Estimating longitudinal hierarchies: A framework for studying the dynamics of dominance

Eli D Strauss[1,2,3]
Kay E Holekamp[1,2,3]

[1]Department of Integrative Biology, Michigan State University, East Lansing, MI, USA
[2]Program in Ecology, Evolutionary Biology, and Behavior, Michigan State University, East Lansing, MI, USA
[3]BEACON Center for the Study of Evolution in Action, Michigan State University, East Lansing, Michigan, USA

Eli Strauss (Corresponding author): straussed@gmail.com
Kay Holekamp: holekamp@msu.edu

1   **Abstract**

2   1. Social inequality is a consistent feature of animal societies, often manifesting in

3   dominance hierarchies, in which each individual is characterized by a dominance rank

4   denoting its place in the network of competitive relationships among group-members.

5   Most studies treat dominance hierarchies as static entities despite their true

6   longitudinal, and sometimes highly dynamic, nature.

7   2. To guide study of the dynamics of dominance, we propose the concept of a longitudinal

8   hierarchy: the characterization of a single, latent hierarchy and it's dynamics over time.

9   Longitudinal hierarchies thus comprise both periodic rank-order estimates and

10   assessment of rank dynamics between these estimates. Methods that estimate

11   longitudinal hierarchies should optimize accuracy of rank dynamics as well as the rank

12   orders themselves. Although ranking algorithms have often been assessed for their

13   accuracy at estimating rank orders, their ability to estimate rank dynamics has never

14   been evaluated.

15   3. Estimation of longitudinal hierarchies requires both incorporation of new individuals

16   and optimal identification of rank dynamics. We propose that prior knowledge of

17   dominance correlates can be used to inform placement of new individuals in the

18   hierarchy, and that an inertial tendency for individuals to maintain their ranks over time

19   in the absence of evidence to the contrary can be used to prevent overestimation of

20   rank dynamics. We suggest a novel matrix-based method ('CRAM') for determining

21   longitudinal hierarchies that incorporates these strategies.

4. Using both a simulated dataset and a long-term empirical dataset from a species with two distinct sex-based dominance structures, we compare the performance of CRAM with four existing methods. We show that incorporation of inertial tendency is vital to avoid overestimation of rank dynamics, and that prior knowledge of dominance correlates improves the accuracy of estimated longitudinal hierarchies. We find that CRAM is the most accurate method for estimating more stable longitudinal hierarchies, like those found in most long-lived vertebrates, and that a modified Elo rating method performs best when estimating highly dynamic hierarchies.

5. This work provides the first explicit framework for studying the dynamics of dominance.

INTRODUCTION

Social inequality is a surprisingly consistent feature of group living. This inequality often manifests as a dominance hierarchy, in which repeated agonistic interactions between group members lead to the emergence of a linear order of individuals from high to low rank (H. de Vries, 1998), with rank position often determining priority of access to resources. Found in a diverse set of organisms including protists (Fortunato, Queller, & Strassmann, 2003), insects (Shimoji, Abe, Tsuji, & Masuda, 2014), fish (Sneddon, Schmidt, Fang, & Cossins, 2011), mammals (Bro Jørgensen & Beeston, 2015), and birds (Schjelderup-Ebbe, 1922), dominance hierarchies promote group stability and reduce the costs of living in an uncertain social environment (Beaulieu, Mboumba, Willaume, Kappeler, & Charpentier, 2014; Mendonça-Furtado et al., 2014). Since the first description of a 'peck order' among chickens (Schjelderup-Ebbe, 1922), dominance hierarchies have proven to be an important force structuring myriad

44   aspects of an organism's biology, including its space use (Boydston, Kapheim, Szykman, &

45   Holekamp, 2003; Murray, Mane, & Pusey, 2007), health (Sapolsky 2005; Flies et al. 2016;

46   Snyder-Mackler et al. 2016), longevity (Swanson, Dworkin, & Holekamp, 2011; Tung, Archie,

47   Altmann, & Alberts, 2016), and reproductive success (D. de Vries, Koenig, & Borries, 2016;

48   Pusey, Williams, & Goodall, 1997).

49   Dominance hierarchies are dynamic systems: the positions of individuals within a

50   hierarchy can change over time due to temporal variation in demography, ecological

51   conditions, ontogenetic development, or the relative body condition of group members

52   (Hasegawa & Kutsukake, 2014; Lea, Learn, Theus, Altmann, & Alberts, 2014; Samuels, Silk, &

53   Altmann, 1987; Sapolsky, 2005). Nevertheless, most studies treat dominance as a constant

54   individual attribute rather than a transient state. This static approach to dominance is

55   appropriate for many studies conducted over short periods of time; consistency, at least over

56   short times scales, is a defining feature of dominance relationships (Drews, 1993). While useful

57   for understanding the consequences of dominance hierarchies, this approach does not allow

58   for the study of system dynamics. Understanding these dynamics will facilitate study of the

59   structure of inequality in animal societies, the relationships between dominance networks and

60   other social networks, and the stability of dominance hierarchies. The rapid recent advances in

61   social network analysis techniques (Farine, 2017; Hobson, Avery, & Wright, 2013; Pinter-

62   Wollman et al., 2014) offer promising methods for assessing the dynamics of dominance

63   hierarchies, but reliable estimation of hierarchies and their dynamics is a requisite first step.

64   Until now, an explicit framework for investigating the dynamics of dominance

65   hierarchies has been lacking. Most studies have estimated the dynamics of dominance systems

66   by dividing the study into periods, determining ranks independently within each period, and

67   inferring changes in rank based on the differences observed between one period and the next

68   (Chapais, Girard, & Primi, 1991; Combes & Altmann, 2001; Hasegawa & Kutsukake, 2014;

69   Higham & Maestripieri, 2010; K E Holekamp, Ogutu, Dublin, Frank, & Smale, 1993; Koenig,

70   Borries, Caselli, & Lu, 2013; Samuels et al., 1987). Others simply report directly observing

71   changes (Dettmer, Woodward, & Suomi, 2015; Ehardt & Bernstein, 1986; Perry, 1998). Most

72   studies focus on rank reversals, in which two individuals reverse their dominance relationship.

73   For some, rank reversals are only identified if they are observed to be persistent (Chikazawa,

74   Gordon, Bean, & Bernstein, 1979; Samuels et al., 1987), but it is often unclear whether or not

75   transient cases are included as rank reversals. The Elo-rating method has been used to calculate

76   changes in numerical dominance scores for individuals by progressively updating scores after

77   each new interaction or after each discreet period (Albers & de Vries, 2001; McDonald & Dillon,

78   2015; Neumann et al., 2011). In light of these varied approaches, it would be useful to have an

79   overarching framework for assessing the dynamics of dominance hierarchies.

80   To guide research on the dynamics of dominance, we propose the concept of a

81   'longitudinal hierarchy', which is the characterization of a single, latent hierarchy and it's

82   dynamics over time (Fig. 1). In this framework, the total length of the study is divided into time

83   periods. These time periods could be long enough to encompass many interactions, shortened

84   to contain only one interaction (i.e., progressively updating ranks after each interaction;

85   Neumann et al., 2011), or nested such that each period includes the previous periods plus some

86   new interactions (e.g., McDonald & Dillon, 2015). Although any period delineation approach is

87   consistent with the longitudinal hierarchy framework, some approaches are likely to work

88   better than others depending on the ranking method used, data availability, and the biology of

89   the study organism. Whereas static hierarchies describe the rank (r) of each individual (i) during

90   a single period (t), longitudinal hierarchies estimate both the rank of each individual in multiple

91   time periods (r_{it}, r_{it+1}, etc.) and the rank dynamics of all individuals between time periods (Δ_{it},

92   Δ_{it+1}). Rank dynamics for any single individual comprise both the sum of its changes in rank due

93   to rank reversals and those that occur in the absence of rank reversals, such as those due to

94   demographic events. The amount of rank change an individual undergoes due to demographic

95   events is thus calculated as the total change by one individual from one period to the next

96   minus the number of group-mates with which that individual reversed its dominance

97   relationship from one period to the next. Hence the key to estimating the dynamics of a latent

98   hierarchy is to estimate the number of rank-reversals taking place.

99   Following from this framework, the best approach to studying the dynamics of

100   dominance is to construct longitudinal hierarchies that most accurately estimate both rank

101   orders of individuals (henceforth, *r*) and rank dynamics (henceforth, Δ). Although rank dynamics

102   are likely to be extremely important to the biology of social animals, a lack of guidelines on how

103   to reliably estimate Δ currently inhibits our ability to study them. Furthermore, whereas much

104   work has investigated the performance of different ranking algorithms at optimally estimating

105   r, no work to date has explicitly investigated the performance of these algorithms at estimating

106   Δ.

107   The primary challenge in estimating longitudinal hierarchies is to ensure the accuracy of

108   estimates of the dynamics of the hierarchy (Δ). An obvious approach to estimating Δ is to

109   estimate ranks for all individuals at time *t* and time *t+1* and then infer Δ from the difference

110   between the two. However, this approach is likely to overestimate the amount of rank change

111   in the hierarchy if the estimates of the rank order at time *t* and *t+1* are calculated

112   independently, because slight inaccuracies in the rank orders at each time point can be

113   amplified to suggest rank reversals that never took place (Fig 2). In light of limited sample size,

114   sampling error, and natural variation in the outcome of dominance interactions, the problem is

115   how best to discriminate between constant and shifting dominance ranks. We propose that an

116   optimal procedure for estimating Δ in a longitudinal hierarchy should incorporate an inertial

117   tendency for individuals to maintain their ranks from one time period to the next unless the

118   data suggest otherwise. This inertial tendency is justified by the very definition of dominance,

119   which includes temporal stability as an essential component (Drews, 1993). In fact, the notion

120   that dominance hierarchies reduce uncertainty about the outcomes of contests between group

121   members (Beaulieu et al., 2014; Mendonça-Furtado et al., 2014) assumes that the state of the

122   hierarchy at a given time is predictive of future interactions. Points-based methods that

123   continuously update scores after new interactions, such as the Elo-rating method, already

124   incorporate this inertial tendency, because the scores of individuals at any time point reflect

125   the outcomes of recent interactions as well as historical interactions. The simplest approach to

126   incorporating inertial tendency into a dynamic ranking method is to use the rank-order from

127   one study period as an initial order for subsequent periods. The rank order during any given

128   study period is thus determined to be the order from the previous study period amended by

129   new data from the current study period.

130   The incorporation of demographic changes poses another critical challenge to the

131   construction of longitudinal dominance hierarchies. Although individuals leaving the group (due

132   to emigration or death) don't pose a major problem, individuals who join the group (due to

133   immigration or birth) need to be included in the hierarchy in appropriate places. One approach

134   is to have individual-specific 'burn-in' periods, where some early data for each new individual

135   are used to assign its starting rank, and that individual is then added to the hierarchy at the end

136   of the burn-in period. This burn-in paradigm is used on the whole hierarchy in the Elo-rating

137   method (Albers & de Vries, 2001; Neumann et al., 2011), and can result in significant amounts

138   of discarded data, which are particularly troubling when interaction data are sparse (Newton-

139   Fisher, 2017). A better approach uses prior knowledge of dominance correlates to inform the

140   placement of new individuals in an existing hierarchy. The success of this approach will vary

141   among species, and it will only be viable for species for which prior research has identified

142   reliable correlates of dominance rank. For example, in species with age- or size-based

143   hierarchies (e.g. Archie et al. 2006), new individuals can be added to the hierarchy based on

144   their body size relative to the sizes of other same-sex group members. Newton-Fisher (2017)

145   used prior information about individual dominance ranks to improve the effectiveness of the

146   Elo-rating in resolving a hierarchy of male chimpanzees. Not only does this approach eliminate

147   data lost from a burn-in period, but it also allows for systematic study of the causes of deviation

148   from expected rank based on rank correlates.

