Title

Social and demographic correlates of male androgen levels in wild white-faced capuchin monkeys (Cebus capucinus)

Journal

American Journal of Primatology, 79(7)

ISSN

0275-2565

Authors

Schaebs, Franka S
Perry, Susan E
Cohen, Don
et al.

Publication Date

2017-07-01

DOI

10.1002/ajp.22653

Peer reviewed

Definition and determination of predictor variables

In the following, we describe how we operationalized several of the key predictors in our models, for which the outcome variable was the log of the immunoreactive fecal androgen (ifA) level. Data processing steps required for this were conducted in R [R Core Team, 2015], version 3.2.3, and the code was written by authors R. Mundry and F.S. Schaebs.

Co-residence score

The purpose of this score was to measure how familiar a particular male (the one whose ifA levels are being measured) was with the other male members of his group. The idea is that unfamiliar males can pose threats to males, and that ifA levels are predicted to be high during the phase of relationship formation when males are working out the nature of their relationships with one another and possibly competing physically to establish their relative ranks. We expected the effect to decay approximately exponentially and estimated two weeks as a plausible half life. We also assumed that by being in the same group on the same day (i.e., by being co-resident), monkeys become more familiar with one another.

Individual male capuchins tend to stay in the same group for extended periods of time. When they move between groups, they often tend to first move back and forth between two or more different social groups, spending some time either alone or in loose male bands, and finally deciding to stay with the initial group or migrate into another social group. Group residence by male capuchin monkeys can thus vary from very stable, remaining many years within the same group, to highly unstable, in which individuals are seen within a group for only a few days at a time. This difference in group membership stability reflects both male immigrations and visits by

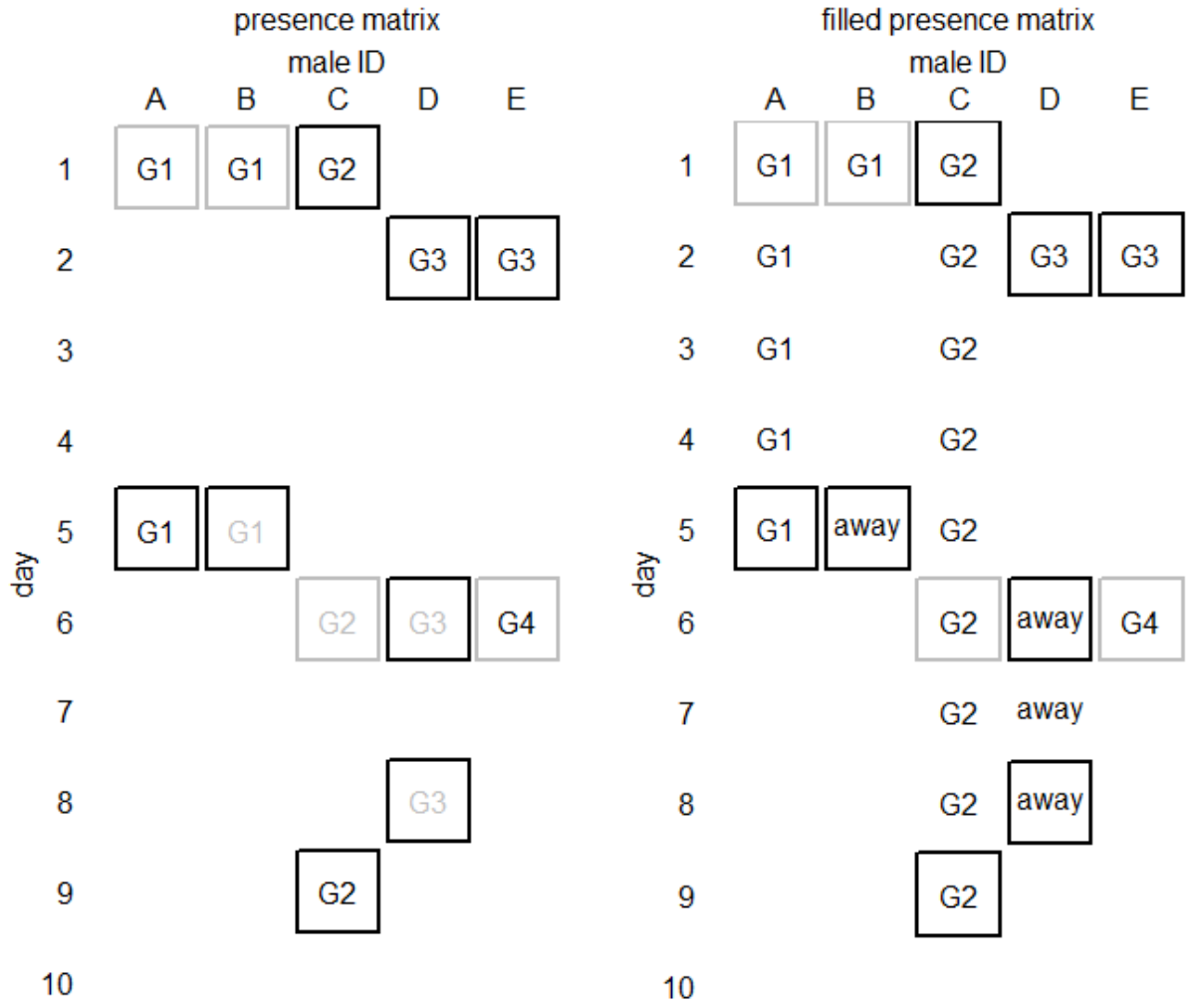
new males, and group membership is more stable the longer the time interval since the last immigration or visiting event.

