

1 **Supplementary Methods**

2 *Study population and grooming data collection*

3 Both yellow and anubis baboon populations are subdivided into stable social groups, which each
4 contain multiple adults and juveniles of both sexes and range in size from approximately 20-100
5 animals. The Amboseli Baboon Research Project monitors multiple such groups ('study groups') in the
6 Amboseli ecosystem of Kenya, although the number of groups under observation varied throughout the
7 study period, from 2-5 groups in any given year. Four of our study subjects were born before
8 continuous observation of our study population began and were first observed as juveniles or young
9 adult females: the ages of these four subjects were estimated to within 2 years based on patterns of
10 growth and development (Altmann et al. 1981). Ages of all other subjects were known to within a few
11 days' error (n=204) or to within a few months' error (n=16). We excluded a few time periods during
12 which rates of behavioral data collection were low because of logistical challenges in the field. We also
13 excluded periods of fission/fusion between social groups because during these times, group
14 membership was uncertain. For individual females and dyads, we excluded years during which the
15 female (for aggregate grooming measures) or the dyad (for dyadic grooming measures) was present in
16 the study population for less than 60 days; these included years in which a female reached maturity or
17 died partway through the year.

18 Our grooming data were collected on all observed grooming events between adult females within
19 the group. However, some grooming events were likely missed, especially in large groups. Uneven
20 sampling of individuals was avoided by collecting the great majority of our grooming data during the
21 course of random-order focal animal sampling on adult females and juveniles. The collection of
22 grooming data was not restricted to grooming by the focal animal because the randomization of
23 observer effort with respect to focal females ensured that observers continually moved to new locations
24 within the group and observed all adult females and juveniles on a regular rotating basis.

25 26 *Relationship between grooming frequency and observer effort*

27 From our observations of grooming behavior, we calculated the daily frequency with which each
28 study subject female was observed giving grooming to any other adult female in the population.
29 Because the number of person-hours we devote to observations of each group does not increase with
30 increasing group size, the apparent per-capita rate of grooming is higher in smaller social groups than in
31 larger groups. This difference arises as a simple artifact of having a constant number of observers
32 regardless of group size. Our grooming data are collected throughout daily observations, during the
33 collection of 10-minute focal animal samples, which we conduct by rotating through adult females in
34 random order, completing each rotation before we begin another. Thus, the number of focal animal
35 samples per female per day changes systematically as a function of group size in a manner that reflects
36 the per capita amount of time we spend directly observing each female. The number of focal animal
37 samples per female per day is therefore a good measure of per-female observer effort (see Archie et. al
38 2014 and Campos et al. 2021 for details)

39 To correct for the fact that the number of observers does not change with group size, we
40 regressed daily rates of grooming given for all adult females alive in the population during a given time
41 period against observer effort, measured as the number of focal animal samples collected in a group per
42 adult female per day. Our grooming index was the residual of this regression. In other words, if we
43 wanted to measure ACA's grooming in a 365-day period, we would determine the grooming frequency
44 and observer effort for all the females alive in our population in that 365-day period. We would use this
45 information to regress grooming frequency on observer effort. ACA's residual from that regression
46 (which reflects whether she groomed more or less than the average expected based on observer effort)
47 would be her grooming index for that time period.

48

49 *Index of Aggregate Grooming Given*

50 To create our index of aggregate grooming given, we calculated the annual mean value of the
51 female's residuals from the regression of daily rates of grooming given on observer effort (see above).
52 This annual mean value was not based on the calendar year, but on each female's birth date and hence
53 age class (e.g., her 6th year of age, 7th year of age, and so on). This approach is computationally
54 expensive. It means that, for example, if we wanted to measure a baboon's grooming between her fifth
55 and sixth birthday, we would determine the grooming frequency and observer effort for all the females
56 alive in our population in that specific 365-day period. We would use these data to regress grooming
57 frequency on observer effort. The only data point we would use from this regression is the residual for
58 our one specific baboon of interest. We calculated data for each female in each year of her life in this
59 manner.

60
61 *Index of dyadic grooming given*

62 To explicitly investigate the role of indirect genetic effects on grooming behavior, we calculated
63 a dyadic grooming index for each pair of adult females that were co-resident in a social group for at least
64 60 days during the study period (January 1983-June 2017) and that had at least one grooming
65 interaction. We used the same set of females and female years as in the aggregate indices. However, for
66 ease of calculation, we calculated the dyadic index for each calendar year instead of for each year of age
67 for a given female. The dyadic grooming index was based on the daily rate of grooming given by the
68 focal individual to a specific partner over the course of a calendar year, considering only the days where
69 both focal and partner were present as adults in the same social group, and thereby available to each
70 other as grooming partners. As with the aggregate indices described above, the dyadic index is the
71 residual of the regression of this daily grooming rate on observer effort. Positive values indicate dyads
72 in which the focal gave above average amounts of grooming to that particular partner (relative to the
73 mean of the population, controlling for observer effort) and negative values indicate dyads where the
74 focal gave less grooming than average to that particular partner (controlling for observer effort) as
75 compared to the entire population. Dyadic grooming indices do not have to be (and overwhelmingly are
76 not) symmetrical between the two females in a dyad. In our mixed effects models, we included only
77 dyads in which both members of the dyad were in our set of study subjects.

