ORIGINAL ARTICLE

Heterospecific SNP diversity in humans and rhesus macaque (*Macaca mulatta*)

Jillian Ng[1], Jessica Satkoski Trask[1,2], David Glenn Smith[1,2] & Sree Kanthaswamy[1,2,3,4]

1 Molecular Anthropology Laboratory, Department of Anthropology, University of California, Davis, CA, USA
2 California National Primate Research Center, University of California, Davis, CA, USA
3 School of Mathematics and Natural Sciences, Arizona State University (ASU) at the West Campus, Glendale, AZ, USA
4 Department of Environmental Toxicology, University of California, Davis, CA, USA

Keywords
genomic comparison – *Homo sapiens* – non-human primates – orthologs

Correspondence
Sree Kanthaswamy, School of Mathematics and Natural Sciences, ASU at the West Campus, 4701 W Thunderbird Rd., Glendale, AZ 85306-4908, USA.
Tel.: (530) 219 2017;
e-mail: sree.kanthaswamy@asu.edu

Accepted April 9, 2015.

Abstract

Background Conservation of single nucleotide polymorphisms (SNPs) between human and other primates (i.e., heterospecific SNPs) in candidate genes can be used to assess the utility of those organisms as models for human biomedical research.
Methods A total of 59,691 heterospecific SNPs in 22 rhesus macaques and 20 humans were analyzed for human trait associations and 4207 heterospecific SNPs biallelic in both taxa were compared for genetic variation.
Results Variation comparisons at the 4207 SNPs showed that humans were more genetically diverse than rhesus macaques with observed and expected heterozygosities of 0.337 and 0.323 vs. 0.119 and 0.102, and minor allele frequencies of 0.239 and 0.063, respectively. In total, 431 of the 59,691 heterospecific SNPs are reportedly associated with human-specific traits.
Conclusion While comparisons between human and rhesus macaque genomes are plausible, functional studies of heterospecific SNPs are necessary to determine whether rhesus macaque alleles are associated with the same phenotypes as their corresponding human alleles.

Introduction

The rhesus macaque (*Macaca mulatta*) is the most common non-human primate model for human biomedical research. With the exception of the chimpanzee, cercopithecoid monkeys, such as the rhesus macaque, provide the most closely related model organism to humans, being phylogenetically separated for about 25 million years [15]. Genomewide comparisons between the two species have identified regions of high genomic orthology and synteny [13, 28, 42]. Both Gibbs et al. [28] and Yan et al. [42] identified regions of human sequence associated with known human traits that are apparently orthologous to regions of the rhesus macaque genome. The rhesus macaque model has been used to study infectious diseases [8] that affect humans, including HIV/AIDS [5] and tuberculosis [32], and complex traits and diseases such as asthma [6], diabetes [3], and aging [29]. As rhesus macaques are invaluable to human biomedical research, more information about orthology between the rhesus macaque and human genomes will improve the utility of this model in experimental studies.

The discovery of human single nucleotide polymorphisms (SNPs) was accomplished through large-scale SNP discovery projects such as the International HapMap (haplotype mapping) Consortium [37–39], which developed high-density SNP maps that include SNPs in many common human disease genes. The publication of the rhesus macaque genome fostered evolutionary and comparative genomic studies of rhesus macaque-specific SNPs [7, 18, 30, 34]. Despite the vast number of SNPs that were subsequently identified for both humans and

rhesus macaques, relatively few are shared between the two species, and, therefore, likely to have similar phenotypic effects.

In an attempt to identify SNPs shared between both humans and rhesus macaques, Kanthaswamy et al. [13] hybridized DNA from two rhesus macaques, one of Indian origin and one of Chinese origin, to the Affymetrix Human SNP 6.0 Array that was designed using 906,000 human SNPs, many of which were derived from the HapMap database [19]. Kanthaswamy et al. [13] reported that approximately 85,473 heterospecific SNPs (i.e., SNPs that are conserved in both humans and rhesus macaques) were shared between humans and rhesus macaques. Of those, over 65,000 SNPs were purported to be conserved in Chinese and Indian rhesus macaque gene orthologs [13]. Because many of the SNPs on the human array found to be heterospecific have been shown to be associated with human traits, the presence of these SNPs in rhesus macaques, with the same reference SNP ID (rs number), may provide essential biomarkers for testing hypotheses that these SNPs foster similar phenotypic effects in both species. To obtain a clearer insight into the species-specific variation and functional potential of heterospecific SNPs in each species, we used comparative genomic and population genetic approaches based on a larger sample set of rhesus macaques (N = 24) and a geographically heterogeneous sample of humans (N = 20). The objectives of this study were to (i) assess the genetic variation within and between rhesus macaques and humans using the heterospecific SNPs and (ii) determine the allelic states of rhesus macaque SNPs in orthologous genomic regions associated with human traits.

