to the significant H-value obtained for this population. The haplotype makes up 88.5% of the Indian population and is defined by the presence of two indels in the second intron, in the homologous region of the STin2 VNTR in humans. The first indel is a 33-bp insertion, which is comprised of two additional repeat units of 17 and 16 bp each, for a total of 7 repeat units in the VNTR. The second is a 12-bp deletion found just downstream of the VNTR region. These two indels co-segregate 100% and are shared by another, closely related haplotype that is found exclusively in the Indian samples (Fig. 4). Thus, these two polymorphic sites are present in 96.2% of the Indian group. By contrast, it is found in only 31.5% of the Chinese group. A chi-square test of independence shows this difference between populations is highly significant ($p < 0.0001$).

DISCUSSION

The gene for the serotonin transporter, *SLC6A4*, has been widely studied for its connection to behavior, particularly in humans and macaques. However, the evolutionary history of this gene is not well understood. Various authors have suggested a range of evolutionary scenarios for this gene, but to date very few studies have examined the molecular variation of the gene itself in an evolutionary context. Here, we examined *SLC6A4* to determine the level of genetic variation within and between macaque species and to explore the possible role of selection on the evolutionary history of this gene.

One of the most heavily investigated polymorphisms for *SLC6A4* is the 5-HTTLPR promoter region polymorphism. Until recently, among macaques this polymorphism was only known within *M. mulatta*, although it is now known to be present in *M. radiata* and *M. munzala* (Sinha et al., 2005; Chakraborty et al., 2010). The analyses presented here indicate the presence of this polymorphism in *M. cyclopis* as well. *M. cyclopis* and *M. mulatta* are very closely related, having only diverged from each other approximately 276 thousand years ago (Vos, 2006), so the presence of this polymorphism is not surprising. However, this sample awaits sequencing to verify the exact nature of this polymorphism. Furthermore, a wider sampling of individuals within *M. cyclopis* is needed, so at present this finding is suggestive but not confirmed.

Within *M. mulatta*, there is a significant difference in the distribution of the short and long alleles across populations. Specifically, the short allele is present in much higher frequency in the Chinese population than in the Indian population. This finding is consistent with studies on Chinese and Indian macaque behavior, which have generally found significant differences in temperament. For example, a previous study on the temperament of *M. mulatta* neonates showed that Chinese-Indian hybrids were more likely to score lower on tests of orientation and sustained attention and were more reactive and irritable than Indian-derived neonates (Champoux et al., 1994); in a similar manner, indices of serotonin turnover differed between the two groups (Champoux et al., 1997). Thus, differences in the prevalence of the short allele may explain some of the behavioral differences seen between these two populations.

Importantly for evolutionary studies on the 5-HTTLPR, there is a very low level of linkage between the polymorphism and the beginning of the coding region of *SLC6A4*. This is in contrast to the results of Claw and colleagues (2010), which showed higher indices of linkage looking at the analogous polymorphism in humans. Given the population history of humans, a higher level of LD might be expected (Pritchard and Przeworski, 2001), which

increases the likelihood that the promoter region and the gene share similar evolutionary histories. For macaques, however, this low level of LD means two things. First, the promoter region and the main coding portion of the gene could have distinct evolutionary histories, and the results we obtained in this study from the sequence data cannot meaningfully be extended to the promoter region. Second, since the studies examining gene-by-environment interactions in macaques are limited to 5-HTTLPR, this means that without further testing we cannot say that variations in the coding regions and their introns show a similar level of environmental interaction.

In fact, in examining *SLC6A4*, we found very little evidence for either positive or balancing selection acting on this gene. Even ω , which is typically low in a functional gene due to purifying selection, is not significantly different than 1. This is not to suggest that purifying selection does not occur on this gene – the overall low levels of mutations that occur in exons compared to introns or the UTR would suggest otherwise. Rather, it indicates that the synonymous sites are not completely neutral and have also been subjected to purifying selection.

The most notable exception to the general results was found in the separate examination of the 5' region. For this region, which included the second intron and flanking areas, we found a significant Fay and Wu's H-value for *M. mulatta*. This seems to be due to a derived haplotype that is found at a high frequency in the Indian population. This could be the result of a bottleneck, which is known to occur in the recent history of Indian *M. mulatta* (Hernandez et al., 2007). Furthermore, given that this was the only test with a significant result, and only when considered in isolation from the other areas sequenced, evidence of selection is weak, and an interpretation of neutrality is a stronger fit for the data.

Regardless of the role of selection on this region, this test highlighted two newly described polymorphisms that are derived and largely separate Indian *M. mulatta* from the other samples. Both of these polymorphisms are located in or very near the STin2 VNTR region. Within hominids, this region is highly variable, with the number of repeat units ranging between 4 (*Pongo pygmaeus*) and 40 (*Gorilla gorilla*); humans possess 9–12 repeat units (Soeby et al., 2005; Inoue-Murayama et al., 2008). In contrast, all Old World monkeys examined prior to this have shown only 5 repeat units (Soeby et al., 2005; Inoue-Murayama et al., 2008). Here we show that *M. mulatta* does show variation in this region, with an alternative pattern of 7 repeat units. It is known that this region has a regulatory effect on *SLC6A4* (Fiskerstrand et al., 1999; MacKenzie and Quinn, 1999; Paredes et al., 2012) and influences behavior (Ogilvie et al., 1996; Vormfelde et al., 2006). While we cannot confirm the phenotypic effect of the variants discovered here, it seems reasonable to predict they will be similar to those seen in other species. Thus, the presence of these polymorphisms may potentially contribute to behavioral differences seen between these two populations, or between *M. mulatta* and other macaque species. Furthermore, it provides additional evidence of genetic convergence between macaques and humans.