78
79 *Rationale for the Grooming Index Approach*

80 Using the residuals of grooming frequency on observer effort to define our phenotypes of
81 interest imposes some limitations on our analysis, but also provides distinct benefits. The primary
82 limitation of using residuals in analyses is that the parameter estimates for any fixed effects that are
83 truly correlated with observer effort will be conservatively biased (Darlington and Smulders 2001). In
84 our case, these fixed effects include group size and ordinal dominance rank. Despite this limitation, we
85 chose to use residuals as our phenotype of interest instead of raw grooming counts because this
86 approach makes our results easier to interpret in two important ways. First, the data are distributed in
87 an intuitive way: positive values indicate more grooming than the population mean for that time period
88 while negative values indicate less grooming than the population mean. Second, the quantitative
89 genetic parameter estimates produced by a linear mixed effects model are significantly easier to
90 interpret than the parameters produced by a generalized linear mixed effects model (which we would
91 need to implement if our response variable was uncorrected counts of grooming bouts) because GLMMs
92 provide inference on a statistically convenient latent scale, while we wish to express quantitative genetic
93 parameters on the scale upon which traits our traits were measured. While methods exist for
94 converting parameters expressed on the latent scale to the observed scale (see de Villedenreuil et al.
95 2016) they are not without complication. Because our primary interest is in understanding genetic

96 contributions to grooming behavior, and not in the effect of group size on grooming behavior, we chose
97 the statistical approach that provides the most easily interpreted quantitative genetic parameters.

98

99 *Appropriateness of the animal model*

100 Pedigree structure

101 In this population, maternities are identified from long-term records of births, and maternities
102 and paternities are verified with genetic parentage analysis, using microsatellite genotypes obtained
103 from DNA derived from fecal samples or, in some cases, blood samples. Specifically, analyses of
104 paternity and relatedness are routinely conducted for the study population (Buchan et al. 2003, Alberts
105 et al. 2006, Charpentier et. al 2008, Tung et. al 2012). For samples extracted from faeces, all apparent
106 homozygous genotypes are reamplified at least four and up to seven additional times to guard against
107 allelic dropout. All genotype data were produced on either an ABI 3700 Sequence Analyzer or an ABI
108 3730xl Sequence Analyzer.

109

110

111 Grooming patterns among kin

112 Most female-female grooming in baboons occurs between maternal relatives (Silk 1987).
113 Interactions between relatives can present challenges for the animal model, particularly in partitioning
114 between direct and indirect sources of genetic variance. Specifically, if interactions occur only between
115 individuals who are equally related (e.g., when interactions occur exclusively within sib-families), the
116 direct and indirect genetic variance are not statistically distinguishable (see Bijma 2014, appendix in
117 Bijma et al. 2007, Cheng et al. 2009). However, this problem does not affect our dataset. Female
118 baboons have strong grooming bonds with close relatives, but they groom with individuals of all levels
119 of relatedness (Silk et al. 2006a,b, Figure S5). Most of the grooming pairs in our study are not closely
120 related, although more grooming events occur per pair in the more closely related dyads (see results in
121 the main text). In our sample, relatedness between grooming pairs ranges from 0 to >0.5 (Figure S5),
122 and a given pair of half siblings often has only partially overlapping sets of grooming partners (Figure
123 S6). Indeed, because of the multi-male, multi-female mating system of baboons, even when maternal
124 siblings share the same grooming partners, they often have different pedigree relationships with their
125 grooming partners (Figure S6). This combination of differential interactions and relatedness with social
126 partners among maternal half siblings allows the animal model to partition between direct and indirect
127 genetic effects.

128 Because female-female grooming is most common between maternal relatives, the presence or
129 absence of these relatives is an important predictor of female grooming. In our aggregate grooming
130 models, we include the number of relatives, type of relatives, and total relatedness to the group as
131 fixed effects (see below and Supplementary Methods for details) and in our dyadic grooming models we
132 include the pedigree relatedness and type of relationship (e.g., mother-daughter pair) as fixed effects.