Materials and methods

The research reported in this manuscript adhered to the approved protocols of the UC Davis Institutional Animal Care and Use Committee (IACUC) and the legal requirements of the United States where the research took place. Twenty-four unrelated Indian rhesus macaques housed at the California National Primate Research Center (CNPRC) under conditions reported in Kanthaswamy et al. [12] were randomly selected for this study. Whole blood was drawn from each individual into EDTA Vacutainer tubes by the CNPRC veterinary staff following standard operating procedures. Individuals showing any adverse effects, such as stress or trauma, from the procedure were immediately treated according to the CNPRC standard operating procedures. Genomic DNA (1.5 μg) was extracted from the whole blood using the Qiagen QIAamp DNA blood mini kit (Valencia, CA, USA).

The DNA samples were hybridized to the NimbleGen AccuSNP (Madison, WI) custom array containing the 85,473 heterospecific SNPs that Kanthaswamy et al. [13] identified from the Affymetrix Genome-Wide Human SNP 6.0 array (Santa Clara, CA, USA), processed, and subjected to general QC at NimbleGen's service laboratory. Of the 24 rhesus macaque samples, 22 were successfully genotyped, but only 20 were genotyped with at least 90% completeness. Genotypes generated for SNPs on the Affymetrix Genome-Wide Human SNP 6.0 array in 20 random human samples, comprised of five unrelated individuals each from the four major HapMap ethnic populations (including CEPH (Utah residents with ancestry from northern and western Europe), Yoruba in Ibadan, Nigeria, Han Chinese in Beijing, and Japanese in Tokyo) were downloaded from the HapMap data repository [38] and compared to the rhesus macaque genotype data. Of the total 85,473 heterospecific SNPs that were interrogated, only 79,404 were successfully genotyped for both data sets. On average, SNPs and individuals were genotyped with 88.84% and 99.32% completeness for rhesus macaque and human samples, respectively.

A total of 59,691 autosomal SNPs were genotyped with at least 90% completeness and queried against the National Center for Biotechnology Information's dbSNP [33] (http://www.ncbi.nlm.nih.gov/projects/SNP/) and Online Mendelian Inheritance in Man [9] (OMIM; http://www.ncbi.nlm.nih.gov/omim) databases as well as the National Human Genome Research Institute's (NHGRI) Catalog of Published Genome-Wide Association Studies [11] (http://www.genome.gov/gwastudies/) to identify possible human disease and trait associations. The results from these databases were compared to identify the total number of SNPs associated with human disease/phenotypes and those that were cross-listed in the three databases. The human ancestral alleles as listed in the databases and their associated human phenotypes were identified for SNPs from the OMIM database results and compared to the rhesus macaque alleles to identify derived alleles.

Observed (OH) and expected (EH) heterozygosities and minor allele frequencies (MAF) for the human and rhesus macaque samples were calculated using only the 4207 biallelic (two alleles per position that are the same in both species) heterospecific SNPs with PLINK v.1.07 [27] (http://pngu.mgh.harvard.edu/purcell/plink/) following the methods of Trask et al. [40]. Statistical significance of differences in estimates of OH, EH, and MAF between the two species were evaluated using a one-tailed Mann–Whitney *U*-test with Bonferroni correction.

Results

Of the 59,691 heterospecific SNPs that were analyzed in this study, only 4207 were biallelic in the samples of both species. The rhesus macaque and human samples exhibited average values of observed heterozygosity of 0.119 and 0.337, average values of expected heterozygosity of 0.102 and 0.323 (Fig. 1), and average MAF values of 0.063 and 0.239 (Fig. 2), respectively. The Mann–Whitney *U*-tests revealed statistically significant differences in OH, EH, and MAF between humans and rhesus macaques ($P << 0.01$ for all three comparisons).

Four hundred and thirty-one of the 59,691 heterospecific SNPs analyzed here were found to have some clinical significance (benign or pathogenic) and associated with human traits and/or diseases. The results of the database query are listed in the Supplementary document and summarized in Table 1. Sixty-one SNPs were linked to verified genetic traits or disorders in the OMIM database, 14 of which involve functional mutations in specific traits. Eighty-six SNPs were not linked to an OMIM entry but contained clinical significance data in the dbSNP database, and the remaining 284 SNPs were determined to be associated with traits and diseases through NHGRI's curation of published whole-genome association studies but have not been verified by OMIM. Rhesus macaques were monomorphic for 400 of the 431 SNPs and biallelic for only three of the 61 OMIM identified SNPs. While almost all of the 400 'fixed' SNPs in the rhesus macaque samples represented the ancestral allele in humans, five SNPs were monomorphic for the human-derived alleles that have been shown to be associated with several phenotypes in humans including stature, obesity, polycythemia vera, thrombocythemia, myelofibrosis, as well as hypercarotenemia and vitamin A deficiency.