149   Here we use the longitudinal hierarchy framework to provide the first methodological

150   investigation of the dynamics of dominance hierarchies. We first describe the Conservative

151   Rank Assignment Method (CRAM), a novel method for estimating longitudinal hierarchies that

152   uses a matrix-reordering approach that incorporates prior knowledge of dominance correlates,

153   as well as an updating algorithm to reliably estimate both r and Δ. We then use both simulated

154   data and empirical data from a long-term study of spotted hyenas (*Crocuta crocuta*) to assess

155   the accuracy of *r* and Δ estimated by our new method in relation to the Elo method, a modified

156   Elo method informed by prior knowledge of dominance correlates, the percolation and

157   conductance method, and the David's score method. To demonstrate the general applicability

158   of using prior knowledge of dominance correlates to inform the creation of longitudinal

159   hierarchies, we conducted tests using dominance hierarchies of both philopatric female and

160   immigrant male hyenas, as these hierarchies are structured by two different dominance

161   correlates. We hope that this longitudinal hierarchy framework and the methodological

162   investigation we present here will encourage future research on the structure and dynamics of

163   non-egalitarian societies.

164

165   **METHODS**

166   **Conservative Rank Assignment Method (CRAM)**

167   *Overview*

168   CRAM is a matrix-reordering approach that overcomes the challenges to estimation of

169   longitudinal hierarchies by incorporating an inertial tendency for ranks to remain stable over

170   time, and by using prior knowledge of dominance correlates to inform the placement of new

171   individuals as they enter the hierarchy. The method's general approach is to divide the data

172   into periods that together span the entire duration of the study (Fig 1). In each period, an initial

173   order is determined that is identical to the order from the previous period, with any new

174   individuals added to the hierarchy according to some criteria determined *a priori,* and based on

175   prior knowledge of the biology of the organism. The outcome of agonistic interactions from the

176 current study period are then converted to a binary outcome matrix (McDonald & Dillon, 2015)

177 and used to identify dyads for which the observations from the current period are inconsistent

178 with the initial order. Akin to the I&SI method (H. de Vries, 1998; Schmid & de Vries, 2013),

179 individuals in inconsistent dyads are then moved iteratively to minimize the number of

180 inconsistencies in the matrix and all minimal-inconsistency orders are saved. Finally, the

181 optimal order for the current study period is selected as the minimal-inconsistency order that is

182 most similar to the initial order.

183

184 *Assumptions*

185   CRAM makes two key assumptions that guide the rank assignment process. First, the

186 hierarchy in question is assumed to be transitive and linear. Although this assumption may not

187 hold for all animal groups, we argue that it is reasonable for the majority of species. In a survey

188 of 172 groups of mammals, insects, fish, and birds, Shizuka and McDonald (2015) found that

189 transitive triadic motifs were common and that non-transitive (i.e., cyclical) triadic motifs were

190 rare, suggesting that transitive hierarchies are normative across phyla. Furthermore, although

191 some have argued for the pervasiveness of non-linear dominance hierarchies (Douglas et al.,

192 2017), the linearity metrics that support this claim are biased towards underestimating linearity

193 (Shizuka & McDonald, 2012), and are highly sensitive to the amount of missing information in

194 the dataset (Klass & Cords, 2011). Finally, the assumption of linearity is made implicitly or

195 explicitly in all ordinal ranking methods, including all those we investigate here. However, to

196 confirm this assumption, we recommend that researchers use the triangle transitivity measure

197   (t_{tri}; Shizuka & McDonald, 2012) to assess the linearity of their aggression networks before

198   proceeding with any linear rank ordering approach, including CRAM.

199   The second assumption made by CRAM is that the true ranks of individuals at

200   consecutive time points are non-independent. This assumption manifests itself in CRAM's

201   procedure for assigning ranks for all study periods after the first. That is, ranks are assumed to

202   have remained the same from one period to the next unless there are observations suggesting

203   that a rank change has taken place. Not only is this assumption justified by the definition of

204   dominance, but it is also made by the Elo method.

205

206   *Pre-processing*

207   Before beginning the main phase of CRAM, a few pre-processing steps must be taken by

208   the investigator. The first step is to divide the total duration of the study into periods.

209   Individuals are assigned a rank within each period, and an individual's rank may change from

210   period to period. The length of each period should depend on the research aim, the biology of

211   the organism, and the frequency of observation of the study organism. Furthermore, the

212   duration of the periods need not be identical; for example, some researchers have found

213   dividing a longitudinal study into periods based on demographic events to be a useful strategy

214   (Neumann, Assahad, Hammerschmidt, Perwitasari-Farajallah, & Engelhardt, 2010; Schülke,

215   Bhagavatula, Vigilant, & Ostner, 2010). Alternatively, to emulate the progressive evaluation

216   implemented in the Elo method, one could divide the study into periods such that each period

217   contained only a single interaction.

218   Next, the researcher must choose an *a priori* ordering rule based on previous research

219   on the study organism. This rule guides the formation of the starting order in the first period,

220   and determines where new individuals are initially added into the order in subsequent periods.

221   An ordering rule might be based on body size, weapon size, tenure in the study group, age/sex

222   class, individual variation in aggressiveness, maternal social position, or any other reliable

223   correlate of social rank in the study species. In the absence of an unambiguous ordering rule,

224   another ranking method could potentially be used to set an initial order.

225

226   *Main phase*

227   CRAM selects an optimal order for each period by modifying the order from the

228   previous period to minimize the proportion of dyadic relationships inconsistent with the data

229   while retaining as much of the structure from the previous period as possible. The algorithm

230   starts by creating an initial order for the current period based on the order from the previous

231   period. Dead or emigrated individuals are removed from the order and new individuals are

232   added to the hierarchy according to the *a priori* ordering rule.

233   Next, the algorithm adjusts this initial order by tabulating all outcomes of observed

234   agonistic interactions in a sociomatrix, and binarizing the matrix such that a 1 appears in cells

235   where the row individual beat the column individual more than vice versa, and a 0 appears in

236   cells where the row individual lost to the column individual more than vice versa. Cells for

237   dyads that don't interact or where both individuals win the same number of interactions also

238   receive 0s. Inconsistencies are identified as dyads with an observed winner that is not

239   consistent with the proposed order. In order to minimize overfitting of the data from a

240   particular study period, inconsistencies are ignored if there are no subsequent observations of

241   the dyad that corroborate the inconsistency. A transient inconsistent outcome in a dyadic fight

242   might occur, for example, when a low-ranking female defends her offspring against a higher-

243   ranking female; the apparent reversal in the rank relationship between the two adult females

244   indicated by their agonistic behavior does not necessarily persist after the specific fight in

245   question. This approach is similar to that taken by other researchers (Koenig et al., 2013;

246   Samuels et al., 1987). Because future observations are required to corroborate any observed

247   inconsistency, studies investigating the dynamics of dominance should exclude the final period

248   from their analysis, as no subsequent observations are available for that period.

249   Once persistent inconsistencies have been identified, the order is optimized in a multi-

250   step reordering process similar to that implemented in the I&SI method (H. de Vries, 1998;

251   Schmid & de Vries, 2013). In the first step the order is iteratively modified by randomly

252   selecting a single individual involved in an inconsistent dyad and moving it to a new location in

253   the hierarchy. All potential new placements are checked, and each new placement that

254   improves the order is saved as a minimally inconsistent order (MIO). This reordering process

255   then repeats 1000 times, each time selecting a new individual involved in an inconsistency from

256   the list of MIOs. Next, step 1 is repeated 100 times, producing a list of orders that each have

257   the minimal number of inconsistencies identified during the reordering processes. Finally, the

258   optimal order is selected from the list of MIOs by comparing the similarity between each MIO

259   and the initial order from the current study period. Ties between equally optimal orders are

260   broken by selecting a MIO that best fits, first, the observations from the study period, then

261   observations from subsequent study periods. If ties remain after these selection criteria are

262   applied, a final order is selected randomly from among the tied orders.

263

264   **Comparison of different rank-ordering methods**

265   *Rank-ordering methods*

266   We tested five rank-ordering methods using both empirical and simulated datasets. We

267   used two versions of the Elo-rating method, a method in which individuals gain or lose points

268   after each interaction according to the outcome of the interaction and the relative ranks of the

269   contestants (Albers & de Vries, 2001; Elo, 1978; Neumann et al., 2011). Here the "Informed Elo"

270   method used prior information about rank correlates to initialize starting scores for new

271   individuals, whereas the "Uninformed Elo" method did not. For both Elo implementations, the k

272   parameter, which determines the magnitude of point gain/loss after each interaction, was set

273   to 200 (following Sánchez-Tójar, Schroeder, & Farine, 2017). The order produced by the Elo

274   method was determined by the final scores from each study period. All individuals began each

275   study period after the first with their final scores from the previous period. Thus, the Informed

276   Elo method, like CRAM, incorporated both inertial tendency and prior knowledge of dominance

277   correlates to optimally estimate r and Δ, whereas the Uninformed Elo method only

278   incorporated the inertial tendency. Here both Informed and Uninformed Elo methods were

279   implemented with a modified version of the *elo_scores* function in the aniDom package in R

280   (Farine & Sanchez-Tojar, 2017; R scripts in Supplemental Materials). We also tested the David's

281   scores method (implemented in the EloRating package, Neumann & Kulik, 2014), which assigns

282   each individual a score based on the proportion of group-members it dominates, corrected for

14

283   the number of observations per dyad (David, 1987; Gammell, de Vries, Jennings, Carlin, &

284   Hayden, 2003). Both Elo and David's scores methods produce numerical rank scores, which are

285   more useful for addressing some research questions than a simple rank-order (Albers & de

286   Vries, 2001; Neumann et al., 2011). Our current longitudinal hierarchy framework doesn't

287   support the use of non-ordinal ranks, however, so we converted scores used with both

288   methods to ordinal ranks. The fourth method we tested was the percolation and conductance

289   (PERC; Fujii et al., 2016) method, a promising new rank-ordering approach that uses transitive

290   network motifs to inform rank assignment. The PERC method was run with a maximum path

291   length of four; if there were no paths of length four, this value was decreased to match the

292   data. Neither PERC nor the David's score method incorporate an inertial tendency for ranks to

293   remain constant from one period to the next, so we expected both to overestimate Δ. Finally,

294   we tested CRAM, which was parameterized to perform 100 attempts at finding a best order for

295   each study period with up to 1000 reordering steps per attempt. We elected not to test the I&SI

296   method in our main study, despite its similarity to CRAM, because it has prohibitively slow

297   runtimes, especially with sparse data, and because our method is a direct extension of I&SI. All

298   methods and statistics were implemented in the statistical software R, version 3.4.0 (R

299   Development Core Team, 2017).

300

301   *Simulated data*

302   To compare performance among these methods in a situation with a known true

303   hierarchy, we simulated 10 longitudinal hierarchies, each comprised of 20 individuals and

304   spanning 20 study periods. We simulated minor shifts in the hierarchy through randomly

305   determined rank reversals. We simulated hierarchies both with moderate amounts of changes

306   of small magnitude and with many changes of large magnitude to assess the performance of

307   each method under varying conditions. In each study period, the number of individuals

308   undergoing rank changes was randomly selected from a Poisson distribution (lambda = {1, 5}).

