Supplement for 'Temporal dynamics and fitness consequences of coalition formation in male primates'

Christof Neumann Lars Kulik Muhammad Agil Antje Engelhardt
Anja Widdig

Contents

Supplementary methods

Study site and subjects

We studied two groups, R1 and PB, of crested macaques in the Tangkoko Nature Reserve, Sulawesi, Indonesia between March 2009 and May 2011 [1,2]. Each group comprised up to 85 individuals, with the number of adult males being present at any given time ranging between 7 and 18. All animals of both groups were completely habituated to human observers and adults were individually recognizable based on size, facial features and natural body markings, e.g., scars or broken limbs.

Behavioral data

We collected behavioural data using focal animal sampling [3] of 37 adult males (mean = 66.1h, range = 0.6 – 130.0h per male, total = 2447.2h). Focal protocols lasted up to 60 minutes during which continuous data on aggressive behaviour involving the focal male and any group member were recorded. We defined a coalition as simultaneously occurring aggression by at least two individuals (participants) directed towards at least one male (target) upon which the target left or fled the participants. This definition includes the possibility that males and females form mixed-sex coalitions, although in this study we only analysed coalitions involving males. Coalitions could also target more than one individual but since we never observed such coalitions, our operationalisation of coalitions is restricted to single targets.

Since rates with which we observed coalitions during focal animal sampling were small compared to other studies on male macaques, we supplemented our data set with ad libitum observations of coalitionary events. Note however, that this does not imply that we saw all coalitionary events taking place within the group. The habitat was generally dense and groups often were spread over more than 100 m (pers. obs.). Therefore, rates of participation in and being targeted by coalitions are based on focal animal data, whereas our main analyses are based on single coalitionary events as unit of analysis (see below).

In total, we observed 128 coalitions between adult males. Data from an additional 9 coalitions had to be discarded because either not all males could be identified or males were not yet considered adult.

Paternity assessment and analysis

We assessed paternity from 19 infants conceived during our study. Assumed period of conception was back calculated from date of birth based on the range of gestation duration in our population (171-185 days, [4]). Details of paternity assignment are described elsewhere [5,6]. In brief, we opportunistically collected up to three fecal samples from identified individuals and stored them following a two-steps alcohol-silica protocol [5]. After standard DNA extraction, all individuals, males, females and infants, were genotyped at at least 39 microsatellite loci with 12 primer pairs following a two-steps multiplex PCR approach (modified from [7], [5]). Sequencing was performed on an ABI 3130xL sequencer and allele size was determined with PeakScanner (Applied Biosystems®).

Genotypes were considered definitive when at least two different samples of the same individual produced the same results in at least four amplifications for heterozygotes and six for homozygotes (multitubes approach, [8]). Following a conservative approach, we assigned paternity only when exclusion and likelihood calculations revealed the same father [5,9]. We assigned paternity following strict criteria: the male that had no mismatches with a given mother–offspring pair across all loci while all other potential sires mismatched the offspring at two or more loci was assigned the father or more relaxed criteria when necessary: the male with no mismatches with a given mother–offspring pair across all loci while one or more males mismatched the offspring at one locus only was assigned the father [5]. We used the program FINDSIRE (https://www.uni-kiel.de/medinfo/mitarbeiter/krawczak/download) to establish paternity exclusion and CERVUS 3.0 to calculate likelihood-odds (LOD) scores and confidence levels and to confirm sires using likelihood analyses [5].

To assess the impact of rank on paternity success, we used a generalized linear mixed model with binomial error structure. Specifically, we modeled whether the Elo-rating during the assumed time around conception predicted the likelihood of siring a given infant for all males that were present during this time. The conception window was calculated as the 11-day period around the likely conception date (day 3 of the four-day fertile phase, [10] and see below), to allow for some estimation error. In this model, we also controlled for individual male age, and on the group level included hierarchy stability and the competition index as control variables (see below).

Variables used in the statistical analysis

Role

We classified individuals either as targets (victims) of coalitions, or as participants (partners) in aggression directed at a target.