Discussion

The use of heterospecific SNPs permitted a direct genomic and population genetic comparison between rhesus macaques and humans, but this approach may have introduced an ascertainment bias that favored greater genetic diversity in humans. Focusing on heterospecific SNPs in rhesus macaques that are associated with known human traits provides relevant information for assessing the suitability of this species as subjects in certain types of biomedical studies. Among the 59,691 human SNPs reported to be heterospecific in rhesus macaques by Kanthaswamy et al. [13] and successfully genotyped in all twenty humans and rhesus macaques, humans exhibited fixed alleles (or monomorphism) at only 20% of the markers (N = 11,777), while 91% (N = 54,393) of the SNPs in rhesus macaques were monomorphic and 18% of the SNPs (N = 10,686) were monomorphic in both species. Seventy-three percent (N = 43,707) of the SNPs in humans, 2% (N = 1091) of those in rhesus macaques, and 7% (N = 4207) of those in both species were biallelic. Among the 4207 SNPs biallelic in humans and rhesus macaques (exhibiting

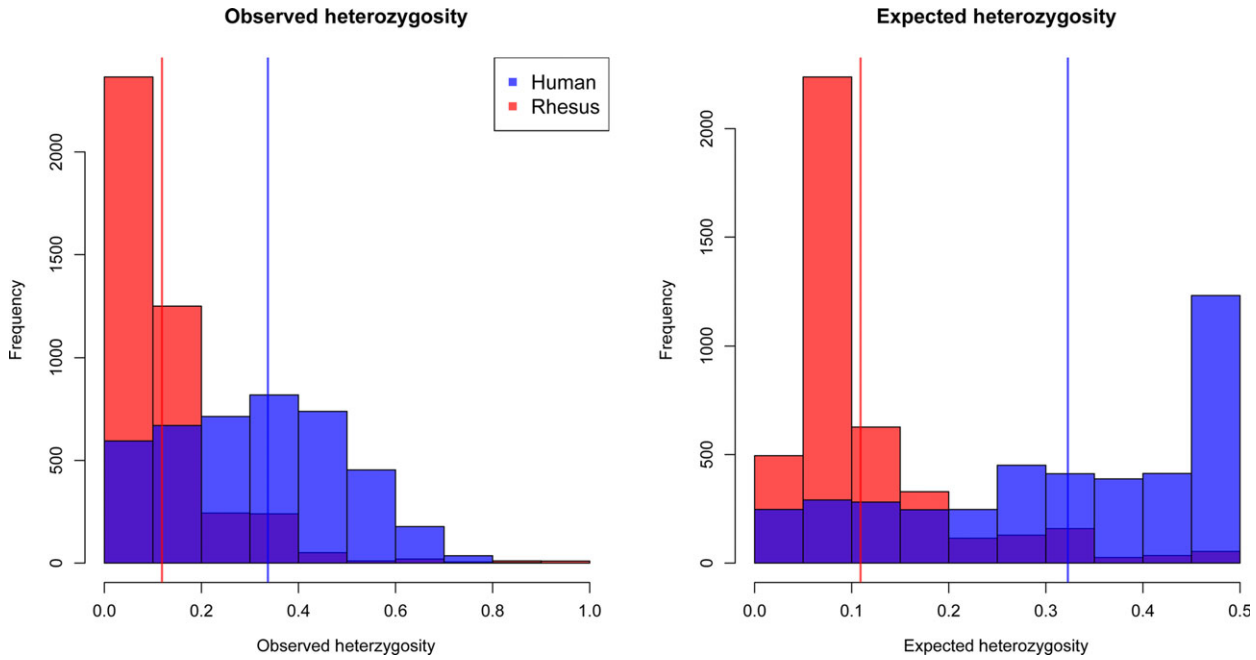

Fig. 1 Observed and expected heterozygosity distribution for rhesus macaque and human with significant mean differences ($P < 0.01$).