309   The number of rank positions each individual moved up or down the hierarchy during a rank

310   change was drawn from an exponential distribution (rate = {1, 0.5}), and a random sample from

311   a binomial distribution (prob = {0.5, 0.5}) determined whether they moved up or down**.**

312   Additionally, we simulated the demographic changes of birth or immigration by including a 20%

313   probability that a new individual would enter the hierarchy at a random location in each study

314   period. We did not simulate emigration or death because individuals who leave the group are

315   no longer assigned ranks.

316   Simulated interactions were generated using the *generate_interactions* function from

317   the *aniDom* package, with parameters a = 15 and b = 0, both ID- and rank-based biases, and

318   generating 20 interactions per individual (Farine & Sanchez-Tojar, 2017). This produced

319   hierarchies similar to those found in gregarious vertebrates, in which (1) higher-ranked

320   individuals were very likely to win interactions, (2) this likelihood increased with the rank

321   difference between the two individuals, and (3) individuals varied in the number of interactions

322   in which they engaged (see Farine & Sanchez-Tojar, 2017; Sánchez-Tójar, Schroeder, & Farine,

323   2017 for details). Because ranking methods are highly sensitive to the proportion of 'unknown'

324   dyads, or dyads for which there are no observed behavioral interactions during the study period

325   (H. de Vries, 1998; Gammell et al., 2003; Klass & Cords, 2011; Neumann et al., 2011), we tested

326   each method with three different proportions of missing data. Henceforth we will refer to this

327 proportion as "*% unknowns*", following Klass & Cords (2011). To do this, we removed

328 observations from randomly selected dyads until the *% unknowns* matched a predetermined

329 value, which we set to low (20%), medium (50%) or high (80%). The simulation occasionally

330 produced more than 20 *% unknowns* before any data were removed, in which case we simply

331 left all data intact in the 'low' condition. For the methods incorporating prior knowledge of

332 dominance correlates, we simulated an arbitrary rank-associated trait (e.g., body size) value

333 between 0 and 100 for each individual such that the Spearman's rank correlation between the

334 trait value and the true ranks matched a specified value, which we tested at r= 0.9, 0.7, 0.5, and

335 0.2. This rank-associated trait was used as the initial order for both CRAM and the Informed Elo

336 method, and was also used in placing new individuals in the hierarchy with these two methods.

337 For the Informed Elo method, the initial scores for individuals were arranged from -1900 to

338 1900 by steps of 200, and ordered by the rank-associated trait. To ensure our results were not

339 biased by the starting scores, we also ran a test with initial scores ranging from -475 to 475 by

340 steps of 50 (Supplemental Materials Figs. S7, S8). For the Uninformed Elo method, each

341 individual was assigned an initial score of 0 and new individuals entering the hierarchy were

342 assigned the overall mean score at that time step. For the Informed Elo method, new

343 individuals entered the hierarchy according to their simulated trait value and with an Elo score

344 set to the mean of the two individuals adjacent in the hierarchy.

345   To assess the performance of each method, we tabulated the number of dyadic

346 relationships identified correctly and incorrectly by each method across study periods, and used

347 a binomial GLMM to test the effect of method on the proportion of dyads correctly identified.

348 We also tabulated the numbers of dyads each method identified as changing with respect to

349   their dominance order and those identified as remaining static, and we used a binomial GLMM

350   to test the effect of method on Δ. Because CRAM uses future observations to corroborate

351   changes observed in any given study period, we excluded the last period from our analyses. The

352   first period was excluded from the analysis of Δ because no rank reversals could take place in

353   the first period. We included random effects for simulation run and study period nested within

354   simulation run for all analyses.

355

356   *Empirical data*

357   We tested the performance of the five ranking methods on 25-year hierarchies of

358   immigrant male and philopatric female spotted hyenas. Male and female hierarchies were

359   analyzed separately because they follow different but predictable patterns, which allowed for

360   testing CRAM and the Informed Elo method with two different dominance correlates. Like the

361   pattern observed in many cercopithecine primates, rank acquisition in female spotted hyenas

362   typically follows a pattern of maternal rank inheritance with youngest ascendancy (Kay E

363   Holekamp & Smale, 1991; Kawamura, 1958), wherein each new female acquires the rank

364   immediately below that of her mother and above her older sisters. In contrast, immigrant male

365   spotted hyenas queue for dominance, such that males with the longest tenure in the group

366   occupy the highest ranks (East & Hofer, 1993; Smale, Nunes, & Holekamp, 1997). We used

367   these two principles to add new females and males to their respective hierarchies with both

368   methods. New females in the Informed Elo method were added to the female hierarchy with a

369   score of one less than their mothers (to mimic maternal rank inheritance with youngest

370   ascendancy), and new males were added to the bottom of the male hierarchy with a score of

371   200 less than the next newest male (to mimic queueing).

372   The empirical data were collected by personnel of the Mara Hyena Project in the Masai

373   Mara National Reserve in southern Kenya (Frank, 1986) from 1989 to 2015. Contest outcomes

374   were recorded using all-occurrence sampling (Altmann, 1974) of agonistic interactions observed

375   during daily morning and evening observation sessions, and genealogical relationships between

376   mothers and daughters were inferred based on genotyping and observations of nursing

377   behavior. Individuals were identified based on their unique spot patterns, and new immigrant

378   males were considered to have immigrated once they had been observed interacting with

379   members of the clan on at least three different occasions. We divided the longitudinal datasets

380   into calendar year-long periods stretching from 1989 to 2013; given that hyenas breed and

381   disperse throughout the year, calendar year represents a fair but arbitrary break point. Average

382   group size was 24.23 (16-52) adult females, and there were on average 4.6 (1-14) females

383   recruited per year and 3.56 (0-13) females lost per year. We calculated ranks for all females at

384   least 1.5 years old at the start of each year and all immigrant males. The 1.5 year age threshold

385   was selected based on previous work indicating that the process of rank acquisition among

386   maturing females is complete at around 1.5 years of age (Smale, Frank, & Holekamp, 1993). For

387   both Elo and CRAM methods, ranks for the first study year were set according to observations

388   made by Laurence Frank, KEH and Laura Smale prior to 1989 using a strategy similar to that

389   described by Martin & Bateson 1993 (Frank, 1986). We observed an average of 10.13

390   interactions per female per year (1.25-26.43), leading to a mean *% unknowns* per year of 65.2

391   (range 44.2-85). There were on average 18.81 (11-26) adult males, and there were on average

392   3.84 (1-9) individuals recruited and 3.84 (0-13) individuals lost per year. Data from males were

393   significantly more sparse, with an average of 3.64 interactions per male per year and a mean *%*

394   *unknowns* per year of 83.05 (65.6-96.67). Although these values are lower than were depicted

395   in a previous study from the same population (Kay E Holekamp & Smale, 1993), the adjacency

396   matrix depicted earlier was created from three years of interactions, whereas the current study

397   separated interactions by year.

398   We assessed performance of each method based on three criteria: (1) the number of

399   rank-reversals from year to year, (2) the proportion of outcomes observed during a particular

400   year that were consistent with the order produced by the method, and (3) the proportion of

401   outcomes observed during the following year that were consistent with the order produced by

402   the method. Based on reports in the literature about the societies of spotted hyenas and other

403   species with maternal rank inheritance, we expect the hierarchy to be highly stable (East &

404   Hofer, 1993; Frank, 1986; K E Holekamp & Smale, 1991). Methods that produce orders that

405   change frequently from year to year and that are inconsistent with future data are likely to

406   suffer from overfitting. To test the amount of change from year to year estimated by each

407   method, we coded each dyad as a Bernoulli trial according to whether the dyadic relationship

408   had changed from the previous year. We used a binomial GLM to test the effect of method on

409   the yearly proportion of dyadic changes in the longitudinal hierarchy. Similarly, we recorded the

410   proportion of aggressive interactions from the current year and from the subsequent year

411   according to whether their outcomes were consistent with the order generated by the current

412   year. Here again, we used a binomial GLM to test the relationship between ranking method and

413   the proportion of data consistent with the rank orders produced. In all analyses we used Tukey

414   post-hoc tests (R package *multicomp,* Bretz et al. 2016) to investigate pairwise differences

415   between specific methods.

416

417   **RESULTS**

418   *Simulated data*

419   Both *% unknowns* and ranking method had strong effects on the accuracy of orders

420   produced (Fig. 3-4) and on the estimated dynamics in the simulated dataset (Fig. 5-6). As the *%*

421   *unknowns* increased, the methods estimated longitudinal hierarchies with more rank reversals

422   (ß = 1.723, p < 0.0001) and less accurate rank orders (ß = -2.134, p < 0.0001). The differences in

423   accuracy of *r* were relatively small among most methods, but CRAM and the Elo methods

424   estimated *r* with the highest accuracy (Fig. 3c, 4c).

425   Although differences in the accuracy of *r* estimates were small, the methods differed

426   importantly in the accuracy of their estimates of Δ. When tested on more stable simulated

427   hierarchies, all methods overestimated the number of rank reversals from one period to the

428   next, and CRAM produced the most accurate estimate of Δ (Fig. 5). Furthermore, this difference

429   was not small; the Informed Elo method did second best, and it was 3.49 times as likely to

430   identify a rank reversal as was actually observed, whereas CRAM was only 1.20 times as likely.

431   When tested on highly unstable simulated hierarchies, the Informed Elo method had the most

432   accurate estimate of Δ, and CRAM underestimated Δ (Fig. 6). In both cases, methods that

433   incorporated inertial tendency estimated Δ most accurately, whereas methods without an

434   inertial tendency consistently overestimated the dynamics of the latent hierarchy. Finally,

435   reducing the informative value of the dominance correlate (by reducing its correlation with the

436   true order from 0.9 to 0.7, 0.5 and 0.2) somewhat reduced the performance of the methods

437   using this information, but did not qualitatively change the pattern of results (Figs. S1-S6),

438   suggesting that the Informed Elo method and CRAM were robust to variation in the accuracy of

439   the dominance correlate.

440

441   *Empirical data*

442   In tests with the empirical dataset on philopatric females, method had strong effects on

443   both the amount of change in the orders identified (Fig. 7) and on the fit between the identified

444   orders and the observed data (Fig. 8). Pairwise comparisons (Table S12) revealed that CRAM

445   produced hierarchies with the fewest rank-reversals from year to year (Fig. 7f; $p < 0.0001$).

446   Furthermore, although the PERC method produced orders that were most consistent with the

447   observations for the current year (Fig. 8a; $p < 0.001$; Table S10), these orders were poorly

448   predictive of the outcome of observations from the following year (Fig. 8b; Table S11), and

449   changed frequently from year to year (Fig. 7d). CRAM produced orders that were consistent

450   with the current data and most consistent with future data (Fig. 8; $p < 0.0001$). As in the

451   simulated study, methods that incorporated inertial tendency estimated significantly fewer

452   rank-reversals than methods without inertial tendency. Although we don't know the true

453   number of rank-reversals taking place in the latent hyena hierarchies, the methods that

454   incorporated inertial tendency produced more biologically realistic estimates of hierarchy

455   dynamics (Fig 7).