133

134 Accounting for genetic admixture

135 As noted in the main text, the study population is admixed, consisting of individuals who tend to
136 have majority yellow baboon ancestry but also carry some introgressed ancestry from neighboring
137 anubis baboon populations (Alberts and Altmann 2001; Tung et al. 2008; Vilgalys, Fogel et al. 2022).
138 Recent work indicates that intermittent gene flow has been occurring between our study population
139 and neighboring anubis populations for hundreds to thousands of generations (Wall et. al, 2016;
140 Vilgalys, Fogel et al. 2022). Using the same 7-14 microsatellite loci that we used to construct the
141 pedigree, we calculated the proportion of recent anubis versus yellow ancestry in each of our study
142 subjects using STRUCTURE 2.3 (Pritchard et al. 2000, Falush et al. 2003; see Tung et al. 2008; Charpentier
143 et al. 2012 for details of its use in the baboon study). Because admixed individuals are fully viable and

144 reproduce freely in Amboseli, the amount of recent admixture varies continuously across individuals in
145 our population from yellow-like to anubis-like (Vilgalys et al. 2022; here we used microsatellite-based
146 scores because resequencing data are not available for many individuals in our sample). In our study
147 subjects, the mean point estimate for the microsatellite-based admixture scores was 0.25 (range 0.024
148 to 0.899). Previous work in our population suggests that admixture score is an important predictor of
149 male-female grooming, so we include admixture score as a fixed effect in order to capture genetic
150 variance in grooming behavior that is explained by anubis versus yellow ancestry. Consequently, our
151 estimates of heritability from models that include fixed effects are an estimate of the proportion of
152 phenotypic variance in our population that is due to genetic variance *independent of* admixture between
153 yellow and anubis genetic backgrounds.

154 *Example Code*

155

156 Covariance between direct and indirect genetic effects with the dyadic index

157 To determine the genetic correlation between the direct genetic effects of grooming given to a
158 partner and the indirect genetic effects of grooming elicited from a partner, we fitted three models: a
159 model in which the genetic covariance was constrained to zero, a model in which the genetic covariance
160 was constrained to +1 and a model in which the genetic covariance was free to vary. The code for these
161 models will be difficult to interpret without some familiarity with Asreml-R. Here we have highlighted
162 the primary differences between the models with red text, please see the [Asreml-R manual](#) for details
163 about the rest of the code. This code is for Asreml-R version 4.

164 *Cartoon model with the DGE/IGE covariance constrained to zero:*

```
165 asreml (fixed=trait1~fixed_effects,  
166         random=~str (~vm (actor, ainv*) +vm (actee, ainv) , ~corgh (2) :vm (actor, ainv) )  
167         +ide (actor) +ide (actee) ,  
168         data=data,  
169         G.param = initial_values.zero**  
170     )
```

171
172 *ainv is determined from the pedigree file

173 **initial_values.zero is a set of parameters with the covariance between genetic variance for actor and genetic
174 variance for actee set to zero

175 *Cartoon model with the DGE/IGE covariance constrained to positive one:*

```
176 asreml (fixed=trait1~fixed_effects,  
177         random=~str (~vm (actor, ainv*) +vm (actee, ainv) , ~corgh (2) :vm (actor, ainv) )  
178         +ide (actor) +ide (actee) ,  
179         data=data,  
180         G.param = initial_values.one***  
181     )
```

182
183 *** initial_values.zero is a set of parameters with the covariance between genetic variance for actor and genetic
184 variance for actee set to one

185 *Cartoon model with DGE/IGE covariance free to vary:*

```
186 asreml (fixed=trait1~fixed_effects,  
187         random=~str (~vm (actor, ainv*) +vm (actee, ainv) , ~corgh (2) :vm (actor, ainv) )  
188         +ide (actor) +ide (actee) ,  
189         data=data,  
190     )
```

191

192 The primary difference in these models is no assignment of fixed values for G.param in the model where
193 the covariance is free to vary.

194

195

196

Table S1. Fixed effects used in models.

	Aggregate Grooming Indices	Dyadic Grooming Index	Variable Type
Individual traits			
Focal individual's ordinal dominance rank ^a	Included	Included	Continuous
Partner individual's ordinal dominance rank ^a	NA	Included	Continuous
Focal-partner rank interaction	NA	Included	Interaction
Focal age at start of observation period ^b	Included	Included	Continuous
Partner age at start of observation period ^b	NA	Included	Continuous
Focal-partner age interaction	NA	Included	Interaction
Family effects			
Proportion of time period co-resident with mother ^c	Included	NA	Continuous
Proportion of time period co-resident with adult daughters ^d	Included	NA	Continuous
Proportion of time period co-resident with adult maternal sisters ^e	Included	NA	Continuous
Total pedigree relatedness to other adult females ^f	Included	NA	Continuous
Mother-daughter pair	NA	Included	Categorical (yes/no)
Maternal sister pair	NA	Included	Categorical (yes/no)
Relatedness to partner ^g	NA	Included	Continuous
Demographic effects			
Group size ^h	Included	Included	Continuous
Sex ratio ⁱ	Included	Included	Continuous
Admixture effects			
Focal admixture score ^j	Included	Included	Continuous
Partner admixture score ^j	NA	Included	Continuous
Focal-partner admixture score interaction	NA	Included	Interaction

198

199

200 ^aSocial dominance rank is calculated on a monthly basis by minimizing entries below the diagonal in agonism
 201 matrices (see Lea et al. 2014).