Current rank, future rank and time distance

We measured dominance rank by calculating Elo-ratings that were based on decided dyadic conflicts and displacements [2,11,12]. Elo-ratings represent a continuous score of individual dominance strength [12], but if desired can be readily transformed into ordinal ranks. In brief, after each dyadic dominance interaction, the Elo-rating procedure updates the ratings of the two individuals involved in way that the winner receives points whereas the loser's points decrease. The amount of points that are won and lost depends on the expected interaction outcome. This expected outcome is based on the rating difference before the interaction took place: if a strong individual wins against a weak one (as expected based on their rating difference), the winner will gain only few points while the loser loses little. If, against such an expectation, a weaker individual wins against a stronger one the points received and lost will be relatively more. In this way, ratings of individuals are updated after each single interaction, allowing for fine-tuned monitoring of rank dynamics over time (for more details see [2]). For the purpose of this study, we used ratings at the end of a given observation day, i.e. if an individual was observed in more than one dyadic interaction on a given day, that day's rating is the one after the last observed interaction. Note, none of the interactions considered as coalitions contributed to Elo-rating estimation as it is based on dyadic interactions only. Elo-ratings have several advantages over commonly used dominance indices [2]. With respect to our study the most important of these advantages are that (1) ratings are continuously updated, which allows estimation of dominance status on small time scales (in our case on a daily basis), (2) ratings are on an interval scale, i.e. differences between ratings are meaningful, and (3) ratings are not affected by changes in the group composition. The parameter k, which determines the maximum number of points that can any individual rating can change after a single interaction, was set to 100 for all Elo-rating calculations [2].

Current rank of a male was defined as his Elo-rating on the date of the observed coalition. Future rank was calculated at time steps of 10 days following the coalition up until 120 days (see data organization, below). This time distance variable was introduced in order to investigate in which time frame coalitions have an effect (see [13] for an example of this approach). We decided to set the upper limit of the time frame to be considered at about four months (i.e. 120 days) because first, previous studies on male macaques indicated that rank changes are generally expected to occur in the magnitude of months, not years (e.g., [2,14,15]), and second, since we consider single events and not rates of coalitions over observation time we expect changes to occur on relatively small time scales.

We only used individual Elo-ratings that were updated within a maximum of five days from the date in question (including the date of the interaction itself). For example, if a coalition occurred on day 0 and we wanted to obtain the ratings of the involved males 20 days later, we would only use ratings of those males on day 20 that had at least one dyadic interaction between days 16 and 20. This ensures that if a rating from a male is identical to the one at the time step before, this is not due to this male not having been observed during a recent dyadic dominance interaction. In this way, we also excluded cases if a male emigrated from the group or died and therefore no current rating was available.

Finally, in order to present results that can be compared to the predictions of the PvS model with regards to what ranks targets and participants held, we transformed Elo-ratings into ordinal ranks. For this, we simply ordered the Elo-ratings of all males present in the group on the day a coalition occurred and assigned the male with the highest Elo-rating rank 1, the male with the second-highest rating was assigned rank 2, and so on.

Configuration

We assigned coalition events one out of three configurations. Coalitions were considered all-down (all participants had higher Elo-ratings than the target), all-up (all participants had lower Elo-ratings than the target), or bridging (at least one participant had a higher Elo-rating than the target and at least one participant had a lower Elo-rating than the target).

Feasibility

We calculated feasibility as the difference between the sum of Elo-ratings of all participants and the Elo-rating of the target, as measured on the day of the coalition (for a comparable continuous assessment of 'asymmetry in strength' see [16]). Prior to this calculation we standardized ratings in such a way that the highest rating of all males present that day was set to 1 and the lowest to 0 while keeping differences between these standardized ratings proportional to the differences between males as measured on the original Elo-rating scale. A positive feasibility value therefore indicates that the combined 'power' of the participants was greater than the 'power' of the target, whereas negative values indicate that combined power of participants was smaller than the one of the target (which can only occur in the all-up configuration).

Control variables

We added several control predictor variables to account for their possible confounding effects on future Elo-rating. Foremost, we added the Elo-rating on the date of the coalition itself. Next, we added age as control factor to account for age related rank trajectories [17–19]. Since birth years of adult males in the study population were unknown, we classified each male into one of three age categories (young ($n = 15$ males), middle ($n = 13$ males) and old ($n = 11$ males)), based on size, appearance and tooth wear. In addition, we added individual aggression rates (aggressive acts received from adult males per hour of focal observation time) into the model to account for individual variation in aggression (e.g. [20]).

The final two control variables measured two characteristics at the group level. We incorporated a quantitative measure of hierarchy stability assessed +/- five days around the date of the coalition (S index, [2]), reasoning that coalitions might affect future rank differentially depending on whether the competitive situation among resident males is stable or unstable. Finally, as a measure of degree of competition between males for access to fertile females we calculated a competition index akin to the operational sex ratio. For the day of the coalition and for each of the five days before and after, we counted the number of fertile females present in the group based on hormonal and morphological markers of a four day fertile phase during which ovulation is most likely to occur (see [10] for details). For each day, this number was then divided by the number of adult males present. A single fertile female present on a given day represents the highest possible degree of competition possible, which then decreases as more females become fertile, leading to the lowest possible degree of competition if there are (at least) as many fertile females as there are males. To account for the possibility that no fertile female was present and therefore the index was zero and faultily indicate highest possible competition, we replaced index values of 0 with 1. This assured that the index captures the similarity (i.e. no competition) between situations regardless of whether there were no fertile females available or as many as there are males. In this way, an index value close to zero represents high competition whereas values approaching 1 indicate less competition. Having calculated this index for each day, we then calculated the mean of these values over the 11-day time window, which we incorporated into the model.