Results of the 5' region notwithstanding, analyses indicate that genetic drift and other non-selective evolutionary forces are the predominant mechanisms shaping genetic variation for *SLC6A4*. This does not rule out *SLC6A4*'s adaptive importance. Indeed, the low level of linkage between this gene and its promoter region leaves open the possibility that selection has acted primarily through regulatory mechanisms. It is feasible that a separate examination of sequence data surrounding 5-HTTLPR would yield different results. In fact, while not explicitly tested, Lindell and colleagues (2012) provide evidence that is suggestive of balancing selection in the promoter region in macaques, in contrast to what we found here. This evidence, combined with our results on the STin2 VNTR region, further stress that selection, if present, is acting on very specific regions of *SLC6A4*. While selection might be

expected to increase LD between the gene and the promoter region (Sabeti et al., 2002), our analyses show that recombination is high for this area of the genome. Thus, unless selection occurred very recently, recombination would quickly break down any LD (Przeworski, 2002).

It has been argued that genes such as *SLC6A4* might be best viewed as plasticity genes (Belsky et al., 2009; Caspi et al., 2010; Homberg and Lesch, 2011); that is, different alleles for *SLC6A4* have varying sensitivity to environmental variation. As mentioned previously, this has only been established for 5-HTTLPR in macaques, and given the absence of LD between this polymorphism and the rest of the gene, we cannot establish that this is true for all polymorphisms. Nevertheless, if we assume that *SLC6A4* as a whole contributes to plasticity (Lesch, 2001), then the expectations for selection on *SLC6A4* are not very straightforward. If behavioral plasticity itself was the target of selection, then we might make clear predictions about the role of positive selection on *SLC6A4*, and the focus of study should be on the degree that an individual responds to environmental variability (see, for example, Bell and Robinson, 2011). However, if plasticity is not being selected for, then the possible impact of selection is much more unclear. If past environmental circumstances were widely variable, as might be expected in the course of macaque evolution (Delson, 1980; Jablonski et al., 2000), we might expect to see balancing selection (Gillespie and Turelli, 1989). However, if environments were not consistently distinct, or if most environments occur where norms of reactions cross such that phenotypic differences between alleles were rarely expressed (Wright, 1931; Gupta and Lewontin 1982; Sultan 1987; Via 1987; see also Lindell et al., 2012), the poor correlation between genotype and phenotype would make selection ineffective at producing genetic evolution, and the gene in question would largely evolve in a neutral fashion. The results shown here support this interpretation and contrast with other hypotheses about the evolution of *SLC6A4*, such as heterozygous advantage (Trefilov et al., 2000) or selection for maintenance of genetic diversity due to variable environments (e.g., Suomi, 2006).

In summary, we have shown that the 5-HTTLPR polymorphism of *SLC6A4* is significantly variable across *M. mulatta* populations and have furthermore discovered its presence in *M. cyclopis*, which has not been described before. We also demonstrate that the 5-HTTLPR polymorphism is not linked to coding regions of *SLC6A4* and may therefore have a distinct evolutionary history. In examining *SLC6A4*, we found almost no evidence of selection. However, we did find two large polymorphisms in the second intron that clearly separate Indian *M. mulatta* from Chinese *M. mulatta* and other macaque species, and establish that variation in the STin2 VNTR region does exist outside of hominids. Finally, we argue that the neutral evolution seen in this gene supports the idea that the plasticity conferred by *SLC6A4* limits the influence of selection. As such, the evolutionary focus for this gene should not only be on specific behaviors, but should also incorporate behavioral plasticity. This shift in focus could help to explain inconsistent results seen across study populations (Belsky et al., 2009; Homberg and Lesch, 2011), and will lead to a more nuanced understanding of behavioral evolution.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

GRANT SPONSORSHIP: National Science Foundation (BCS-0925458)

We would like to thank S. Williams, A. Bell, K. Milich, M. Grabowski, P. Jelinek, and two anonymous reviewers for helpful feedback on the manuscript. This research was funded by a grant from the National Science Foundation

(BCS-0925458) and a portion of the samples provided by the National Non-Human Primate DNA Bank (NIH grant RR000163).