In order to estimate familiarity we used census data to create a calendar which lists, for each day and each monkey, all of the groups he was seen in, as well as records of him being seen alone or with particular other migrating males. We used census data starting in January 2004, i.e., 32 months prior to the onset of hormonal sampling, to ensure that we were accurately assessing the degree of familiarity among males at the time ifA levels were measured. It was possible, though rare, for a monkey to be seen in more than one group on the same day and in this case an arbitrary choice was made, which resulted in about 0.25% of all observations during a census being discarded. Most groups were not observed on any given day, and the step above left many blank dates on the calendar. However, monkeys tended to stay in the same group for extended periods of time. In the census data for the interval 2004-2011, 96.6% (54600 out of 56529) of the entries for males who had been seen before were seen in the same group as the previous time. Therefore, a reasonable approximation was to fill in blank dates with the group where the monkey was last seen, given its next sighting occurred in the same respective group.

However, we also made use of evidence that a monkey was not in a group on a given day. This was based on "observation time" data, a dataset indicating the duration of observation per group and day (at most one time for a given group on a given day; see Supplementary table SI 1 for median observation times per social group and day). If a group was observed for at least 6 hours on a given day without seeing a monkey who was last observed in that group, we put "away" (i.e., definitely not in that group) in the calendar for that monkey on that day. The intuition, though approximate, was that if the group was observed for 6 hours without seeing that

monkey then he was probably not there. If the group was observed for less than 6 hours we assumed that the male might have been there but just not seen.

Finally, we used the strong tendency for monkeys to stay in the same groups, but used it both forward and backward in time: if a day on the calendar was blank but the previous non-blank day contained the same group (including "away") as the following non-blank day, we filled in that blank with the data from the surrounding non-blank days. Note the gap could be more than one day, e.g., if a monkey were in group G on day 1 and day 10, and days 2-9 were blank, then for each of days 2-9 the previous filled in day would be day 1 and the following filled in day would be day 10, so all 8 of the intervening days would be assigned group G. The days still remaining blank were those for which the previous non-blank entry differed from the subsequent non-blank entry for the monkey in question. SI figure 1 demonstrates these rules pictorially. With these rules we were, to our knowledge, able to cover all possible constellations of gaps in the census data and observation times.