456   Despite the fact that male and female hierarchies in spotted hyena clans are organized

457   based on entirely different rules (tenure and maternal rank inheritance, respectively), results

458 from the immigrant male data corroborated our findings from females. Here again, CRAM

459 produced the hierarchy with fewest rank-reversals (Fig. 9; p < 0.001; Table S15). The patterns of

460 consistency between current and future interactions were similar to those observed in females;

461 however, not all difference that were significant in females were significant in males (Fig. S9),

462 and CRAM and the Informed Elo method were tied as the methods that produced orders that

463 were most consistent with future interactions (p = 0.979; Table S14). Furthermore, the

464 sparseness of the data used to derive the male hierarchies limits the conclusions that can be

465 drawn from this analysis.

466

467 **DISCUSSION**

468 Data from empirical and simulated datasets suggest that CRAM and the Informed Elo

469 method produce the most accurate longitudinal hierarchies. We found that CRAM was the most

470 accurate at estimating longitudinal hierarchies from less dynamic latent hierarchies (Figs. 3,5),

471 whereas the Informed Elo method did better when estimating longitudinal hierarchies from

472 highly dynamic latent hierarchies (Figs. 4,6). Results from our empirical study confirm that

473 CRAM produces a more conservative estimate of hierarchy dynamics than the Informed Elo

474 rating. Because we expect the latent hierarchy of spotted hyenas to have few rank reversals

475 (Hofer & East, 2003; Kay E Holekamp, Smith, Strelioff, Van Horn, & Watts, 2012), CRAM's

476 estimate of the hierarchy dynamics in this case may be more accurate. However, our lack of

477 knowledge of the true state of the latent hierarchy prevents us from being certain.

478 Most generally, we found that the incorporation of inertial tendency is vital to the

479 accurate estimation of longitudinal hierarchies because methods that calculate orders *de novo*

480   each period tend to overestimate the dynamics of the latent hierarchy (Figs 2, 4, 6, 7). We

481   conclude that the use of methods without an inertial tendency to characterize longitudinal

482   hierarchies is inappropriate. Furthermore, the fact that the two best methods incorporated

483   prior information on dominance correlates suggests that this information also significantly

484   improves the ability to estimate accurate longitudinal hierarchies. Obviously, this strategy is

485   dependent upon a previously identified dominance correlate or acquisition rule (e.g., maternal

486   rank inheritance with youngest ascendancy), and some organisms may not yet be sufficiently

487   well-studied to identify such a pattern, or no such pattern may exist at all. In these cases, use of

488   CRAM or the Informed Elo method is not feasible, so we expect the standard Elo method to be

489   most appropriate. However, we foresee the incorporation of prior knowledge of dominance

490   correlates to be useful for a wide variety of organisms; hierarchies based on status badges (e.g.,

491   Tibbetts & Dale 2004), tenure- or inheritance-based rank acquisition (Lea, Learn, Theus,

492   Altmann, & Alberts, 2014), or hierarchies based on size, age or weapon size (Archie et al., 2006;

493   Côté, 2000) may all provide the basis for such an approach. Finally, as found elsewhere (Klass &

494   Cords, 2011), the *% unknowns* negatively influenced the performance of each method. Because

495   of the sensitivity of ranking methods to missing information, it is important to report the *%*

496   *unknowns* from datasets used to derive dominance measures, and caution should be taken

497   when interpreting output from methods supplied with very sparse data.

498   What explains the different estimates of hierarchy dynamics between CRAM and the

499   Informed Elo method, given that both incorporate inertial tendency and prior information of

500   dominance correlates? One possible explanation is that points-based methods like Elo-rating

501   suffer from conflation of dominance and aggression, which could lead to overestimation of

502   rank-reversals. Although these two behavioral traits are highly related, they are not identical,

503   and aggressive tendencies can vary independent of dominance status (Buwalda, Koolhaas, & de

504   Boer, 2017; Kay E Holekamp & Strauss, 2016; Searcy & Wingfield, 1980). Consider two

505   individuals who have adjacent dominance ranks but vary in their aggressiveness. The more

506   aggressive individual may engage in more contests with lower-ranked group members, and

507   consequently gain more points than the less aggressive individual, even though the less

508   aggressive individual may have a higher dominance rank. This is especially likely in species

509   where group members are spatiotemporally dispersed, as in fission-fusion societies or in groups

510   that aggregate loosely. Consequences of individual differences in aggressive tendencies may be

511   magnified when individuals have more choice regarding which group mates become their

512   contest opponents.

513   The fact that the structure of the latent hierarchy influenced whether CRAM or the

514   Informed Elo method was the most accurate highlights a major issue facing the study of

515   dominance hierarchies: the optimal methods for estimating the structure of hierarchies depend

516   upon that structure. For example, Sánchez-Tójar, Schroeder, & Farine, (2017) found that the

517   optimal method for estimating rank orders of a static hierarchy depends on hierarchy steepness

518   (i.e., the strength of the effect of rank difference on the probability of dominants winning

519   contests), and we have yet to evaluate the effect of hierarchy steepness on the efficacy of

520   methods for estimating longitudinal hierarchies. Given that properties of latent hierarchies can

521   only be measured by ranking algorithms whose performance depends upon these properties,

522   how does one know which method to use for a given dataset? One approach to solving this

523   problem may be to identify hierarchical measurements that could provide a preliminary

524   assessment of the latent hierarchy and thus guide selection of an optimal ranking algorithm.

525   For example, Neumann et al., (2011) introduced a measurement of the overall stability of a

526   hierarchy, *S*, that could be used to determine whether a hierarchy of interest was more

527   appropriately estimated using the Informed Elo method or CRAM. However, further research is

528   needed to explore which measurements are reliable estimates of latent hierarchical properties

529   and how optimal ranking algorithms map to values of these measurements. For now, previous

530   reports specific to the study organism could be used to provide a best guess about these

531   properties. For example, most vertebrate hierarchies are stable most of the time, although

532   periods of extreme instability can arise as a result of new individuals entering the group

533   (Broom, Koenig, & Borries, 2009). It may be that periods of upheaval and periods of stability are

534   best measured by different methods.

535   Assigning dominance ranks is a data-hungry and complex optimization problem, one

536   exacerbated when using long-term data. In this paper, we advance the concept of a longitudinal

537   hierarchy as a framework for investigating the dynamics of dominance hierarchies (Fig 1). A

538   longitudinal hierarchy characterizes a latent hierarchy through estimation of both rank orders

539   during distinct time periods and rank dynamics between time periods. Consequently, effective

540   methods for estimating longitudinal hierarchies should be evaluated in terms of their accuracy

541   at estimating both rank orders and rank dynamics. Despite its challenges, the estimation of

542   longitudinal dominance hierarchies allows for greater understanding of the forces that have led

543   to the evolution of unequal societies across taxa, and is thus worthy of further study. Long-

544   term data provide the opportunity to detect patterns that operate on a scale larger than would

545   to be detectable by short-term studies (Smith, Lehmann, Montgomery, Strauss, & Holekamp,

546 2017), and thus a reliable system for extracting dominance information from longitudinal data

547 will enable new perspectives on dominance. Our future work will use this framework to

548 investigate the dynamic structure of inequality and the role of individual traits, socioecological

549 factors, and social bonds in shaping animal societies.

550

551 **AUTHOR CONTRIBUTIONS**

552 EDS led the study design, the analyses, and the writing of the manuscript. KEH collected the

553 data and contributed significantly to study design and manuscript preparation. Both authors

554 give final approval for publication.

555

556 **DATA ACCESSIBILITY**

557 All data for this paper were either generating using the accompanying R scripts (online at

558 https://github.com/straussed/longitudinal_hierarchies) or will be made available on dryad

