Acceleration of Olfactory Receptor Gene Loss in Primate Evolution: Possible Link to Anatomical Change in Sensory Systems and Dietary Transition

Yoshihito Niimura,*[,1,2,3] Atsushi Matsui,[1,2] and Kazushige Touhara*[,1,2]

[1]Department of Applied Biological Chemistry, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan

[2]ERATO Touhara Chemosensory Signal Project, JST, The University of Tokyo, Tokyo, Japan

[3]Lead Contact

*Corresponding authors: E-mails: aniimura@mail.ecc.u-tokyo.ac.jp; ktouhara@mail.ecc.u-tokyo.ac.jp.

Associate editor: Yoko Satta

Abstract

Primates have traditionally been regarded as vision-oriented animals with low olfactory ability, though this "microsmatic primates" view has been challenged recently. To clarify when and how degeneration of the olfactory system occurred and to specify the relevant factors during primate evolution, we here examined the olfactory receptor (OR) genes from 24 phylogenetically and ecologically diverse primate species. The results revealed that strepsirrhines with curved noses had functional OR gene repertoires that were nearly twice as large as those for haplorhines with simple noses. Neither activity pattern (nocturnal/diurnal) nor color vision system showed significant correlation with the number of functional OR genes while phylogeny and nose structure (haplorhine/strepsirrhine) are statistically controlled, but extent of folivory did. We traced the evolutionary fates of individual OR genes by identifying orthologous gene groups, demonstrating that the rates of OR gene losses were accelerated at the ancestral branch of haplorhines, which coincided with the acquisition of acute vision. The highest rate of OR gene loss was observed at the ancestral branch of leaf-eating colobines; this reduction is possibly linked with the dietary transition from frugivory to folivory because odor information is essential for fruit foraging but less so for leaf foraging. Intriguingly, we found accelerations of OR gene losses in an external branch to every hominoid species examined. These findings suggest that the current OR gene repertoire in each species has been shaped by a complex interplay of phylogeny, anatomy, and habitat; therefore, multiple factors may contribute to the olfactory degeneration in primates.

Key words: **primate evolution, olfactory receptor, gene loss, multigene family, sensory ecology, chemical sense.**

Introduction

Olfaction, or the sense of smell, is essential for the survival of most if not all mammals. It is used for foraging, communicating with conspecifics, and recognizing predators (Wyatt 2014). Diverse odor molecules in the environment are detected by olfactory receptors (ORs) expressed in the olfactory epithelium of the nasal cavity (Buck and Axel 1991; Nei et al. 2008; Touhara and Vosshall 2009; Niimura 2012). OR genes form the largest multigene family in mammals, accounting for ~5% of the entire genes. Many mammalian species examined to date, such as mice, rats, dogs, cows, and horses, have 800–1,200 functional OR genes (Niimura et al. 2014). Interestingly, African elephants have a surprisingly large number of functional OR genes, being ~2,000 (Niimura et al. 2014). On the other hand, the number of functional OR genes in humans, chimpanzees, orangutans, rhesus macaques, and marmosets were reported to be 300–400 (Matsui et al. 2010), which is considerably smaller than those for most nonprimate mammals. Traditionally, primates have been regarded as vision-oriented animals with low olfactory ability, and the degenerated OR gene repertoires in primates are consistent with this traditional view. However, this "microsmatic primates" view has been challenged recently; numerous studies showed reliance on olfaction and excellent olfactory ability in humans and nonhuman primates (Laska et al. 2000; Shepherd 2004; Porter et al. 2007; Drea 2015; Wackermannova et al. 2016; McGann 2017).

The order Primates is classified into two suborders based on nostril shape: one is strepsirrhines, meaning "curved nose," and the other is haplorhines, meaning "simple nose." The former consists of lorisiformes (lorises and galagos) and lemuriforms (lemurs), while the latter consists of tarsiiforms (tarsiers), New World monkeys (NWMs), Old World monkeys (OWMs), and hominoids (humans and apes). The phylogenetic position of tarsiers was controversial for a long time, because tarsiers share morphological features to both strepsirrhines and anthropoids (including NWMs, OWMs, and hominoids) (Fleagle 2013). Traditionally, a closer relationship of tarsiers to strepsirrhines, that is, the prosimian monophyly, was hypothesized (Murphy et al. 2001; Chatterjee et al. 2009), but recent molecular studies unequivocally resolved a close relationship to anthropoids, that is, the haplorhine

Article

monophyly (Schmitz et al. 2001; Matsui et al. 2009; Perelman et al. 2011; Springer et al. 2012). One of the most striking differences between strepsirrhines and haplorhines is the presence and absence of the rhinarium, respectively (Martin 1990; Smith et al. 2007). The rhinarium is a moist and hairless skin surface around the frontal part of the nose, which is also present in many mammals such as cats and dogs, but the function of the rhinarium is unknown in most cases (Glaser and Kroger 2017). Among strepsirrhines, all lorises and galagos are nocturnal, while lemurs have diverse activity patterns including diurnality, nocturnality, and various types of cathemerality (i.e., being active both during the day and at night) (Donati et al. 2013; Valenta et al. 2016). On the other hand, haplorhines are predominantly diurnal with two exceptional taxa having nocturnality: the tarsiers, and the owl monkey *Aotus*, a genus of NWMs.

Primates are exceptional among placental mammals in having trichromatic color vision. Reconstruction of ancestral placental mammals showed them to have been dichromatic based on having the two cone opsin genes (*OPN1SW* and *OPN1LW*) that encode the short- (SWS1) and the middle-to-long-wavelength sensitive (M/LWS) photopigments (Jacobs 2009). Among primates, only catarrhines (OWMs and hominoids) and the howler monkey *Alouatta*, a genus of NWMs, have routine trichromacy, which was attained by gene duplication of the M/LWS opsin gene. In many NWMs and some strepsirrhine species, the X-linked M/LWS opsin gene is polymorphic; therefore, heterozygous females are trichromatic, whereas homozygous females and all males are dichromatic. Similar allelic variation was reported for several diurnal or cathemeral strepsirrhine species (Veilleux and Bolnick 2009; Jacobs et al. 2017). The adaptive nature for the evolutionary persistence of these polymorphisms is controversial (Kawamura et al. 2012; Veilleux et al. 2016). On the other hand, some nocturnal strepsirrhine species and the owl monkey *Aotus*, a genus of NWMs, are monochromatic, having only the M/LWS opsin gene due to the loss of a functional SWS1 opsin gene (Kawamura and Kubotera 2004; Tan et al. 2005; Veilleux et al. 2013).

To determine a possible link between routine trichromacy and olfactory degeneration, Gilad et al. (2004) examined the sequences of 100 OR genes randomly chosen from each of 19 primate species. They reported that fractions of OR pseudogenes are significantly higher in catarrhines and howler monkey than in the other NWMs, strepsirrhines, and mice, concluding that the degeneration of OR genes had occurred concomitantly with the acquisition of routine trichromacy. However, our previous analyses using the whole genome sequences of five anthropoid species indicated no significant differences in the number of intact OR genes between catarrhines and NWMs (Matsui et al. 2010). Moreover, the analyses did not show an acceleration of OR gene losses in the ancestral branch of catarrhines where the acquisition of routine trichromacy occurred (Matsui et al. 2010). Therefore, the drastic reduction of OR genes in higher primates cannot be explained by the acquisition of routine trichromatic vision.

To date, entire OR gene repertoires have been reported for only five anthropoid primates (Matsui et al. 2010). Accurate numbers of OR genes in strepsirrhine species have not yet been reported. Dong et al. (2010) reported that mouse lemurs have fewer than 400 functional OR genes; however, this number is inaccurate due to the use of the low quality genome sequences available at that time (Matsui et al. 2010). To trace the degeneration of olfaction during primate evolution, we need to investigate phylogenetically and ecologically diverse species. The purpose of this study is to examine the timing and extent of OR gene losses during primate evolution and to specify relevant factors in OR gene degeneration. For this purpose, we analyzed the whole-genome sequences of 24 diverse primate species in terms of phylogenetic distribution, activity pattern, color vision system, and feeding habitat (extent of folivory and frugivory). Notably, the 24 species include nocturnal strepsirrhine, diurnal strepsirrhine, nocturnal haplorhine, and diurnal haplorhine species, and they included monochromatic, dichromatic, polymorphic, and trichromatic species. We also used two nonprimate species for comparison: the colugo in Dermoptera and the treeshrew in Scandentia. Primates, Dermoptera, and Scandentia form the superordinal group Euarchonta. The phylogenetic relationship among the three orders has long been debated, but a recent genomic study overwhelmingly supported a sister-group relationship between primates and colugos, to the exclusion of treeshrews (Mason et al. 2016).

As a result, we found that acceleration of OR gene losses has occurred in the ancestral branch of haplorhines and that of colobines, the folivorous (leaf-eating) OWMs. The former appears to be linked to the acquisition of acute vision, while the latter would be associated with changes in feeding habitats. In addition, unexpectedly, we found independent acceleration of OR gene losses in each hominoid lineage leading to the extant species.