LITERATURE CITED

- Andolfatto P. Controlling type-I error of the McDonald-Kreitman test in genomewide scans for selection on noncoding DNA. *Genetics*. 2008; 180:1767–1771. [PubMed: 18791238]
- Anisimova M, Nielsen R, Yang Z. Effect of recombination on the accuracy of the likelihood method for detecting positive selection at amino acid sites. *Genetics*. 2003; 164:1229–1236. [PubMed: 12871927]
- Barr CS, Newman TK, Becker ML, Champoux M, Lesch KP, Suomi SJ, Goldman D, Higley JD. Serotonin transporter gene variation is associated with alcohol sensitivity in rhesus macaques exposed to early-life stress. *Alcohol Clin Exp Res*. 2003; 27:812–817. [PubMed: 12766626]
- Battersby S, Ogilvie AD, Smith CA, Blackwood DH, Muir WJ, Quinn JP, Fink G, Goodwin GM, Harmar AJ. Structure of a variable number tandem repeat of the serotonin transporter gene and association with affective disorder. *Psychiatr Genet*. 1996; 6:177–181. [PubMed: 9149321]
- Bell AM, Robinson GE. Behavior and the dynamic genome. *Science*. 2011; 332:1161–1162. [PubMed: 21636765]
- Belsky J, Jonassaint C, Pluess M, Stanton M, Brummett B, Williams R. Vulnerability genes or plasticity genes? *Mol Psychiatry*. 2009; 14:746–754. [PubMed: 19455150]
- Bennett AJ, Lesch KP, Heils A, Linnoila M. Serotonin transporter gene variation, CSF 5-HIAA concentrations, and alcohol-related aggression in rhesus monkeys (*Macaca mulatta*). *Am J Primatol*. 1998; 45:168–169.
- Bennett AJ, Lesch KP, Heils A, Long JC, Lorenz JG, Shoaf SE, Champoux M, Suomi SJ, Linnoila MV, Higley JD. Early experience and serotonin transporter gene variation interact to influence primate CNS function. *Mol Psychiatry*. 2002; 7:118–122. [PubMed: 11803458]
- Bethea CL, Streicher JM, Coleman K, Pau FK, Moessner R, Cameron JL. Anxious behavior and fenfluramine-induced prolactin secretion in young rhesus macaques with different alleles of the serotonin reuptake transporter polymorphism (5HTTLPR). *Behav Genet*. 2004; 34:295–307. [PubMed: 14990868]
- Brent LJN, Heilbronner SR, Horvath JE, Gonzalez-Martinez J, Ruiz-Lambides AV, Robinson A, Skene JHP, Platt ML. Genetic origins of social networks in rhesus macaques. *Sci Rep*. 2013; 3:1042. [PubMed: 23304433]
- Brummett BH, Boyle SH, Siegler IC, Kuhn CM, Ashley-Koch A, Jonassaint CR, Züchner S, Collins A, Williams RB. Effects of environmental stress and gender on associations among symptoms of depression and the serotonin transporter gene linked polymorphic region (5-HTTLPR). *Behav Genet*. 2008; 38:34–43. [PubMed: 17955359]
- Canli T, Lesch KP. Long story short: the serotonin transporter in emotion regulation and social cognition. *Nat Neurosci*. 2007; 10:1103–1109. [PubMed: 17726476]
- Caspi A, Hariri AR, Holmes A, Uher R, Moffitt TE. Genetic sensitivity to the environment: the case of the serotonin transporter gene and its implications for studying complex diseases and traits. *Am J Psychiatry*. 2010; 167:509–527. [PubMed: 20231323]
- Caspi A, Sugden K, Moffitt TE, Taylor A, Craig IW, Harrington H, McClay J, Mill J, Martin J, Braithwaite A, Poulton R. Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science*. 2003; 301:386–389. [PubMed: 12869766]
- Chakraborty S, Chakraborty D, Mukherjee O, Jain S, Ramakrishnan U, Sinha A. Genetic polymorphism in the serotonin transporter promoter region and ecological success in macaques. *Behav Genet*. 2010; 40:672–679. [PubMed: 20411319]
- Champoux M, Bennett A, Shannon C, Higley JD, Lesch KP, Suomi SJ. Serotonin transporter gene polymorphism, differential early rearing, and behavior in rhesus monkey neonates. *Mol Psychiatry*. 2002; 7:1058–1063. [PubMed: 12476320]
- Champoux M, Higley JD, Suomi SJ. Behavioral and physiological characteristics of Indian and Chinese-Indian hybrid rhesus macaque infants. *Dev Psychobiol*. 1997; 31:49–63. [PubMed: 9222116]

