An evolutionary model explaining the Neolithic transition from egalitarianism to leadership and despotism

Simon T. Powers*

Laurent Lehmann†

July 13, 2014

*Corresponding author. Address: Department of Ecology & Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland. Tel: +41 21 692 4267. Email: Simon.Powers@unil.ch

†Address: Department of Ecology & Evolution, University of Lausanne, CH-1015 Lausanne, Switzerland. Tel: +41 21 692 4183. Email: Laurent.Lehmann@unil.ch

Abstract

The Neolithic was marked by a transition from small and relatively egalitarian groups, to much larger groups with increased stratification. But the dynamics of this remain poorly understood. It is hard to see how despotism can arise without coercion, yet coercion could not easily have occurred in an egalitarian setting. Using a quantitative model of evolution in a patch-structured population, we demonstrate that the interaction between demographic and ecological factors can overcome this conundrum. We model the co-evolution of individual preferences for hierarchy alongside the degree of despotism of leaders, and the dispersal preferences of followers. We show that voluntary leadership without coercion can evolve in small groups, when leaders help to solve coordination problems related to resource production. An example is coordinating construction of an irrigation system. Our model predicts that the transition to larger despotic groups will then occur when: 1. surplus resources lead to demographic expansion of groups, removing the viability of an acephalous niche in the same area and so locking individuals into hierarchy; 2. high dispersal costs limit followers' ability to escape a despot. Empirical evidence suggests that these conditions were likely met for the first time during the subsistence intensification of the Neolithic.

Keywords: despotism — dispersal — egalitarian — hierarchy — leadership — Neolithic

Introduction

Understanding how leadership and dominance behaviours in humans have changed over evolutionary time is relevant to both biology and the social sciences. What drove the transition from largely egalitarian hunter-gatherer groups, where leadership was facultative and dominance attenuated [1], to the hereditary and more despotic forms of leadership that arose during the Neolithic [2, 3]?

On the one hand, "coercive" (or "agency") theories have focused on the development of inequality that was made possible with the origin of food storage and agriculture, allowing dominant individuals to build up resource surpluses that could be used to consolidate their power [4, 5, 6]. On the other hand, "functional" (or "integrative") theories have addressed the benefits that leaders provide to other group members. In particular, as human group size increased during the Neolithic [7, 8], the resulting scalar stress [9] would have necessitated increased hierarchy in order to solve various coordination and collective action problems [10, 11, 12, 13, 14, 15, 16]. Leadership could have been favoured to solve problems including the coordinated harvesting of marine resources [17, 18, 19], the construction of irrigation systems [20, 21, 22, 23], and defensive warfare [24, 25].

But when considered alone as competing theories, both coercive and functional models struggle to explain the transition to despotism seen during the Neolithic. Purely coercive theories cannot explain why individuals would initially choose to follow a despot [16, 26]. Boehm [1] presents evidence suggesting that present-day hunter-gatherers actively form coalitions to suppress would-be dominants, and argues that pre-historic hunter-gatherers did likewise. Moreover, the advent of projectile weapons is likely to have made such coalitions particularly effective [27], tipping the balance of power away from an individual dominant. Thus the question is, why would individuals not continue to prevent despotic behaviour? But if individuals are unconstrained in their choice of leader, then it is difficult to see how despotism could develop.

Several authors have argued that an adequate model of the origin of increased social stratification must incorporate both functional and coercive aspects [22, 28, 15]. There is evidence that aspiring leaders drove the development of technology that increased subsistence intensification and raised population carrying capacity [22, 17]. For example, construction of irrigation systems would have allowed more land to be used for agriculture, providing an incentive for individuals to follow the leader. This fits with functional theories. On the other hand, the surplus resources that this provided could then be appropriated by leaders to further their own ends and consolidate their power. This is particularly the case given that irrigation farmers would be tied to the system, making dispersal away from a despot difficult. Spencer [22] developed a verbal model of this for the case of irrigation systems in prehispanic Mexico, and warfare in prehispanic Venezuela. However, the feedbacks between population size, functional aspects of leadership, and the development of despotism remain poorly understood and are difficult to capture with verbal models.