Data organization

Since we were interested in the effects of coalition characteristics on future male rank, we organized our data set in such a way that each coalitionary event is linked to different future time points. To do so, each individual involved (target and participants) in a given event was initially represented as a single case, that is, a single line, in our data set. For example, a male coalition with two participants and one target was represented as three lines in our data set. For each coalition event, we then expanded the data set such that each original coalition event was represented 12 times (once for each time point). In this way, the example coalition from above resulted in $3 \times 12 = 36$ lines in the data set. After applying this approach to all events, the data set then comprised 4920 lines, representing 128 coalition events. The different lines differed with regard to future rank and the time distance variable. As stated above, 1290 lines (26.2%) were excluded because either future ratings were not available due to emigration/death or if ratings were not recently updated due to the lack of recent observations of dyadic dominance interactions.

The final data set used for the analysis comprised 3630 data points, representing 127 events and 28 males (17 targets and 27 participants).

Statistical analysis

Our general approach was to model future male rank at different time distances from the original event as function of several predictor variables that related to single coalitionary events. We built linear mixed effects models (LMMs) in R 4.1.1 [21] using the `lmer` function in the `lme4` package (version 1.1-28, [22]). Prior to analysis all numeric variables were checked for symmetric distributions and transformed if necessary.

Following this step, we standardized all variables to a mean of 0 and a standard deviation of 1 [23].

In addition to the main effects described above, we added several additional terms to the models. First, considering the time distance variable, we hypothesized that the effect might be non-linear, i.e. the effect of time distance on future rank might show a local optimum at some intermediate time distance value. We therefore added the quadratic term of time distance as additional fixed effect to allow for such a potential local optimum. Second, we added group (R1 or PB) as additional fixed predictor because with two categories fitting it as random effect is ineffective. Third, because we expected rank consequences to differ between participants and targets, we added interaction terms between all variables and role in the coalition. Fourth, we included the three-way interaction between role, configuration and feasibility. As above, we hypothesized that the effects of configuration and feasibility differ between participants and targets. Furthermore, we expected that effects of feasibility might differ depending on configuration. Specifically, all-down and bridging configurations have, by definition, relatively large feasibilities and hence we expected feasibility to play no major role here. In contrast, all-up coalitions can either be feasible or not and hence we expected larger changes in future ranks depending on feasibility in this configuration.

In order to account for repeated observations, we added male ID and coalition event ID as random intercepts. In addition we fitted uncorrelated random slopes for the most important predictors in the following way. For male ID, we included random slope terms for those fixed terms that varied within male. Specifically, we included current rating (i.e. Elo-rating on day of coalition), feasibility and configuration and the two-way interactions of these terms with role. We also included the three-way interaction between role, feasibility and configuration as random slope in male ID. With regards to event ID, we included uncorrelated random slopes for time distance and its squared term. We refrained from fitting a more complex random structure out of concern for overfitting our model.

Finally, we controlled for temporal autocorrelation, which is likely present in our data in the following way: first, we built our full model with all main, random and interaction terms as described above. From these models, we derived the residuals. For each data point separately, we calculated the weighted average of the residuals of all other data points of the same male, with the weights being equal to the time lag to that particular data point. The weight followed a normal distribution with its standard deviation determined by minimizing the AIC of the model including this autocorrelation term as additional fixed effect (R. Mundry, pers. comm., see also [24] and [25]).

The initial model was then checked for whether the assumptions of linear models were met [26]. Several diagnostic tools indicated violations of these assumptions. Particularly, Cook's measure of influence (assessed with the R package `influence.ME`, [27]) indicated a range of influential cases ($n = 94$, 0.26%). Inspection of these cases revealed that most of them could be attributed to data of two specific males (IM, WJ). Both males underwent severe rank drops in January 2010, likely due to severe injuries they suffered, the origins of which are not completely known to us. In one case (IM), we suspect that the injuries were the consequence of a dyadic fight between IM and OM, both of which were seen with fresh and in the case of IM severe wounds within a few minutes from each other. We did, however, not observe the actual event that led to these injuries. IM, in contrast to OM, dropped to the bottom of the hierarchy and was repeatedly harassed by the group, including being targeted by adult coalitions in the days following his injury. With WJ, the circumstances are less clear. He also suffered severe injuries, though no other male showed obvious signs of having been involved in a fight with WJ on the day we detected the injuries. Unlike IM, WJ left the group for seven days before rejoining the group at the bottom of the hierarchy. We decided to exclude these influential cases because we suspect that these rank drops were not the immediate consequence of coalitions, but more likely were the consequence of dyadic fights between males and the associated injuries. In addition, in this study, we are interested in the gradual changes in dominance relationships, while these two cases were abrupt changes in dominance.