J Med Primatol **44** (2015) 194–201

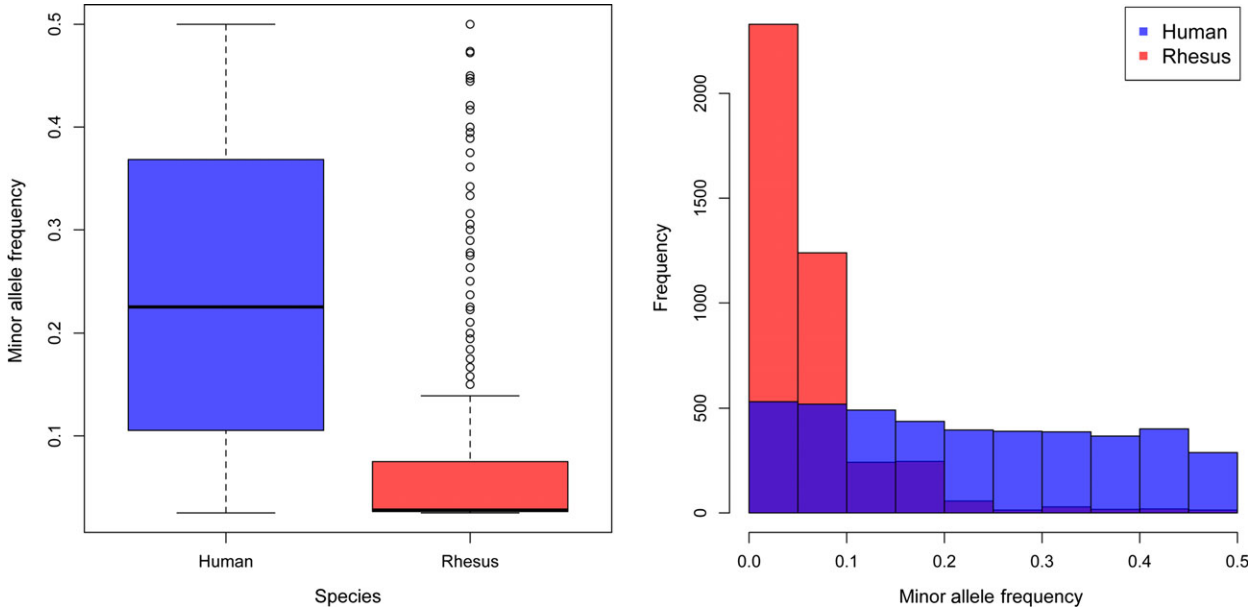

Fig. 2 Minor allele frequency distribution for rhesus macaque and human with significant mean MAF value differences ($P < 0.01$).

Table 1 The resulting number of total hits and unique SNPs from querying against the NCBI OMIM and dbSNP databases and the NHGRI Catalog of Published Genome-Wide Association Studies database

Query Results	Unique SNPs	Total Hits
OMIM	61	73
dbSNP	30,234	36,693
With clinical significance	99	585
NHGRI	303	387
Comparisons	**Unique SNPs**	**Shared SNPs**
OMIM	42	
dbSNP with clinical significance		13
NHGRI		19
NHGRI	284	
dbSNP with clinical significance		1
OMIM		19
dbSNP with clinical significance	86	
NHGRI		1
OMIM		13
Across all three databases	430	1

average pairwise distances of 615.90 and 632.45 kb, respectively), humans exhibited far greater genetic diversity than rhesus macaques. The interrogation of the 4207 biallelic SNP positions revealed significant interspecies differences and highlights the potential effects on biomedical research when rhesus macaques are chosen as models for studying human traits.

Strong positive selection on derived SNP alleles in humans has resulted in an increase of phenotypic variation. A higher average genomewide mutational bias in humans may also have contributed to the higher level of genetic diversity in humans in the present study. Nguyen et al. [23] compared the genes within known human and mouse CNVs and showed that human CNVs were often associated with genes that have relatively elevated ratios of non-synonymous to synonymous substitution rates. This may be seen as evidence for positive selection on CNVs during the evolution of humans, a relaxation of selective pressure or the exertion of purifying selection against CNVs in other gene types and families. One other possible example of selection in humans is the selection for risk variants for diabetes mellitus hypothesized to have occurred during early human evolution as a beneficial adaptive response to store more energy with which to endure environment changes [4, 21, 31]. As humans evolved from early hunter-gatherer to more agrarian lifestyles, food became more readily available, but the once beneficial trait, the 'thrifty genotype', of storing had not co-evolved with it, resulting in hyperglycemic disorders today.