58

59*<u>Supplementary figure SI 1:</u> Illustration of the process of producing the presence matrix and the*
60*filled presence matrix according to the rules explained in the text. A group name (G1, G2 etc.)*
61*printed in black text means the male was seen in that group on that day. A grey text group name*
62*means the male was not seen that day in that group. A black-bordered box means that the group*
63*was observed >6 hours; a grey border of a box means that the group was observed <6 hours.*

64

65

Next we computed a daily "unfamiliarity" score for each pair of monkeys (dyad) based on the assumption that being in the same group on the same day makes them more familiar with each other, and that spending time away from one another reduces familiarity. The unfamiliarity ranges from 1 (totally unfamiliar) to close to zero for highly familiar. Modeling this as a continuous score (ranging from 0 to 1) rather than a binary (0/1) score enabled us to allow the males to slowly get (un-) accustomed to one other. Each dyad was assigned a measure of 1 on the first day they were seen together. Recall that we started computing familiarity 32 months before the first fecal samples were collected so as to have accurate estimates of familiarity by the time of the first ifA measurements. On each subsequent day, the measure was updated according to the following rules:

- If the two males were in the same group (not including “away” or “blank”) then the score was multiplied by 0.95, decreasing the score. This means that a dyad reached a score of ca. 0.05 when continuously staying together for ca. 60 days.

- If the two were in different groups (not including “away” or “blank”), the score was divided by 0.95, increasing the score.

- If one male was in a group (not “away” or “blank”) and the other was “away”, they were treated as if in two different groups.

- Otherwise (i.e., when both were “away” or at least one was “blank”), the score was left as it was on the previous day.

This way of assigning the score could lead to values larger than one, and also values being essentially zero. Therefore, after determining the score for a given day, it was set to one when it was larger than one, and to 0.001 (a value reached after ca. 135 consecutive days of being in the same group) when it was smaller than 0.001. One reason for not letting it go any

lower was that our intuition told us that there should be no real difference between 20 weeks and more than 20 weeks of co-residence. More importantly, we wanted an absence of more than 20 weeks to restore the initial unfamiliarity, regardless of how long the pair was together before that.

Finally, the dyadic unfamiliarity scores were used to compute the dyadic co-residence scores as follows: The co-residence score for a sample taken from a given monkey, M, on a given day, D, was computed by first finding all the males over 7 years of age other than M in the same group as M on day D. The average of their unfamiliarity scores with M on day D was subtracted from 1 to get the co-residence score for the sample. If, for example a male co-resided in a group with two other males with which it had dyadic unfamiliarity scores of 0.3 and 0.4, respectively, their average would be 0.35 and the resulting co-residence score would be 0.65. Note that this average would be undefined if there were no other males over 7 years of age in the same group as M on day D. Such cases were excluded from the data used in this study.

This way of calculating the co-residence score was chosen to model expected impacts of changes in the social stability of a group in a biologically meaningful way, and higher scores indicate longer co-residence, i.e., greater social stability. We only included dyads with males that were at least seven years of age, assuming that this is the youngest age males become serious contenders for breeding positions; our demographic data show that the youngest male ever to become alpha and breed as an alpha was 7.7 years old, and the youngest male to breed as a subordinate male was 6.25 years old.

Note that the co-residence score was z-transformed before use in the models. However, this is a linear transformation, so it only affects the units of the estimates in Table 1. Instead of a difference of 1 between minimum and maximum scores we now had a difference of about 6 for

the co-residence score (and 4 for the lowest co-residence score, which is described in the next section).

Lowest co-residence score in the group

The lowest co-residence score is similar to the average score described above, but it represents a different hypothesis of how the presence of unfamiliar males might affect ifA levels. This score is simply the lowest dyadic co-residence score between any two males > age 7 years in the same group as the sample donor on the day of the sample. The lowest co-residence score and the individual (average) co-residence score were highly correlated (Pearson's r=0.79; N=577), but variance inflation factors indicated no severe issues with collinearity (see the section headed ´Further details about implementation of the statistical analysis´ below).