After exclusion of the influential cases, we ran model diagnostics again. Residuals were homogenously and approximately normally distributed. Variance inflation factors were calculated with the function `vif` from the `car` package [28] and were all below 10 (max: 1.89), indicating collinearity between predictor variables not to be of concern [26].

To assess the statistical significance of our results, we used the following approach. First, we calculated likelihood ratio tests (LRT, [26]) to assess whether the full model (with all main effects and interactions) was different from its respective null model [29]. This null model contained the same random structure as the full model, but comprised as fixed effects only aggression rate, hierarchy stability, competition index, individual age, current status, group and the auto-correlation term. In other words, in order to derive the null model we removed from the full model those terms that we were primarily interested in. Next, we assessed the significance of the interaction terms by using LRTs to compare the model including the interaction to the model without the interaction. If an interaction was not significant at $p < 0.05$ as inferred from the LRT, we removed the interaction to be able to interpret the respective higher-order/main effects.

Supplementary results

General characteristics of coalitions

We observed a total of 137 coalitionary events between males, all of which were targeted at single adult males. The average number of participants was 2.2 (median = 2, range: 2 – 6), hence mainly reflecting triadic coalitions. For 128 coalitions (93%) we were able to identify the target and all participants, and were therefore able to determine the configuration. Targets of all-up coalitions had an average ordinal rank derived from Elo-rating of 3.8 (range: 1 – 8), while the rank of the highest rated participant was on average 6.5 (range: 4 – 14). The average size of these all-up coalitions was 2.3 (range: 2 – 3). Targets of bridging coalitions had an average ordinal rank of 7.8 (range: 2 – 14), while the highest-ranking participant had an average ordinal rank of 3.6 (range: 1 – 6). On average, 2.6 males (range: 2 – 6) participated in bridging coalitions.

During the cumulative focal animal sampling time of 2447.2h on 37 males, we observed 81 coalitions, i.e. 0.03 events per hour, or one event about every 30h. Rates per 10 hours observation time were 0.09 (median: 0.00, range: 0.00 – 1.97) and 0.11 (median: 0.00, range: 0.00 – 0.86) for targets and participants, respectively.

We did not observe levelling coalitions where a target male is displaced from monopolizing a fertile female to gain direct mating access.

Coalition dynamics

Here we present the results for the model testing the consequences of coalition events on male dominance trajectories (fixed effects in table S1, random structure in table S2).

Table S1: Fixed effects of the coalition LMM. Interactions are depicted by colons. Test levels for categorical predictors are in parentheses.

	full model			null model		
	β	SE	t	β	SE	t
intercept	-0.19	0.29	-0.67	-0.25	0.27	-0.93
role (target)	-2.16	0.41	-5.33			
time since coalition	-0.01	0.01	-0.46			
time since coalition (squared)	0.00	0.01	0.35			
configuration (bridge)	-0.10	0.14	-0.68			
configuration (all-down)	-0.23	0.12	-1.82			
feasibility	-0.06	0.08	-0.76			
aggression rate	-0.05	0.08	-0.62	-0.01	0.08	-0.09
stability index	0.09	0.02	3.85	0.07	0.02	2.96
competition index	-0.09	0.02	-3.81	-0.06	0.02	-3.00
age (middle)	0.25	0.27	0.94	0.18	0.28	0.62
age (old)	-0.33	0.22	-1.50	-0.36	0.24	-1.52
current Elo	0.56	0.06	9.54	0.51	0.06	9.08
group (R1)	0.58	0.17	3.44	0.51	0.16	3.13
auto-correlation	0.32	0.01	59.42	0.32	0.01	58.94
role (target) : time since coalition	0.02	0.01	2.31			
role (target) : time since coalition (squared)	0.01	0.01	1.33			
role (target) : configuration (bridge)	0.74	0.12	6.36			
role (target) : configuration (all-down)	0.91	0.11	8.05			
role (target) : feasibility	-0.17	0.05	-3.26			
role (target) : aggression rate	-0.13	0.09	-1.39			
role (target) : stability index	-0.06	0.01	-4.69			
role (target) : competition index	0.08	0.01	5.39			
role (target) : age (middle)	1.22	0.39	3.11			
role (target) : age (old)	0.89	0.36	2.49			
role (target) : current Elo	-0.02	0.07	-0.22			
role (target) : group (R1)	0.11	0.21	0.53			
configuration (bridge) : feasibility	0.04	0.09	0.41			
configuration (all-down) : feasibility	0.07	0.09	0.75			
role (target) : configuration (bridge) : feasibility	0.29	0.05	5.97			
role (target) : configuration (all-down) : feasibility	0.16	0.05	3.11			

