biology letters

Mating system and *avpr1a* promoter variation in primates

Lia Rosso, Laurent Keller, Henrik Kaessmann and Robert L Hammond

Biol. Lett. 2008 **4**, 375-378
doi: 10.1098/rsbl.2008.0122

biology letters
Evolutionary biology

Biol. Lett. (2008) **4**, 375–378
doi:10.1098/rsbl.2008.0122
Published online 22 April 2008

Mating system and *avpr1a* promoter variation in primates

Lia Rosso[1], Laurent Keller[2], Henrik Kaessmann[1] and Robert L. Hammond[2,*]

[1]*Center for Integrative Genomics, University of Lausanne, Génopode, 1015 Lausanne, Switzerland*
[2]*Department of Ecology and Evolution, University of Lausanne, Biophore, 1015 Lausanne, Switzerland*
**Author and address for correspondence: Department of Biological Sciences, University of Hull, Hull HU6 7RX, UK (r.hammond@hull.ac.uk).*

It has been suggested that primate mating and social behaviours may be influenced by variation in promoter region repetitive DNA of the vasopressin receptor 1a gene (*avpr1a*). We show that male mating behaviour does not covary in a simple way with promoter repetitive DNA in 12 Old World primates. We found that one microsatellite (−553 bp upstream) was present in all species, irrespective of their behaviour. By contrast, two microsatellites (−3956 and −3625 bp upstream) were present only in some species, yet this variation did not correlate with behaviour. These findings agree with a recent comparative analysis of voles and show that the variation in repetitive DNA in the *avpr1a* promoter region does not generally explain variation in male mating behaviour. Phylogenetic analysis revealed a GAGTA motif that has been independently deleted three times and involved in another larger deletion. Importantly, the presence/absence of this GAGTA motif leads to changes in predicted transcription factor-binding sites. Given the repeated loss of this motif, we speculate that it might be of functional relevance. We suggest that such non-repetitive variation, either in indels or in sequence variation, are likely to be important in explaining interspecific variation in *avpr1a* expression.

Keywords: primates; mating behaviour; microsatellites; vasopressin; *avpr1a*; voles

1. INTRODUCTION

There is great interest in how much variation in social and mating behaviours is caused by genetic variation (Robinson *et al.* 2005), particularly by genes of large effect. Possibly, the best example is the vasopressin receptor 1a gene (*avpr1a*), which codes for the receptor (AVPR1A) of the nonapeptide arginine vasopressin. This gene has been shown in a series of elegant studies to have a major role explaining both inter- and intraspecific differences in the social and mating behaviours of voles (Young *et al.* 1997, 1999; Lim *et al.* 2004; Hammock & Young 2005). Data suggest that differences in receptor distribution and behaviour are caused by variation in repetitive DNA approximately 500 bp upstream of the *avpr1a* transcription start site (Hammock & Young 2005). In monogamous 'social' vole species (prairie and pine voles), where males pair bond with their mating partners, there are larger blocks of repetitive DNA upstream of the transcription start site compared with polygamous 'asocial' vole species (e.g. montane and meadow voles), where males do not pair bond.

Hammock & Young (2005) recently suggested that, as in voles, differences in repetitive non-coding DNA may explain behavioural differences among primate species. This suggestion was motivated by their discovery that humans and bonobos (*Pan paniscus*) differ from chimpanzees (*Pan troglodytes*) in the number of microsatellites in a non-coding region upstream of *avpr1a*. Humans and bonobos have two microsatellites (−3956 and −3625 bp upstream of the start site; Thibonnier *et al.* 2000), whereas a 360 bp region, including the −3625 bp microsatellite, is absent in the chimpanzee (*P. troglodytes*). However, testing this idea needs data from more species with well-defined differences in social and mating behaviours and known phylogenetic relationships. Furthermore, the −3956 and −3625 bp microsatellites are both dinucleotide repeats and are much further upstream from the transcription start site than the tetranucleotide repeat microsatellite that influences mating and social behaviours in voles (Fink *et al.* 2006). There is a third microsatellite, not considered by Hammock and Young, 553 bp upstream of *avpr1a*, which has the same repeat type (tetranucleotide) and is in a similar position to the vole microsatellite (Thibonnier *et al.* 2000). *A priori*, the −553 bp microsatellite is perhaps a better candidate than either the −3956 or the −3625 bp microsatellite for influencing male behaviour.