Number of males

The number of males was calculated as the sum of subadult and adult males that were present in a group on a given day using the filled presence matrix. We only used males older than seven years, because males younger than this are not serious competitors for breeding opportunities and hence we did not expect them to have an impact on male androgen levels. Using the filled presence matrix was justified, as there were so few long observation gaps, i.e., consecutive censuses during which a male was seen in two different groups (for the individual co-residence, 94% of the gap durations were zero days, 96% were ≤ 1 day, and 98% were ≤10 days; for the lowest co-residence 92% of the gap durations were zero days, 94% were ≤1 and 97% were ≤10 days). Furthermore, the filled presence matrix was considered better than raw census data because it represents our best estimate of who was in each group on each day. All the samples were collected on

days when the group containing the donor was under observation, of course, and for these groups on these days the only difference between census data and the filled presence matrix is that if the group was observed for less than 6 hours, the filled presence matrix (but not census data) included monkeys who were not seen but still assumed to be present.

Detailed information on the median number of observation hours per group and day, and on the number of days each social group was observed between January 2004 and December 2011, are given in Supplementary table SI 1.

Number of potentially fertile females

The presence of potentially fertile females is expected to have a large effect on adult male androgen levels. However, since we did not have female hormonal data to determine female fertile periods, we designed a proxy. We considered a female to be potentially fertile (i.e., having regular ovarian cycles) for a period of three months prior to her conception date. Conception dates were inferred by counting 159 days (i.e., the estimated gestation length [Nagle & Denari, 1983; Carnegie et al. 2011]) back from the birth of the last infant for each female group member. Out of 146 births that happened during our study period, 28 were known with the exact date, 102 birth dates had an uncertainty of not more than one week, 143 births had an uncertainty of not more than one month, and no birth date was less precise than three months. This added some imprecision to the estimated conception dates, but we considered these imprecisions to be inherent limits of field data. The number of potentially fertile females for a given sample was the number of females considered potentially fertile on the day the sample was collected and in the same group as the donor of the sample on that day.

Alpha male tenure duration

Social instability is expected to correlate positively with ifA because it indicates competition among males for rank, which in turn is very highly correlated with reproductive success. Change in the alpha male position is the most prominent example of social instability in a capuchin social group. In addition to clearly indicating instability on the day of the change, further such changes are much more likely when there was a recent change than when there was not, which justifies expecting ifA levels to be high in response to a recent turnover, and gradually decrease with time. (Of the 93 alpha tenures in our records that have ended, 32 lasted less than 10 days (34%), and 55 (59%) less than 30 days. Of the remaining 38, 16 lasted between 30 days and one year, 10 between one and three years, and 12 longer than 3 years, up to 16 years.)

To account for potential instabilities caused by recent changes in the alpha male rank position, we calculated the time interval since the last change in the alpha position of the group. This value was then log transformed because the expected impact on androgen levels of an alpha turnover was expected to decrease with time, e.g., the difference between the beginning and end of the second week is larger than the difference between the beginning and end of the third week.

Threat of Infanticide

Direct observations of infanticidal events are rare in the wild, which makes it hard to assess their impact on capuchin male androgen levels. However, as we expected that the threat of infanticide leads to an increase in male androgen levels, we included the number of dependent offspring (i.e., those being most vulnerable) present in the group as a proxy for the threat of infanticide. Specifically, to assess the maximum age at which an infant would be likely to be an infanticidal victim, we used empirical data on 13 cases of infanticidal events that occurred at Lomas

Barbudal between 1999 and 2013. From these, we know that the oldest infant that was victim of infanticidal killing was about 1 year of age, and that younger infants are more susceptible to infanticidal killing (69% of infanticides happened within the first 100 days of life, representing nine of the 13 reported cases). Therefore, to assess the impact of dependent offspring on male androgen levels, we used the sum of the inverse of the ages (using the inverse to put more emphasis on younger infants, which would be even more susceptible to be victim of infanticidal killing) of the infants which were younger than 1 year (precisely, $\sum(1/(1+\text{age in days}))$, as a proxy for the threat of infanticide.