Table S2: Random effects of the coalition LMM. Interactions are depicted by colons. Test levels for categorical predictors are in parentheses. Note that the structure of the random effects is the same for full and null model.

		full model	null model
level	component	SD	SD
residual		0.253	0.259
subject ID	intercept	0.427	0.463
subject ID	role (target)	0.389	0.608
subject ID	current Elo	0.219	0.213
subject ID	feasibility	0.088	0.098
subject ID	configuration (bridge)	0.293	0.346
subject ID	configuration (all-down)	0.165	0.176
subject ID	role (target) : current Elo	0.170	0.174
event ID	intercept	0.238	0.236
event ID	time since coalition	0.116	0.115
event ID	time since coalition (squared)	0.049	0.047

Age effects and aggression rates

Across ages, we found that participants of coalitions had higher future Elo-ratings than coalition targets (figure S1). For middle-aged males, this advantage of participating as opposed to being targeted was smallest, i.e. middle-aged participants had only slightly higher future ratings than targets, while young and old participants benefited from participating relatively more than being targeted. Furthermore, the figure reveals that old males' ratings generally decrease, but less so if they participated in coalitions. Finally, for young males, participation in coalitions was associated with an increase in ratings whereas young males that were targeted decreased in Elo-rating.

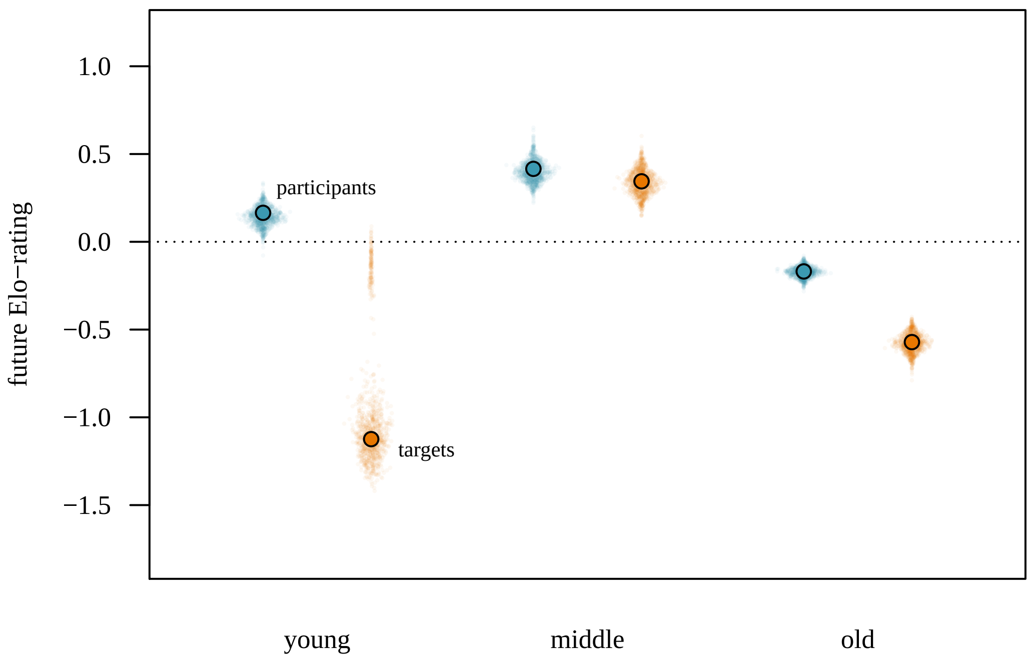

Figure S1: Coalitions add to age-related dominance trajectories. The large circles show the predicted future Elo-rating of coalition participants and targets across three age categories. The small, jittered circles represent predicted values from 1,000 bootstraps.