Results

Numbers of OR Genes in 24 Primate and 2 Nonprimate Species

We identified OR genes from 24 primates and two nonprimate species for which the whole-genome sequences are available, and we classified them into three categories according to Niimura and Nei (2007) as shown in figure 1: intact genes, truncated genes, and pseudogenes (supplementary tables S1 and S2, Supplementary Material online). An intact gene is putatively functional, while a truncated gene may become intact if the quality of the genome sequence is improved. The fraction of truncated genes (T%) is generally low (<10%); therefore, the number of intact genes can be regarded as a good estimate of the number of functional OR genes for most species examined. However, for some species, for example, hamadryas baboon or colugo, the qualities of genome assemblies are low, and the T% values are high. The T% values showed a strong positive correlation with the reciprocals of N50 values of genome assemblies (supplementary fig. S1, Supplementary Material online). This observation is presumably due to the fact that the more incomplete or fragmented the genome is, the more truncated genes will be found. Therefore, we consider the following two

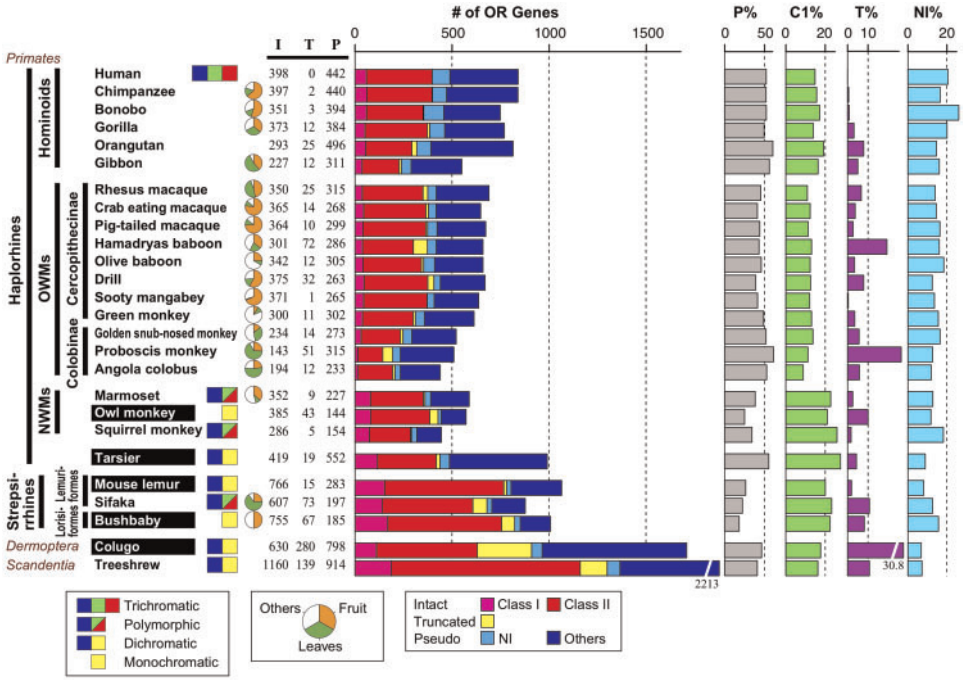

Fig. 1. Numbers of OR genes in 24 primate and two nonprimate species. "I," "T," and "P" represent the number of intact genes, truncated genes, and pseudogenes, respectively. An intact gene is defined as a sequence starting from an initiation codon and ending with a stop codon without any gaps in conserved regions. A pseudogene is a sequence containing interrupting stop codons, frameshifts, or deletions in conserved regions. A truncated gene is a partial intact sequence that is located at a contig end (Niimura and Nei 2007; Niimura et al. 2014). In the bar graph, intact genes are classified into Class I and Class II genes, and pseudogenes are further divided into nearly intact (NI) genes and other pseudogenes. P%, C1%, T%, and NI% indicate the fraction (%) of pseudogenes, that of Class I genes, that of truncated genes, and that of NI genes, respectively, and were calculated as $P/(I+T+P)$, $C_1/(C_1+C_2)$, $T/(I+T)$, and N/P, respectively, where I, T, P, C_1, C_2, and N represent the number of intact genes, that of truncated genes, that of pseudogenes, that of intact Class I genes, that of intact Class II genes, and that of NI genes, respectively. A species name set on a black bar indicates a nocturnal species. The color vision system for each species is shown by color-coded boxes (Jacobs 2009; Moritz et al. 2013; Melin et al. 2016). All hominoid and OWM species are trichromatic. Percentage of feeding time or records on fruit and leaves is depicted in a pie chart for 18 primate species. See supplementary tables S2 and S3, Supplementary Material online, for details.

values as an estimate of the number of functional OR genes in each species: 1) F_{sum}, the sum of the number of intact genes and that of truncated genes, and 2) F_{cal}, the estimated number of functional OR genes obtained by calibration under the assumption that the fraction of missing OR genes is proportional to the reciprocal of the N50 value (supplementary table S2, Supplementary Material online).

The number of functional OR genes in strepsirrhines (600–800) is almost twice that in haplorhines (200–400), and the fraction of OR pseudogenes (P%) is smaller in strepsirrhines (mean, 22.5%) than in haplorhines (47.5%; fig. 1 and supplementary table S2, Supplementary Material online). Notably, the number of functional OR genes in tarsiers is similar to that in anthropoids, but is considerably smaller than that in strepsirrhines, highlighting a gap between strepsirrhines and haplorhines. Treeshrews have a much larger number of OR genes than any primate species do, while the number of functional OR genes in colugos is unclear due to the large T% value. Nocturnal species tend to have a larger repertoire of intact OR genes than their diurnal close relatives (i.e., owl monkey vs. marmoset or squirrel monkey; mouse lemur or bushbaby vs. sifaka); however, the difference in the number of functional OR genes is much smaller than that between strepsirrhines and haplorhines. Colobines have the smallest numbers of functional OR genes among the primates examined (Zhou, Wang, et al. 2014). Among the three major groups in haplorhines, hominoids, OWMs, and NWMs, the fraction of pseudogenes is the highest in hominoids (mean, 54.1%), followed by OWMs (47.2%) and NWMs (32.8%). Mammalian OR genes can be clearly classified into two groups, Class I and Class II, based on sequence similarity (Glusman et al. 2000; Niimura 2012). The fraction of Class I OR genes among intact genes is the highest in NWMs (mean, 23.0%), followed by hominoids (16.0%) and OWMs (12.0%).

In this study, we newly defined a nearly intact (NI) pseudogene, a pseudogene sequence that contains either one interrupting stop codon or one frameshift and is otherwise the same as an intact gene. An NI pseudogene is expected to have been generated recently. Hominoids have a larger number of NI pseudogenes than the other primate species examined (fig. 1), suggesting that hominoids have lost more OR genes than other primates in recent times (see below).

Phylogenetic Generalized Least-Squares Regression Analyses

To investigate which ecological/anatomical factors affect the number of functional OR genes while statistically controlling

Table 1. Results of PGLS Regression for Number of Functional OR Genes versus Anatomical/Ecological Factors.

Response Variable	*n*	Predictor Variable	Slope	*P*	R^2	λ^a
F_{sum}	24	Nose structure	−401	*0.0014***	0.376	0.996
F_{sum}	24	Color vision	−70.1	*0.025**	0.209	1.000
F_{sum}	24	Activity pattern	−130	*0.044**	0.172	1.000
F_{sum}	18	Frugivory	1.10	0.136	0.133	1.000
F_{sum}	18	Folivory	−0.601	0.414	0.042	0.998
F_{cal}	24	Nose structure	−386	*0.0014***	0.377	0.970
F_{cal}	24	Color vision	−64.3	*0.039**	0.180	0.982
F_{cal}	24	Activity pattern	−120	0.064	0.148	0.983
F_{cal}	18	Frugivory	1.37	0.084	0.175	0.995
F_{cal}	18	Folivory	−1.05	0.204	0.099	0.974

[a]Pagel's λ for phylogenetic signal (Pagel 1999).
*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

Table 2. Results of PGLS Multivariate Regression for Number of Functional OR Genes versus Nose Structure and Ecological Factors.

Response Variable	*N*	Predictor Variables	Slope	*P*	R^2	λ^a
F_{sum}	24	Nose structure	−408	*0.031**	0.503	0.990
		Color vision	−48.2	0.083		
F_{sum}	24	Nose structure	−375	*0.0027***	0.508	0.990
		Activity pattern	−107	0.075		
F_{sum}	18	Nose structure	−299	*0.037**	0.885	0.000
		Frugivory	2.10	*0.0092***		
F_{sum}	18	Nose structure	−423	*<0.0001****	0.916	0.000
		Folivory	−2.66	*0.00095****		
F_{cal}	24	Nose structure	−439	*0.020**	0.491	0.962
		Color vision	−39.2	0.15		
F_{cal}	24	Nose structure	−377	*0.0022***	0.494	0.961
		Activity pattern	−86.0	0.15		
F_{cal}	18	Nose structure	−212	0.39	0.546	0.974
		Frugivory	1.21	0.068		
F_{cal}	18	Nose structure	−411	*<0.0001****	0.911	0.000
		Folivory	−2.84	*0.00058****		

[a]Pagel's λ for phylogenetic signal (Pagel 1999).
*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

for phylogeny, we performed phylogenetic generalized least-squares (PGLS) regression analyses (Symonds and Blomberg 2014). We selected the ecological/anatomical factors as predictor variables for this analysis, namely, nose structure (strepsirrhine/haplorhine), color vision system (monochromatic/dichromatic/polymorphic/trichromatic), activity pattern (nocturnal/diurnal), and the extent of frugivory and folivory (percentage of feeding time or event counts for fruits and leaves, respectively, in total diet) (fig. 1 and supplementary table S3, Supplementary Material online). According to the results from PGLS regression analyses, the nose structure showed the highest correlation with the number of functional OR genes among the examined predictor variables ($P < 0.01$; table 1). In addition, color vision and activity pattern (for F_{sum}) also showed significant correlation ($P < 0.05$). We then conducted multivariate PGLS regression analyses to access the influence of each ecological factor on the number of functional OR genes that is independent from nose structure. The results showed that the extent of folivory has a highly significant effect (with a negative slope) on the number of OR genes ($P < 0.001$; table 2). The effect was also observed for the extent of frugivory though to a lesser magnitude and with a positive slope ($P < 0.01$ for F_{sum}). These observations suggest that folivorous primates tend to have less (and frugivorous primates tend to have more) functional OR genes. On the other hand, color vision system and activity pattern did not show significant correlations when phylogeny and difference in nose structure were statistically controlled ($P > 0.05$). Taken together, the PGLS analyses demonstrated that nose structure and extent of folivory are important factors affecting the number of functional OR genes. The results for F_{sum} and F_{cal} as a response variable were similar (tables 1 and 2).

We also examined the correlation between the number of functional OR genes and the relative size of the olfactory bulb (OB) to that of the total brain using the data from Heritage (2014) (supplementary table S4, Supplementary Material online). We found a clear linear relationship between both F_{sum} and F_{cal} and relative OB size (fig. 2). The PGLS regression analyses revealed significant positive correlations ($P = 0.0085$ for F_{sum} and $P = 0.011$ for F_{cal}; supplementary table S5, Supplementary Material online). On the other hand, we found no significant correlation between F_{sum} (or F_{cal}) and the absolute OB size, nor between the fraction of F_{sum} (or F_{cal}) to the total number of OR genes and the relative OB size (supplementary fig. S2 and table S5, Supplementary Material online).

Loss of OR Genes in Primate Evolution

Next, we investigated when the rapid loss of OR genes occurred in primate evolution. For this purpose, we identified orthologous gene groups (OGGs) among primates to trace the evolutionary fates of individual OR genes. An OGG is defined as a group of genes that originated from a single gene in the most recent common ancestor (MRCA) of the species under consideration (Niimura et al. 2014). In this case, we are considering the MRCA of all primates. By identifying OGGs and by examining the presence/absence of a given OGG, we can specify when a given gene in the primate MRCA was lost in primate evolution.