The results of this study are consistent with the identification by the Rhesus Macaque Genome Sequencing and Analysis Consortium [28] of 229 wild-type ancestral variants, such as the 'thrifty genotype' still retained in chimpanzees and rhesus macaques that would cause inherent disease in humans. Besides divergent mutation rates, population bottlenecks and expansions in the recent human and rhesus macaque histories may have distorted measures of human diversity, influencing

interspecific differences. The significant difference in genomewide variation in both these species is not consistent with the findings of Yuan et al. [43] who reported that rhesus macaques exhibit three times more SNP diversity than humans. This discrepancy is undoubtedly due to the ascertainment bias of identifying heterospecific SNPs among known human SNPs as opposed to selecting SNPs for analysis in both species simultaneously as in Yuan et al.'s [43] study. Interestingly, despite reporting greater diversity in rhesus macaques across all SNPs, Yuan et al. [43] observed that the number of deleterious SNPs (i.e., those that exhibit damaging effects on protein functions) are similar in both humans and rhesus macaques.

While many of the human SNPs in this study are more variable than their orthologs in rhesus macaques, suggesting that more mutations have occurred at those SNPs in the lineage leading to humans, the species-specific difference observed in this study probably stems from the selection of highly variable human SNPs as genetic markers for the target probe. The reliance on SNPs selected from an array containing probes designed to interrogate human diversity (i.e., SNPs first discovered and subsequently developed for maximizing coverage and variation in humans) necessarily diminishes the level of variation discovered using these probes for genotyping rhesus macaques [24, 25]. A higher level of genetic diversity in rhesus macaques than in humans would have undoubtedly been discovered had SNPs originally identified in rhesus macaques been screened to identify heterologous SNPs in humans.

The interspecies allelic differences revealed here using heterospecific SNPs in humans and rhesus macaques should be determined *a priori* to assess the suitability of rhesus macaques as models for specific human traits. The ancestral but minor human allele A (MAF = 0.19 from dbSNP) at SNP rs3807218 causes the overexpression of the gene *DPP6*, which Alders et al. [1] proposed as the pathogenic mechanism for familial idiopathic ventricular fibrillation. While the ancestral A allele is present in both humans and rhesus macaques, its frequency has changed drastically between the two species such that the allele is completely fixed in rhesus macaques but relatively rare in humans. However, the derived human allele confers susceptibility to disease for the majority of the heterospecific SNPs associated with human disease leading to fixation in rhesus macaques of ancestral alleles associated with the absence of human-specific phenotypes. For example, SNP rs5917 controls the formation of Pen human platelet alloantigens, the G allele for Pen^{a} and A allele for Pen^{b}, with Pen^{b} the suspected cause of the immunopathologic conditions post-transfusion purpura and neonatal alloimmune thrombocytopenic purpura [41]. Rhesus macaques are fixed for the G allele, which would result in the species having only Pen^{a} alloantigens if the phenotypes are conserved.

As conservation of SNPs does not necessarily imply identity by descent, subsequent DNA and protein analyses of these trait-associated SNPs should be performed in rhesus macaques to determine whether the presence or absence of the alleles at those human SNPs actually yield the same phenotypic or functional effects. Functional studies that interrogate the effects of specific alleles at SNPs associated with various human neuropsychiatric behavioral conditions have been previously conducted in mice [2, 14, 17]. Humanized mice have also been created by inserting human genes and even entire chromosomes into the mouse genome [22]. Although the use of rhesus macaques in these types of studies is currently limited by costs and difficulties involving the construction of transgenic rhesus macaque lines, they may be unnecessary if the same phenotypes occur naturally in rhesus macaques due to the common descent of alleles or have arisen through mutations in the gene orthologs [20, 26].

Identifying pathogenic or disease-associated human SNPs in model organisms can also be used as the starting points to investigate mechanisms that result in disease. By examining orthologous regions in mice where pathogenic human SNPs are located, Lee at al. [16] and Sur et al. [35] identified regulatory regions involved in inflammation and tumorigenesis, respectively. Alternatively, SNPs affecting human disease, especially extremely rare ones, may first be identified in model organisms and then mapped in humans. For example, a SNP identified in dogs that is associated with obsessive–compulsive disorder was subsequently shown to result in differential gene expression and reduced DNA protein binding in human cells [36].

More comprehensive studies of both human-specific and rhesus macaque-specific SNPs associated with specific phenotypes can improve our understanding of human diseases and better evaluate the use of rhesus macaques as model organisms. Similar studies involving Chinese rhesus macaques ought to be carried out because Indian rhesus macaques are more highly derived than Chinese rhesus macaques and probably experienced a genetic bottleneck during which many low-frequency alleles still present in Chinese rhesus macaques, and phenotypes associated with them, were lost [10]. Chinese, but not Indian, rhesus macaques provide a useful animal model for the study of phenotypes associated with such SNPs identified as heterologous in humans. Other heterologous SNPs found in Indian, but not Chinese, rhesus macaques will have been derived since the split of Indian rhesus macaques from Chinese rhesus macaques, making Indian, but not Chinese,

J Med Primatol **44** (2015) 194–201

rhesus macaques a suitable animal model for the study of associated human phenotypes.