As expected, males in our study attained higher ranks after participating in coalitions as compared to when males were targeted by coalitions. However, only young and middle aged participants did indeed rise in rank, i.e. they had higher rank after a coalition as compared to before (figure S1). Old participants, on the other hand, dropped in rank, but they dropped less as compared to when being targets. This illustrates the interplay between age, dominance status and the moderating effects of coalitions. Male crested macaques as many male primates in general follow fairly predictable rank trajectories over their life-time, increasing in rank during young adulthood, reaching their highest status at an intermediate age, and as they age usually rank declines again [18,19]. Coalitions therefore seem to have additive effects on dominance trajectories. Given that there remains some variance in rank unexplained by age alone – for example, not all middle aged males present at one time in a group can have the highest rank – factors other than age must contribute to the ability of males to achieve the highest rank possible. Our results indicate that coalitions can act as such a means to give males an advantage during competition with other males, independent of age. As such, perhaps the term 'rank-changing' is better replaced by 'rank-moderating' given that for old males keeping their status is already beneficial, which may be indicated by the near-flat trajectories of old males participating in coalitions.

In line with the results obtained for age, we also found expected results with respect to aggression rates of males (figure S2). Males that received more aggression had generally lower future ratings than males receiving less aggression. Note that coalition events were *not* included in the calculation of aggression rates. Again, this relationship can be split between coalition targets and coalition participants. The relationship between aggression received and future ratings was more pronounced for targets compared to participants.

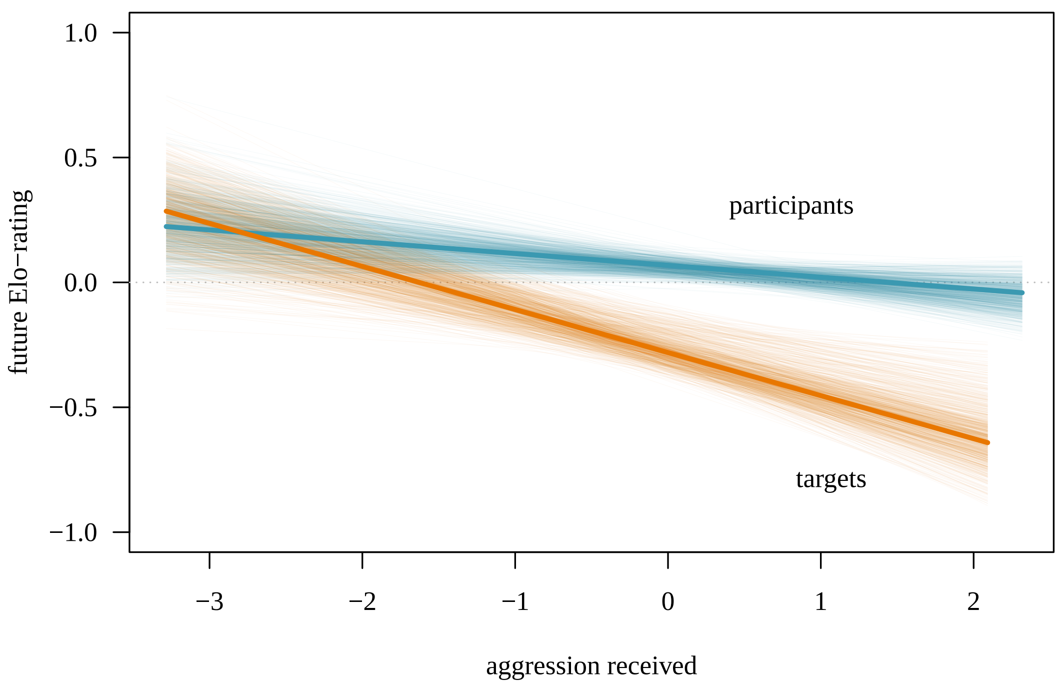

Figure S2: Aggression rates are also related to dominance trajectories. The thick lines respresent the fitted model. The thin lines represent predicted values from 1,000 bootstraps.

In sum, the results presented here support the notion that coalitions are one, albeit potentially important, factor that contributes to dominance trajectories in an additive way with other factors like age and non-coalitionary aggression, and potentially other factors that we did not include here, for example tenure, body size and personality [30].

Please note here that neither of the two variables presented here as examples (age and aggression rate) were directly related to our research questions. Rather, we considered them as 'control variables' that are not of primary importance but plausibly relevant to explain variation in the response variable [18–20]. As such, these variables were part of the null model and hence we refrained from formally testing their statistical significance post-hoc.

Paternities

Here we present the results for the model testing paternity success (fixed effects in table S3, random structure in table S4). The model is based on 19 conceptions by 17 mothers, and contains 31 adult males as potential fathers.