As a result, 8,948 intact and 539 truncated OR genes identified from 24 primate species were classified into 680 primate-OGGs. Therefore, the MRCA of all primates was estimated to have had 680 functional OR genes. This number is smaller than the estimate of 781 for the MRCA of placental mammals (Niimura et al. 2014). We then estimated the number of OGG losses in each branch in primate evolution and calculated the rate of OGG losses per million years (My) in each branch (fig. 3A). Here, we used the total of intact and truncated genes for analyses, but the results are essentially the same when only intact genes are used (supplementary fig. S3, Supplementary Material online). The results showed that the rate of OGG losses has highly fluctuated during primate evolution (fig. 3B). On an average, the OGG loss rate is the highest in hominoids, followed by OWMs and NWMs, and the lowest in strepsirrhines. We found that the branch between the ancestral node of OWMs and that of colobines showed the highest OGG loss rate (23.0; fig. 3A; very short branches such as those leading to hamadryas baboon or to olive

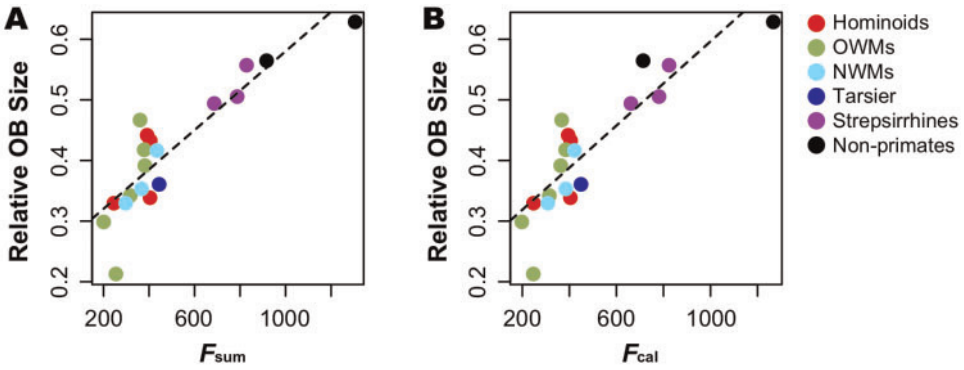

FIG. 2. Correlation between the relative OB size and the number of functional OR genes for F_{sum} (*A*) and F_{cal} (*B*). The vertical axis indicates $\log_{10}$(OB volume)/$\log_{10}$(total brain volume). OB and total brain volumes were taken from Heritage (2014). When the same species as that used in this analysis was not found in Heritage (2014), the data for the closest species were used (supplementary table S4, Supplementary Material online). A dashed line represents the regression line. Pearson's correlation coefficients $r = 0.886$, Spearman's correlation coefficients $r_S = 0.850$ (*A*); $r = 0.874$, $r_S = 0.850$ (*B*).

baboon were ignored). Acceleration of OGG loss rates was also observed at the branch between the ancestral node of all primates and that of haplorhines (11.5) and the branch leading to each of the six hominoid species examined (10.8–16.7). In strepsirrhines, the number of OGG losses is generally much smaller and the number of OR gene duplications is much larger than in haplorhines.

We performed the same analyses for Class I and Class II OR genes separately (supplementary figs. S4 and S5, Supplementary Material online). The overall pattern of OGG losses for Class II genes is the same as that in figure 3A. For Class I genes, however, the highest OGG loss rate was observed at the ancestral branch of colobines, followed by the ancestral branch of OWMs. The high rate of OGG losses at the OWM ancestor is consistent with the lower fraction of Class I genes in OWMs than in hominoids and NWMs (fig. 1).

The OGG loss rate assigned to the exterior branch connecting to each extant species is expected to reflect the number of pseudogenes that have been generated recently in the species. We therefore investigated the correlation between the OGG loss rate in each exterior branch connecting to a given species and the number of NI pseudogenes in the species (fig. 1). As expected, the OGG loss rate at an exterior branch showed a significant positive correlation with the number of NI pseudogenes (fig. 3C and supplementary fig. S6, Supplementary Material online). Therefore, the species with a larger number of NI pseudogenes tends to show a higher OGG loss rate at its exterior branch, supporting the independent acceleration of OGG loss rates in each hominoid lineage observed above.

Evolution of OR Genes in Hominoids, OWMs, and NWMs

Next, we compared the evolutionary dynamics of OR genes among three major haplorhine groups, hominoids, OWMs, and NWMs. The MRCAs of hominoid, OWM, and NWM species examined in this study were present at ~19.4, 17.6, and 18.4 Ma, respectively (supplementary fig. S7, Supplementary Material online), which show similar timing. Figure 3A suggested that the MRCA of hominoids (460) had a larger repertoire of OR genes than that of OWMs (408) or NWMs (438). However, note that these numbers do not include the gene duplications that had occurred between the MRCA of primates and each of the MRCAs of hominoids/OWMs/NWMs. Therefore, to investigate the difference in the evolutionary dynamics of OR genes among the three groups, we identified hominoid-OGGs, OWM-OGGs, and NWM-OGGs.

The results showed a sharp contrast in OR gene evolution between hominoids and OWMs/NWMs (fig. 4 and supplementary fig. S8, Supplementary material online). The estimated number of functional OR genes in the MRCA of hominoids (528) is considerably larger than that in the MRCA of OWMs (458) or NWMs (463). The mean OGG loss rate was suggested to be much faster in hominoids (13.4) than in OWMs (9.4) or NWMs (6.7) (fig. 4*B*). Moreover, the estimated numbers of OR gene gains are generally larger in hominoids (mean, 27.3) than in OWMs (19.5) or NWMs (21.0). Therefore, although the MRCA of hominoids was estimated to have a larger repertoire of OR genes than that for OWMs/NWMs, faster OR gene losses in the hominoid lineages than in the OWM/NWM lineages have resulted in similar sizes of OR gene repertoires in extant species for hominoids (mean, 348.8), OWMs (326.6), and NWMs (360.0).

Discussion

In this study, we examined OR gene repertoires in 24 primate species with diverse eye and nose morphologies, color vision systems, activity patterns, and feeding habitats. PGLS regression analyses demonstrated that 1) haplorhines have a significantly smaller number of functional OR genes, and 2) the extent of folivory is significantly negatively correlated with the number of functional OR genes. In addition, we traced the evolutionary fates of individual OR genes by identifying OGGs and found acceleration of OR gene loss rates in the following: 1) the ancestral branch of haplorhines, 2) the ancestral branch of colobines, and 3) the exterior branch leading to each hominoid species examined (fig. 5). The results from PGLS