Here we present an evolutionary model of the dynamics of the transition from small-scale egalitarian to larger-scale hierarchical groups, which integrates both functional and coercive aspects of leadership. We use a demographically explicit model of a patch-structured population, in which surplus resources translate into increased reproductive output for those who receive them, as has been common throughout human history [29, 30]. Unlike previous work, this allows us to capture the ecological and demographic interactions between subsistence intensification, dispersal costs, and the evolution of despotic behaviour.

The model

Life cycle and social traits

We consider a population that is subdivided into a finite number, N_{p}, of patches, which are subjected to local stochastic demography (as per [31, 32]). The lifecycle consists of discrete and non-overlapping generations, as follows. (1) Social interactions occur on each patch with its members possibly choosing a leader, who may affect local resource production. (2) Each individual on a patch has a Poisson distributed number of offspring, with the mean determined by the outcome of social interactions and local resource abundance (defined explicitly below). (3) Adults of the previous generation perish. (4) Individuals of the descendant generation may disperse, conditional on the result of the stage of social interactions. Dispersing individuals suffer a cost C_{D}, such that individuals survive dispersal with probability $1-C_{\mathrm{D}}$, and then enter a patch taken at random from the population (excluding the natal patch).

Each individual in this population carries a cultural trait, h. This takes the value zero or one, and determines whether the carrier has a preference for hierarchy ($h = 1$) or acephalous ($h = 0$) social organisation. In each generation and for each patch, one individual is chosen at random from the subset of individuals with a preference for hierarchy ($h = 1$) to act as the leader (this could be an individual with unusual characteristics such as strong organisational abilities). There are then up to three classes of individuals on a patch: (i) the individual chosen as the leader (l class); (ii) the remaining individuals with $h = 1$ that act as followers (f class); (iii) acephalous individuals (with $h = 0$) that choose not to have a leader (a class).

When in the role of a leader, an individual is assumed to expresses a culturally inherited trait, z, which represents the proportion of the surplus it generated that it keeps for itself. This is a continuous variable between zero and one. Offspring of the leader are assumed to remain philopatric, but offspring of followers or acephalous individuals may disperse. We denote by d_{f} the conditional dispersal strategy of the offspring of a follower. Specifically, d_{f} is the maximum proportion of the surplus that an individual will tolerate the leader of the parental generation taking, and is thus continuous between zero and one. This assumption accords with evidence from social psychology that individuals tend to disperse from groups with autocratic leaders [33]. Finally, d_{a} determines the unconditional dispersal probability of the offspring of an acephalous individual, which is independent of the outcome of social interactions. The assumption that the offspring of a leader remain philopatric is appropriate in this model, since by remaining philopatric they increase the probability that one of their lineage will be chosen as leader on that patch in the next generation. Moreover, since offspring inherit the z trait of their parent, it is less biologically realistic that an individual would disperse based on how much of the surplus their parent took, when they themselves would take the same amount. We have, however, also investigated the effects of relaxing the assumption that the offspring of a leader must remain philopatric (Appendix S3).

Each individual carries all four cultural traits (h, z, d_{f}, and d_{a}) which are all assumed to be transmitted vertically from parent to offspring [34] with independent probability $1 - \mu$. When a mutation occurs at trait h (probability μ), an offspring adopts the opposite trait. When a mutation occurs at the three remaining continuous traits, Gaussian mutation is performed by addition of a truncated Gaussian distributed random variable centred around the current trait value, with variance 0.1.

We stress that our model aims to capture qualitative behavioural trends. A more quantitatively accurate model would include individual and social learning of behavioural traits within a generation. For example, hierarchy preference could be a continuous trait updated by an individual's estimate of the likely payoff from following a leader, and from copying the behaviour of more successful individuals. However, these processes would largely result in the same qualitative outcome as the vertical transmission with differential reproduction that we model, apart from the fact that they operate on a much shorter timescale.