Table S3: Model results of GLMM for paternity success.

	full model			null model		
	β	SE	z	β	SE	z
intercept	-2.78	0.51	-5.47	-2.32	0.66	-3.50
Elo-rating	1.43	0.34	4.20			
competition index	0.18	0.28	0.64	0.09	0.27	0.34
stability index	-0.26	0.29	-0.89	0.04	0.27	0.15
age (old)	-0.31	0.55	-0.57	-0.71	0.78	-0.91
age (young)	-1.85	1.09	-1.69	-2.12	1.24	-1.71

Table S4: Random effects of the paternity GLMM. Note that the structure of the random effects is the same for full and null model.

		full model	null model
level	component	SD	SD
group	intercept	0.000	0.000
infant/group	intercept	0.000	0.000
male	intercept	0.000	1.117

Note that this analysis represents a novel approach to quantify the relationship between dominance status and reproductive success. As unit of analysis we used the conception of an individual infant. This differs from traditional approaches that usually take the rank as unit of analysis, i.e. how many infants did a certain rank sire, which could have been held by one or several males (e.g. [14,31,32]). Instead, we modeled directly the probability that a male, given his status at the time of conception, would father an infant. This approach has at least four advantages. First, it accommodates systems with frequently changing male cohorts, i.e. we can investigate directly whether it was usually the highest-rated male, and not any low-rated male, that sired an infant taking into account all males that were present at the time of conception. This approach therefore does not confound rank/status and male identity. Second, by doing so, we make the analysis independent of the number of males. Third, with this focus on conception *events*, we circumvent the need to temporarily aggregate data into time blocks of arbitrary length. This is especially advantageous in systems as ours where group composition changes frequently (see above) and where rank-changes are common [2]. Finally, this approach allows to statistically adjust for other variables, both on the individual level (e.g. age) and the group level (e.g. sex ratio).

References

1. Duboscq J, Neumann C, Agil M, Perwitasari-Farajallah D, Thierry B, Engelhardt A. 2017 Degrees of freedom in social bonds of crested macaque females. *Animal Behaviour* **123**, 411–426. (doi:10.1016/j.anbehav.2016.11.010)

2. Neumann C, Duboscq J, Dubuc C, Ginting A, Irwan AM, Agil M, Widdig A, Engelhardt A. 2011 Assessing dominance hierarchies: Validation and advantages of progressive evaluation with Elo-rating. *Animal Behaviour* **82**, 911–921. (doi:10.1016/j.anbehav.2011.07.016)

3. Altmann J. 1974 Observational study of behavior: Sampling methods. *Behaviour* **49**, 227–266. (doi:10.1163/156853974X00534)

4. Kerhoas D, Perwitasari-Farajallah D, Agil M, Widdig A, Engelhardt A. 2014 Social and ecological factors influencing offspring survival in wild macaques. *Behavioral Ecology* **25**, 1164–1172. (doi:10.1093/beheco/aru099)

5. Engelhardt A, Muniz L, Perwitasari-Farajallah D, Widdig A. 2017 Highly polymorphic microsatellite markers for the assessment of male reproductive skew and genetic variation in critically endangered crested macaques (*Macaca nigra*). *International Journal of Primatology* **38**, 672–691. (doi:10.1007/s10764-017-9973-x)

6. Kerhoas D, Kulik L, Perwitasari-Farajallah D, Agil M, Engelhardt A, Widdig A. 2016 Mother-male bond, but not paternity, influences male-infant affiliation in wild crested macaques. *Behavioral Ecology and Sociobiology* **70**, 1117–1130. (doi:10.1007/s00265-016-2116-0)

7. Arandjelovic M, Guschanski K, Schubert G, Harris TR, Thalmann O, Siedel H, Vigilant L. 2009 Two-step multiplex polymerase chain reaction improves the speed and accuracy of genotyping using DNA from noninvasive and museum samples. *Molecular Ecology Resources* **9**, 28–36. (doi:10.1111/j.1755-0998.2008.02387.x)

8. Taberlet P, Griffin S, Goossens B, Questiau S, Manceau V, Escaravage N, Waits LP, Bouvet J. 1996 Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research* **24**, 3189–3194. (doi:10.1093/nar/24.16.3189)

9. Widdig A, Muniz L, Minkner M, Barth Y, Bley S, Ruiz-Lambides A, Junge O, Mundry R, Kulik L. 2017 Low incidence of inbreeding in a long-lived primate population isolated for 75 years. *Behavioral Ecology and Sociobiology* **71**, 18. (doi:10.1007/s00265-016-2236-6)

10. Higham JP, Heistermann M, Saggau C, Agil M, Perwitasari-Farajallah D, Engelhardt A. 2012 Sexual signalling in female crested macaques and the evolution of primate fertility signals. *BMC Evolutionary Biology* **12**, 89. (doi:10.1186/1471-2148-12-89)