- Champoux M, Suomi SJ, Schneider ML. Temperament differences between captive Indian and Chinese-Indian hybrid rhesus macaque neonates. *Lab Anim Sci.* 1994; 44:351–357. [PubMed: 7983847]
- Claw KG, Tito RY, Stone AC, Verrelli BC. Haplotype structure and divergence at human and chimpanzee serotonin transporter and receptor genes: implications for behavioral disorder association analyses. *Mol Biol Evol.* 2010; 27:1518–1529. [PubMed: 20118193]
- Cleveland A, Westergaard GC, Trenkle MK, Higley JD. Physiological predictors of reproductive outcome and mother-infant behaviors in captive rhesus macaque females (*Macaca mulatta*). *Neuropsychopharmacology.* 2004; 29:901–910. [PubMed: 14666122]
- Collier DA, Arranz MJ, Sham P, Battersby S, Vallada H, Gill P, Aitchison KJ, Sodhi M, Li T, Roberts GW, Smith B, Morton J, Murray RM, Smith D, Kirov G. The serotonin transporter is a potential factor for bipolar affective disorder. *Neuroreport.* 1996; 7:1675–1679. [PubMed: 8904780]
- Deinard A, Smith DG. Phylogenetic relationships among the macaques: evidence from the nuclear locus *NRAMP1*. *J Hum Evol.* 2001; 41:45–59. [PubMed: 11414773]
- Delson, E. Fossil macaques, phyletic relationships and a scenario of deployment. In: Lindburg, DG., editor. *The macaques: studies in ecology, behavior and evolution*. New York: Van Nostrand Reinhold; 1980. p. 10-30.
- Eley TC, Sugden K, Corsico A, Gregory AM, Sham P, McGuffin P, Plomin R, Craig IW. Gene-environment interaction analysis of serotonin system markers with adolescent depression. *Mol Psychiatry.* 2004; 9:908–915. [PubMed: 15241435]
- Excoffier L, Laval G, Schneider S. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol Bioinform Online.* 2005; 1:47–50. [PubMed: 19325852]
- Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics.* 1992; 131:479–491. [PubMed: 1644282]
- Fay JC, Wu CI. Hitchhiking under positive Darwinian selection. *Genetics.* 2000; 155:1405–1413. [PubMed: 10880498]
- Ferguson B, Capitanio J, Folks T, Hotchkiss C, Johnson Z, Kean L, Kubisch HM, Lank S, Lyons L, Miller GM, Nylander J, O'Connor D, Vallender EJ, Wiseman R. Resource brief: The National Non-Human Primate DNA Bank. *Methods.* 2009; 49:3–4. [PubMed: 19706346]
- Fiskerstrand CE, Lovejoy EA, Quinn JP. An intronic polymorphic domain often associated with susceptibility to affective disorders has allele dependent differential enhancer activity in embryonic stem cells. *FEBS Lett.* 1999; 458:171–174. [PubMed: 10481059]
- Flory JD, Newcorn JH, Miller C, Harty S, Halperin JM. Serotonergic function in children with attention-deficit hyperactivity disorder: relationship to later antisocial personality disorder. *Br J Psychiatry.* 2007; 190:410–414. [PubMed: 17470955]
- Gerald MS, Higley JD. Evolutionary underpinnings of excessive alcohol consumption. *Addiction.* 2002; 97:415–425. [PubMed: 11964058]
- Gerald MS, Higley S, Lussier ID, Westergaard GC, Suomi SJ, Higley JD. Variation in reproductive outcomes for captive male rhesus macaques (*Macaca mulatta*) differing in CSF 5-hydroxyindoleacetic acid concentrations. *Brain Behav Evol.* 2002; 60:117–124. [PubMed: 12373062]
- Gillespie JH, Turelli M. Genotype-environment interactions and the maintenance of polygenic variation. *Genetics.* 1989; 121:129–138. [PubMed: 17246488]
- Greenberg BD, Tolliver TJ, Huang SJ, Li Q, Bengel D, Murphy DL. Genetic variation in the serotonin transporter promoter region affects serotonin uptake in human blood platelets. *Am J Med Genet.* 1999; 88:83–87. [PubMed: 10050973]
- Gupta AP, Lewontin RC. A study of reaction norms in natural populations of *Drosophila pseudoobscura*. *Evolution.* 1982; 36:934–948.
- Hariri AR, Holmes A. Genetics of emotional regulation: the role of the serotonin transporter in neural function. *Trends Cogn Sci.* 2006; 10:182–191. [PubMed: 16530463]
- Heck A, Lieb R, Ellgas A, Pfister H, Lucae S, Roeske D, Putz B, Muller-Myhsok B, Uhr M, Holsboer F, Ising M. Investigation of 17 candidate genes for personality traits confirms effects of the HTR2A gene on novelty seeking. *Genes Brain Behav.* 2009; 8:464–472. [PubMed: 19566713]