202 ^bFemale age can be determined with a high degree of certainty, because in most cases we know female birthdates
 203 to within just a few days' error.

204 ^cThe proportion of the year that an individual was co-resident with her mother. Co-residency means the mother
 205 and daughter were alive and in the same social group

206 ^{d,e}The proportion of the year that an individual was co-resident with her adult daughter^e or maternal sisters^e. If an
 207 individual was co-resident with more than one adult daughter or adult maternal sister we summed the percentage
 208 of time spent with each daughter or sister, meaning these values can exceed 1.

209 ^fThe sum of the simple pedigree relatedness values between the focal individual and all other adult females in the
 210 group (except for her mother, adult daughters and adult maternal sisters, who were modeled separately). This
 211 sum was weighted by how many days she spent with each individual in the given time period.

212 ^gPedigree relatedness between the dyad

213 ^hThe average number of adults (females who have reached menarche plus males who have achieved adult
 214 dominance rank) in the social group over the year

215 ⁱThe average number of adult females present over the course of the year divided by the average number of adult
216 males present over the course of the year. Higher values indicated time periods where the group was female-
217 biased.

218 ^jAs noted in the main text, the study population has majority yellow ancestry, with some contribution from anubis
219 baboons (Alberts and Altmann 2001; Tung et al. 2008). We have calculated admixture scores based on
220 microsatellite data, where higher scores represent more anubis-like ancestry (see Tung et al. 2008, Tung et al.
221 2012 for details).

222

223
224

Table S2: Test statistics and degrees of freedom for fixed effects in aggregate and dyadic models. We report the denominator degrees of freedom and F statistic from conditional Wald tests for every fixed

	Aggregate Grooming Index	Dyadic Grooming Index
Individual traits		
Focal individual's ordinal dominance rank ^a	denDF= 584.8 F.con= 3.886	denDF=590.6 F.con=3.517
Partner individual's ordinal dominance rank ^a	NA	denDF=123.4 F.con=0.329
Focal-partner rank interaction	NA	denDF=5797.7 F.con=52.81
Focal age at start of observation period ^b	denDF=1806.2 F.con=191.300	denDF=2377.6 F.con=196.2
Partner age at start of observation period ^b	NA	denDF=1542.7 F.con3.320
Focal-partner age interaction	NA	denDF=8283.1 F.con2.111
Family effects		
Proportion of time period co-resident with mother ^c	denDF=1237.3 F.con=0.273	NA
Proportion of time period co-resident with adult daughters ^d	denDF=1813.9 F.con=113.600	NA
Proportion of time period co-resident with adult maternal sisters ^e	denDF=469.3 F.con=0.554	NA
Total pedigree relatedness to other adult females ^f	denDF=660.4 F.con=0.273	NA
Mother-daughter pair	NA	denDF=1942.6 F.con=314.000
Maternal sister pair	NA	denDF=1711.3 F.con=84.45
Relatedness to partner ^g	NA	denDF=2093.0 F.con=14.19
Demographic effects		
Group size ^h	denDF=1167.1 F.con=1.036	denDF=1751.8 F.con=14.19
Sex ratio ⁱ	denDF=1820.4 F.con=3.418	denDF=10362.1 F.con=1.331
Admixture effects		
Focal admixture score ^j	denDF=143.6 F.con=3.234	denDF=138.4 F.con=2.631
Partner admixture score ^j	NA	denDF=123.4 F.con=0.329
Focal-partner admixture score interaction	NA	denDF=2068.2 F.con=1.119

225 effect in both of the best models (Heritability model for Aggregate Index, IGE model for Dyadic Index).

226

227

228

229

Table S3. Heritability estimates from models without fixed effects, and estimates with fixed effects (which are also reported in Tables 2 and 3 in the main text)

	Additive genetic contribution without fixed effects (h^2), with standard error	Additive genetic contribution from models that include fixed effects (h^2), with standard error (also reported in Tables 2 and 3)
AGGREGATE INDEX		
Grooming given	0.30 (0.07)	0.22 (0.048)
DYADIC INDEX		
Focal (Direct Effects)	0.03 (0.01)	0.048 (0.015)
Partner (Indirect Effects)	0.01 (0.005)	0.020 (0.005)