We have expanded upon Hammock & Young's speculations by comparing the structure of all three microsatellites (−553, −3625 and −3956 bp) in a sample of 12 primates, comprising humans, five great apes (bonobo, two subspecies of chimpanzee, gorilla and orang-utan), three species of gibbon (lar, siamang and crested) and three Old World monkeys (green monkey, hamadryas baboon and rhesus macaque). Importantly, these species show great variation in the mating behaviour of males, particularly the duration and exclusivity of male–female post-mating affiliations, and the main topology of the phylogeny is known without ambiguity (except gibbon relationships). Male humans and gibbons are capable of forming long-term pair bonds with female mates, whereas male chimpanzees, bonobos, macaques and green monkeys show only weak non-exclusive associations with individual female mates. Male hamadryas baboons and gorillas have strong, but non-exclusive, associations with female mates while orang-utans form brief consortships (see electronic supplementary material for support for these classifications). Given this variation in social behaviour, we asked the following questions:

(i) does variation in any of the microsatellites upstream of *avpr1a* covary with social behaviour in a predictable way? For example, do pair-bonding gibbons and humans have larger regions of repetitive DNA than species that do not pair bond? and
(ii) can we identify, using a phylogenetic approach, regions other than the microsatellites that may be functionally important?

Electronic supplementary material is available at http://dx.doi.org/10.1098/rsbl.2008.0122 or via http://journals.royalsociety.org.