- Heils A, Mossner R, Lesch KP. The human serotonin transporter gene polymorphism- basic research and clinical implications. *J Neural Transm.* 1997; 104:1005–1014. [PubMed: 9503253]
- Heils A, Teufel A, Petri S, Stober G, Riederer P, Bengel D, Lesch KP. Allelic variation of human serotonin transporter gene expression. *J Neurochem.* 1996; 66:2621–2624. [PubMed: 8632190]
- Hernandez RD, Hubisz MJ, Wheeler DA, Smith DG, Ferguson B, Rogers J, Nazareth L, Indap A, Bourquin T, McPherson J, Muzny D, Gibbs R, Nielsen R, Bustamante CD. Demographic histories and patterns of linkage disequilibrium in Chinese and Indian rhesus macaques. *Science.* 2007; 316:240–243. [PubMed: 17431170]
- Hey J, Wakeley J. A coalescent estimator of the population recombination rate. *Genetics.* 1997; 145:833–846. [PubMed: 9055092]
- Higley JD, Linnoila M. Low central nervous system serotonergic activity is traitlike and correlates with impulsive behavior. A nonhuman primate model investigating genetic and environmental influences on neurotransmission. *Ann NY Acad Sci.* 1997; 836:39–56. [PubMed: 9616793]
- Higley JD, Mehlman PT, Higley SB, Fernald B, Vickers J, Lindell SG, Taub DM, Suomi SJ, Linnoila M. Excessive mortality in young free-ranging male nonhuman primates with low cerebrospinal fluid 5-hydroxyindoleacetic acid concentrations. *Arch Gen Psychiatry.* 1996; 53:537–543. [PubMed: 8639037]
- Higley JD, Suomi SJ, Linnoila M. A nonhuman primate model of type II excessive alcohol consumption? Part 1. Low cerebrospinal fluid 5-hydroxyindoleacetic acid concentrations and diminished social competence correlate with excessive alcohol consumption. *Alcohol Clin Exp Res.* 1996; 20:629–642. [PubMed: 8800378]
- Hoffman JB, Kaplan JR, Kinkead B, Berga SL, Wilson ME. Metabolic and reproductive consequences of the serotonin transporter promoter polymorphism (5-HTTLPR) in adult female rhesus monkeys (*Macaca mulatta*). *Endocrine.* 2007; 31:202–211. [PubMed: 17873333]
- Homberg JR, Lesch KP. Looking on the bright side of serotonin transporter gene variation. *Biol Psychiatry.* 2011; 69:513–519. [PubMed: 21047622]
- Howell S, Westergaard G, Hoos B, Chavanne TJ, Shoaf SE, Cleveland A, Snoy PJ, Suomi SJ, Dee Higley J. Serotonergic influences on life-history outcomes in free-ranging male rhesus macaques. *Am J Primatol.* 2007; 69:851–865. [PubMed: 17330868]
- Hudson RR, Kreitman M, Aguade M. A test of neutral molecular evolution based on nucleotide data. *Genetics.* 1987; 116:153–159. [PubMed: 3110004]
- Inoue-Murayama M, Hibino E, Iwatsuki H, Inoue E, Hong K-W, Nishida T, Hayasaka I, Ito S, Murayama Y. Interspecies and intraspecies variations in the serotonin transporter gene intron 3 VNTR in nonhuman primates. *Primates.* 2008; 49:139–142. [PubMed: 18204817]
- Jablonski NG, Whitfort MJ, Roberts-Smith N, Qinqi X. The influence of life history and diet on the distribution of catarrhine primates during the Pleistocene in eastern Asia. *J Hum Evol.* 2000; 39:131–157. [PubMed: 10968926]
- Kimura M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature.* 1977; 267:275–276. [PubMed: 865622]
- Krackow S, König B. Microsatellite length polymorphisms associated with dispersal-related agonistic onset in male wild house mice (*Mus musculus domesticus*). *Behav Ecol Sociobiol.* 2008; 62:813–820.
- Kruesi MJ, Rapoport JL, Hamburger S, Hibbs E, Potter WZ, Lenane M, Brown GL. Cerebrospinal fluid monoamine metabolites, aggression, and impulsivity in disruptive behavior disorders of children and adolescents. *Arch Gen Psychiatry.* 1990; 47:419–426. [PubMed: 1691910]
- Lesch KP. Variation of serotonergic gene expression: neurodevelopment and the complexity of response to psychopharmacologic drugs. *Eur Neuropsychopharm.* 2001; 11:457–474.
- Lesch KP, Bengel D, Heils A, Sabol SZ, Greenberg BD, Petri S, Benjamin J, Muller CR, Hamer DH, Murphy DL. Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. *Science.* 1996; 274:1527–1531. [PubMed: 8929413]
- Lesch KP, Meyer J, Glatz K, Flugge G, Hinney A, Hebebrand J, Klauck SM, Poustka A, Poustka F, Bengel D, Mossner R, Riederer P, Heils A. The 5-HT transporter gene-linked polymorphic region (5-HTTLPR) in evolutionary perspective: alternative biallelic variation in rhesus monkeys. *J Neural Transm.* 1997; 104:1259–1266. [PubMed: 9503271]

- Lewontin RC, Kojima K. The evolutionary dynamics of complex polymorphisms. *Evolution*. 1960; 14:458–472.
- Li J, Han K, Xing J, Kim HS, Rogers J, Ryder OA, Disotell T, Yue B, Batzer MA. Phylogeny of the macaques (Cercopithecidae: *Macaca*) based on *Alu* elements. *Gene*. 2009; 448:242–249. [PubMed: 19497354]
- Li N, Stephens M. Modeling linkage disequilibrium and identifying recombination hotspots using single-nucleotide polymorphism data. *Genetics*. 2003; 165:2213–2233. [PubMed: 14704198]
- Lindell SG, Yuan Q, Zhou Z, Goldman D, Thompson RC, Lopez JF, Suomi SJ, Higley JD, Barr CS. The serotonin transporter gene is a substrate for age and stress dependent epigenetic regulation in rhesus macaque brain: potential roles in genetic selection and gene $\times$ environment interactions. *Dev Psychopathol*. 2012; 24:1391–1400. [PubMed: 23062305]
- Linnoila M, Virkkunen M, George T, Higley D. Impulse control disorders. *Int Clin Psychopharmacol*. 1993; 8(Suppl 1):53–56. [PubMed: 7504705]
- MacKenzie A, Quinn J. A serotonin transporter gene intron 2 polymorphic region, correlated with affective disorders, has allele-dependent differential enhancer-like properties in the mouse embryo. *PNAS*. 1999; 96:15251–15255. [PubMed: 10611371]
- Manuck SB, Flory JD, Ferrell RE, Muldoon MF. Socio-economic status covaries with central nervous system serotonergic responsivity as a function of allelic variation in the serotonin transporter gene-linked polymorphic region. *Psychoneuroendocrinology*. 2004; 29:651–668. [PubMed: 15041087]
- McDonald JH, Kreitman M. Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature*. 1991; 351:652–654. [PubMed: 1904993]
- Mehlman PT, Higley JD, Faucher I, Lilly AA, Taub DM, Vickers J, Suomi SJ, Linnoila M. Low CSF 5-HIAA concentrations and severe aggression and impaired impulse control in nonhuman primates. *Am J Psychiatry*. 1994; 151:1485–1491. [PubMed: 7522411]
- Mehlman PT, Higley JD, Fernald BJ, Sallee FR, Suomi SJ, Linnoila M. CSF 5-HIAA, testosterone, and sociosexual behaviors in free-ranging male rhesus macaques in the mating season. *Psychiatry Res*. 1997; 72:89–102. [PubMed: 9335200]
- Melnick DJ, Hoelzer GA, Absher R, Ashley MV. mtDNA diversity in rhesus monkeys reveals overestimates of divergence time and paraphyly with neighboring species. *Mol Biol Evol*. 1993; 10:282–295. [PubMed: 8487631]
- Morales JC, Melnick DJ. Phylogenetic relationships of the macaques (Cercopithecidae: *Macaca*), as revealed by high resolution restriction site mapping of mitochondrial ribosomal genes. *J Hum Evol*. 1998; 34:1–23. [PubMed: 9467779]
- Nakamura M, Ueno S, Sano A, Tanabe H. The human serotonin transporter gene linked polymorphism (5-HTTLPR) shows ten novel allelic variants. *Mol Psychiatry*. 2000; 5:32–38. [PubMed: 10673766]
- Nei M, Gojobori T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol*. 1986; 3:418–426. [PubMed: 3444411]
- Nei, M.; Kumar, S. Molecular evolution and phylogenetics. New York: Oxford University Press; 2000.
- Newman TK, Syagailo YV, Barr CS, Wendland JR, Champoux M, Graessle M, Suomi SJ, Higley JD, Lesch KP. Monoamine oxidase A gene promoter variation and rearing experience influences aggressive behavior in rhesus monkeys. *Biol Psychiatry*. 2005; 57:167–172. [PubMed: 15652876]
- Nielsen R, Yang Z. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics*. 1998; 148:929–936. [PubMed: 9539414]
- Ogilvie AD, Battersby S, Bubbs VJ, Fink G, Harmar AJ, Goodwin GM, Smith CA. Polymorphism in serotonin transporter gene associated with susceptibility to major depression. *Lancet*. 1996; 347:731–733. [PubMed: 8602004]
- Paredes UM, Bubbs VJ, Haddley K, Macho GA, Quinn JP. Intronic tandem repeat in the serotonin transporter gene in Old World monkeys: a new transcriptional regulator? *J Mol Neurosci*. 2012; 47:401–407. [PubMed: 22038691]
- Potts R. Evolution and climate variability. *Science*. 1996; 273:922–923.
- Pritchard JK, Przeworski M. Linkage disequilibrium in humans: models and data. *Am J Hum Genet*. 2001; 69:1–14. [PubMed: 11410837]