230

231

232

233

234

235

236

237

238

239

240

241

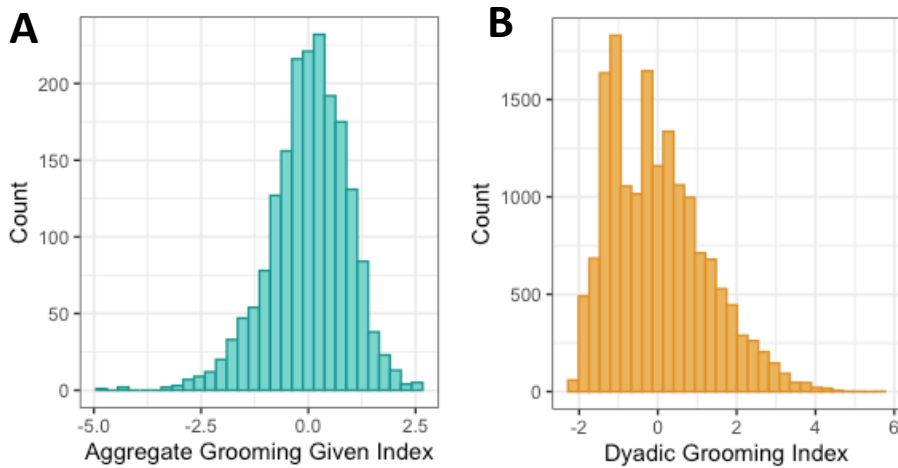
242

243

244

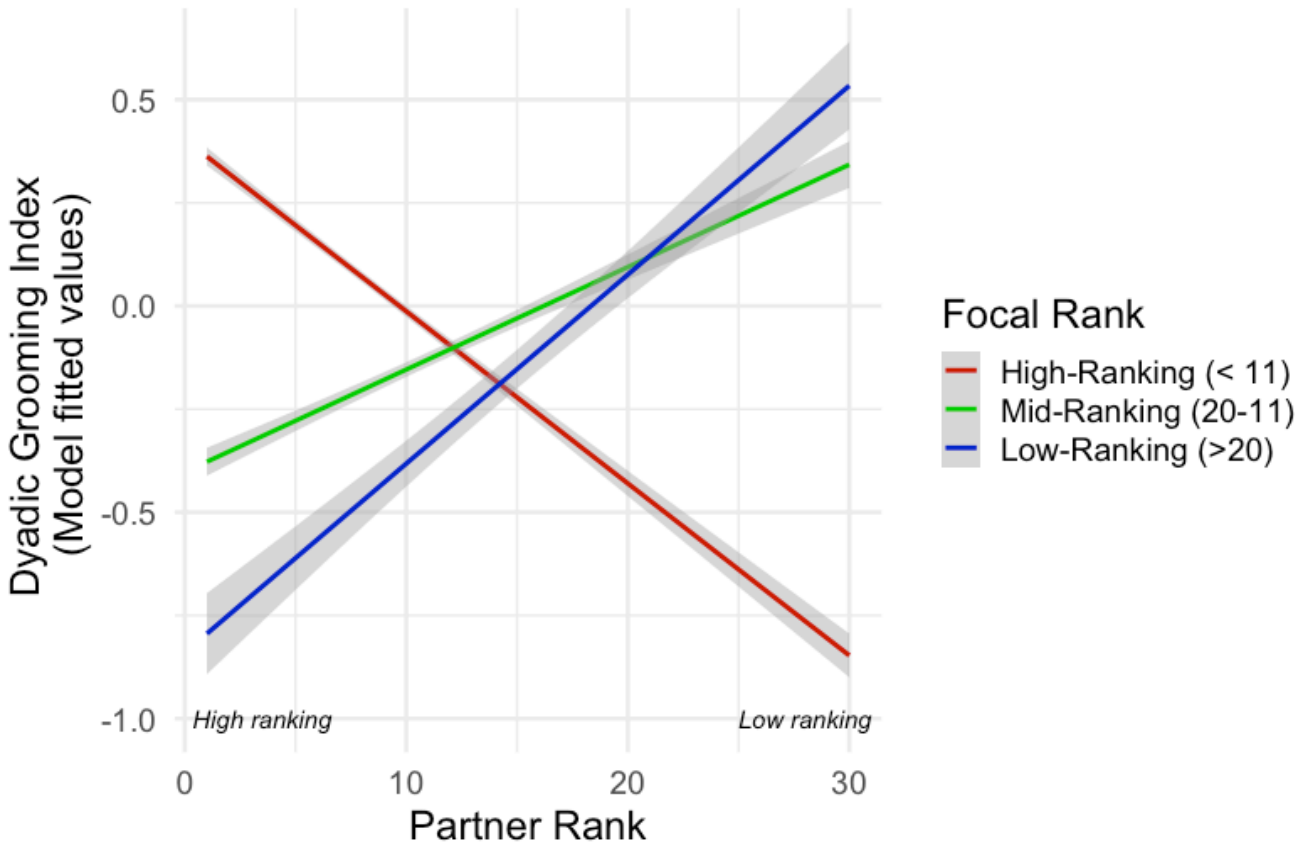
245

246



249
250
251
252
253
254
255
256
257
258
259
260

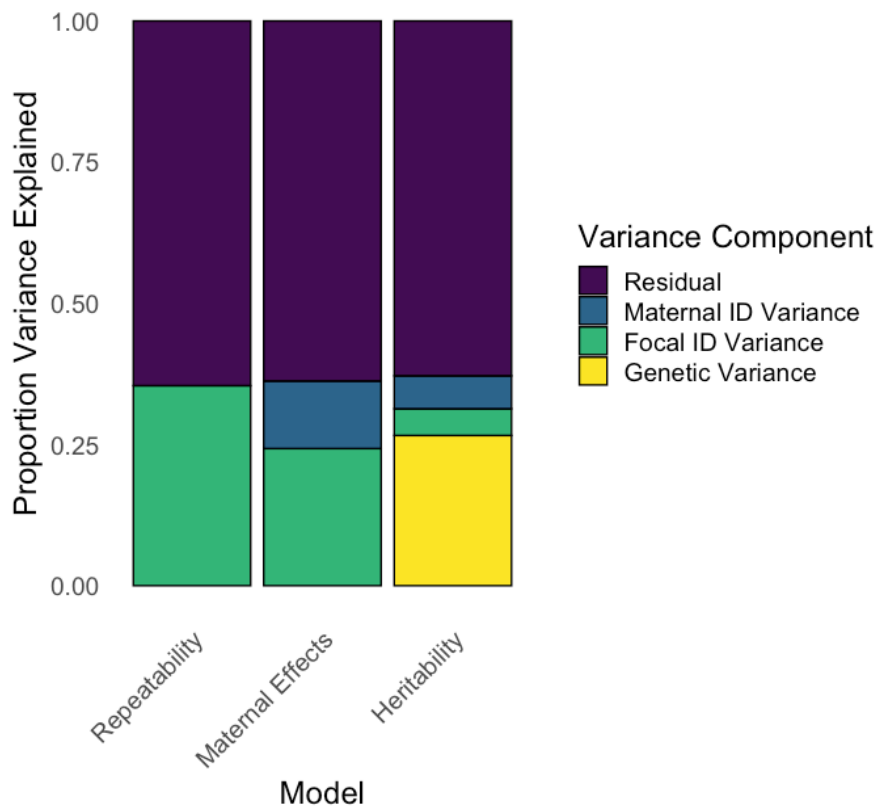
Figure S1. Distributions of grooming index values. (A) The index of *aggregate grooming given* reflects the frequency with which a given female in a given year of life groomed other adult females, relative to the grooming given by all other adult females alive in the same year. (B) The *dyadic grooming index* was calculated for each adult female dyad for each calendar year in which both the focal and partner were present as adults in the same social group and thereby available to each other as grooming partners. Positive values of the dyadic index indicate cases in which an adult female gave high frequencies of grooming to a specific partner relative to all other partner pairs in the population for that year, while negative values indicate cases in which an adult female gave relatively low frequencies of grooming to a specific partner.



261

262

263 **Figure S2.** The dyadic grooming index is partially predicted by an interaction between the ordinal
 264 dominance ranks of the focal female and her partner. High-ranking females (red line) gave more
 265 grooming to high-ranking partners (those with lower ordinal rank numbers), and the effect was in the
 266 opposite direction when focal females were middle or low-ranking (blue and green lines). Gray shaded
 267 areas represent 95% confidence intervals. Ranks are presented as categories for the purposes of
 268 visualization, but in the underlying model both focal and partner rank were modeled as continuous
 269 variables.