Reproduction

The mean number of offspring produced (of the Poisson distribution in stage two of the life cycle) by individuals within patches is assumed to follow a Beverton-Holt model, with two niches (e.g., [31, 35]). The two niches correspond to either having a leader (individuals of the l and f classes), which we refer to as the hierarchical niche H, or remaining acephalous (acephalous niche A, containing individuals of class a). The degree of competition between the niches is set by two parameters, α_{AH} and α_{HA}, which represent the per capita effects of individuals in the hierarchical niche on those in the acephalous niche, and vice versa, respectively. The total number of individuals in the hierarchical niche (leader plus followers) on patch j at time t is denoted by $n_{\mathrm{H}j}(t)$, and the number of individuals in the acephalous niche by $n_{\mathrm{A}j}(t)$.

According to these assumptions, we write the mean number of offspring produced, respectively, by a leader, a follower, and an acephalous individual on patch j at time t as

$$\begin{aligned} w_{\mathrm{l}j}(t) &= \frac{r_{\mathrm{l}j}(t)}{1 + n_{\mathrm{H}j}(t)/K_{\mathrm{H}j}(t) + \alpha_{\mathrm{HA}} n_{\mathrm{A}j}(t)} \\ w_{\mathrm{f}j}(t) &= \frac{r_{\mathrm{b}}}{1 + n_{\mathrm{H}j}(t)/K_{\mathrm{H}j}(t) + \alpha_{\mathrm{HA}} n_{\mathrm{A}j}(t)} \\ w_{\mathrm{a}j}(t) &= \frac{r_{\mathrm{b}}}{1 + n_{\mathrm{A}j}(t)/K_{\mathrm{A}j}(t) + \alpha_{\mathrm{AH}} n_{\mathrm{H}j}(t)}. \end{aligned} \tag{1}$$

The numerator in each expression can be thought of as the maximal birth rate of an individual in the corresponding class. For followers and acephalous individuals, this is given by a constant r_{b}, while for the leader this depends upon the outcome of surplus production, as defined below. The denominator in each expression can be thought of as the intensity of density-dependent competition. This depends on a time dependent variable $K_{ij}(t)$, which is a proxy for the carrying capacity of niche i on patch j (maximum population size). The exact carrying capacity in the Beverton-Holt model is a function of all fitness parameters, but increases directly with K. In the classical one niche deterministic case, K gives the carrying capacity when $r_{\mathrm{b}} = 2$, which is a value we use throughout. Hence, we refer (loosely) to K as the "carrying capacity". $K_{ij}(t)$ is affected by surplus resource production (detailed below), which allows for local demographic expansions due to social interactions [31, 32].

Surplus production

In each patch individuals take part in a social enterprise, which may generate surplus resources for their niche. Individuals may also fail to produce this surplus, and to capture these two cases in a probabilistic way we let

$$\phi_{\tau j}(t) = \begin{cases} 1 & \text{with probability } s(n_{\tau j}(t), g_\tau) \\ 0 & \text{otherwise,} \end{cases}$$

where $\phi_{\tau j}(t)$ is the indicator random variable taking the value one if the surplus is produced in niche $\tau \in \{\mathrm{A}, \mathrm{H}\}$ on patch j at time t, zero otherwise. Surplus production occurs with probability

$$s(n_{\tau j}(t), g_\tau) = \exp(-g_\tau n_{\tau j}(t)), \tag{2}$$

where g_τ is a parameter giving the gradient of how the probability of surplus generation changes with the number of individuals in the niche ("social group size"). We assume that g_τ is positive, such that the probability of success decreases with increasing social group size. This represents the effects of scalar stress. We further assume that $g_{\mathrm{H}} < g_{\mathrm{A}}$, such that the success probability declines at a slower rate with increasing group size in the presence of a leader, and that for a given group size, groups with a leader are more likely to generate the surplus.