- Przeworski M. The signature of positive selection at randomly chosen loci. *Genetics*. 2002; 160:1179–1189. [PubMed: 11901132]
- Ramboz S, Oosting R, Amara DA, Kung HF, Blier P, Mendelsohn M, Mann JJ, Brunner D, Hen R. Serotonin receptor 1A knockout: an animal model of anxiety-related disorder. *Proc Natl Acad Sci USA*. 1998; 95:14476–14481. [PubMed: 9826725]
- Rhesus Macaque Genome Sequencing and Analysis Consortium. Evolutionary and biomedical insights from the rhesus macaque genome. *Science*. 2007; 316:222–234. [PubMed: 17431167]
- Rogers J, Kaplan J, Garcia R 4th, Shelledy W, Nair S, Cameron J. Mapping of the serotonin transporter locus (SLC6A4) to rhesus chromosome 16 using genetic linkage. *Cytogenet Genome Res*. 2006; 112:341A.
- Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R. DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*. 2003; 19:2496–2497. [PubMed: 14668244]
- Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol*. 2000; 132:365–386. [PubMed: 10547847]
- Sabeti PC, Reich DE, Higgins JM, Levine HZ, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ, McDonald GJ, Ackerman HC, Campbell SJ, Altshuler D, Cooper R, Kwiatkowski D, Ward R, Lander ES. Detecting recent positive selection in the human genome from haplotype structure. *Nature*. 2002; 419:832–837. [PubMed: 12397357]
- Satkoski JA, Malhi RS, Kanthaswamy S, Tito RY, Malladi VS, Smith DG. Pyrosequencing as a method for SNP identification in the rhesus macaque (*Macaca mulatta*). *BMC Genomics*. 2008; 9:256. [PubMed: 18510772]
- Satkoski-Trask JAS, Malhi RS, Kanthaswamy S, Johnson J, Garnica SW, Malladi VS, Smith DG. The influence of SNP discovery method and sample size on the estimation of population genetic parameters in Chinese and Indian rhesus macaques (*Macaca mulatta*). *Primates*. 2011; 52:129–138. [PubMed: 21207104]
- Serretti A, Calati R, Mandelli L, De Ronchi D. Serotonin transporter gene variants and behavior: a comprehensive review. *Curr Drug Targets*. 2006; 7:1659–1669. [PubMed: 17168841]
- Shattuck, MR. PhD dissertation. University of Illinois; Urbana-Champaign: 2011. The molecular evolution of the serotonin system in macaques (*Macaca*): a detailed survey of four serotonin-related genes.
- Sinha A, Datta A, Madhusudan MD, Mishra C. *Macaca munzala*: a new species from western Arunachal Pradesh, northeastern India. *Int J Primatol*. 2005; 26:977–989.
- Slatkin M. Linkage disequilibrium in growing and stable populations. *Genetics*. 1994; 137:331–336. [PubMed: 8056320]
- Slatkin M, Excoffier L. Testing for linkage disequilibrium in genotypic data using the expectation-maximization algorithm. *Heredity*. 1996; 76:377–383. [PubMed: 8626222]
- Smith DG. Genetic characterization of Indian-origin and Chinese-origin rhesus macaques (*Macaca mulatta*). *Comp Med*. 2005; 55:227–230. [PubMed: 16089169]
- Smith GS, Lotrich FE, Malhotra AK, Lee AT, Ma Y, Kramer E, Gregersen PK, Eidelberg D, Pollock BG. Effects of serotonin transporter promoter polymorphisms on serotonin function. *Neuropsychopharmacology*. 2004; 29:2226–2234. [PubMed: 15354180]
- Smith DG, McDonough J. Mitochondrial DNA variation in Chinese and Indian rhesus macaques (*Macaca mulatta*). *Am J Primatol*. 2005; 65:1–25. [PubMed: 15645455]
- Smith DG, McDonough JW, George DA. Mitochondrial DNA variation within and among regional populations of longtail macaques (*Macaca fascicularis*) in relation to other species of the *fascicularis* group of macaques. *Am J Primatol*. 2007; 69:182–198. [PubMed: 17177314]
- Soeby K, Larsen SA, Olsen L, Rasmussen HB, Werge T. Serotonin transporter: evolution and impact of polymorphic transcriptional regulation. *Am J Med Genet B Neuropsychiatr Genet*. 2005; 136B: 53–57. [PubMed: 15858819]
- Stephens M, Donnelly P. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *Am J Hum Genet*. 2003; 73:1162–1169. [PubMed: 14574645]
- Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet*. 2001; 68:978–989. [PubMed: 11254454]