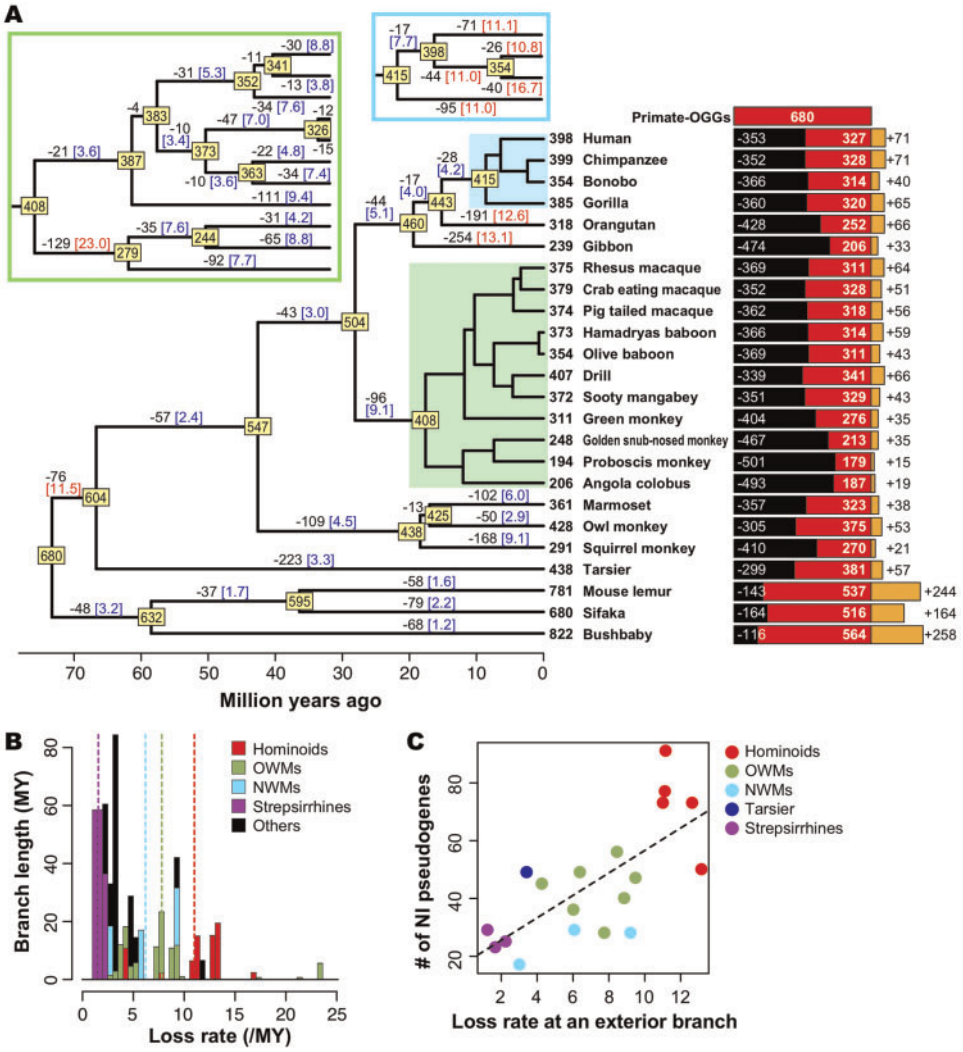

Fig. 3. Loss of OR genes in primate evolution inferred from intact and truncated genes. The result estimated from intact genes is shown in supplementary figures S3 and S6, Supplementary Material online. (*A*) A number with a minus sign at each branch indicates the number of primate-OGGs that were estimated to have been lost at the branch. Numbers in brackets represent the rate of OGG losses per million years (Mys). Each branch length was calculated from divergence times obtained from TimeTree (Hedges et al. 2006) (supplementary fig. S7, Supplementary Material online). A loss rate $\geq$10 and that $<$10 are shown in red and blue, respectively. The loss rates with a branch length shorter than 2 My are omitted. Numbers in yellow boxes indicate the number of OGGs that were estimated to have been present at each ancestral node. Green and blue boxes correspond to the green and blue insets, respectively. The number at each species name shows the total number of intact and truncated genes. In the bar graph at the right, the black bar represents the number of OGG losses in each species' lineage, which is equal to the summation of the numbers of OGG losses at respective branches from the MRCA of primates to each species. Red and orange bars show the number of OGGs that remained in the extant species and the number of OR genes gained in each lineage, respectively. For example, humans have lost a total of 353 OGGs among 680 OGGs, but 71 gene gains have occurred, resulting in the repertoire of 398 intact OR genes. (*B*) Distribution of the OGG loss rates for branches in the primate evolution calculated from (*A*). The height of a bar indicates the length of the corresponding branch. A bar corresponding to the branch in hominoids, OWMs, NWMs, and strepsirrhines are colored by red, green, sky blue, and purple, respectively. Other branches are shown in black. The means of the loss rates for all branches in hominoids, OWMs, NWMs, and strepsirrhines are 11.0, 7.8, 6.2, and 1.6, respectively, which are shown by a dashed line in the same color as a bar. (*C*) Correlation between the OGG loss rate at exterior branches and the numbers of NI pseudogenes. OGG loss rates were inferred from intact and truncated genes. To avoid noise in OGG loss rates due to short branch lengths, we only considered exterior branches longer than 5 Mys. For species with short exterior branches, we chose a representative species with the lowest T% among the closely related species. For example, the exterior branch connecting to the chimpanzee is 2.4 My long; therefore, we joined the branch from the ancestral node of the human–chimpanzee MRCA to that of the chimpanzee–bonobo MRCA (4.0 My) with the exterior branch and regarded it as one branch of 6.4 My long with 70 ($=44+26$) OGG losses, and we did not use the exterior branch connecting to bonobo. We excluded bonobos, rhesus macaques, crab-eating macaques, hamadryas baboons, and drills from the analysis and used the 19 remaining primate species. A dashed line represents the regression line. $r=0.711$ ($P=0.00065$, $df=17$), $r_S=0.736$ ($P=0.00033$). Note that all data are independent because the exterior branches examined do not share any interior branches.

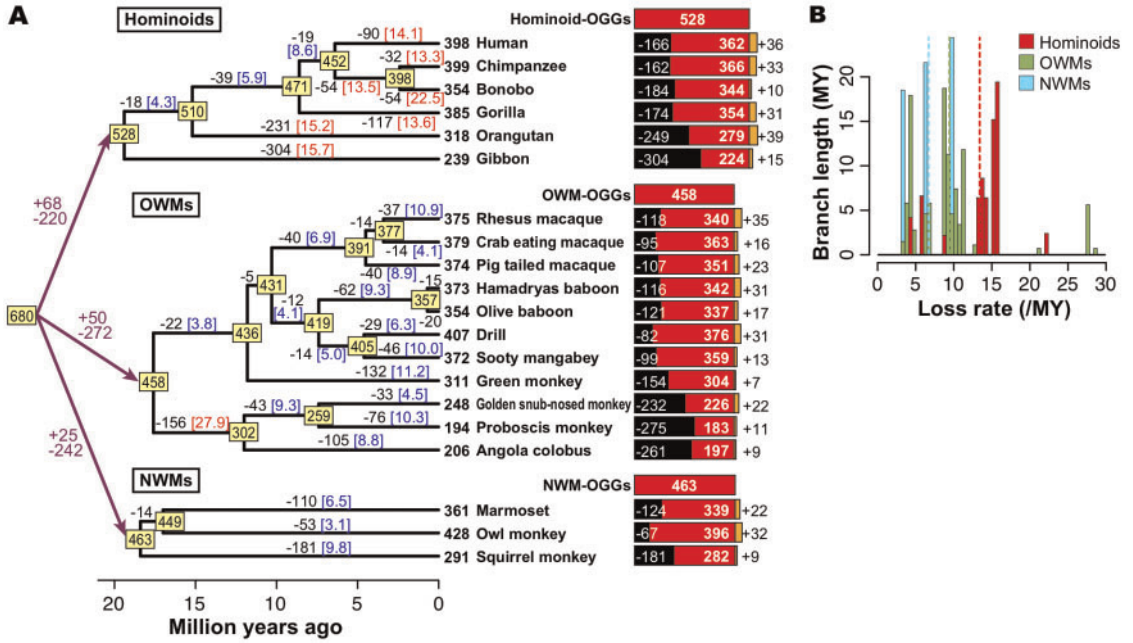

Fig. 4. Loss of OR genes in the evolution of hominoids, OWMs, and NWMs inferred from intact and truncated genes. The result estimated from intact genes only is shown in supplementary figure S8, Supplementary Material online. (*A*) A number with a minus sign at each branch indicates the number of hominoid-OGGs (top), OWM-OGGs (middle), and NWM-OGGs (bottom) that were estimated to have been lost at the branch. The number in brackets represents the rate of OGG losses per million years. The loss rate ≥12.5 and that <12.5 are shown in red and blue, respectively. Here, we used 12.5 rather than 10 (as in fig. 3) as a threshold because there is a gap at this value in the distribution of OGG loss rates as shown in (*B*). The loss rates for short branches (<2 My) are omitted. Among 680 primate-OGGs, 460 OGGs survived at the MRCA of hominoids (see fig. 3*A*); on the other hand, the identification of hominoid-OGGs suggested that the number of OR genes in the MRCA of hominoids is 528. Therefore, it is inferred that, between the MRCA of primates and that of hominoids, 68 (= 528−460) gene gains have occurred and 220 (= 680−460) primate-OGGs have been lost. Other values shown in purple were estimated in the same manner. For further explanation, see the legend of figure 3. (*B*) Distribution of OGG loss rates for branches in the evolution of hominoids, OWMs, and NWMs calculated from (*A*). The height of a bar indicates the length of the corresponding branch. The means of the OGG loss rates for all branches in hominoids, OWMs, and NWMs are 13.4, 9.4, and 6.7, respectively, which are shown by a dashed line in the same color as a bar.

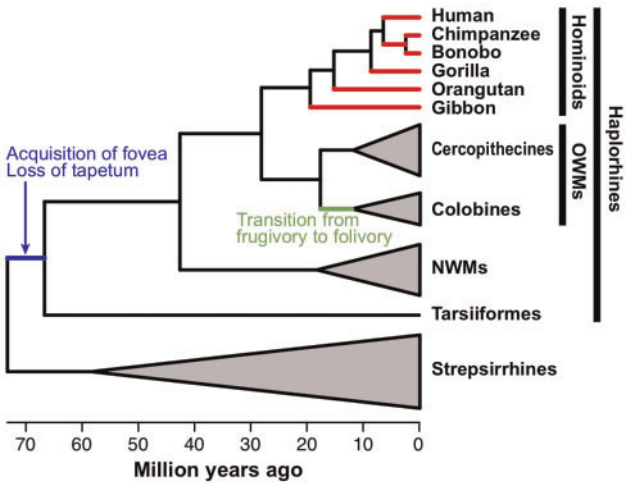

Fig. 5. Schematic illustration of the degeneration of OR gene repertoires in primate evolution. Colored branches indicate accelerated loss of OGGs as follows: blue, the ancestral branch of haplorhines; green, the ancestral branch of colobines; and red, an exterior branch leading to each hominoid species.

regression analyses and those from OGG loss rates are consistent with each other. We will discuss the evolutionary significance of the acceleration of OR gene losses as follows.

Ancestral Branch of Haplorhines: Acquisition of Acute Vision

As mentioned in the Introduction, the phylogenetic position of tarsiers among primates has long been debated, owing to tarsiers showing a mixture of strepsirrhine and anthropoid features (Fleagle 2013). The results in figure 1 support a haplorhine–strepsirrhine dichotomy rather than a prosimian-anthropoid dichotomy in terms of the number of functional OR genes. The relative OB size is 0.360 in tarsiers, which is very close to the mean of the relative OB sizes among the 13 anthropoid species (0.366) examined here and is much smaller than that among the three strepsirrhine species (0.518; see supplementary table S4, Supplementary Material online). Barton (2006) also reported a difference in relative OB size between haplorhines and strepsirrhines. All of these observations support the presence of a gap in olfactory ability between strepsirrhines and haplorhines, but not between prosimians and anthropoids.

Tarsiers also have many distinctive features all their own. The most striking feature of tarsiers is the size of their eyes, each of which is actually larger than the brain (Fleagle 2013). Tarsiers possess a fovea (a pit in the central retina), giving an enhanced visual acuity, and lack a tapetum (a reflecting layer of the retina) in their eyes, both traits being unusual for

nocturnal mammals. These and other lines of evidence suggest the hypothesis that the common ancestor of haplorhines is diurnal (or crepuscular, i.e., active under dim light conditions; Melin et al. 2013) and tarsiers are secondarily adapted to nocturnality (Tan et al. 2005; Williams et al. 2010; Joffe et al. 2014). It is probable that most of the adaptations for high visual acuity found in anthropoids evolved in the haplorhine ancestor, and that these features are partly obscured in tarsiers through a secondary adaptation to nocturnality. According to this hypothesis, the acquisition of a fovea and the loss of a tapetum occurred in the common ancestor of haplorhines in the diurnal setting. Recent discovery of the fossils of basal haplorhines also supported the hypothesis of diurnality of the haplorhine ancestor (Ni et al. 2013).

Figure 3 suggests a radical loss of OR genes in the ancestral branch of haplorhines and a reduced loss rate in the ancestral branch of anthropoids. In fact, the OGG loss rate in the ancestral branch of anthropoids (2.4 in fig. 3A) is five times slower than that in the ancestral branch of haplorhines (11.5). These observations are consistent with the evolutionary scenario of primates mentioned earlier; that is, the degeneration of olfactory ability occurred concomitantly with the acquisition of acute visual sense in the common ancestor of haplorhines. We therefore suggest that the radical loss of OR genes in the ancestral branch of haplorhines can best be explained by the acquisition of acute visual sense.