How surplus affects acephalous individuals

We relate surplus production to the "carrying capacity" of acephalous individuals by assuming that

$$K_{\mathrm{A}j}(t) = K_{\mathrm{b}} + \phi_{\mathrm{A}j}(t)\beta_{\mathrm{k}}\left(1 - \exp\left[-\gamma_{\mathrm{k}} n_{\mathrm{A}j}\left(t\right)\right]\right) \\ + \left[1 - \phi_{\mathrm{A}j}(t)\right]\left(1 - \epsilon\right)\left(K_{\mathrm{A}j}\left(t-1\right) - K_{\mathrm{b}}\right), \quad (3)$$

where K_{b} is the baseline capacity. If the surplus is generated, this is then increased by $\beta_{\mathrm{k}}\left(1 - \exp\left[-\gamma_{\mathrm{k}} n_{\mathrm{A}j}\left(t\right)\right]\right)$, which is a positive concave function of $\gamma_{\mathrm{k}} n_{\mathrm{A}j}$ (entailing diminishing returns), where γ_{k} sets the gradient of the carrying capacity increase, and $n_{\mathrm{A}j}$ is taken as the amount of surplus resource produced. Alternatively the surplus can be thought of as proportional to population size, with conversion factor γ_{k} (this is assumed to hold for both niches). The parameter β_{k} sets the maximum possible increase in carrying capacity. If the surplus is not successfully generated, the carrying capacity is then given by $(1 - \epsilon)K_{\mathrm{b}} + \epsilon K_{\mathrm{A}j}\left(t-1\right)$, where ϵ is the surplus decay rate from one generation to the next, and $K_{\mathrm{A}j}(0) = K_{\mathrm{b}}$ (if $\epsilon < 1$, there is some ecological inheritance of modified carrying capacity).

How surplus affects leaders and followers

For individuals in the hierarchical niche, the leader keeps a proportion of any surplus for itself, as given by the value of its z-trait. Let $z_{\mathrm{l}j}(t)$ denote the z-trait of the leader on patch j at time t, then the carrying capacity of individuals in the hierarchical niche is given by an analogous expression to that of acephalous individuals (eq. 3); namely,

$$K_{\mathrm{H}j}(t) = K_{\mathrm{b}} + \phi_{\mathrm{H}j}(t)\beta_{\mathrm{k}}\left(1 - \exp\left[-\gamma_{\mathrm{k}}\{1 - z_{\mathrm{H}j}(t)\} n_{\mathrm{H}j}\left(t\right)\right]\right) + \\ \left(1 - \phi_{\mathrm{H}j}(t)\right)\left(1 - \epsilon\right)\left[K_{\mathrm{H}j}\left(t-1\right) - K_{\mathrm{b}}\right], \quad (4)$$

where $\{1 - z_{\mathrm{H}j}(t)\} n_{\mathrm{H}j}\left(t\right)$ is the amount of surplus used to increase the carrying capacity of the leader and its followers. The remainder $z_{\mathrm{l}j}(t) n_{\mathrm{H}j}\left(t\right)$ of the surplus is retained by the leader and used to increase its own birth rate (which has occurred throughout human history [29, 30]) as follows

$$r_{lj}(t) = r_b + \phi_{Hj}(t)\beta_r \left(1 - \exp\left[-\gamma_r z_{lj}(t) n_{Hj}(t)\right]\right), \quad (5)$$

where γ_r gives the gradient of the increase in birth rate with respect to the absolute magnitude of the surplus that the leader takes. The parameter β_r gives the maximal possible increase in the leader's birth rate. This represents the maximum degree of despotism that it is possible for a leader to exert. This will depend upon both ecological and social factors, and in particular, on the degree to which followers are able to resist coercion. Where followers have little power to resist the leader, then we would expect a large value of β_r. Conversely, if followers are able to resist coercion to a large degree, for example by forming coalitions, then a smaller value of β_r would be more plausible.

