Supplemental Information:

Chimpanzee culture extends beyond matrilineal family units

Edwin J. C. van Leeuwen, Roger Mundry, Katherine A. Cronin, Mark Bodamer, and Daniel B. M. Haun

Supplemental Methods

Study years

In our original study [S1], we incorporated data from 2 years (2010 and 2011). Here, we added another study year (2007) to increase our power to identify any matrilineal effect on palm-to-palm clasping (PPC) tendencies. The data from 2007 were collected by EJCvL and MB. The data collection in 2007 differed from the data collection in 2010 and 2011 with respect to the sampling techniques and observation windows. Whereas in 2007 data were collected by means of *opportunistic* all-occurrence sampling between the hours of 8AM and 5PM (for roughly 60 days between May and August), the data in 2010 and 2011 were *systematically* collected only between 8AM and 11:45AM (for 10 days each year between May and July). Importantly, though, the coding schemes for 2007 and 2010–2011 were exactly the same: that is, grooming handclasp (GHC)-style in all years was operationalized for each clasping individual as either 'palm', 'wrist', 'forearm', or 'other'). Hence, for the purpose of the current study, we combined all data. Note that for the central analysis of the effect of matrilineal relationship on PPC grooming, we only included individuals with known and GHC-active matrilineal kin.

Statistical analysis

To investigate whether matrilines could explain our group-level results for relative convergence on the PPC style, we used Generalized Linear Mixed Models with binomial error structure ('PPC', yes/no) and logit link function [S2]. These models allow for testing multiple predictors simultaneously and control for repeated observations of, for instance, individuals and dyads.

First, we re-analyzed our original dataset (42 individuals and 619 bouts of grooming; see [S1]) including an appropriate random-slopes structure for all models [S3,S4] (not reported in the main text). To identify the appropriate random effect structure [S3–S5], we checked for each possible combination of fixed and random effect whether the inclusion of a respective random slope (of the fixed effect within the random effect) would be needed (that is, whether the random slope would be identifiable [S3]). Based on this check, in addition to the original random intercept terms (subject, partner, dyad, and date), we now included random slope terms for year (one manually dummy-coded and then centered variable) within subject, partner and dyad, to account for different effects of time on PPC tendencies between individuals and across dyads. The full model comprised the fixed effects 'group' and 'year', including their interaction. The null model comprised only the fixed effect of ‘year’ and, furthermore, the same random-effects structure as the full model. A reduced model built for the purpose of validating the original group differences result comprised ‘group’ and ‘year’ without their interaction (and the same random-effects structure as the full model). Our original results were reconfirmed (Likelihood Ratio Test [S6]): for the full-null model comparison, $\chi^2 = 7.4$, df = 2, p = 0.025; for the reduced-null model comparison, $\chi^2 = 6.07$, df = 1, p = 0.014. PPC did not change differently over the study period for the two groups; for the full-reduced model comparison, $\chi^2 = 1.34$, df = 1, p = 0.250.

Second, for the current study, in order to assess the effects of matrilineal relationships, we added the following terms to our model: the random intercepts of subject matriline and partner matriline, and the random slopes of year (two manually dummy-coded and then centered variables) within subject matriline and partner matriline [S3,S4]. Here, as before, to identify the appropriate random effect structure [S3–S5], we checked whether the variation of a particular fixed effect (such as ‘year’) within a particular random effect (such as ‘subject matriline’) was sufficiently large [S3] (that is, for most subjects there where observations available in all three years). If this was the case, we included the respective random-slopes component, allowing us to assess whether the influence of the fixed effect on the response variable varied between the levels of the random effect (for example, some matrilines may change differently in their yes/no PPC grooming across years than other matrilines). We did not include the correlations among the random intercepts and slopes, but opted to analyze the model with a close to maximal random-effects structure given that i) there is solid evidence that not accounting for random slopes increases the probability of Type I errors for tests of the fixed effects (that is, increased probability of erroneously significant results, see [S3–S5]), and ii) our dataset was sufficiently large for the number of parameters assessed in order to converge on meaningful estimates (12.1 observations per parameter in the smallest dataset and 47.0 observations per parameter in the largest). Importantly, to avoid pseudo-replication, we randomly selected one of the two GHC participants as the ‘subject’ and the other as its ‘partner’ per bout. We conducted 1,000 such random selections, ran the respective GLMM for each data configuration, and averaged the results (this procedure was applied for all reported analyses). Crucially, we tested for the influence of matrilineal relationships on PPC engagement by checking the metrics for subject matriline and partner matriline. Given that reliable p-values for random effects are difficult to obtain in the context of Generalized Linear Mixed Models, especially from model comparisons with Δdf > 1 [S7], we opted to apply a similar permutation test as reported in the original study by Wrangham *et al.* [S8]. In more detail, we permuted matrilines across subjects and also across partners within groups (as no matriline occurred in both groups). We then extracted the likelihood ratio test statistic (which would be chi-square distributed if there were no issue with the degrees of freedom associated with tests of random effects [S7]) obtained by comparing the full model and the null model lacking all random effects of subject matriline and partner matriline. We conducted one permutation for each of the 1,000 random assignments of the two GHC participants as subject and partner whereby for each permuted data configuration, we also extracted the likelihood ratio test statistic comparing the full and the null model lacking the random effects of matriline. The overall p value for the random effects of matriline was then determined as the proportion of permuted data sets revealing a likelihood ratio test statistic at least as large as the average of the original data. In addition to the p values obtained from the permutation tests, we opted to report the standard deviations estimated for the respective random effects.