Activity Patterns and Color Vision Systems

We found that nocturnal species tend to have a larger number of functional OR genes than their diurnal relative species within the same taxonomic group (e.g., NWMs or strepsirrhines) (fig. 1); however, PGLS regression analyses revealed that the difference between nocturnality and diurnality is not significantly correlated with the number of functional OR genes after controlling for phylogeny and the difference between strepsirrhines and haplorhines (table 2). Lemurs show highly diverse activity patterns, including diurnality, nocturnality (scotopic preference and mesopic preference), and various types of cathemerality (Donati et al. 2013; Valenta et al. 2016). Therefore, an examination of more closely related species within lemurs having subtlely different activity patterns would provide us information on the possible association between activity patterns and the numbers of OR genes.

The color vision system also does not significantly affect the number of functional OR genes when phylogeny and difference between strepsirrhines and haplorhines are statistically controlled. As was previously reported using five anthropoid species (Matsui et al. 2010), we found that the number of OR genes in OWMs and hominoids with routine trichromacy was not smaller than in NWMs with partial trichromacy by M/L opsin polymorphism (fig. 1). Moreover, we should note that the OGG loss rate at the ancestral branch of catarrhines, at which the acquisition of routine trichromacy occurred, is very low (3.0; fig. 3A), which is again consistent with our previous results (Matsui et al. 2010). In this regard, a comparison of the OR gene repertoire in the howler monkey, a NWM species having routine trichromacy, with those in other NWMs would be important to reveal the interplay between color vision and olfaction in more detail.

Traditionally, it was assumed that color vision is irrelevant for nocturnal species. However, several recent studies have suggested that nocturnal color vision may be adaptive under certain conditions, though the ecological factors for differential selective pressure on the SWS1 opsin gene in nocturnal primates is unclear (Perry et al. 2007; Melin et al. 2012; Veilleux et al. 2013). Our data showed that the number of functional OR genes in monochromatic bushbabies and that in dichromatic mouse lemurs is similar (fig. 1). It therefore appears that the maintenance of color vision in a nocturnal condition would hardly contribute to the reduction of OR genes. This study provides another line of evidence that color vision and the number of OR genes are not correlated in the context of nocturnality.

Ancestral Branch of Colobines: Adaptation to Folivory

Our analysis revealed that colobines had the smallest numbers of functional OR genes among the primates (~200; fig. 1), and the ancestral branch of colobines showed the most accelerated rate of OR gene losses among those examined (fig. 3A). The mean relative OR size of colobines (0.255; supplementary table S4, Supplementary Material online) is considerably smaller than the mean among cercopithecines (0.404), the other family of OWMs. These observations suggest a reduced reliance on olfaction in colobines.

Colobines are unique primates in that they rely on leaves and seeds rather than fruits and insects as their major food source (Fleagle 2013). Like cows, colobines have a complex, sacculated stomach that enables them to maintain bacterial colonies for digesting cellulose. Thanks to their specialized digestive systems, they can feed on mature leaves and unripe fruits that are not edible by cercopithecines and apes (Dominy et al. 2001).

There are a number of researchers who suggest that frugivorous primates use olfactory cues in fruits foraging (Hiramatsu et al. 2009; Melin et al. 2009; Valenta et al. 2013; Nevo and Heymann 2015). For example, Nevo et al. (2015) demonstrated that captive spider monkeys can discriminate the odor mixtures of ripe fruits from those of unripe fruits of two plant species consumed by the monkeys. Fruit volatile compounds are indicative of the fruits' nutrient content (Goff and Klee 2006), and they were suggested to be a signal that coevolved with the primates that disperse their seeds (Nevo et al. 2016). Moreover, behavioral tests showed that spider monkeys have a higher olfactory sensitivity to odorants commonly present in fruits than do nonprimate mammals such as rats and dogs do (Hernandez Salazar et al. 2003; Laska et al. 2006).

On the other hand, several studies indicated that olfaction is less important in leaf foraging than in fruit foraging (Nevo and Heymann 2015). A study in three captive strepsirrhine species with different feeding ecologies reported that folivorous Coquerel's sifakas used visual rather than olfactory cues to select food, while frugivorous ruffled lemurs equally used both cue types (Rushmore et al. 2012). Field observation of wild proboscis monkeys suggested that the strongest

explanatory factor for their leaf preference was the spatial distribution pattern of plant species, measured as abundance; in other words, proboscis monkeys do not selectively but rather opportunistically feed (Matsuda et al. 2017). Moreover, Barton (2006) reported that folivorous primates have a significantly smaller relative OB size than frugivorous primates. It is therefore likely that folivorous primates rely less on olfaction for foraging than frugivorous primates. Instead, they may use other sensory cues such as location, texture (leaf toughness) (Matsuda et al. 2017), and leaf color (Dominy and Lucas 2001; Melin et al. 2017). PGLS analyses revealed that the extent of folivory is significantly negatively correlated with the number of functional OR genes, and the opposite pattern was observed for frugivory to a lesser magnitude (table 2). Taken together, we suggest that the dietary transition from frugivory to folivory caused a radical loss of OR genes in the ancestor of colobines.

Exterior Branch Leading to Each Hominoid Species

We also found that the rates of OR gene loss were accelerated in the exterior branch leading to each of the six hominoid species examined. The qualities of genome assemblies are very high for Homininae (human, chimpanzee, and gorilla; N50 > 50,000; see supplementary table S1, Supplementary Material online). Therefore, the estimated number of OGG losses at an exterior branch for Homininae is expected to be accurate, though that for orangutan and that for gibbon may slightly vary when the qualities of their genome sequences are improved. Moreover, the observation of accelerated OR gene losses in hominoids is not likely to be an artifact for the following reasons.

First, the number of NI pseudogenes in each genome was positively correlated with the rate of OGG losses at an external branch for the species (fig. 3C). This observation is expected, because NI pseudogenes are assumed to have been generated recently in evolution. The presence of a large number of NI pseudogenes in hominoids is consistent with the accelerated rate of OR gene losses in the branch leading to each hominoid species. It was reported that hominoid genomes contain a specific group of OR pseudogenes, named H* or 7E pseudogenes (Newman and Trask 2003; Niimura and Nei 2005; Niimura et al. 2014), which were generated from one ancestral gene, *OR7E24*, and are scattered throughout the genome. We identified 89, 93, 64, 60, 75, and 36 H* pseudogenes from human, chimpanzee, bonobo, gorilla, orangutan, and gibbon genomes, respectively, but no NI pseudogenes were found among the H* pseudogenes, except for one NI pseudogene from each gorilla and orangutan genomes (supplementary table S2, Supplementary Material online). Therefore, the presence of a large number of NI pseudogenes in hominoids is not due to the presence of H* pseudogenes.

Second, the number of functional OR genes in the hominoid MRCA (528) is considerably larger than those in the OWM and NWM MRCAs (~460; fig. 4A). This observation is consistent with the rapid losses in each hominoid branch, because hominoids, OWMs, and NWMs currently possess similar numbers of OR genes. Therefore, intriguingly, it is possible that our hominoid ancestor may have had a better sense of smell than the OWM or NWM ancestors.

Third, OR gene repertoires in hominoids are more diversified to one another compared with those in OWMs/NWMs. For example, when we compare the repertoires of primate-OGGs between humans and chimpanzees, which diverged 6.4 Ma (supplementary fig. S7, Supplementary Material online), 19.1% of the repertoire is species-specific (supplementary fig. S9, Supplementary Material online). On the other hand, the species-specificity in primate-OGGs between pig-tailed macaques and sooty mangabeys, which diverged 10.3 Ma, is only 11.3%. Therefore, though the evolutionary relationship between humans and chimpanzees is closer than that between pig-tailed macaques and sooty mangabeys, the OR gene repertoires are more diversified for the former pair of species than the latter. This observation is consistent with the faster OR gene loss rates in hominoids than in OWMs.

It has been reported that loss-of-function variants are significantly enriched at OR gene loci (MacArthur et al. 2012). Therefore, it is possible that some OR gene loss represents a segregating pseudogene, which has both functional and nonfunctional alleles segregating in a population. In a search of the 1000 Genomes Project data, Olender et al. (2012) identified 26 segregating OR pseudogenes that are annotated as pseudogenes in the reference genome. When we considered the 26 segregating pseudogenes in Olender et al. as functional, the estimated number of OGG losses at the external branch leading to humans decreased from 71 (loss rate, 11.1) to 54 (8.4), while those at the other branches did not change at all. When a given species has a segregating pseudogene, it is highly likely that a closely related species retains its functional ortholog; for this reason, inclusion of segregating pseudogenes in a given species into a functional gene category is expected to almost exclusively affect the number of gene losses at the external branch leading to the species and to hardly affect those at the other branches. Therefore, the rate of OGG losses at an external branch may be affected by the presence of segregating pseudogenes, but our conclusion of accelerated gene losses at the ancestral branch of haplorhines and that of colobines does not change.

Figure 3 showed that the acceleration of OR gene losses have occurred independently at every branch leading to an extant hominoid species. The factors that caused the acceleration might differ among the branches, and further studies should be necessary to specify these factors.

Number of OR Genes and the Correlation with the OB Size

"Olfactory ability" contains two different aspects: one is the sensitivity to odors, and the other is the discrimination of odors, the breadth of odors that can be discriminated. It is unclear which aspect of olfactory ability is represented in the number of OR genes, but it is likely that the number is correlated with odor discrimination rather than odor sensitivity, because ORs with different amino acid sequences are expected to bind different sets of odorants. On the other hand, sensitivity to a particular odor may be determined by the absolute amount of expressed ORs that bind to the odor.

In fact, behavioral tests indicated that the discrimination ability for structurally related odorants such as enantiomers are similar among humans, pigtail macaques, and squirrel monkeys, while Asian elephants and mice performed markedly better than the three primate species (Laska et al. 2005; Rizvanovic et al. 2013). The results from the behavioral test are consistent with the numbers of OR genes identified in this study (fig. 1) and a previous study (Niimura et al. 2014).