Received 4 March 2008
Accepted 2 April 2008

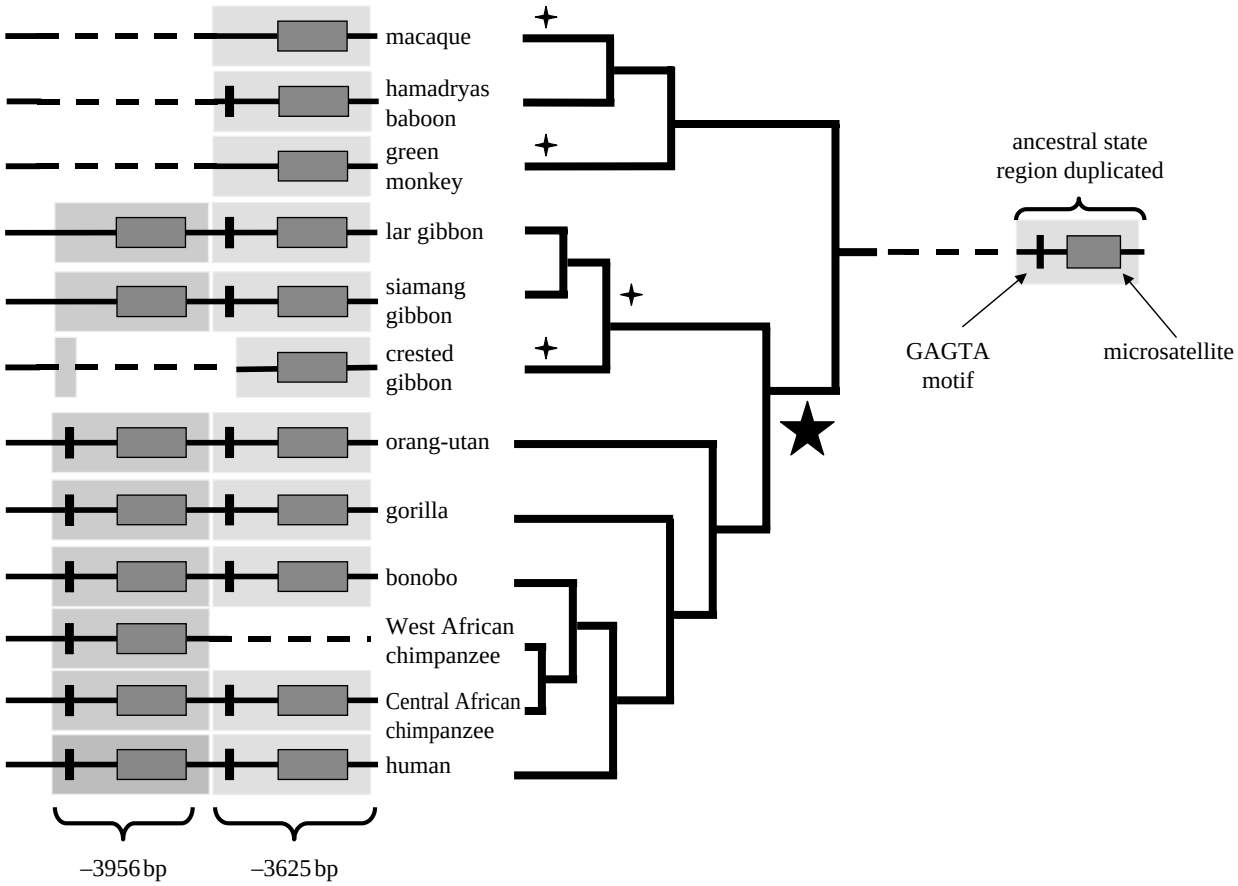

Figure 1. *avpr1a* dinucleotide microsatellite structure mapped onto the phylogeny for 12 Old World primates. Shaded rectangles labelled −3956 and −3625 bp show duplicated regions containing microsatellites (dark-shaded rectangles) and the GAGTA motif (black vertical bars). Broken lines indicate sequence absences either owing to deletions or because they are basal to the −3625 bp duplication event. Five pointed star shows position of the −3625 bp duplication that gave rise to the −3956 bp region, four pointed stars show losses of GAGTA motif.

2. MATERIAL AND METHODS

(a) *Sequences*

From a sample of 12 Old World primate taxa (see electronic supplementary material), we amplified and sequenced two regions upstream of the *avpr1a* transcription start site that contained up to three repetitive DNA elements. We amplified and sequenced the 'dinucleotide region' (including −3956 and −3625 bp dinucleotide microsatellites, Thibonnier *et al.* 2000) using primers (5′ ClaI site removed) described in Hammock & Young (2005), and the 'tetranucleotide region' (−553 bp tetranucleotide microsatellite) using primers from Kim *et al.* (2002). PCR products were cloned into pGEM Easy T (Promega) plasmids. We sequenced three to six clones per individual in both directions.

(b) *Alignments*

We arbitrarily chose the allele with the longest repeat at the −3625 bp microsatellite (except in the West African chimpanzee where we chose the longest −3956 bp microsatellite allele) and we aligned all sequences by eye.

(c) *Identification of transcription factor-binding sites*

Mapping aligned sequences on to the phylogeny identified a number of insertion/deletion events that may have caused changes in transcription factor-binding sites. We identified potential transcription factor-binding sites using the program ALIBABA v. 2.1 (Grabe 2000, http://www.gene-regulation.com/pub/programs/alibaba2/index.htm) to search the Transfac v. 4.0 database (Wingender *et al.* 2000).

3. RESULTS

We found considerable variation among species in the −3956 and −3625 bp repeat regions (figure S1 in the electronic supplementary material). The three most basal species in the phylogeny (hamadryas baboon, macaque and green monkey) had the −3625 bp repeat region, but lacked the −3956 bp region (figure 1). In seven out of nine ape taxa, both the −3956 and −3625 bp microsatellites were present, but the −3956 bp repeat region was missing in the black-crested gibbon and the −3625 bp repeat region was missing in the West African chimpanzee (Hammock & Young 2005; figure 1). Interestingly, both the −3956 and −3625 bp microsatellites were present in chimpanzees from Central Africa, showing that chimpanzees are polymorphic for the deletion of the −3625 bp microsatellite. This pattern of change, mapped onto the phylogeny, suggests that the −3956 bp microsatellite arose by tandem duplication of the −3625 bp region in the ancestor to all apes (figure 1). Subsequently, there have been two independent losses of a repeat region in the black-crested gibbon and in West African chimpanzees. In contrast to the evolutionary lability of the −3956 and −3625 bp repeat regions, the −553 bp tetranucleotide microsatellite was present in all 12 taxa (figure S2 in the electronic supplementary material).