559 upon publication.
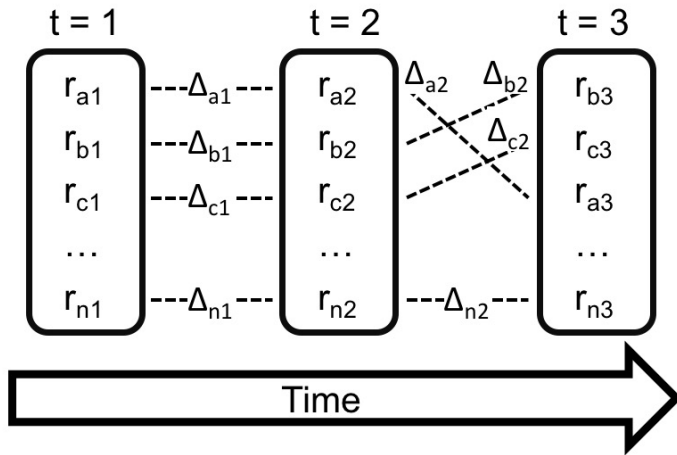
560

561 **ACKNOWLEDGEMENTS**

569

570

Figure 1. A longitudinal hierarchy characterizes the rank orders and dynamics of a single latent hierarchy over time and is comprised of rank orders at various time points and the dynamics between time points. Rank changes between time points can be caused by rank reversals, demographic events or both.
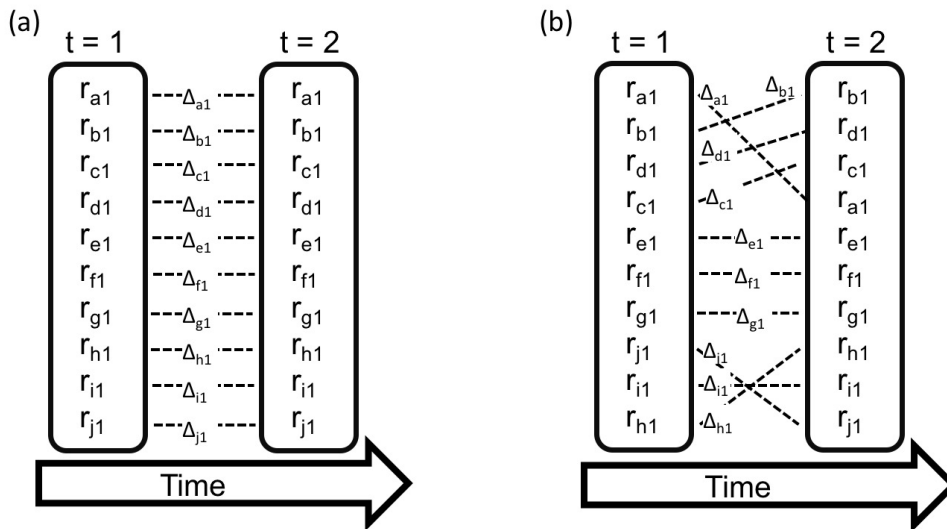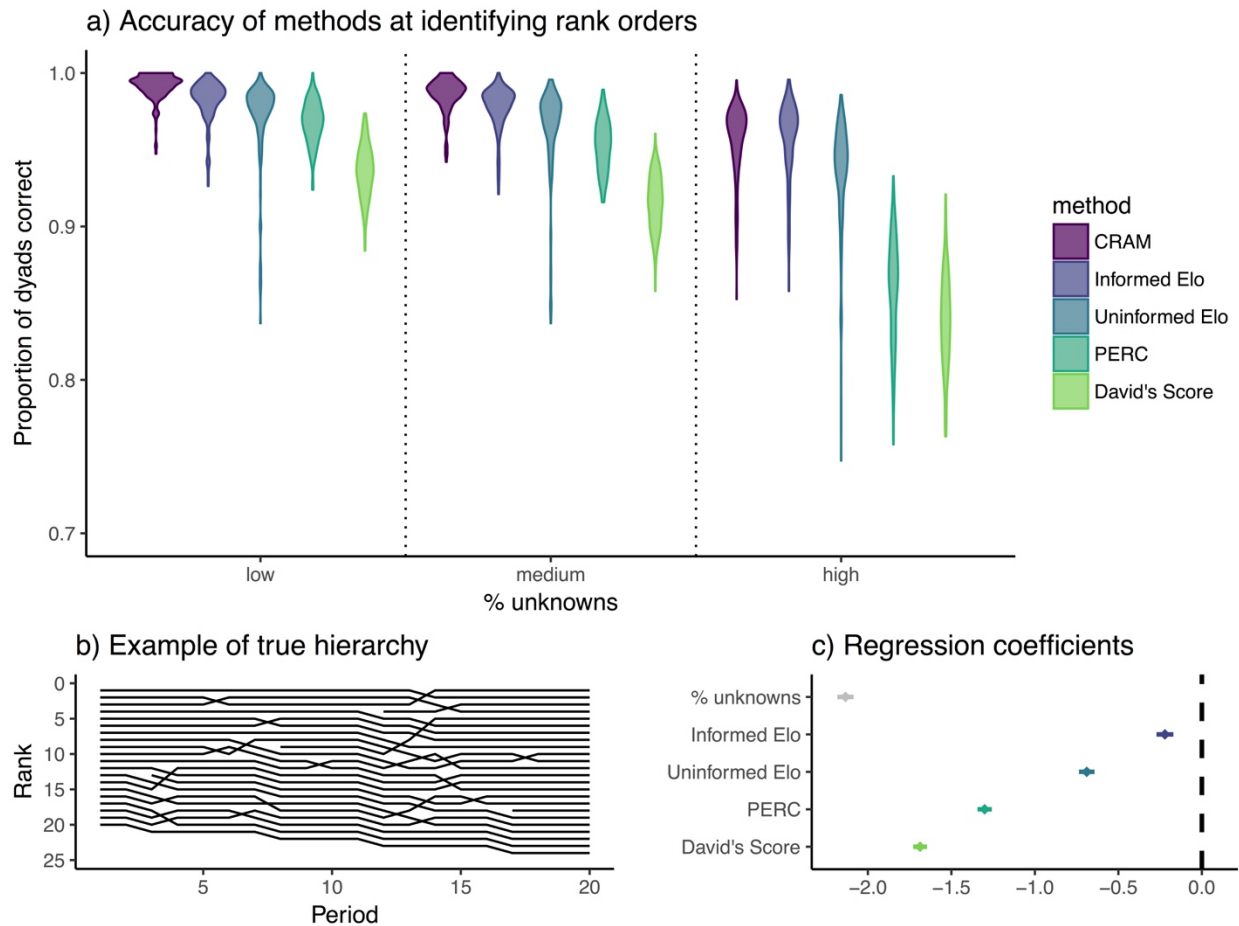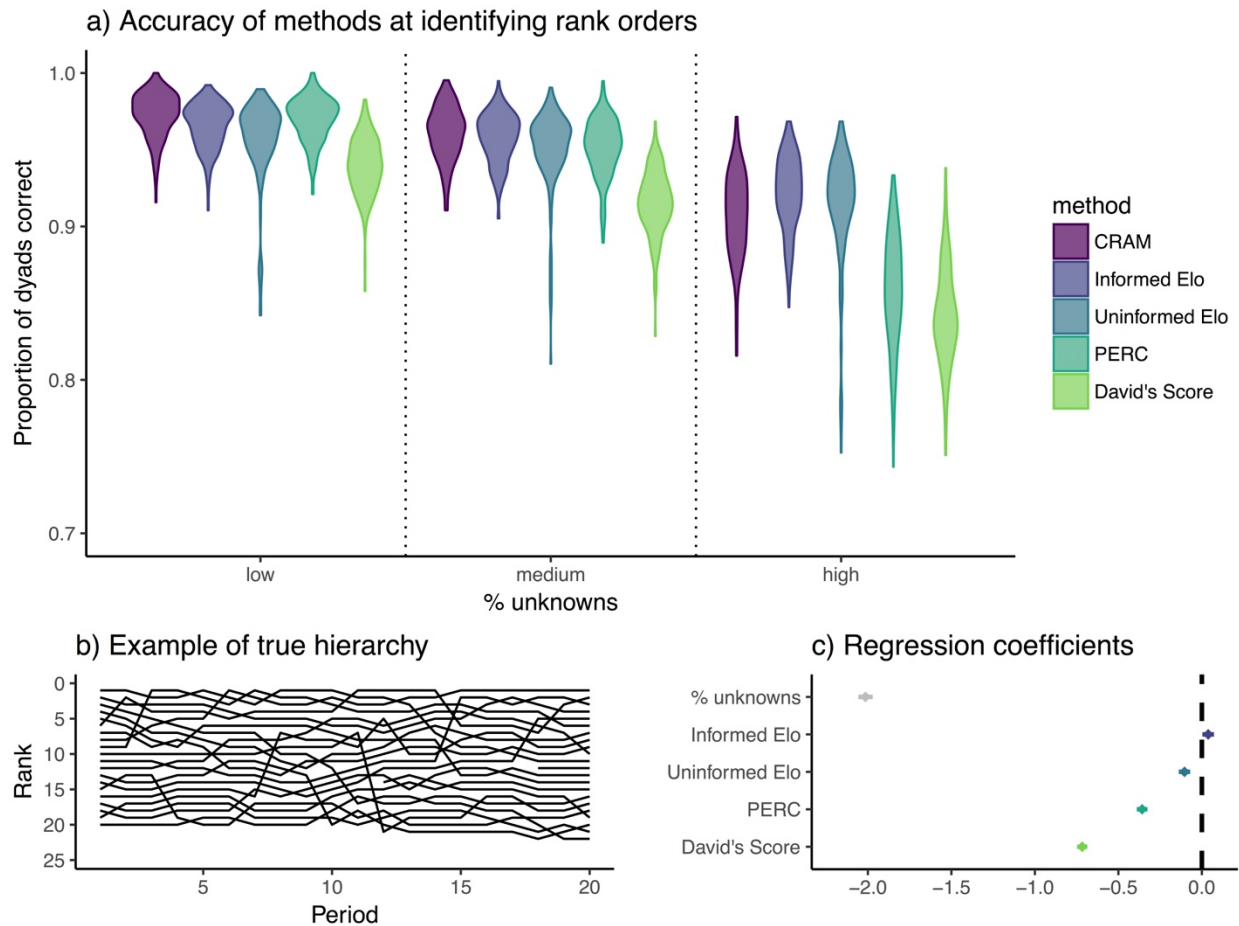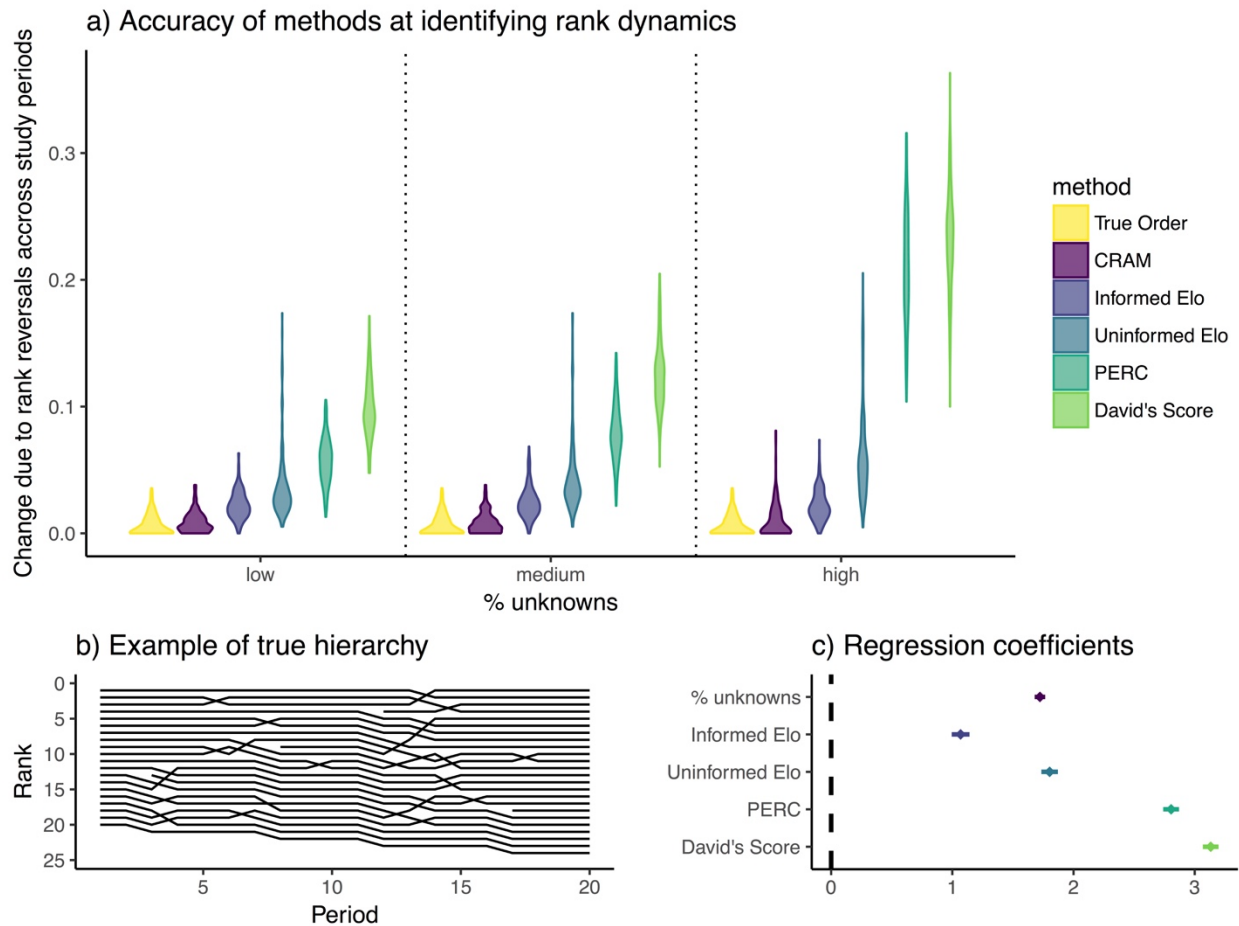
Figure 2. The true state of the hierarchy (a) and an imperfect longitudinal hierarchy estimating it (b). Rank orders in (b) have a Pearson's rank correlation with the true orders of 0.9. The longitudinal hierarchy in (b) overestimates the dynamics in the hierarchy as a result of small inaccuracies in rank orders at both t1 and t2.