We demonstrated that the number of functional OR genes, but not the fraction of functional OR genes, is significantly positively correlated with the relative size of the OB to the total brain (fig. 2; supplementary fig. S2 and table S5, Supplementary Material online), suggesting that the number of functional OR genes in a genome, rather than the fraction of them, is a good proxy of a species' olfactory ability (Steiger et al. 2008, 2009; Niimura et al. 2014). We also showed that the number of OR genes is significantly correlated with the relative OB size but not with the absolute OB size (supplementary fig. S2 and table S5, Supplementary Material online). Some researchers have argued that the olfactory sensitivity is better correlated with the absolute OB size than the relative OB size (Smith and Bhatnagar 2004; Pihlstrom et al. 2005). Therefore, it may be assumed that the relative OB size tends to reflect the olfactory discrimination, while the absolute OB size tends to reflect the olfactory sensitivity (but see Garrett and Steiper 2014). The OB is composed of glomeruli, and each glomerulus is regarded to converge inputs from one type of OR (Mori et al. 1999); however, the relationship between the number of glomeruli and that of functional OR genes is unclear (Moriya-Ito et al. 2015). To elucidate the connection between the genomic and anatomical aspects of the olfactory system in more detail, further study using a broader range of species should be necessary.

Perspectives

In this study, we demonstrated that accelerations of OR gene losses have occurred at several lineages during the primate evolution, suggesting that the current OR gene repertoire in each species has been shaped by a complex interplay of anatomical and ecological factors. The function of ORs that has been lost in a particular lineage, however, remains elusive at this stage. In this regard, the OGGs identified in this study will be useful to elucidate the link between an environmental factor (e.g., folivory) and the function of OR genes. They may also provide the basis for identifying particular amino acid sites that are responsible to a specific function in a further study.

Materials and Methods

Data

The latest genome assemblies for human (GRCh38), gibbon (nomLeu3), olive baboon (papAnu2), marmoset (calJac3), squirrel monkey (saiBol1), and bushbaby (otoGar3) were obtained from the UCSC Genome Browser (http://genome.ucsc.edu/). The genome sequence for hamadryas baboon (Pham_1.0) were downloaded from the Ensemble genome browser (http://www.ensembl.org). The latest genome assemblies for the other 19 species, chimpanzee (Pan_tro_3.0), bonobo (panpan1.1), gorilla (gorGor4), orangutan (P_pygmaeus_2.0.2), rhesus macaque (Mmul_8.0.1), crab eating macaque (Macaca_fascicularis_5.0), pig-tailed macaque (Mnem_1.0), drill (Mleu.le_1.0), sooty mangabey (Caty_1.0), green monkey (Chlorocebus_sabeus_1.1), golden snub-nosed monkey (Rrox_v1), proboscis monkey (Charlie1.0), Angola colobus (Cang.pa_1.0), owl monkey (Anan_1.0), tarsier (Tarsius_syrichta-2.0.1), mouse lemur (Mmur_2.0), sifaka (Larkin et al. 2007), colugo (G_variegatus-3.0.2), and treeshrew (TupChi_1.0) were downloaded from GenBank (https://www.ncbi.nlm.nih.gov/genbank/). See supplementary table S1, Supplementary Material online, for details.

Identification of OR Genes from Genome Sequences

OR genes were identified from whole-genome sequences by using the method described previously (Niimura 2013) with slight modification (described in supplementary note and supplementary fig. S10, Supplementary Material online). After the identification of OR genes, we eliminated genes that are duplicated possibly due to a low quality of genome assembly in the following way. We excluded OR genes encoded in a short contig for which the entire sequence is embedded in a larger contig/scaffold. We first calculated amino acid sequence identities for all possible pairwise comparisons among all OR genes identified from each species' genome on the basis of pairwise alignments generated by ClustalW2 (Larkin et al. 2007). We then extracted OR gene pairs showing a >99% amino acid sequence identity without any gaps. For each OR gene pair, a nucleotide sequence identities was calculated between the two contigs/scaffolds containing the OR genes. When a shorter contig was embedded in a longer contig/scaffold with a >99% nucleotide sequence identity and with no gaps, the OR gene encoded in a shorter contig was removed. Amino acid and nucleotide sequences of the OR genes are available in supplementary data S1 and S2, Supplementary Material online, respectively, for 24 primate species and one colugo and one treeshrew species.

Estimation of the Number of Functional or Genes

F_{cal} is calculated as follows. Let the number of intact OR genes be I and that of missing intact OR genes due to incompleteness of a genome assembly be M. We assume that the fraction of M is obtained from the regression line $y = 2.073x$ shown in supplementary figure S1, Supplementary Material online, following the equation $M/(I + M) = 2.073/N_{50}$, where N_{50} is the N50 value in kilobases. F_{cal} is equal to $I + M$.

Phylogenetic Generalized Least-Squares Regression Analysis

We used PGLS regression (Symonds and Blomberg 2014) to investigate the relationship between the number of functional OR genes and ecological/anatomical factors while statistically controlling for phylogeny. PGLS regression analyses were performed using R with the caper packages (https://cran.r-project.org/web/packages/caper/index.html). The phylogeny of primates was obtained from Perelman et al. (2011)

and Springer et al. (2012), and divergence times were taken from TimeTree (http://www.timetree.org/; supplementary fig. S7, Supplementary Material online) (Hedges et al. 2006). For dietary data, we compiled the percentage of feeding time or records on fruit and leaf consumption for 18 primate species (supplementary table S3, Supplementary Material online): gorilla (Doran-Sheehy et al. 2009); rhesus macaque (Zhou, Wei, et al. 2014); golden snub-nosed monkey (Li 2007); proboscis monkey (Matsuda et al. 2009); marmoset (Amora et al. 2013); sifaka (Lewis and Kappeler 2005), and 12 other primate species (Campbell et al. 2010). When multiple data were available for one species, we took the mean. Categorical data are incorporated into the model according to the following coding scheme: strepsirrhine (0), haplorhine (1); monochromatic (0), dichromatic (1), polymorphic (2), trichromatic (3); and nocturnal (0), diurnal (1).

PGLS regression analyses were also performed to examine the correlation between the number/fraction of functional OR genes and the relative/absolute OB size.

Construction of Phylogenetic Trees

Each neighbor-joining (NJ) tree (Saitou and Nei 1987) was constructed with Poisson correction (PC) distance using the LINTREE program (Takezaki et al. 1995) (http://www.personal.psu.edu/nxm2/software.htm). Multiple alignments of translated amino acid sequences were generated by the program MAFFT (Katoh and Standley 2013) (http://mafft.cbrc.jp/alignment/software/).

Identification of OGGs among Primates

Intact and truncated OR genes identified from the genomes of 24 primate species were classified into OGGs among primates (primate-OGGs). For this purpose, we used 781 OGGs among placental mammals (placental-OGGs) identified in Niimura et al. (2014). An OGG is defined as a group of genes originated from a single gene in the MRCA of the species considered. Note that genes included in an OGG vary depending on the MRCA. If gene duplication(s) had occurred in the branch between the MRCA of placental mammals and that of primates, one placental-OGG corresponds to two or more primate-OGGs. Therefore, the basic strategy is to assign each primate OR gene to one of the placental-OGGs and then to divide placental-OGGs into primate-OGGs.

To assign primate OR genes to placental-OGGs, we first performed BLASTP searches (Altschul et al. 1997) using each of 8,948 intact OR genes identified from the 24 primate species as a query against 10,659 intact OR genes from 13 placental mammals identified in Niimura et al. (2014). We then extracted the top five OGGs that showed the lowest to the fifth lowest *E*-values for each query primate OR gene. For each query gene, an NJ phylogenetic tree was constructed using all intact genes from the 13 placental mammals belonging to the five OGGs (Niimura et al. 2014) together with the query gene examined. We then extracted the smallest clade that contained the query gene and all member genes belonging to an OGG or a set of OGGs (supplementary fig. S11, Supplementary Material online). In this way, each of the 8,948 query genes were assigned to one or a set of the 781 placental-OGGs.

Next, we constructed an NJ phylogenetic tree for each set of OGGs using all member genes from the 13 placental mammals and all primate query genes assigned to the set of OGGs. If a query is assigned to a set of OGGs and another query is assigned to a different set of OGGs, but some OGG(s) are common to both sets, then all OGGs involved are considered. For example, if query A is assigned to OGG1 and OGG2, and the query B is assigned to OGG1 and OGG3, then a phylogenetic tree is constructed using all genes belonging to OGG1, OGG2, and OGG3, together with queries A and B. By visual inspection, a phylogenetic tree containing multiple OGGs was subdivided into placental-OGGs. In this process, some mis-annotated OGGs in Niimura et al. (2014) were corrected. In this way, each query gene was assigned to one of the (revised version of) placental-OGGs.

We then constructed an NJ phylogenetic tree for each placental-OGGs using all query genes from the 24 primate species assigned to the OGG. As mentioned earlier, a placental-OGG may correspond to two or more primate-OGGs. All the trees were visually inspected. When a tree contained two clades such that 1) each clade contains both strepsirrhine genes and haplorhine genes, 2) some species are common to both clades, and 3) the separation of the two clades is supported with a $>70\%$ bootstrap value, then the two clades were separated into two OGGs. This process was repeated until no more clades could be separated. As a result, we obtained 680 primate-OGGs.

Truncated genes were assigned to the 680 primate-OGGs identified above in the following way. We performed BLASTP searches (Altschul et al. 1997) using each of the 539 truncated genes identified from the 24 primate species against 8,948 intact OR genes from the 24 primate species. Each truncated gene was assigned to the primate-OGG containing the intact gene showing the smallest *E*-value among the 8,948 intact genes.

In the same way, 7,333 pseudogenes from the 24 species were also assigned to the 680 primate-OGGs. We defined the pseudogenes assigned to OGG2-3 (Niimura et al. 2014), which contains *OR7E24*, as H* (7E) pseudogenes (Newman and Trask 2003; Niimura and Nei 2005).

Lists of the OR genes included in each of the 680 primate-OGGs are provided in supplementary data S3, Supplementary Material online.

Identification of Hominoid-OGGs, OWM-OGGs, and NWM-OGGs

We identified hominoid-OGGs, OWM-OGGs, and NWM-OGGs by further separating primate-OGGs. Among the 680 primate-OGGs, 455, 395, and 435 OGGs contained at least one hominoid, OWM, and NWM genes, respectively. For each of those OGGs, we constructed an NJ phylogenetic tree using all member genes from the 24 primate species belonging to the primate-OGG. Among them, 103 trees showed paraphyly for either hominoids, OWMs, or NWMs. Suppose that a tree contains two hominoid-specific clades. If the topology of the tree suggested that the separation of the two clades was more

ancient than the hominoid/NWM divergence, for example, one hominoid-specific clade forms a monophyletic clade with an OWM-specific clade with a high bootstrap value, then the two clades were separated into two hominoid-OGGs. In a similar manner, OWM-OGGs and NWM-OGGs were also identified for the 103 primate-OGGs.