Mapping indels onto the phylogeny suggested that non-repetitive regions of the *avpr1a* promoter may be functionally important. For example, before the −3956 and −3625 bp microsatellites, there is a GAGTA motif that is either present or absent (figure S1 in the electronic supplementary material: positions 94–98 and 445–449). All great apes, except the western

chimpanzee, have two GAGTA motifs, a pattern that suggests that the ancestral state was hamadryas-like. Given this, there have been at least three independent losses of the GAGTA motif: in macaques, in green monkeys, and in the ancestor to gibbons. In contrast to these repeated losses, the GAGTA motif is embedded in a 35–40 bp block of sequence (figure S1 in the electronic supplementary material: positions 76–110 and 427–466) that is conserved, with only single substitutions separating all gibbons (figure S1 in the electronic supplementary material: −3956 bp region: position 79), or the siamang (figure S1 in the electronic supplementary material: −3625 bp region: position 454), from all other taxa.

4. DISCUSSION

The variation in the presence/absence of dinucleotide repeat regions (−3956 and −3625 bp) does not covary in a simple way with male behaviour (figure 1). For example, differences in repeat structure between chimpanzee subspecies and among gibbons do not correspond to large differences in male behaviour. Likewise, species with dissimilar male behaviours (e.g. orang-utans, bonobos and lar/symphalangus gibbons) have the same repeat region structure. We also found no evidence that variation in repetitive DNA approximately 500 bp upstream of the transcription start site explained variation in male mating behaviour, as the −553 bp tetranucleotide repeat region was present in all taxa irrespective of their behaviour (figure S2 in the electronic supplementary material). Furthermore, our prediction that monogamous gibbons would have more repetitive DNA than non-monogamous taxa was not supported by our data. It might be argued that all primates studied here are capable of forming strong social bonds in contrast to voles where males are either able (e.g. prairie and pine voles) or unable (e.g. montane and meadow voles) to do so. This re-categorization of the behaviour of the primates still leads to the conclusion that variation in repetitive DNA in the −3956 and −3625 bp microsatellite regions does not covary simply with behaviour, as the lack of variation in behaviour (all primates can form social bonds) would be at odds with the high variation in promoter repetitive DNA we found. We suggest, therefore, that our findings agree with a recent study that found no general relationship between *avpr1a* repeat structure and male mating behaviour across a large number of *Microtus* voles (Fink *et al.* 2006).

Although our results, and those of Fink *et al.* (2006), clearly show no association of repeat structure and behaviour, it is important to stress that neither our study, nor that of Fink *et al.* (2006), tested whether AVPR1A receptor distribution in the brains of rodents and primates covaries with behaviour. This is still a viable hypothesis that needs to be tested in both voles and primates (Young & Hammock 2007). It is, therefore, premature to conclude that there is no general relationship between the gene *avpr1a* and male behaviour.

One possibility is that non-repetitive, as well as repetitive, regions in the promoter influence the expression of *avpr1a* (Hammock & Young 2002). For example, our phylogenetic analysis showed the repeated loss of the same GAGTA motif and the possible role of the GAGTA repeat in the large deletion in the black-crested gibbon (the −3956 bp repeat region deletion spans the two GAGTA motifs). Although repeated losses might be interpreted as a lack of conservation, the repeated losses of the GAGTA motif contrasts with the conservation of the surrounding 40 bp, suggesting that this region might have some functional role in the expression of *avpr1a*. In support of this notion, searches of the Transfac database (Wingender *et al.* 2000) with the program ALIBABA v. 2.1 (Grabe 2000) show that the loss of the GAGTA motif alters predicted transcription factor-binding sites, for example, changing *HNF-3b* to *MEF2* (figure S3 in the electronic supplementary material), with the latter being important in neuronal development and survival (Shalizi & Bonni 2005). Additionally, there are other indels and single nucleotide variation (see electronic supplementary material) that may be important in regulating cell-type-dependent expression of *avpr1a*.