Further details about implementation of the statistical analysis:

To achieve normally distributed and homogeneous residuals (checked by visual inspection of a qqplot and residuals plotted against fitted values) we log transformed immunoreactive fecal androgen (ifA) levels. Under the assumption that there is one seasonal peak per year for ifA levels, season was included in the models by coding Julian date as a circular variable. To do so, day of sampling was turned into radians by dividing it by the average duration of a year (365.25 days) and then multiplying it by $2*\pi$ and finally including the sine and cosine of the resulting variable into the model. All fixed effects, except dominance status and season were z-transformed to a mean of zero and a standard deviation of one [Schielzeth, 2010]. The models were fitted using the function lmer of the R-package lme4 [Bates et al., 2015] using Maximum Likelihood. We established model stability by excluding levels of the random effects one at a time and comparing the estimates derived for the fixed effects with those derived based on all data. This revealed no influential random effects levels to exist. To check whether collinearity was an issue we determined Variance Inflation Factors (VIF, [Field, 2009]) for a standard linear

model excluding the random effects and the interactions. This revealed collinearity to be no issue (maximum VIF: 3.33). VIF values were calculated using the function vif of the R-package car [Fox & Weisberg, 2011].To establish the significance of the full model as compared to the null model [Forstmeier & Schielzeth, 2011] we used a likelihood ratio test [Dobson, 2002]. The significance of the individual effects was determined using likelihood ratio tests [Barr et al., 2013]. In case the full null model comparison revealed significance and an interaction included in the full model did not, we removed it to enable easier interpretation of the respective main effects [Schielzeth, 2010].

Supplementary table SI 1: Median daily observation times per group and total observation days per group between January 2004 and December 2011, ca. 2914 days. Indicated in bold are the social groups that were included in the hormonal study and for which all the predictor variables defined were calculated. The other social groups were necessary to consider as males immigrated into or emigrated from them.

social group	median daily observation time (h)	total observation days
GRR	**12.5**	**1210**
GAA	**12.5**	**1217**
GFF	**12.5**	**1154**
GFL	**12**	**565**
GRF	**12**	**359**
GPR	11.4	20
GLB	10.125	150
GMK	**10.1**	**451**
GNM	**8.33**	**289**
GCU	**7.75**	**190**
GSP	**7**	**300**
GSO	2	27
GAO	2	3
GBD	1.83	15
GBL	1.25	1
GSR	0.5	6
GCN	0.6	5
GES	0.5	6
Other unhabituated groups	0.635	38

Supplementary table SI 2: Summary of interaction terms for statistical analysis. The last row indicates whether the interaction was included into the respective model or not.

Interaction	**prediction tested**	**included**
number of males * number of potentially fertile females	Male ifA levels increase with the number of males more steeply when more potentially fertile females are present	yes
dominance status * infanticide susceptibility	ifA levels of alpha and subordinate males are affected differently by the risk of infanticide, with alpha males being more affected, because they are siring the majority of offspring	yes
dominance status * lowest co-residence score in the group	ifA levels of alpha and subordinate males are differentially affected by incoming males, with alpha males showing a stronger reaction as their breeding position is challenged directly	yes
dominance status * log of alpha male tenure duration	ifA levels of alpha and subordinate males change differentially over the course of an alpha male's tenure duration. While alpha males are expected to show decreasing ifA levels, subordinate males are expected to show increasing ifA levels with longer alpha male tenure duration. This is expected as subordinates are likely to have more breeding opportunities as soon as the alpha males´ female offspring mature.	yes
infanticide susceptibility * lowest co-residence score in the group	ifA levels of males are affected differently by incoming males, depending on how susceptible the	no

	respective group is to the risk of infanticide. Specifically, incoming males are expected to have a higher impact on male androgen levels when younger offspring are present	

237

Supplementary table SI 3: Summary of results for the full model of the impact of social and demographic factors on immunoreactive fecal androgen levels in adult males.