11. Elo AE. 1978 *The rating of chess players, past and present.* New York: Arco.

12. Albers PCH, de Vries H. 2001 Elo-rating as a tool in the sequential estimation of dominance strengths. *Animal Behaviour* **61**, 489–495. (doi:10.1006/anbe.2000.1571)

13. Gruber T, Zuberbühler K, Neumann C. 2016 Travel fosters tool use in wild chimpanzees. *eLife* **5**, e16371. (doi:10.7554/eLife.16371)

14. Schülke O, Bhagavatula J, Vigilant L, Ostner J. 2010 Social bonds enhance reproductive success in male macaques. *Current Biology* **20**, 2207–2210. (doi:10.1016/j.cub.2010.10.058)

15. Silk JB. 1993 Does participation in coalitions influence dominance relationships among male bonnet macaques? *Behaviour* **126**, 171–189. (doi:10.1163/156853993X00100)

16. Bissonnette A, Lange E, van Schaik CP. 2009 A cardinal measure of competitive ability in Barbary macaque males (*Macaca sylvanus*). *Ethology* **115**, 671–681. (doi:10.1111/j.1439-0310.2009.01652.x)

17. Alberts SC, Watts HE, Altmann J. 2003 Queuing and queue-jumping: Long-term patterns of reproductive skew in male savannah baboons, *Papio cynocephalus*. *Animal Behaviour* **65**, 821–840. (doi:10.1006/anbe.2003.2106)

18. Bissonnette A, de Vries H, van Schaik CP. 2009 Coalitions in male Barbary macaques, *Macaca sylvanus*: Strength, success and rules of thumb. *Animal Behaviour* **78**, 329–335. (doi:10.1016/j.anbehav.2009.05.010)

19. Sprague DS. 1998 Age, dominance rank, natal status, and tenure among male macaques. *American Journal of Physical Anthropology* **105**, 511–521. (doi:10.1002/(SICI)1096-8644(199804)105:4<511::AID-AJPA8>3.0.CO;2-Q)

20. Gilby IC, Brent LJN, Wroblewski EE, Rudicell RS, Hahn BH, Goodall J, Pusey AE. 2013 Fitness benefits of coalitionary aggression in male chimpanzees. *Behavioral Ecology and Sociobiology* **67**, 373–381. (doi:10.1007/s00265-012-1457-6)

21. R Core Team. 2021 R: A language and environment for statistical computing.

22. Bates DM, Mächler M, Bolker BM, Walker SC. 2015 Fitting linear mixed-effects models using `lme4`. *Journal of Statistical Software* **67**, 1–48. (doi:10.18637/jss.v067.i01)

23. Schielzeth H. 2010 Simple means to improve the interpretability of regression coefficients. *Methods in Ecology and Evolution* **1**, 103–113. (doi:10.1111/j.2041-210X.2010.00012.x)

24. Barelli C, Reichard UH, Mundry R. 2011 Is grooming used as a commodity in wild white-handed gibbons, *Hylobates lar*? *Animal Behaviour* **82**, 801–809. (doi:10.1016/j.anbehav.2011.07.012)

25. Fürtbauer I, Mundry R, Heistermann M, Schülke O, Ostner J. 2011 You mate, I mate: Macaque females synchronize sex not cycles. *PLoS ONE* **6**, e26144. (doi:10.1371/journal.pone.0026144)

26. Quinn GP, Keough MJ. 2002 *Experimental design and data analysis for biologists*. Cambridge: Cambridge University Press.

27. Nieuwenhuis R, te Grotenhuis M, Pelzer B. 2012 `influence.ME`: Tools for detecting influential data in mixed effects models. *The R Journal* **4**, 38–47.

28. Fox J, Weisberg S. 2011 *An R companion to applied regression*. 2nd edn. Thousand Oaks: Sage.

29. 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. (doi:10.1007/s00265-010-1038-5)

30. Hinde RA, Datta SB. 1981 Dominance: An intervening variable. *Behavioral and Brain Sciences* **4**, 442. (doi:10.1017/S0140525X00009614)

31. Alberts SC, Buchan JC, Altmann J. 2006 Sexual selection in wild baboons: From mating opportunities to paternity success. *Animal Behaviour* **72**, 1177–1196. (doi:10.1016/j.anbehav.2006.05.001)

32. Engelhardt A, Heistermann M, Hodges JK, Nürnberg P, Niemitz C. 2006 Determinants of male reproductive success in wild long-tailed macaques (*Macaca fascicularis*)—male monopolisation, female mate choice or post-copulatory mechanisms? *Behavioral Ecology and Sociobiology* **59**, 740–752. (doi:10.1007/s00265-005-0104-x)