Figure 3. a) The similarity between the true order from the simulated data and the orders identified by each method. Five rank-ordering methods were tested under conditions of low (20%), medium (50%), and high (80%) proportions of dyads with missing data. b) An example of one of the 10 true latent hierarchies, which were simulated to have few rank reversals of small magnitude. c) Regression coefficients and 95% confidence intervals from a binomial GLMM modeling the proportion of dyadic relationships identified correctly by the different methods for each study period, with data under three conditions of *% unknowns* (n = 2850 periods). CRAM is the reference level in this analysis, so estimates above or below 0 indicate methods that performed better or worse than CRAM.
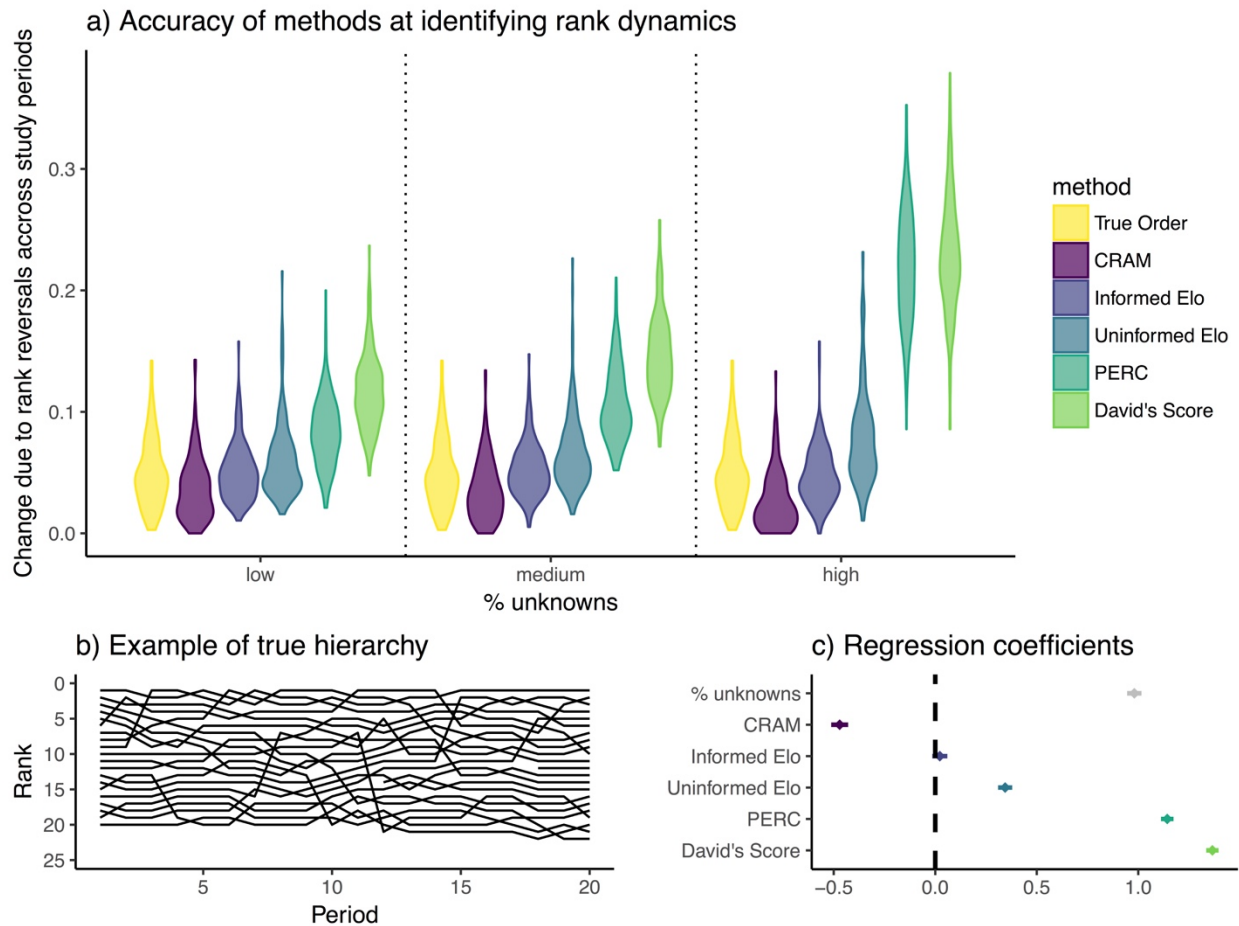
583

584

585

Figure 4. a) The similarity between the true order from the simulated data and the orders identified by each method. Five rank-ordering methods were tested under conditions of low (20%), medium (50%), and high (80%) proportions of dyads with missing data. b) An example of one of the 10 true latent hierarchies, which were simulated to have many rank reversals of large magnitude. c) Regression coefficients and 95% confidence intervals from a binomial GLMM modeling the proportion of dyadic relationships identified correctly by the different methods for each study period, with data under three conditions of *% unknowns* (n = 2850 periods). CRAM is the reference level in this analysis, so estimates above or below 0 indicate methods that performed better or worse than CRAM.

586

a) Accuracy of methods at identifying rank dynamics

b) Example of true hierarchy
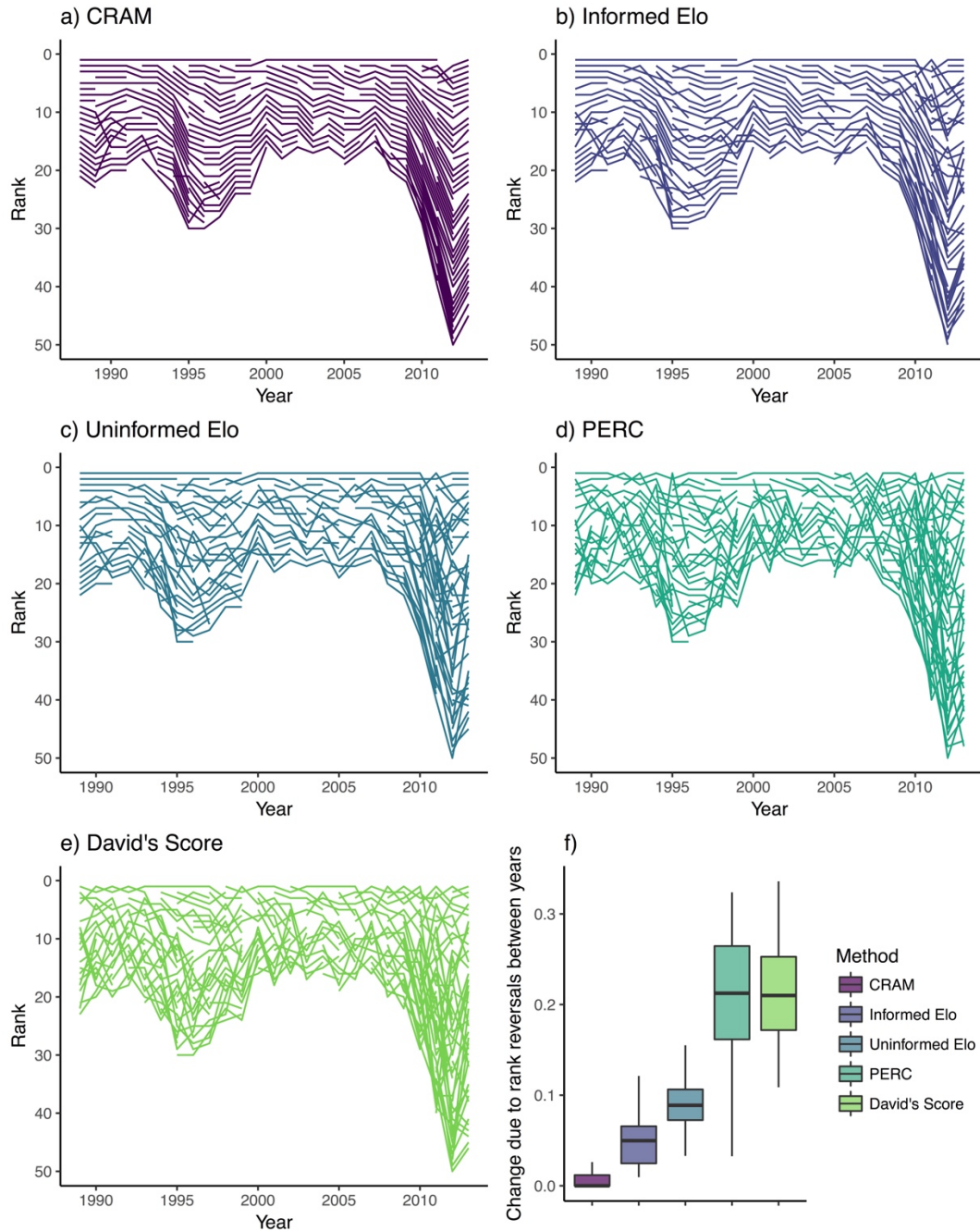
c) Regression coefficients

587

Figure 5. a) The similarity between the proportion of dyadic relationships that changed in the simulated data for each period and the proportions dyads that changed identified by each method. Five rank-ordering methods were tested under conditions of low (20%), medium (50%), and high (80%) proportions of dyads with missing data. b) An example of one of the 10 true latent hierarchies, which were simulated to have few rank reversals of small magnitude. c) Regression coefficients and 95% confidence intervals from a binomial GLMM modeling the proportion of dyads that changed identified by the different methods, with data under three conditions of *% unknowns* (n = 3240 periods). The True Order is the reference level in this analysis, so estimates above or below 0 indicate that methods over- or underestimated the dynamics of the hierarchy.

588

a) Accuracy of methods at identifying rank dynamics

b) Example of true hierarchy

c) Regression coefficients

589

Figure 6. a) The similarity between the proportion of dyadic relationships that changed in the simulated data for each period and the proportions dyads that changed identified by each method. Five rank-ordering methods were tested under conditions of low (20%), medium (50%), and high (80%) proportions of dyads with missing data. b) An example of one of the 10 true latent hierarchies, which were simulated to have many rank reversals of large magnitude. c) Regression coefficients and 95% confidence intervals from a binomial GLMM modeling the proportion of dyads that changed identified by the different methods, with data under three conditions of *% unknowns* (n = 3240 periods). The True Order is the reference level in this analysis, so estimates above or below 0 indicate that methods over- or underestimated the dynamics of the hierarchy.

590

591

592

Figure 7. (a-e) Visual depiction of the dominance hierarchies identified by each method from the 25-year empirical dataset from a wild population of philopatric female spotted hyenas. Each line represents the rank assigned to a single adult female over the course of its lifetime. Crossing lines indicate an identified rank reversal, whereas non-crossing lines indicate stability in the identified longitudinal hierarchy. Numbers of rank positions varied with size of the group over time. (f) The number of rank reversals across years in the hierarchies identified by each method shown in a-e, as measured by Kendall's rank correlation between orders in consecutive years.