- Sultan, SE. Evolutionary implications of phenotypic plasticity in plants. In: Hecht, MK.; Wallace, B.; Prance, GT., editors. *Evolutionary biology*. Vol. 21. New York: Plenum Press; 1987. p. 127-178.
- Suomi SJ. Risk, resilience, and gene× environment interactions in rhesus monkeys. *Ann N Y Acad Sci*. 2006; 1094:52–62. [PubMed: 17347341]
- Swanson WJ, Nielsen R, Yang Q. Pervasive adaptive evolution in mammalian fertilization proteins. *Mol Biol Evol*. 2003; 20:18–20. [PubMed: 12519901]
- Tajima F. Evolutionary relationship of DNA sequences in finite populations. *Genetics*. 1983; 105:437–460. [PubMed: 6628982]
- Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 1989; 123:585–595. [PubMed: 2513255]
- Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol*. 2007; 24:1596–1599. [PubMed: 17488738]
- Teacher AG, Griffiths DJ. HapStar: automated haplotype network layout and visualization. *Mol Ecol Resour*. 2011; 11:151–153. [PubMed: 21429113]
- Thompson JD. Phenotypic plasticity as a component of evolutionary change. *TREE*. 1991; 6:246–249. [PubMed: 21232470]
- Tosi AJ, Morales JC, Melnick DJ. Paternal, maternal, and biparental molecular markers provide unique windows onto the evolutionary history of macaque monkeys. *Evolution*. 2003; 57:1419–1435. [PubMed: 12894949]
- Trefilov A, Berard J, Krawczak M, Schmidtke J. Natal dispersal in rhesus macaques is related to serotonin transporter gene promoter variation. *Behav Genet*. 2000; 30:295–301. [PubMed: 11206084]
- Vallender EJ, Priddy CM, Hakim S, Yang H, Chen GL, Miller GM. Functional variation in the 3' untranslated region of the serotonin transporter in human and rhesus macaque. *Genes Brain Behav*. 2008; 7:690–697. [PubMed: 18445138]
- Via, S. Genetic constraints on the evolution of phenotypic plasticity. In: Loeschcke, V., editor. *Genetic constraints on adaptive evolution*. Berlin: Springer-Verlag; 1987. p. 47-71.
- Virkkunen M, Rawlings R, Tokola R, Poland RE, Guidotti A, Nemeroff C, Bissette G, Kalogeras K, Karonen SL, Linnoila M. CSF biochemistries, glucose metabolism, and diurnal activity rhythms in alcoholic, violent offenders, fire setters, and healthy volunteers. *Arch Gen Psychiatry*. 1994; 51:20–27. [PubMed: 7506515]
- Vormfelde SV, Hoell I, Tzvetkov M, Jamrozinski K, Sehr D, Brockmoller J, Leibing E. Anxiety- and novelty seeking-related personality traits and serotonin transporter gene polymorphisms. *J Psychiatr Res*. 2006; 40:568–576. [PubMed: 16313923]
- Vos, RA. *Inferring large phylogenies: The big tree problem*. Simon Fraser University; 2006.
- Watterson GA. On the number of segregating sites in genetical models without recombination. *Theor Popul Biol*. 1975; 7:256–276. [PubMed: 1145509]
- Weir, BS. *Genetic data analysis II*. Sunderland, MA: Sinauer Associates; 1996.
- Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. *Evolution*. 1984; 38:1358–1370.
- Wendland JR, Lesch KP, Newman TK, Timme A, Gachot-Neveu H, Thierry B, Suomi SJ. Differential functional variability of serotonin transporter and monoamine oxidase a genes in macaque species displaying contrasting levels of aggression-related behavior. *Behav Genet*. 2006; 36:163–172. [PubMed: 16402281]
- Westergaard GC, Cleveland A, Trenkle MK, Lussier ID, Higley JD. CSF 5-HIAA concentration as an early screening tool for predicting significant life history outcomes in female specific-pathogen-free (SPF) rhesus macaques (*Macaca mulatta*) maintained in captive breeding groups. *J Med Primatol*. 2003a; 32:95–104. [PubMed: 12823632]
- Westergaard GC, Suomi SJ, Chavanne TJ, Houser L, Hurley A, Cleveland A, Snoy PJ, Higley JD. Physiological correlates of aggression and impulsivity in free-ranging female primates. *Neuropsychopharmacology*. 2003b; 28:1045–1055. [PubMed: 12700686]
- Williams RB, Marchuk DA, Gadde KM, Barefoot JC, Grichnik K, Helms MJ, Kuhn CM, Lewis JG, Schanberg SM, Stafford-Smith M, Suarez EC, Clary GL, Svenson IK, Siegler IC. Central

nervous system serotonin function and cardiovascular responses to stress. *Psychosom Med.* 2001; 63:300–305. [PubMed: 11292279]