Lastly, in light of the minimal random-effects structure model, called “parsimonious” in the response by Wrangham and colleagues [S9], we additionally fitted two more series of analyses with an increasingly minimal random-effects structure (each series consisted of one across-group test and two separate within-group tests to assess the effects of matrilineal relationship). Keeping our fixed-effects structure constant throughout (except for ‘group’, which was logically only assessed in the across-group tests), the first series comprised our primary model excluding the random-slope terms within subject matriline and partner matriline. The second series comprised models including only random *intercept* (this model was advanced by Wrangham *et al.* as the most adequate one, see [S9]) of ‘subject’, ‘partner’, ‘dyad’, ‘subject matriline’, ‘partner matriline’, and ‘date’. These steps toward minimal random-effects structures increased power: in the first series, the smallest dataset contained 15.3 observations per parameter and the largest dataset contained 57.4 observations per parameter; in the second series, there were 25.6 observations per parameter in the smallest dataset and 86.1 observations per parameter in the largest dataset). As before, here, we based inference on p values derived from the permutation tests and the estimated standard deviations for the random effects. Importantly, these auxiliary analyses yielded similar results as our own analytical approach (see main text). We chose to fit these additional models in order to preclude potentially unwarranted dismissal of matrilineal effects on PPC grooming, and based on arguments against using a (close to) maximal random-effects structure ([S9] referring to [S10]).

All models were fitted in R (version 3.3.1; [S11]) using the function glmer of the R package lme4 (version 1.1-12; [S12]).

Supplemental References

S1. van Leeuwen, E.J.C., Cronin, K.A.C., Haun, D.B.M., Mundry, R. and Bodamer, M.D. (2012). Neighbouring chimpanzee communities show different preferences in social grooming behaviour. Proc. Biol. Sci. *279*, 4362–4367.

S2. Baayen R.H. (2008). Analyzing linguistic data. A practical introduction to statistics. (Cambridge: Cambridge University Press).

S3. Barr, D.J., Levy, R., Scheepers, C. and Tily, H.J. (2013). Random effects structure in mixed-effects models: keep it maximal. J. Mem. Lang. *68*, 255–278.

S4. Schielzeth, H. and Forstmeier, W. (2009). Conclusions beyond support: overconfident estimates in mixed models. Behav. Ecol. *20*, 416–420.

S5. Aarts, E., Dolan, C.V., Verhage, M. and van der Sluis, S. (2015). Multilevel analysis quantifies variation in the experimental effect while optimizing power and preventing false positives. BMC Neurosci. *94*, 1–15.

S6. Dobson, A.J. (2002). An Introduction to Generalized Linear Models. (Boca Raton, FL: Chapman & Hall/CRC).

S7. Bolker, B.M., Brooks, M.E., Clark, C.J., Geange, S.W., Poulsen, J.R., Stevens, M.H.H. and White, J.-S.S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. Trends Ecol. Evol. *24*, 127–135.

S8. Wrangham, R.W., Koops, K., Machanda, Z.P., Worthington, S., Bernard, A.B., Brazeau, N.F., Donovan, R., Rosen, J., Wilke, C., Otali, E. and Muller, M.N. (2016). Distribution of a chimpanzee social custom is explained by matrilineal relationship rather than conformity. Curr. Biol. *26*, 3033–3037.

S9. Wrangham, R.W., Worthington, S., Bernard, A.B., Koops, K., Machanda, Z.P. and Muller, M.N. (2017). Response to: Chimpanzee culture extends beyond matrilineal family units. Curr. Biol. *27*, R590-R591.

S10. Bates, D., Kliegl, R., Vasishth, S., and Baayen, R.H. (2015). Parsimonious mixed models. arXiv:1506.04967.

S11. R Core Team (2016). R: A language and environment for statistical computing. (Vienna: R Foundation for Statistical Computing).

S12. Bates, D., Mächler, M., Bolker, B. and Walker, S. (2014). Fitting linear mixed-effects models using lme4. J. Stat. Soft. *67*, 1–48.