593
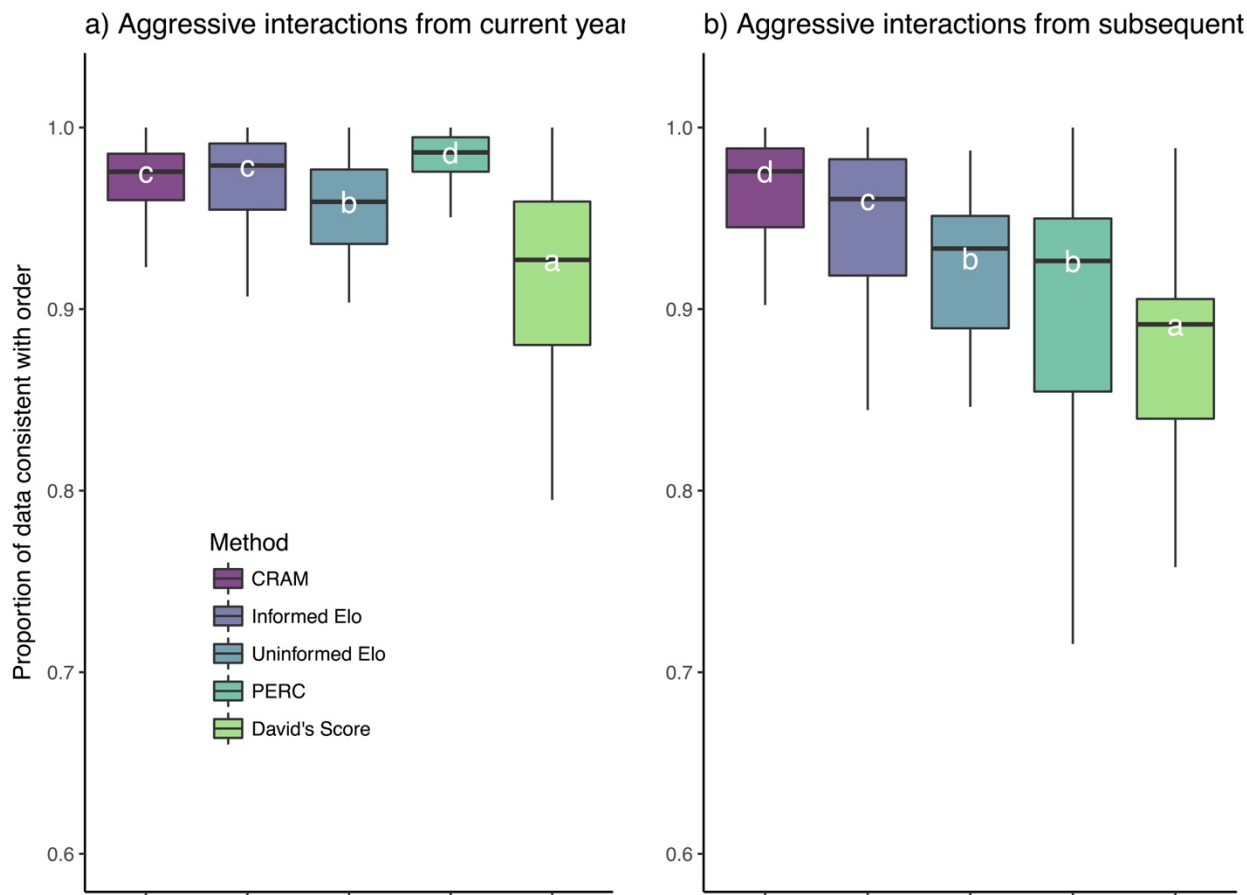
Figure 8. The proportion of observed outcomes of agonistic interactions between philopatric adult female spotted hyenas and the rank orders identified by each method. Data from (a) were from a current year in the study and were used to produce the order. Data in (b) were from the following study year. Disparities between the fit in (a) and (b) indicate overfitting. Letters indicate significantly different groups.

594

595
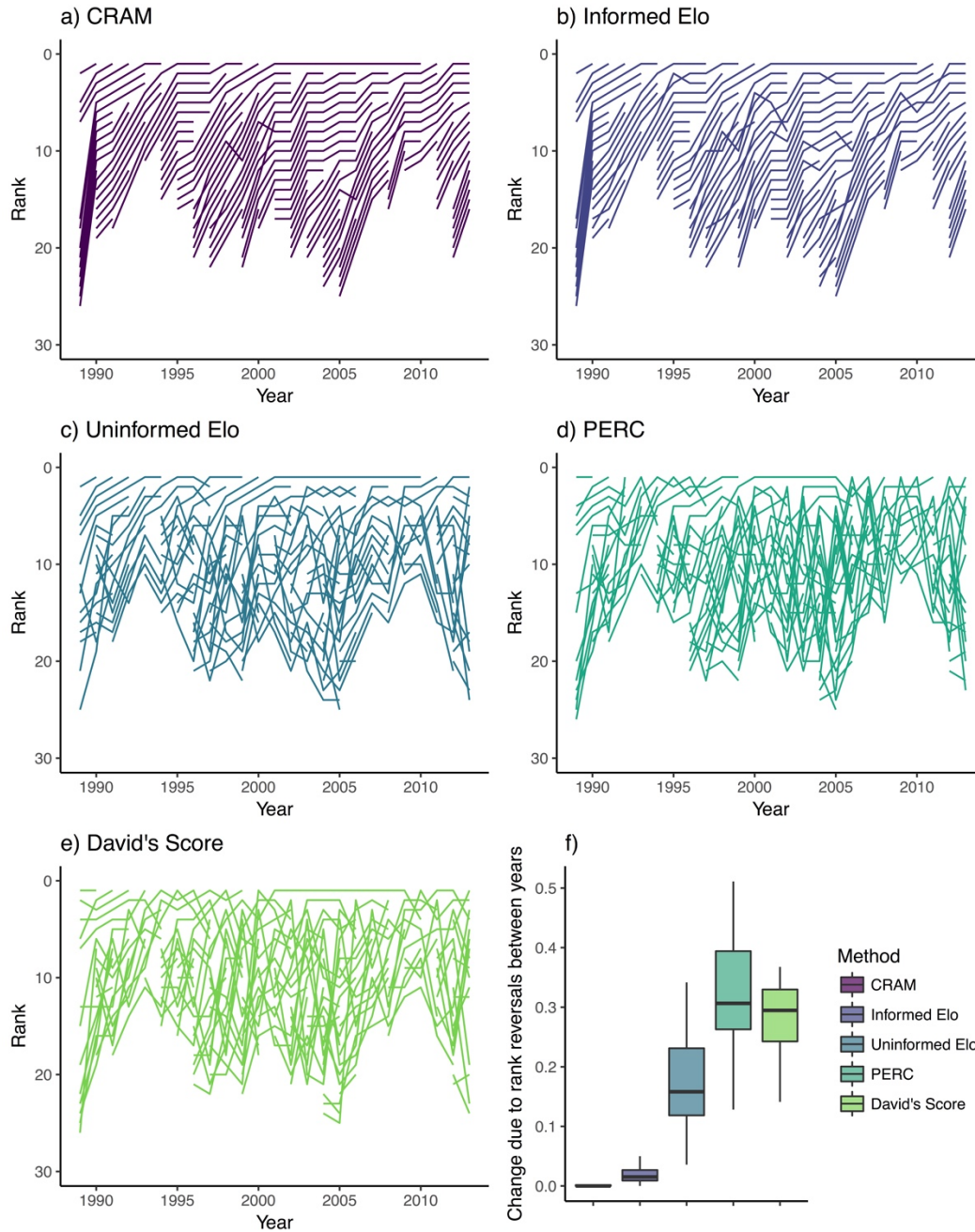
596

Figure 9. (a-e) Visual depiction of the dominance hierarchies identified by each method from the 25-year empirical dataset from a wild population of immigrant male spotted hyenas. Each line represents the rank assigned to a single adult male over the course of its lifetime. Crossing lines indicate an identified rank reversal, whereas non-crossing lines indicate stability in the identified longitudinal hierarchy. Numbers of rank positions varied with size of the group over time. (f) The amount of change across years in the hierarchies identified by each method shown in a-e, as measured by Kendall's rank correlation between orders in consecutive years.

597

598   **REFERENCES**

599   Albers, P. C. H., & de Vries, H. (2001). Elo-rating as a tool in the sequential estimation of
600   dominance strengths. *Animal*, *61*, 489–495. doi:10.1006/anbe.2000.1571
601   Altmann, J. (1974). Observational study of behavior: sampling methods. *Behaviour*, *49*(3), 227–
602   266.
603   Archie, E. A., Morrison, T. A., Foley, C. A. H., Moss, C. J., & Alberts, S. C. (2006). Dominance rank
604   relationships among wild female African elephants,< i> Loxodonta africana</i>. *Animal*
605   *Behaviour*, *71*(1), 117–127.
606   Beaulieu, M., Mboumba, S., Willaume, E., Kappeler, P. M., & Charpentier, M. J. E. (2014). The
607   oxidative cost of unstable social dominance. *Journal of Experimental Biology*, *217*(Pt 15),
608   2629–2632.
609   Boydston, E. E., Kapheim, K. M., Szykman, M., & Holekamp, K. E. (2003). Individual variation in
610   space use by female spotted hyenas. *Journal of Mammalogy*, *84*(3), 1006–1018.
611   Bretz, F., Westfall, P., Heiberger, R. M., Schuetzenmeister, A., & Scheibe, S. (2016). Package "
612   multcomp ."
613   Bro Jørgensen, J., & Beeston, J. (2015). Multimodal signalling in an antelope: fluctuating
614   facemasks and knee-clicks reveal the social status of eland bulls. *Animal Behaviour*, *102*,
615   231–239.
616   Broom, M., Koenig, A., & Borries, C. (2009). Variation in dominance hierarchies among group-
617   living animals: modeling stability and the likelihood of coalitions. *Behavioral Ecology*, *20*(4),
618   844–855.
619   Buwalda, B., Koolhaas, J. M., & de Boer, S. F. (2017). Physiology & Behavior Trait aggressiveness
620   does not predict social dominance of rats in the Visible Burrow System. *Physiology &*
621   *Behavior*, *178*, 134–143. doi:10.1016/j.physbeh.2017.01.008
622   Chapais, B., Girard, M., & Primi, G. (1991). Non-kin alliances, and the stability of matrilineal
623   dominance relations in Japanese macaques. *Animal Behaviour*.
624   Chikazawa, D., Gordon, T. P., Bean, C. A., & Bernstein, I. S. (1979). Mother-daughter dominance
625   reversals in rhesus monkeys (Macaca mulatta). *Primates*, *20*(2), 301–305.
626   Combes, S. L., & Altmann, J. (2001). Status change during adulthood: life-history by-product or
627   kin selection based on reproductive value? *Proceedings of the Royal Society B: Biological*
628   *Sciences*, *268*(1474), 1367–1373.
629   Côté, S. D. (2000). Dominance hierarchies in female mountain goats: stability, aggressiveness
630   and determinants of rank. *Behaviour*, *137*, 1541–1566.
631   David, H. A. (1987). Ranking from unbalanced paired-comparison data. *Biometrika*, *74*(2), 432–
632   436.
633   de Vries, D., Koenig, A., & Borries, C. (2016). Female reproductive success in a species with an
634   age-inversed hierarchy. *Integrative Zoology*.
635   de Vries, H. (1998). Finding a dominance order most consistent with a linear hierarchy: a new
636   procedure and review. *Animal Behaviour*, 1–17.
637   Dettmer, A. M., Woodward, R. A., & Suomi, S. J. (2015). Reproductive Consequences of a
638   Matrilineal Overthrow in Rhesus Monkeys, *352*(July 2014), 346–352.
639   doi:10.1002/ajp.22350
640   Douglas, P. H., Ngonga Ngomo, A. C., & Hohmann, G. (2017). A novel approach for dominance