Acknowledgments

This study was supported by the California National Primate Research Center (CNPRC) base grant (No. RR000169-48), ARRA supplement grant RR018144-07 to S.K. and Nick Lerche (CNPRC), and NIH grants RR005090 and RR025871 to D.G.S. The authors like to thank the staff of the Molecular Anthropology Laboratory at University of California – Davis for their assistance with sample preparation.

References

1 Alders M, Koopmann TT, Christiaans I, Postema PG, Beekman L, Tanck MW, Zeppenfeld K, Loh P, Koch KT, Demolombe S, Mannens MM, Bezzina CR, Wilde AA: Haplotype-sharing analysis implicates chromosome 7q36 harboring DPP6 in familial idiopathic ventricular fibrillation. *Am J Hum Genet* 2009; **84**:468–76.

2 Blendy JA: Modeling neuropsychiatric disease-relevant human SNPs in mice. *Neuropsychopharmacology* 2011; **36**:364–5.

3 Bremer AA, Stanhope KL, Graham JL, Cummings BP, Wang W, Saville BR, Havel PJ: Fructose-fed rhesus monkeys: a nonhuman primate model of insulin resistance, metabolic syndrome, and type 2 diabetes. *Clin Transl Sci* 2011; **4**:243–52.

4 Brown EA: Genetic explorations of recent human metabolic adaptations: hypotheses and evidence. *Biol Rev Camb Philos Soc* 2012; **87**:838–55.

5 Carlsson HE, Schapiro SJ, Farah I, Hau J: Use of primates in research: a global overview. *Am J Primatol* 2004; **63**:225–37.

6 Coffman RL, Hessel EM: Nonhuman primate models of asthma. *J Exp Med* 2005; **201**:1875–9.

7 Ferguson B, Street SL, Wright H, Pearson C, Jia Y, Thompson SL, Allibone P, Dubay CJ, Spindel E, Norgren RB Jr: Single nucleotide polymorphisms (SNPs) distinguish Indian-origin and Chinese-origin rhesus macaques (*Macaca mulatta*). *BMC Genom* 2007; **8**:43.

8 Gardner MB, Luciw PA: Macaque models of human infectious disease. *ILAR J* 2008; **49**:220–55.

9 Hamosh A, Scott AF, Amberger J, Bocchini C, Valle D, McKusick VA: Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic Acids Res* 2002; **30**:52–5.

10 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**:240–3.

11 Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA: Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 2009; **106**:9362–7.

12 Kanthaswamy S, Kou A, Smith DG: Population genetic statistics from Rhesus Macaques (*Macaca mulatta*) in three different housing configurations at the California national primate research center. *J Am Assoc Lab Anim Sci* 2010; **49**:598–609.

13 Kanthaswamy S, Ng J, Ross CT, Trask JS, Smith DG, Buffalo VS, Fass JN, Lin D: Identifying human-rhesus macaque gene orthologs using heterospecific SNP probes. *Genomics* 2013; **101**:30–7.

14 Kas MJ, Gelegen C, Schalkwyk LC, Collier DA: Interspecies comparisons of functional genetic variations and their implications in neuropsychiatry. *Am J Med Genet B Neuropsychiatr Genet* 2009; **150B**:309–17.

15 Kumar S, Hedges SB: A molecular timescale for vertebrate evolution. *Nature* 1998; **392**:917–20.

16 Lee JC, Espeli M, Anderson CA, Linterman MA, Pocock JM, Williams NJ, Roberts R, Viatte S, Fu B, Peshu N, Hien TT, Phu NH, Wesley E, Edwards C, Ahmad T, Mansfield JC, Gearry R, Dunstan S, Williams TN, Barton A, Vinuesa CG, Consortium UIG, Parkes M, Lyons PA, Smith KG: Human SNP links differential outcomes in inflammatory and infectious disease to a FOXO3-regulated pathway. *Cell* 2013; **155**:57–69.

17 Mague SD, Isiegas C, Huang P, Liu-Chen LY, Lerman C, Blendy JA: Mouse model of OPRM1 (A118G) polymorphism has sex-specific effects on drug-mediated behavior. *Proc Natl Acad Sci USA* 2009; **106**:10847–52.

18 Malhi RS, Sickler B, Lin D, Satkoski J, Tito RY, George D, Kanthaswamy S, Smith DG: MamuSNP: a resource for Rhesus Macaque (*Macaca mulatta*) genomics. *PLoS ONE* 2007; **2**:e438.