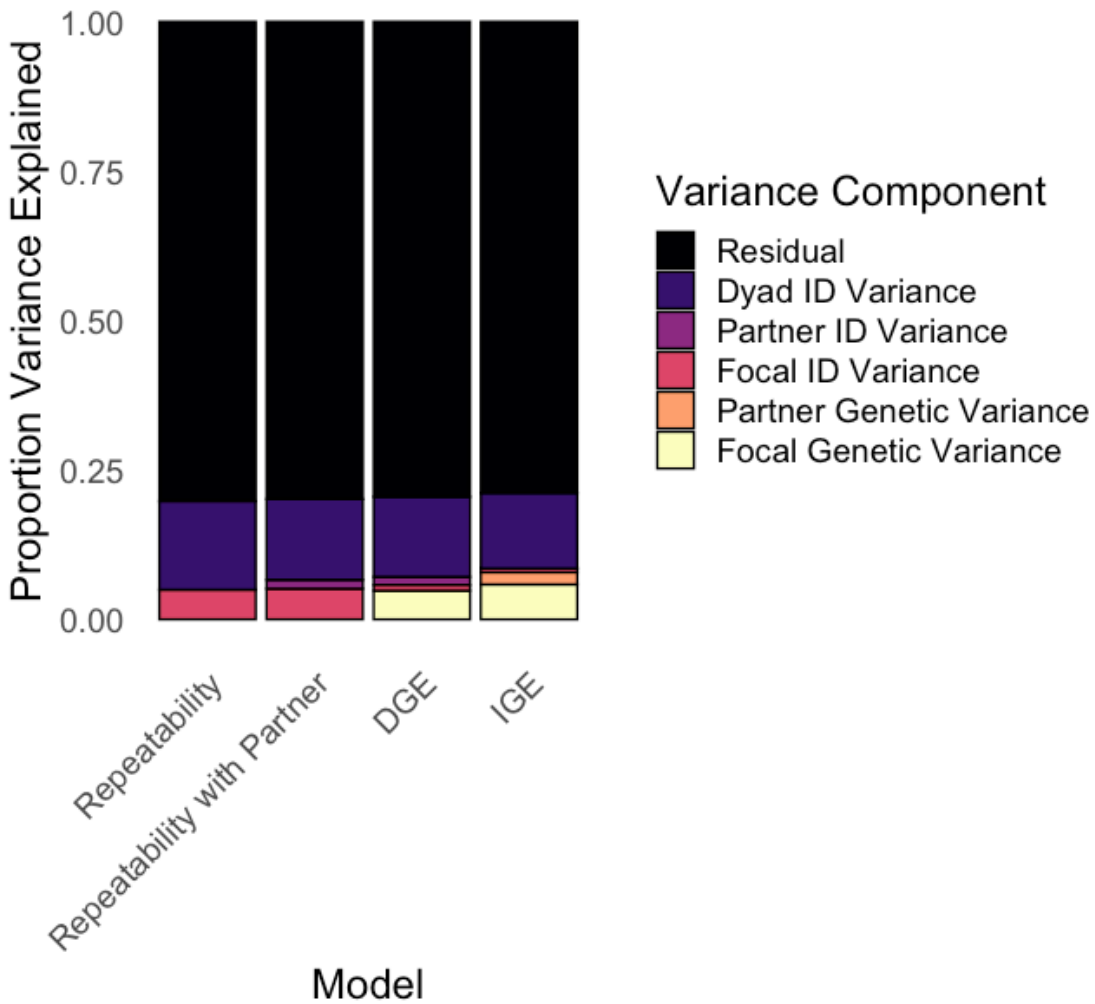
270

271

272 Figure S3. Visualization of the models of aggregate grooming given depicted as the proportion of
 273 variance explained by the following random effects: direct genetic effects (additive genetic variance),
 274 the identity of a female’s mother (maternal ID variance), focal permanent environment (focal ID
 275 variance), and residual variance. For each aggregate metric, the proportion of variance explained by
 276 additive genetic variance in the “heritability” models represents the heritability. The heritability model
 277 was the best model as determined by a likelihood ratio test.