		Estimate	SE	LRT	Df	P	min[3]	max [3]
Intercept	Intercept	8.691	0.130	[1]	[1]	[1]	8.639	8.814
Test predictors	# potentially fertile females	0.070	0.060	[1]	[1]	[1]	0.056	0.189
	# males	0.256	0.105	[1]	[1]	[1]	0.208	0.319
	Dominance status [2]	-1.680	0.189	[1]	[1]	[1]	-1.816	-1.463
	Infanticide susceptibility	0.055	0.084	[1]	[1]	[1]	0.011	0.085
	Lowest co-residence in the group	0.110	0.113	[1]	[1]	[1]	0.065	0.174
	Log α tenure duration	-0.083	0.091	[1]	[1]	[1]	-0.186	-0.019
	Individual co-residence score	-0.014	0.099	0.019	1	0.89	-0.072	0.030
	# potentially fertile females * # males	0.029	0.052	0.186	1	0.67	-0.052	0.054
	Dominance status * Infanticide susceptibility	-0.094	0.120	0.610	1	0.44	-0.167	-0.010
	Dominance status * Lowest co-residence in the group	-0.054	0.125	0.172	1	0.68	-0.125	0.043
	Dominance status * Log α tenure duration	-0.174	0.216	0.601	1	0.44	-0.276	0.071
Control predictors	Time at sampling	-0.035	0.046	0.584	1	0.45	-0.053	-0.017
	Individual age	0.090	0.080	0.729	1	0.39	0.017	0.213
	Sine (season)	0.639	0.085	15.975	1	0.000	0.560	0.708
	Cosine (season)	0.276	0.075	10.385	1	0.001	0.239	0.298

[1]: not indicated because P-values of predictors which are included in interactions have a very limited interpretation [Schielzeth, 2010]

[2]: alpha=0, subordinate=1

[3]: minimum and maximum of model estimates obtained from subsets of the data, derived by excluding levels of random effects, group ID and male ID, one at a time

References

Barr DJ, Levy R, Scheepers C, Tily HJ. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language 68:255–278.

Bates D, Meachler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software, 67 (1): 1-48.

Carnegie SD, Fedigan LM, Melin AD. 2011. Reproductive seasonality in female capuchins (*Cebus capucinus*) in Santa Rosa (Area de Conservación Guanacaste), Costa Rica. International Journal of Primatology 32:1076–1090.

Dobson AJ. 2002. An Introduction to generalized linear models. Boca Raton: Chapman & Hall/CRC. Available from: http://www.sci-technetbase.com/ejournals/books/book_km.asp?id=635

Field AP. 2009. Discovering statistics using SPSS: (and sex, drugs and rock “n” roll). 3rd ed. Los Angeles: SAGE Publications. 821 p.

Forstmeier W, Schielzeth H. 2011. Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner’s curse. Behavioral Ecology and Sociobiology 65:47–55.

Fox J, Weisberg S. 2011. An R companion to applied regression. 2nd ed. Thousand Oaks, Calif: SAGE Publications. 449 p.

268Nagle CA, Denari JH. 1983. The cebus monkey (*Cebus apella*). In: Hearn JP, editor.
269 Reproduction in new world primates. Dordrecht: Springer Netherlands. p 39–67.
270 Available from: http://www.springerlink.com/index/10.1007/978-94-009-7322-0_2

271R Core Team. 2015. R: A language and environment for statistical computing. Vienna, Austria: R
272 Foundation for Statistical Computing. Available from: http://www.R-project.org

273Schielzeth H. 2010. Simple means to improve the interpretability of regression coefficients:
274 Interpretation of regression coefficients. Methods in Ecology and Evolution 1:103–113.