In conclusion, it appears that there is no simple relationship between *avpr1a* promoter region repetitive DNA and male mating behaviour. This does not mean, however, that variation in expression of *avpr1a* is necessarily unimportant in explaining interspecific differences in mammalian mating/social behaviours, but that the regulation of expression is more complex than previously thought (Young & Hammock 2007).

We thank C. Roos and S. Pääbo for primate DNA samples, the Swiss National Science Foundation (funds to H.K. and L.K.) for financial support and two anonymous reviewers for their comments.

Fink, S., Excoffier, L. & Heckel, D. 2006 Mammalian monogamy is not controlled by a single gene. *Proc. Natl Acad. Sci. USA* **103**, 10 956–10 960. (doi:10.1073/pnas.0602380103)

Grabe, N. 2000 AliBaba2: context specific identification of transcription factor binding sites. *In Silico Biol.* **2**, S1–S15.

Hammock, E. A. D. & Young, L. J. 2002 Variation in the vasopressin V1a receptor promoter and expression: implications for inter- and intraspecific variation in social behaviour. *Eur. J. Neurosci.* **16**, 399–402. (doi:10.1046/j.1460-9568.2002.02083.x)

Hammock, E. A. D. & Young, L. J. 2005 Microsatellite instability generates diversity in brain and sociobehavioral traits. *Science* **308**, 1630–1634. (doi:10.1126/science.1111427)

Kim, S. J. *et al.* 2002 Transmission disequilibrium testing of arginine vasopressin receptor 1A (AVPR1A) polymorphisms in autism. *Mol. Psychiat.* **7**, 503–507. (doi:10.1038/sj.mp.4001125)

Lim, M. M., Wang, Z., Olazábal, D. E., Xianghui, R., Terwilliger, E. F. & Young, L. J. 2004 Enhanced partner preference in a promiscuous species by manipulating the expression of a single gene. *Nature* **429**, 754–757. (doi:10.1038/nature02539)

Robinson, G. E., Grozinger, C. M. & Whitfield, J. 2005 Sociogenomics: social life in molecular terms. *Nat. Rev. Genet.* **6**, 257–270. (doi:10.1038/nrg1575)

Shalizi, A. K. & Bonni, A. 2005 Brawn for brains: the role of MEF2 proteins in the developing nervous system. *Curr. Top. Dev. Biol.* **69**, 239–266. (doi:10.1016/S0070-2153(05)69009-6)

Thibonnier, M., Graves, M. K., Wagner, M. S., Chatelain, N., Soubrier, F., Corvol, P., Willard, H. F. & Jeunemaitre, X. 2000 Study of V_1-vascular vasopressin receptor gene microsatellite polymorphisms in human essential hypertension. *J. Mol. Cell. Cardiol.* **32**, 557–564. (doi:10.1006/jmcc.2000.1108)

Wingender, E. *et al.* 2000 TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Res.* **28**, 316–319. (doi:10.1093/nar/28.1.316)

Young, L. J. & Hammock, E. A. D. 2007 On switches and knobs, microsatellites and monogamy. *Trends Genet.* **23**, 209–212. (doi:10.1016/j.tig.2007.02.010)

Young, L. J., Winslow, J. T., Nilsen, R. & Insel, T. R. 1997 Species differences in V(1)a receptor gene expression in monogamous and nonmonogamous voles: behavioral consequences. *Behav. Neurosci.* **111**, 599–605. (doi:10.1037/0735-7044.111.3.599)

Young, L. J., Nilsen, R., Waymire, K. G., MacGregor, G. R. & Insel, T. R. 1999 Increased affiliative response to vasopressin in mice expressing the V_{1a} receptor from a monogamous vole. *Nature* **400**, 766–768. (doi:10.1038/23650)