The remaining phylogenetic trees were examined in the following way. When a tree contained two clades that met the conditions, 1) each clade contains both gibbon genes and nongibbon genes, 2) some species are common to the two clades, and 3) the separation of the two clades is supported with a >70% bootstrap value, then the primate-OGG was separated into two hominoid-OGGs, and this procedure was repeated iteratively. In a similar manner, when a tree contained two clades that met the conditions, 1) each clade contains both cercopithecine genes and colobine genes, and 2) and 3) are as mentioned earlier, then the primate-OGG was separated into two OWM-OGGs. For NWMs, the conditions are as follows: 1) each clade contains the genes from all three species examined (marmoset, owl monkey, and squirrel monkey), and 2) and 3) are as mentioned earlier.

There were 54, 254, and 57 truncated genes identified from hominoids, OWMs, and NWMs; they were assigned to hominoid-OGGs, OWM-OGGs, and NWM-OGGs, respectively, in the following way. We performed BLASTP searches (Altschul et al. 1997) using each of these truncated genes as a query against 8,948 intact OR genes from the 24 primate species. 1) For a given hominoid query gene, if the blast best-hit with the smallest *E*-value is also a hominoid gene, then the query was assigned to the hominoid-OGG containing the best-hit gene. OWM and NWM queries were also assigned to an OWM-OGG and an NWM-OGG, respectively, in the same way. 2) For a given hominoid query gene, if the blast best-hit is not a hominoid gene and the primate-OGG containing the best-hit does not contain any hominoid genes, then the query was assigned to a new hominoid-OGG. OWM and NWM queries were treated in the same way. 3) For a given hominoid query gene, if the blast best-hit is not a hominoid gene and the primate-OGG containing the best-hit contains hominoid gene(s), then an NJ phylogenetic tree was constructed by using the query truncated gene and all member genes included in the primate-OGG. We visually inspected the phylogenetic tree to assign the query to the most appropriate hominoid-OGG. OWM and NWM queries were treated in the same way.

Lists of the OR genes in each of the hominoid-OGGs, OWM-OGGs, and NWM-OGGs are provided in supplementary data S4–S6, Supplementary Material online.

Estimation of the Number of OGG Losses

The number of OGG losses in each branch in the primate evolution was estimated by examining the presence/absence of a gene in each OGG for a given species under the Dollo parsimonious principle (Farris 1977), in which gene losses are considered to be irreversible (Matsui et al. 2010). The OGG loss rate at each branch was calculated from the number of losses and the divergence times obtained from TimeTree (supplementary fig. S7, Supplementary Material online) (Hedges et al. 2006).

Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

Acknowledgments

We are grateful to Dr Archana Singh for useful discussion. This work was supported by the Exploratory Research for Advanced Technology (ERATO) Touhara Chemosensory Signal Project from Japan Science and Technology Agency (grant number JPMJER1202).

References

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25(17):3389–3402.

Amora TD, Beltrao-Mendes R, Ferrari SF. 2013. Use of alternative plant resources by common marmosets (*Callithrix jacchus*) in the semi-arid caatinga scrub forests of northeastern Brazil. *Am J Primatol.* 75(4):333–341.

Barton RA. 2006. Olfactory evolution and behavioral ecology in primates. *Am J Primatol.* 68(6):545–558.

Buck L, Axel R. 1991. A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* 65(1):175–187.

Campbell CJ, Fuentes A, MacKinnon KC, Bearder SK, Stumpf RM. 2010. Primates in perspective. 2nd ed. New York: Oxford University Press.

Chatterjee HJ, Ho SY, Barnes I, Groves C. 2009. Estimating the phylogeny and divergence times of primates using a supermatrix approach. *BMC Evol Biol.* 9:259.

Dominy NJ, Lucas PW. 2001. Ecological importance of trichromatic vision to primates. *Nature* 410(6826):363–366.

Dominy NJ, Lucas PW, Osorio D, Yamashita N. 2001. The sensory ecology of primate food perception. *Evol Anthropol.* 10(5):171–186.

Donati G, Santini L, Razafindramanana J, Boitani L, Borgognini-Tarli S. 2013. Un-expected nocturnal activity in "Diurnal" *Lemur catta* supports cathemerality as one of the key adaptations of the lemurid radiation. *Am J Phys Anthropol.* 150(1):99–106.

Dong D, He G, Zhang S, Zhang Z. 2010. Evolution of olfactory receptor genes in primates dominated by birth-and-death process. *Genome Biol Evol.* 1(0):258–264.

Doran-Sheehy D, Mongo P, Lodwick J, Conklin-Brittain NL. 2009. Male and female western gorilla diet: preferred foods, use of fallback resources, and implications for ape versus old world monkey foraging strategies. *Am J Phys Anthropol.* 140(4):727–738.

Drea CM. 2015. D'scent of man: a comparative survey of primate chemosignaling in relation to sex. *Horm Behav.* 68:117–133.

Farris JS. 1977. Phylogenetic analysis under Dollo's law. *Syst Zool.* 26(1):77–88.

Fleagle JG. 2013. Primate adaptation and evolution. 3rd ed. San Diego (CA): Academic Press.

Garrett EC, Steiper ME. 2014. Strong links between genomic and anatomical diversity in both mammalian olfactory chemosensory systems. *Proc Biol Sci R Soc.* 281(1783):20132828.

Gilad Y, Wiebe V, Przeworski M, Lancet D, Pääbo S. 2004. Loss of olfactory receptor genes coincides with the acquisition of full trichromatic vision in primates. *PLoS Biol.* 2(1):E5.

Glaser N, Kroger RHH. 2017. Variation in rhinarium temperature indicates sensory specializations in placental mammals. *J Thermal Biol.* 67:30–34.

Glusman G, Bahar A, Sharon D, Pilpel Y, White J, Lancet D. 2000. The olfactory receptor gene superfamily: data mining, classification, and nomenclature. *Mamm Genome* 11(11):1016–1023.