19 McCarroll SA, Kuruvilla FG, Korn JM, Cawley S, Nemesh J, Wysoker A, Shapero MH, de Bakker PIW, Maller JB, Kirby A, Elliott AL, Parkin M, Hubbell E, Webster T, Mei R, Veitch J, Collins PJ, Handsaker R, Lincoln S, Nizzari M, Blume J, Jones KW, Rava R, Daly MJ, Gabriel SB, Altshuler D: Inte-

grated detection and population-genetic analysis of SNPs and copy number variation. *Nat Genet* 2008; **40**:1166–74.
20 Miller GM, Bendor J, Tiefenbacher S, Yang H, Novak MA, Madras BK: A mu-opioid receptor single nucleotide polymorphism in rhesus monkey: association with stress response and aggression. *Mol Psychiatry* 2004; **9**:99–108.
21 Neel JV: Diabetes mellitus: a "thrifty" genotype rendered detrimental by "progress"? *Am J Hum Genet* 1962; **14**:353–62.
22 Nguyen D, Xu T: The expanding role of mouse genetics for understanding human biology and disease. *Dis Model Mech* 2008; **1**:56–66.
23 Nguyen D-Q, Webber C, Ponting CP: Bias of selection on human copy-number variants. *PLoS Genet* 2006; **2**:e20.
24 Nielsen R: Population genetic analysis of ascertained SNP data. *Hum Genomics* 2004; **1**:218–24.
25 Nielsen R, Signorovitch J: Correcting for ascertainment biases when analyzing SNP data: applications to the estimation of linkage disequilibrium. *Theor Popul Biol* 2003; **63**:245–55.
26 Premasuthan A, Kanthaswamy S, Satkoski J, Smith DG: A simple multiplex polymerase chain reaction to determine ABO blood types of rhesus macaques (*Macaca mulatta*). *Tissue Antigens* 2011; **77**:584–8.
27 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC: PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am J Hum Genet* 2007; **81**:559–75.
28 Rhesus Macaque Genome Sequencing Analysis Consortium, Gibbs RA, Rogers J, Katze MG, Bumgarner R, Weinstock GM, Mardis ER, Remington KA, Strausberg RL, Venter JC, Wilson RK, Batzer MA, Bustamante CD, Eichler EE, Hahn MW, Hardison RC, Makova KD, Miller W, Milosavljevic A, Palermo RE, Siepel A, Sikela JM, Attaway T, Bell S, Bernard KE, Buhay CJ, Chandrabose MN, Dao M, Davis C, Delehaunty KD, Ding Y, Dinh HH, Dugan-Rocha S, Fulton LA, Gabisi RA, Garner TT, Godfrey J, Hawes AC, Hernandez J, Hines S, Holder M, Hume J, Jhangiani SN, Joshi V, Khan ZM, Kirkness EF, Cree A, Fowler RG, Lee S, Lewis LR, Li Z, Liu YS, Moore SM, Muzny D, Nazareth LV, Ngo DN, Okwuonu GO, Pai G, Parker D, Paul HA, Pfannkoch C, Pohl CS, Rogers YH, Ruiz SJ, Sabo A, Santibanez J, Schneider BW, Smith SM, Sodergren E, Svatek AF, Utterback TR, Vattathil S, Warren W, White CS, Chinwalla AT, Feng Y, Halpern AL, Hillier LW, Huang X, Minx P, Nelson JO, Pepin KH, Qin X, Sutton GG, Venter E, Walenz BP, Wallis JW, Worley KC, Yang SP, Jones SM, Marra MA, Rocchi M, Schein JE, Baertsch R, Clarke L, Csuros M, Glasscock J, Harris RA, Havlak P, Jackson AR, Jiang H, Liu Y, Messina DN, Shen Y, Song HX, Wylie T, Zhang L, Birney E, Han K, Konkel MK, Lee J, Smit AF, Ullmer B, Wang H, Xing J, Burhans R, Cheng Z, Karro JE, Ma J, Raney B, She X, Cox MJ, Demuth JP, Dumas LJ, Han SG, Hopkins J, Karimpour-Fard A, Kim YH, Pollack JR, Vinar T, Addo-Quaye C, Degenhardt J, Denby A, Hubisz MJ, Indap A, Kosiol C, Lahn BT, Lawson HA, Marklein A, Nielsen R, Vallender EJ, Clark AG, Ferguson B, Hernandez RD, Hirani K, Kehrer-Sawatzki H, Kolb J, Patil S, Pu LL, Ren Y, Smith DG, Wheeler DA, Schenck I, Ball EV, Chen R, Cooper DN, Giardine B, Hsu F, Kent WJ, Lesk A, Nelson DL, O'Brien WE, Prufer K, Stenson PD, Wallace JC, Ke H, Liu XM, Wang P, Xiang AP, Yang F, Barber GP, Haussler D, Karolchik D, Kern AD, Kuhn RM, Smith KE, Zwieg AS: Evolutionary and biomedical insights from the rhesus macaque genome. *Science* 2007; **316**:222–34.
29 Roth GS, Mattison JA, Ottinger MA, Chachich ME, Lane MA, Ingram DK: Aging in Rhesus Monkeys: relevance to human health interventions. *Science* 2004; **305**:1423–6.
30 Satkoski J, Malhi R, Kanthaswamy S, Tito R, Malladi V, Smith D: Pyrosequencing as a method for SNP identification in the rhesus macaque (*Macaca mulatta*). *BMC Genom* 2008; **9**:256.
31 Segurel L, Austerlitz F, Toupance B, Gautier M, Kelley JL, Pasquet P, Lonjou C, Georges M, Voisin S, Cruaud C, Couloux A, Hegay T, Aldashev A, Vitalis R, Heyer E: Positive selection of protective variants for type 2 diabetes from the Neolithic onward: a case study in Central Asia. *Eur J Hum Genet* 2013; **21**:1146–51.
32 Sharpe SA, McShane H, Dennis MJ, Basaraba RJ, Gleeson F, Hall G, McIntyre A, Gooch K, Clark S, Beveridge NER, Nuth E, White A, Marriott A, Dowall S, Hill AVS, Williams A, Marsh PD: Establishment of an aerosol challenge model of tuberculosis in rhesus macaques and an evaluation of endpoints for vaccine testing. *Clin Vaccine Immunol* 2010; **17**:1170–82.
33 Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K: dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res* 2001; **29**:308–11.
34 Street S, Kyes R, Grant R, Ferguson B: Single nucleotide polymorphisms (SNPs) are highly conserved in rhesus (*Macaca mulatta*) and cynomolgus (*Macaca fascicularis*) macaques. *BMC Genom* 2007; **8**:480.
35 Sur IK, Hallikas O, Vähärautio A, Yan J, Turunen M, Enge M, Taipale M, Karhu A, Aaltonen LA, Taipale J: Mice lacking a Myc enhancer that includes human SNP rs6983267 Are resistant to intestinal tumors. *Science* 2012; **338**:1360–3.