278

279

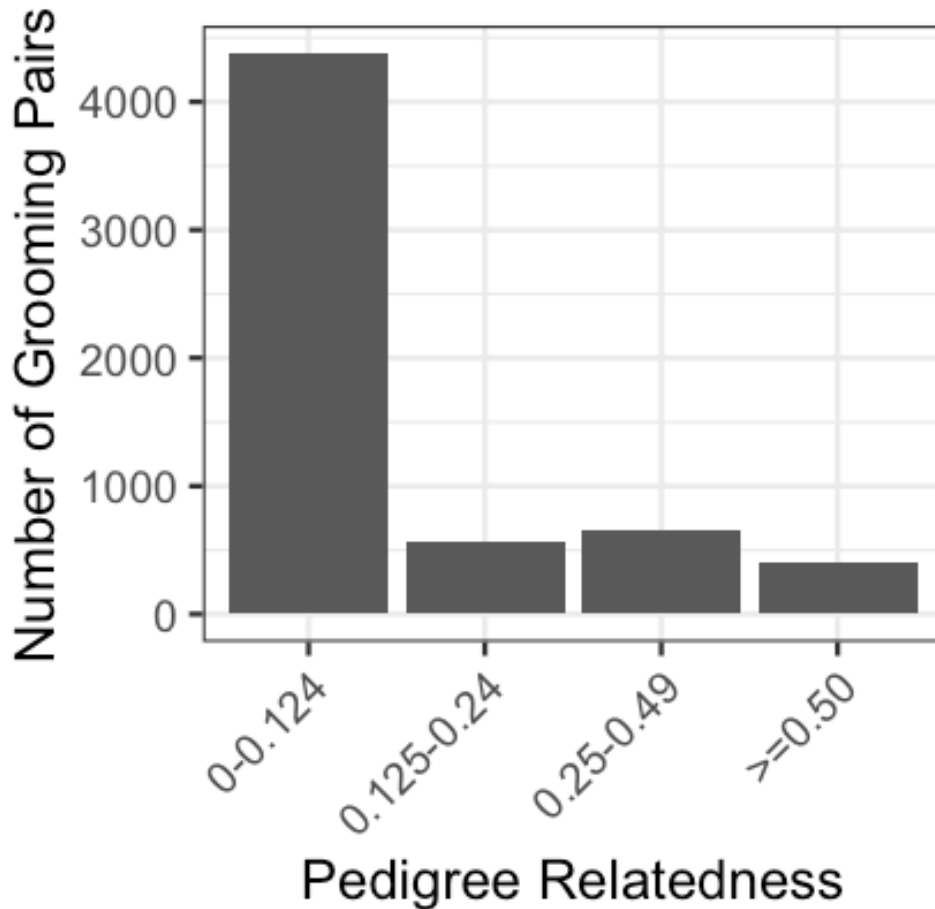


280

281

282 Figure S4. Visualization of the models of dyadic grooming depicted as the proportion of variance
 283 explained by the following random effects: direct genetic effects (focal genetic variance), indirect
 284 genetic effects (partner genetic variance), focal permanent environment effects (focal ID variance),
 285 partner permanent environment effects (partner ID variance), dyad ID variance and residual variance.
 286 The proportion of variance explained by focal genetic variance in the DGE and IGE models represents the
 287 contribution of direct genetic effects to phenotypic variance in the dyadic metric. The proportion of
 288 variance explained by partner genetic variance in the IGE model represents the contribution of indirect
 289 genetic effects to phenotypic variance in the dyadic metric. The IGE model was the best model as
 290 determined by a likelihood ratio test.

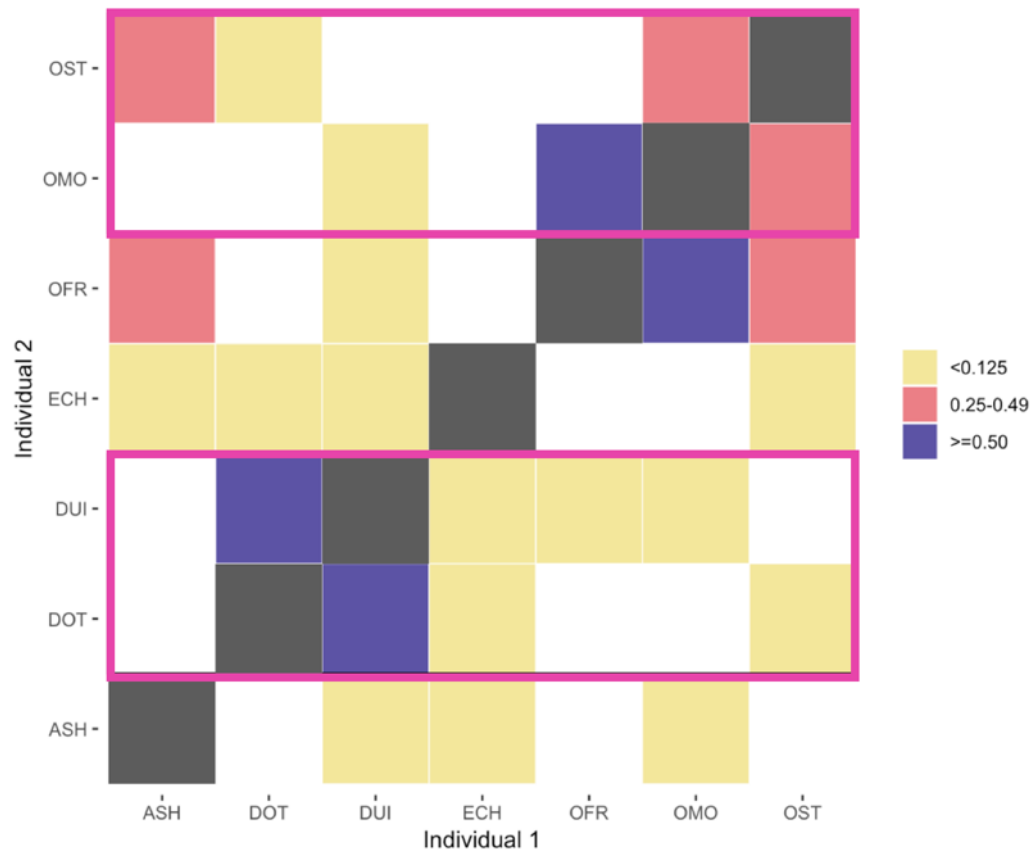
291



292
293

294 **Figure S5.** Distribution of grooming pairs across relatedness categories. Because our pedigree is up to six
295 generations deep and we have both maternal and paternal connections, some pairs of individuals fall in
296 between the typical categories of pedigree relatedness. For instance, we have 544 pairs with $r=0.25$ and
297 397 pairs with $r=0.5$, and we also have 105 pairs for which r falls between 0.25 and 0.5, and 6 pairs for
298 which $r>0.5$.

299
300
301



302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313

Figure S6. Maternal siblings have only partially overlapping grooming partners. The matrix shows all of the adult females present in one small social group in one year, which represents a subset of our grooming dataset. Colored boxes represent pairs of individuals who groomed each other, and the color illustrates their pedigree relatedness; dark gray boxes on the diagonal would correspond to the individual self-grooming, a behavior that is not recorded. White boxes represent dyads who were never observed to groom in that year. Maternal siblings are grouped by pink outlines, illustrating that maternal siblings have only partially overlapping sets of grooming partners. It is not illustrated in this example, but our data also demonstrates that maternal siblings are not always equally related to a given grooming partner.