641   assessment in gregarious species: ADAGIO. *Animal Behaviour*, *123*, 21–32.
642   doi:10.1016/j.anbehav.2016.10.014
643 Drews, C. (1993). The concept and definition of dominance in animal behaviour. *Behaviour*,
644   283–313.
645 East, M. L., & Hofer, H. (1993). Male spotted hyenas (*Crocuta crocuta*) queue for status in social
646   groups dominated by females. *Behavioral Ecology*, *12*(5), 558–68.
647   doi:10.1093/beheco/12.5.558
648 Ehardt, C. L., & Bernstein, I. S. (1986). Matrilineal overthrows in rhesus monkey groups.
649   *International Journal of Primatology*, *7*(2), 157–181. doi:10.1007/BF02692316
650 Elo, A. E. (1978). *The rating of chessplayers, past and present*. New York: Arco Pub.
651 Farine, D. R. (2017). When to choose dynamic versus static social network analysis. *Journal of*
652   *Animal Ecology*, (June), 1–11. doi:10.1111/1365-2656.12764
653 Farine, D. R., & Sanchez-Tojar, A. (2017). aniDom: Inferring Dominance Hierarchies and
654   Estimating Uncertainty. R package version 0.1.2.
655 Flies, A. S., Mansfield, L. S., Flies, E. J., Grant, C. K., & Holekamp, K. E. (2016). Socioecological
656   predictors of immune defences in wild spotted hyenas. *Functional Ecology*, *30*(9), 1549–
657   1557.
658 Fortunato, A., Queller, D. C., & Strassmann, J. E. (2003). A linear dominance hierarchy among
659   clones in chimeras of the social amoeba Dictyostelium discoideum. *Journal of Evolutionary*
660   *Biology*, *16*(3), 438–445.
661 Frank, L. G. (1986). Social organization of the spotted hyaena< i> Crocuta crocuta</i>. II.
662   Dominance and reproduction. *Animal Behaviour*, *34*(5), 1510–1527.
663 Fujii, K., Jin, J., Shev, A., Beisner, B., Mccowan, B., & Fushing, H. (2016). Using Percolation and
664   Conductance to Find Information Flow Cerntainty in a Directed Network.
665 Gammell, M. P., de Vries, H., Jennings, D. J., Carlin, C. M., & Hayden, T. J. (2003). David's score:
666   a more appropriate dominance ranking method than Clutton-Brock et al.'s index. *Animal*
667   *Behaviour*, *66*(3), 601–605.
668 Hasegawa, M., & Kutsukake, N. (2014). Bayesian competitiveness estimation predicts
669   dominance turnover among wild male chimpanzees. *Behavioral Ecology and Sociobiology*,
670   *69*(1), 89–99.
671 Higham, J., & Maestripieri, D. (2010). Revolutionary coalitions in male rhesus macaques.
672   *Behaviour*, *147*(13), 1889–1908.
673 Hobson, E. A., Avery, M. L., & Wright, T. F. (2013). An analytical framework for quantifying and
674   testing patterns of temporal dynamics in social networks. *Animal Behaviour*, *85*(1), 83–96.
675   doi:10.1016/j.anbehav.2012.10.010
676 Hofer, H., & East, M. L. (2003). Behavioral processes and costs of co-existence in female spotted
677   hyenas: a life history perspective. *Evolutionary Ecology*.
678 Holekamp, K. E., Ogutu, J. O., Dublin, H. T., Frank, L. G., & Smale, L. (1993). Fission of a spotted
679   hyena clan: consequences of prolonged female absenteeism and causes of female
680   emigration. *Ethology*, *93*(4), 285–299.
681 Holekamp, K. E., & Smale, L. (1991). Dominance acquisition during mammalian social
682   development: the "inheritance" of maternal rank. *American Zoologist*, *31*(2), 306–317.
683 Holekamp, K. E., & Smale, L. (1991). Dominance acquisition during mammalian social
684   development: the "inheritance" of maternal rank. *American Zoologist*.

685   Holekamp, K. E., & Smale, L. (1993). Ontogeny of dominance in free-living spotted hyaenas:
686   juvenile rank relations with other immature individuals. *Animal Behaviour*, *46*(3), 451–466.
687   Holekamp, K. E., Smith, J. E., Strelioff, C. C., Van Horn, R. C., & Watts, H. E. (2012). Society,
688   demography and genetic structure in the spotted hyena. *Molecular Ecology*, *21*(3), 613–
689   632.
690   Holekamp, K. E., & Strauss, E. D. (2016). Aggression and dominance: an interdisciplinary
691   overview. *Current Opinion in Behavioral Sciences*.
692   Kawamura, S. (1958). The matriarchal social order in the Minoo-B group. *Primates*, *1*(2), 149–
693   156.
694   Klass, K., & Cords, M. (2011). Effect of unknown relationships on linearity, steepness and rank
695   ordering of dominance hierarchies: simulation studies based on data from wild monkeys.
696   *Behavioural Processes*, *88*(3), 168–176.
697   Koenig, A., Borries, C., Caselli, A., & Lu, A. (2013). Effects of age, reproductive state, and the
698   number of competitors on the dominance dynamics of wild female Hanuman langurs.
699   *Behaviour*, *150*(5), 485–523.
700   Lea, A. J., Learn, N. H., Theus, M. J., Altmann, J., & Alberts, S. C. (2014). Complex sources of
701   variance in female dominance rank in a nepotistic society. *Animal Behaviour*, *94*, 87–99.
702   Martin, P., & Bateson, P. (1993). *Measuring behaviour: an introductory guide*. Cambridge, UK:
703   Cambridge University Press.
704   McDonald, D. B., & Dillon, M. E. (2015). Temporal changes in dominance networks and other
705   behaviour sequences. In J. Krause, R. James, D. W. Franks, & D. P. Croft (Eds.), *Animal
706   Social Networks* (pp. 61–72). Oxford, UK: Oxford University Press, USA.
707   Mendonça-Furtado, O., Edaes, M., Palme, R., Rodrigues, A., Siqueira, J., & Izar, P. (2014). Does
708   hierarchy stability influence testosterone and cortisol levels of bearded capuchin monkeys
709   (Sapajus libidinosus) adult males? A comparison between two wild groups. *Behavioural
710   Processes*, *109*, 79–88.
711   Murray, C. M., Mane, S. V, & Pusey, A. E. (2007). Dominance rank influences female space use
712   in wild chimpanzees , Pan troglodytes : towards an ideal despotic distribution. *Animal
713   Behaviour*, *74*, 1795–1804. doi:10.1016/j.anbehav.2007.03.024
714   Neumann, C., Assahad, G., Hammerschmidt, K., Perwitasari-Farajallah, D., & Engelhardt, A.
715   (2010). Loud calls in male crested macaques, Macaca nigra: a signal of dominance in a
716   tolerant species. *Animal Behaviour*, *79*(1), 187–193. doi:10.1016/j.anbehav.2009.10.026
717   Neumann, C., Duboscq, J., Dubuc, C., Ginting, A., Irwan, A. M., Agil, M., … Engelhardt, A. (2011).
718   Assessing dominance hierarchies: validation and advantages of progressive evaluation with
719   Elo-rating. *Animal Behaviour*, *82*(4), 911–921.
720   Neumann, C., & Kulik, L. (2014). EloRating: Animal Dominance Hierarchies by Elo Rating. *R
721   Package Version 0.43.*
722   Newton-Fisher, N. E. (2017). Modeling Social Dominance : Elo-Ratings , Prior History , and the
723   Intensity of Aggression. *International Journal of Primatology*, *38*, 427–447.
724   doi:10.1007/s10764-017-9952-2
725   Perry, S. (1998). A Case Report of a Male Rank Reversal in a Group of Wild White-faced
726   Capuchins, *39*(January), 51–70. doi:https://doi-
727   org.proxy2.cl.msu.edu/10.1007/BF02557743
728   Pinter-Wollman, N., Hobson, E. A., Smith, J. E., Edelman, A. J., Shizuka, D., de Silva, S., …

729   McDonald, D. B. (2014). The dynamics of animal social networks: analytical, conceptual,
730   and theoretical advances. *Behavioral Ecology*, *25*(2), 242–255. doi:10.1093/beheco/art047
731   Pusey, A., Williams, J., & Goodall, J. (1997). The influence of dominance rank on the
732   reproductive success of female chimpanzees. *Science*, *277*(5327), 828–831.
733   R Development Core Team. (2017). R: A language and environment for statistical computing.
734   Vienna, Austria: R Foundation for Statistical Computing.
735   Samuels, A., Silk, J. B., & Altmann, J. (1987). Continuity and change in dominance relations
736   among female baboons. *Animal Behaviour*, *35*(3), 785–793. doi:10.1016/S0003-
737   3472(87)80115-X
738   Sánchez-Tójar, A., Schroeder, J., & Farine, D. R. (2017). A practical guide for inferring reliable
739   dominance hierarchies and estimating their uncertainty. *Journal of Animal Ecology*, 1–15.
740   doi:10.1111/1365-2656.12776
741   Sapolsky, R. M. (2005). The influence of social hierarchy on primate health. *Science*, *308*(5722),
742   648–652.
743   Schjelderup-Ebbe, T. (1922). Contributions to the social psychology of the domestic chicken.
744   *Reprinted from Zeitschrift Fuer Psychologie*, *88*, 225–252.
745   Schmid, V. S., & de Vries, H. (2013). Finding a dominance order most consistent with a linear
746   hierarchy: an improved algorithm for the I&SI method. *Animal Behaviour*, *83*, 1097–1105.
747   Schülke, O., Bhagavatula, J., Vigilant, L., & Ostner, J. (2010). Social bonds enhance reproductive
748   success in male macaques. *Current Biology*, *20*(24), 2207–2210.
749   doi:10.1016/j.cub.2010.10.058
750   Searcy, A., & Wingfield, J. C. (1980). The Effects of Androgen and Antiandrogen on Dominance
751   and Aggressiveness in Male Red-Winged Blackbirds. *Hormones and Behavior*, *14*, 126–135.
752   Shimoji, H., Abe, M. S., Tsuji, K., & Masuda, N. (2014). Global network structure of dominance
753   hierarchy of ant workers. *Journal of the Royal Society, Interface / the Royal Society*, *11*(99),
754   20140599.
755   Shizuka, D., & McDonald, D. B. (2012). A social network perspective on measurements of
756   dominance hierarchies. *Animal Behaviour*, *83*(4), 925–934.
757   Shizuka, D., & McDonald, D. B. (2015). The network motif architecture of dominance
758   hierarchies. *Journal of the Royal Society, Interface / the Royal Society*, *12*(105), 20150080.
759   Smale, L., Frank, L. G., & Holekamp, K. E. (1993). Ontogeny of dominance in free-living spotted
760   hyaenas: juvenile rank relations with adult females and immigrant males. *Animal
761   Behaviour*, *46*(3), 467–477.
762   Smale, L., Nunes, S., & Holekamp, K. E. (1997). Sexually Dimorphic Dispersal in Mammals:
763   Patterns, Causes, and Consequences. In P. J. B. Slater, J. S. Rosenblatt, C. T. Snowdon, & M.
764   Milinski (Eds.), *Advances in the Study of Behavior, Vol. 26* (pp. 181–250). Academic Press.
765   Smith, J. E., Lehmann, K. D. S., Montgomery, T. M., Strauss, E. D., & Holekamp, K. E. (2017).
766   Insights from long - term field studies of mammalian carnivores. *Journal of Mammalogy*,
767   *98*(3), 631–641. doi:10.1093/jmammal/gyw194
768   Sneddon, L. U., Schmidt, R., Fang, Y., & Cossins, A. R. (2011). Molecular Correlates of Social
769   Dominance: A Novel Role for Ependymin in Aggression. *PloS One*, *6*(4), e18181.
770   Snyder-Mackler, N., Sanz, J., Kohn, J. N., Brinkworth, J. F., Morrow, S., Shaver, A. O., … Tung, J.
771   (2016). Social status alters immune regulation and response to infection in macaques.
772   *Science*, *354*(6315), 1041–1046.

773   Swanson, E. M., Dworkin, I., & Holekamp, K. E. (2011). Lifetime selection on a hypoallometric
774   size trait in the spotted hyena. *Proceedings of the Royal Society B: Biological Sciences*,
775   *278*(1722), 3277–3285.
776   Tibbetts, E. A., & Dale, J. (2004). A socially enforced signal of quality in a paper wasp. *Nature*,
777   *432*(7014), 218–222.
778   Tung, J., Archie, E. A., Altmann, J., & Alberts, S. C. (2016). Cumulative early life adversity predicts
779   longevity in wild baboons. *Nature Communications*, *7*, 1–7. doi:10.1038/ncomms11181
780