36 Tang R, Noh H, Wang D, Sigurdsson S, Swofford R, Perloski M, Duxbury M, Patterson E, Albright J, Castelhano M, Auton A, Boyko A, Feng G, Lindblad-Toh K, Karlsson E: Candidate genes and functional noncoding variants identified in a canine model of obsessive-compulsive disorder. *Genome Biol* 2014; **15**:R25.
37 The International HapMap C: A haplotype map of the human genome. *Nature* 2005; **437**:1299–320.
38 The International HapMap C: The International HapMap Project. *Nature* 2003; **426**:789–96.
39 The International HapMap C: A second generation human haplotype map of over 3.1 million SNPs. *Nature* 2007; **449**:851–61.
40 Trask JS, Garnica WT, Kanthaswamy S, Malhi RS, Smith DG: 4040 SNPs for genomic analysis in the rhesus macaque (*Macaca mulatta*). *Genomics* 2011; **98**:352–8.
41 Wang R, Furihata K, McFarland JG, Friedman K, Aster RH, Newman PJ: An amino acid polymorphism within the RGD binding domain of platelet membrane glycoprotein IIIa is responsible for the formation of the Pena/Penb alloantigen system. *J Clin Investig* 1992; **90**:2038–43.
42 Yan G, Zhang G, Fang X, Zhang Y, Li C, Ling F, Cooper DN, Li Q, Li Y, van Gool AJ, Du H, Chen J, Chen R, Zhang P, Huang Z, Thompson JR, Meng Y, Bai Y, Wang J, Zhuo M, Wang T, Huang Y, Wei L, Li J, Wang Z, Hu H, Yang P, Le L, Stenson PD, Li B, Liu X, Ball EV, An N, Huang Q, Zhang Y, Fan W, Zhang X, Li Y, Wang W, Katze MG, Su B, Nielsen R, Yang H, Wang J, Wang X, Wang J: Genome sequencing and comparison of two nonhuman primate animal models, the cynomolgus and Chinese rhesus macaques. *Nat Biotech* 2011; **29**:1019–23.
43 Yuan Q, Zhou Z, Lindell S, Higley J, Ferguson B, Thompson R, Lopez J, Suomi S, Baghal B, Baker M, Mash D, Barr C, Goldman D: The rhesus macaque is three times as diverse but more closely equivalent in damaging coding variation as compared to the human. *BMC Genet* 2012; **13**:52.