Wright S. Evolution in Mendelian populations. *Genetics.* 1931; 16:97–159. [PubMed: 17246615]

Yang Z. Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Mol Biol Evol.* 1998; 15:568–573. [PubMed: 9580986]

Yang Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 2007; 24:1586–1591. [PubMed: 17483113]

Yang Z, Nielsen R, Goldman N, Pedersen AM. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics.* 2000; 155:431–449. [PubMed: 10790415]

Figure 1.

A schematic of the *SLC6A4* gene. The vertical lines represent coding regions, the closed boxes represent UTR regions, and the shaded box represents the promoter region, including 5-HTTLPR. The three regions sequenced are outlined by solid black lines. The promoter region (dashed outline) was assayed for presence of long or short allele.

Figure 2.

Linkage disequilibrium (LD) plot between different polymorphic sites in *M. mulatta* (N=20, or 40 haplotypes). Dark gray boxes designate significant LD ($p \leq 0.05$). Black outlines designate locations of polymorphic sites (see Fig. 1).

Figure 3.

Minimum spanning tree of haplotypes for all regions sequenced. Circles represent haplotypes, with size proportional to frequency of haplotype. The lengths of the lines connecting the circles are proportional (according to the scale provided) to the number of mutations that separate each haplotype. Numbers listed in the legend indicate total number of haplotypes (2N) for each species.

Figure 4.

Haplotype network for the 5' region of SLC6A (reduced median method). Circles represent haplotypes with size proportional to frequency of the haplotype. The lengths of the lines connecting the circles are proportional (according to the scale provided) to the number of mutations that separate each haplotype. Numbers listed in the legend indicate total number of haplotypes (2N) for each species. The two mutations that separate most Indian *M. mulatta* from the other species are indicated (**a**: 33 bp insertion; **b**: 12 bp deletion).

Table 1

Indices of within-species genetic diversity found in *SLC6A4*.

	<i>M. mulatta</i>		<i>M. fascicularis</i>		<i>M. fuscata</i>	<i>M. nemestrina</i>
	India	China	Total			
Haplotype number (2N)	24	16	40	14	18	10
Polymorphic sites	40	41	53	30	12	30
SNP	33	36	45	19	12	29
Indel	7	5	8	1	0	1
UTR	9	12	14	5	2	4
Introns	30	27	36	15	10	23
Exons	1	3	3	0	0	3
NS polymorphisms	1	1	1	0	0	1
Theta (S)	8.84	10.85	10.58	6.29	3.49	10.25
Theta (Pi)	8.71	11.93	11.05	7.67	3.54	11.82

Numbers based on all areas sequenced. The polymorphisms found within each species are separated into two types: SNPs and indels. In addition, the location of the polymorphisms is indicated (UTR, Exon, or Intron), as are the number of nonsynonymous (NS) polymorphisms.

Table 2

Genotypic and allele frequency for the polymorphic region of the promoter in two populations of *M. mulatta*.

	China N=37	India N=29
Long allele (L)	0.4189	0.6034
Short allele (S)	0.5811	0.3966
LL genotype	0.1892	0.3793
LS genotype	0.4595	0.4483
SS genotype	0.3514	0.1724

Table 3Results of five selection tests for *SLC6A4*.

Species (2N)	H	p-value	D	p-value	MK	p-value	ω	p-value	HKA	p-value
<i>M. mulatta</i> (40)	0.049	0.424	-0.459	0.386	0.288	0.591	0.442	0.280	11.471	0.322
Indian (24)	0.015	0.418	-0.557	0.341	0.079	0.779	0.513	0.295	11.216	0.341
Chinese (16)	2.183	0.720	-0.244	0.418	0.156	0.693	0.356	0.194	13.842	0.180
<i>M. fascicularis</i> (14)	0.545	0.548	0.746	0.814	NA	---	0.324	0.257	11.731	0.303
<i>M. fuscata</i> (18)	2.013	0.942	0.050	0.600	NA	---	0.162	0.113	8.733	0.558
<i>M. nemestrina</i> (10)	2.133	0.669	0.641	0.748	0.067	0.795	0.324	0.217	8.742	0.557

Tests are based on all areas sequenced. Parenthetical numbers refer to number of haplotypes used for analysis (2N). MK: McDonald-Kreitman test. The number for MK represents the G-value obtained by the trapezoidal method of numerical integration in DNAsp. The numbers for the HKA test represent the sum of deviations calculated in the HKA program. Where relevant, *M. sylvanus* was used as the outgroup.

Table 4

Results of Fay and Wu's H looking at only the 5' region.

Species (2N)	Fay and Wu's H	p-value
<i>M. mulatta</i> (42)	-1.882	0.030
Indian (26)	-1.711	0.008
Chinese (16)	-0.283	0.246
<i>M. fascicularis</i> (22)	0.294	0.600
<i>M. fuscata</i> (22)	0.416	0.821
<i>M. nemestrina</i> (12)	1.455	0.888

Parentnetical numbers refer to number of haplotypes used for analysis (2N; see also Table S2).