Goff SA, Klee HJ. 2006. Plant volatile compounds: sensory cues for health and nutritional value? *Science* 311(5762):815–819.
Hedges SB, Dudley J, Kumar S. 2006. TimeTree: a public knowledge-base of divergence times among organisms. *Bioinformatics* 22(23):2971–2972.
Heritage S. 2014. Modeling olfactory bulb evolution through primate phylogeny. *PLoS One* 9(11):e113904.
Hernandez Salazar LT, Laska M, Rodriguez Luna E. 2003. Olfactory sensitivity for aliphatic esters in spider monkeys (*Ateles geoffroyi*). *Behav Neurosci.* 117(6):1142–1149.
Hiramatsu C, Melin AD, Aureli F, Schaffner CM, Vorobyev M, Kawamura S. 2009. Interplay of olfaction and vision in fruit foraging of spider monkeys. *Anim Behav.* 77(6):1421–1426.
Jacobs GH. 2009. Evolution of colour vision in mammals. *Philos Trans R Soc Lond B Biol Sci.* 364(1531):2957–2967.
Jacobs RL, MacFie TS, Spriggs AN, Baden AL, Morelli TL, Irwin MT, Lawler RR, Pastorini J, Mayor M, Lei R. 2017. Novel opsin gene variation in large-bodied, diurnal lemurs. *Biol Lett.* 13:20170050.
Joffe B, Peichl L, Hendrickson A, Leonhardt H, Solovei I. 2014. Diurnality and nocturnality in primates: an analysis from the rod photoreceptor nuclei perspective. *Evol Biol.* 41:1–11.
Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 30(4):772–780.
Kawamura S, Hiramatsu C, Melin AD, Schaffner CM, Aureli F, Fedigan LM. 2012. Polymorphic color vision in primates: evolutionary considerations. In: Hirai H, Imai H, Go Y, editors. Post-Genome Biology of Primates. Tokyo: Springer. p. 93–120.
Kawamura S, Kubotera N. 2004. Ancestral loss of short wave-sensitive cone visual pigment in lorisiform prosimians, contrasting with its strict conservation in other prosimians. *J Mol Evol.* 58(3):314–321.
Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, et al. 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23(21):2947–2948.
Laska M, Genzel D, Wieser A. 2005. The number of functional olfactory receptor genes and the relative size of olfactory brain structures are poor predictors of olfactory discrimination performance with enantiomers. *Chem Senses* 30(2):171–175.
Laska M, Rivas Bautista RM, Hernandez Salazar LT. 2006. Olfactory sensitivity for aliphatic alcohols and aldehydes in spider monkeys (*Ateles geoffroyi*). *Am J Phys Anthropol.* 129(1):112–120.
Laska M, Seibt A, Weber A. 2000. 'Microsmatic' primates revisited: olfactory sensitivity in the squirrel monkey. *Chem Senses* 25(1):47–53.
Lewis RJ, Kappeler PM. 2005. Seasonality, body condition, and timing of reproduction in *Propithecus verreauxi* verreauxi in the Kirindy Forest. *Am J Primatol.* 67(3):347–364.
Li Y. 2007. Terrestriality and tree stratum use in a group of Sichuan snub-nosed monkeys. *Primates* 48(3):197–207.
MacArthur DG, Balasubramanian S, Frankish A, Huang N, Morris J, Walter K, Jostins L, Habegger L, Pickrell JK, Montgomery SB, et al. 2012. A systematic survey of loss-of-function variants in human protein-coding genes. *Science* 335(6070):823–828.
Martin RD. 1990. Primate origins and evolution: a phylogenetic reconstruction. London: Chapman and Hall.
Mason VC, Li G, Minx P, Schmitz J, Churakov G, Doronina L, Melin AD, Dominy NJ, Lim NT, Springer MS, et al. 2016. Genomic analysis reveals hidden biodiversity within colugos, the sister group to primates. *Sci Adv.* 2(8):e1600633.
Matsuda I, Clauss M, Tuuga A, Sugau J, Hanya G, Yumoto T, Bernard H, Hummel J. 2017. Factors affecting leaf selection by foregut-fermenting proboscis monkeys: new insight from in vitro digestibility and toughness of leaves. *Sci Rep.* 7:42774.
Matsuda I, Tuuga A, Higashi S. 2009. The feeding ecology and activity budget of proboscis monkeys. *Am J Primatol.* 71(6):478–492.
Matsui A, Go Y, Niimura Y. 2010. Degeneration of olfactory receptor gene repertories in primates: no direct link to full trichromatic vision. *Mol Biol Evol.* 27(5):1192–1200.
Matsui A, Rakotondraparany F, Munechika I, Hasegawa M, Horai S. 2009. Molecular phylogeny and evolution of prosimians based on complete sequences of mitochondrial DNAs. *Gene* 441(1–2):53–66.
McGann JP. 2017. Poor human olfaction is a 19th-century myth. *Science* 356(6338):eaam7263.
Melin AD, Fedigan LM, Hiramatsu C, Hiwatashi T, Parr N, Kawamura S. 2009. Fig foraging by dichromatic and trichromatic *Cebus capucinus* in a tropical dry forest. *Int J Primatol.* 30(6):753–775.
Melin AD, Khetpal V, Matsushita Y, Zhou K, Campos FA, Welker B, Kawamura S. 2017. Howler monkey foraging ecology suggests convergent evolution of routine trichromacy as an adaptation for folivory. *Ecol Evol.* 7(5):1421–1434.
Melin AD, Matsushita Y, Moritz GL, Dominy NJ, Kawamura S. 2013. Inferred L/M cone opsin polymorphism of ancestral tarsiers sheds dim light on the origin of anthropoid primates. *Proc Biol Sci R Soc.* 280(1759):20130189.
Melin AD, Moritz GL, Fosbury RA, Kawamura S, Dominy NJ. 2012. Why aye-ayes see blue. *Am J Primatol.* 74(3):185–192.
Melin AD, Wells K, Moritz GL, Kistler L, Orkin JD, Timm RM, Bernard H, Lakim MB, Perry GH, Kawamura S, et al. 2016. Euarchontan opsin variation brings new focus to primate origins. *Mol Biol Evol.* 33(4):1029–1041.
Mori K, Nagao H, Yoshihara Y. 1999. The olfactory bulb: coding and processing of odor molecule information. *Science* 286(5440):711–715.
Moritz GL, Lim NT, Neitz M, Peichl L, Dominy NJ. 2013. Expression and evolution of short wavelength sensitive opsins in colugos: a nocturnal lineage that informs debate on primate origins. *Evol Biol.* 40:542–553.
Moriya-Ito K, Tanaka I, Umitsu Y, Ichikawa M, Tokuno H. 2015. The olfactory bulb and the number of its glomeruli in the common marmoset (*Callithrix jacchus*). *Neurosci Res.* 93:158–163.
Murphy WJ, Eizirik E, Johnson WE, Zhang YP, Ryder OA, O'Brien SJ. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature* 409(6820):614–618.
Nei M, Niimura Y, Nozawa M. 2008. The evolution of animal chemosensory receptor gene repertoires: roles of chance and necessity. *Nat Rev Genet.* 9(12):951–963.
Nevo O, Heymann EW. 2015. Led by the nose: olfaction in primate feeding ecology. *Evol Anthropol.* 24(4):137–148.
Nevo O, Heymann EW, Schulz S, Ayasse M. 2016. Fruit odor as a ripeness signal for seed-dispersing primates? A case study on four neotropical plant species. *J Chem Ecol.* 42(4):323–328.
Nevo O, Orts Garri R, Hernandez Salazar LT, Schulz S, Heymann EW, Ayasse M, Laska M. 2015. Chemical recognition of fruit ripeness in spider monkeys (*Ateles geoffroyi*). *Sci Rep.* 5:14895.
Newman T, Trask BJ. 2003. Complex evolution of 7E olfactory receptor genes in segmental duplications. *Genome Res.* 13(5):781–793.
Ni X, Gebo DL, Dagosto M, Meng J, Tafforeau P, Flynn JJ, Beard KC. 2013. The oldest known primate skeleton and early haplorhine evolution. *Nature* 498(7452):60–64.
Niimura Y. 2012. Olfactory receptor multigene family in vertebrates: from the viewpoint of evolutionary genomics. *Curr Genomics* 13(2):103–114.
Niimura Y. 2013. Identification of olfactory receptor genes from mammalian genome sequences. *Methods Mol Biol.* 1003:39–49.
Niimura Y, Matsui A, Touhara K. 2014. Extreme expansion of the olfactory receptor gene repertoire in African elephants and evolutionary dynamics of orthologous gene groups in 13 placental mammals. *Genome Res.* 24(9):1485–1496.
Niimura Y, Nei M. 2005. Comparative evolutionary analysis of olfactory receptor gene clusters between humans and mice. *Gene* 346:13–21.
Niimura Y, Nei M. 2007. Extensive gains and losses of olfactory receptor genes in mammalian evolution. *PLoS One* 2(8):e708.
Olender T, Waszak SM, Viavant M, Khen M, Ben-Asher E, Reyes A, Nativ N, Wysocki CJ, Ge D, Lancet D. 2012. Personal receptor repertoires: olfaction as a model. *BMC Genomics* 13:414.
Pagel M. 1999. Inferring the historical patterns of biological evolution. *Nature* 401:877–884.

Perelman P, Johnson WE, Roos C, Seuanez HN, Horvath JE, Moreira MA, Kessing B, Pontius J, Roelke M, Rumpler Y, et al. 2011. A molecular phylogeny of living primates. *PLoS Genet.* 7(3):e1001342.

Perry GH, Martin RD, Verrelli BC. 2007. Signatures of functional constraint at aye-aye opsin genes: the potential of adaptive color vision in a nocturnal primate. *Mol Biol Evol.* 24(9):1963–1970.

Pihlstrom H, Fortelius M, Hemila S, Forsman R, Reuter T. 2005. Scaling of mammalian ethmoid bones can predict olfactory organ size and performance. *Proc Biol Sci R Soc.* 272(1566):957–962.

Porter J, Craven B, Khan RM, Chang S-J, Kang I, Judkewitz B, Judkewicz B, Volpe J, Settles G, Sobel N. 2007. Mechanisms of scent-tracking in humans. *Nat Neurosci.* 10(1):27–29.

Rizvanovic A, Amundin M, Laska M. 2013. Olfactory discrimination ability of Asian elephants (*Elephas maximus*) for structurally related odorants. *Chem Senses* 38(2):107–118.

Rushmore J, Leonhardt SD, Drea CM. 2012. Sight or scent: lemur sensory reliance in detecting food quality varies with feeding ecology. *PLoS One* 7(8):e41558.

Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 4(4):406–425.

Schmitz J, Ohme M, Zischler H. 2001. SINE insertions in cladistic analyses and the phylogenetic affiliations of *Tarsius bancanus* to other primates. *Genetics* 157(2):777–784.

Shepherd GM. 2004. The human sense of smell: are we better than we think? *PLoS Biol.* 2(5):E146.

Smith TD, Bhatnagar KP. 2004. Microsmatic primates: reconsidering how and when size matters. *Anat Rec B New Anat.* 279(1):24–31.

Smith TD, Rossie JB, Bhatnagar KP. 2007. Evolution of the nose and nasal skeleton in primates. *Evol Anthropol.* 16:132–146.

Springer MS, Meredith RW, Gatesy J, Emerling CA, Park J, Rabosky DL, Stadler T, Steiner C, Ryder OA, Janečka JE, et al. 2012. Macroevolutionary dynamics and historical biogeography of primate diversification inferred from a species supermatrix. *PLoS One* 7(11):e49521.

Steiger SS, Fidler AE, Kempenaers B. 2009. Evidence for increased olfactory receptor gene repertoire size in two nocturnal bird species with well-developed olfactory ability. *BMC Evol Biol.* 9:117.

Steiger SS, Fidler AE, Valcu M, Kempenaers B. 2008. Avian olfactory receptor gene repertoires: evidence for a well-developed sense of smell in birds? *Proc Biol Sci R Soc.* 275(1649):2309–2317.

Symonds MRE, Blomberg SP. 2014. A primer on phylogenetic generalised least squares. In: Garamszegi LZ, editor. Modern phylogenetic comparative methods and their application in evolutionary biology. Heidelberg: Springer.

Takezaki N, Rzhetsky A, Nei M. 1995. Phylogenetic test of the molecular clock and linearized trees. *Mol Biol Evol.* 12(5):823–833.

Tan Y, Yoder AD, Yamashita N, Li WH. 2005. Evidence from opsin genes rejects nocturnality in ancestral primates. *Proc Natl Acad Sci U S A.* 102(41):14712–14716.

Touhara K, Vosshall LB. 2009. Sensing odorants and pheromones with chemosensory receptors. *Annu Rev Physiol.* 71:307–332.

Valenta K, Burke RJ, Styler SA, Jackson DA, Melin AD, Lehman SM. 2013. Colour and odour drive fruit selection and seed dispersal by mouse lemurs. *Sci Rep.* 3:2424.

Valenta K, Edwards M, Rafaliarison RR, Johnson SE, Holmes SM, Brown KA, Dominy NJ, Lehman SM, Parra EJ, Melin AD. 2016. Visual ecology of true lemurs suggests a cathemeral origin for the primate cone opsin polymorphism. *Funct Ecol.* 30(6):932–942.

Veilleux CC, Bolnick DA. 2009. Opsin gene polymorphism predicts trichromacy in a cathemeral lemur. *Am J Primatol.* 71(1):86–90.

Veilleux CC, Louis EE Jr, Bolnick DA. 2013. Nocturnal light environments influence color vision and signatures of selection on the OPN1SW opsin gene in nocturnal lemurs. *Mol Biol Evol.* 30(6):1420–1437.

Veilleux CC, Scarry CJ, Di Fiore A, Kirk EC, Bolnick DA, Lewis RJ. 2016. Group benefit associated with polymorphic trichromacy in a Malagasy primate (*Propithecus verreauxi*). *Sci Rep.* 6(1):38418.

Wackermannova M, Pinc L, Jebavy L. 2016. Olfactory sensitivity in mammalian species. *Physiol Res.* 65(3):369–390.

Williams BA, Kay RF, Kirk EC. 2010. New perspectives on anthropoid origins. *Proc Natl Acad Sci U S A.* 107(11):4797–4804.

Wyatt TD. 2014. Pheromones and animal behavior: chemical signals and signatures. Cambridge: Cambridge University Press.

Zhou Q, Wei H, Tang H, Huang Z, Krzton A, Huang C. 2014. Niche separation of sympatric macaques, *Macaca assamensis* and *M. mulatta*, in limestone habitats of Nonggang, China. *Primates* 55(1):125–137.

Zhou X, Wang B, Pan Q, Zhang J, Kumar S, Sun X, Liu Z, Pan H, Lin Y, Liu G, et al. 2014. Whole-genome sequencing of the snub-nosed monkey provides insights into folivory and evolutionary history. *Nat Genet.* 46:1303–1310.