Conditional dispersal of followers

To close the model, it only remains to specify how offspring of followers disperse conditionally on leader behaviour (offspring of acephalous individuals disperse unconditionally, and the offspring of the leader remain philopatric). Denoting by $d_{f,ij}(t)$ the dispersal preference of follower offspring i on patch j at time t, that offspring is assumed to disperse if:

$$z_{lj}(t) > d_{f,ij}(t),$$

that is, if the leader of its parent took more than its threshold value.

The model defines a stochastic process for the four evolving traits (h, z, d_a, d_f), the number of individuals in each niche (n_A, n_H), and their respective carrying capacities (K_A, K_H). Because of the non-linearities of the model, which result from the interactions of all of these variables, we analyse it using individual-based simulations.

Results

We focus on the effect that the following demographic and ecological parameters have on the transition to despotism: (i) the effect that a leader has on surplus generation (g_{A} relative to g_{H}); (ii) the degree to which surplus resources produce demographic expansion (β_{k}); (iii) the cost of dispersal (C_{D}). The other parameters used in the simulations, unless otherwise specified, are: $K_{\mathrm{b}} = 20$, $r_{\mathrm{b}} = 2$, $\gamma_{\mathrm{k}} = 0.05$, $\gamma_{\mathrm{r}} = 0.1$, $g_{\mathrm{H}} = 0.01$, $\alpha_{\mathrm{AH}} = \alpha_{\mathrm{HA}} = 0.03$, $\epsilon = 0.1$, $\mu = 0.01$, $N_{\mathrm{p}} = 50$.

The voluntary creation of hierarchy through cultural evolution

Figures 1a and 1c illustrate that when leaders confer a large advantage in surplus generation (g_{A} is large relative to g_{H}), hierarchical individuals can invade a population of acephalous individuals. This is because for a given group size, hierarchical individuals are more likely to produce a surplus than acephalous individuals on their patch (eq. 2). Individuals that receive surplus resources then enjoy a fitness increase, mediated by a reduction in the intensity of density-dependent competition in their niche. Consequently, they produce more offspring than individuals that do not receive a surplus. In this way, when leaders increase the likelihood of surplus generation, and share some of this surplus with their followers, then hierarchical individuals can outcompete acephalous individuals.

Crucially, this can occur even when leaders evolve to retain a large proportion of the surplus for themselves (Fig. 1c). This is because even when leaders retain some of the surplus, followers can still each receive more extra resource than they would in acephalous groups, where the surplus would be generated less frequently. This demonstrates the voluntary creation of hierarchy, where individuals that accept inequality in their groups are better off than those that remain egalitarian. Whether or not this is the case depends upon the magnitude of the advantage that leaders confer in surplus generation.

Figures. 1e and 1g illustrate the case where leaders do not provide much advantage in surplus generation. In this situation, acephalous individuals each receive on average a larger amount of surplus resources than followers of a leader. This is because acephalous groups are almost as likely to generate the surplus as hierarchal groups, but all of the surplus is shared amongst themselves rather than some being retained by a leader. Consequently, hierarchy is not favoured, and the unconditional dispersal probability trait of acephalous individuals, d_{a}, depends mainly on the dispersal cost and decreases as the cost increases (figs. 1e and 1g, further discussion in Appendix S1). We discuss the conditional dispersal trait of followers, and its co-evolution with the proportion of surplus that leaders retain, below.

The co-evolution of group size and hierarchy

When individuals receive surplus resources, this leads to a reduction in competition for resources with other individuals on the patch in their niche. As a result, their niche can support a larger number of individuals (eqs. 3 and 4), leading to an increase in group size. Figures 1b and 1d illustrate that when hierarchy invades, it drives an increase in group size. For example, in Fig. 1b the population initially starts out fixed for acephalous individuals, who produce some surplus. This surplus drives an increase in their local number from the base value of 20, to around 40. But because of the problems of coordinating in large groups without a leader (represented by a large value of g_A), they are unable to reliably generate the surplus in groups above this size. Thus, their group size stabilises around this value. However, as hierarchy invades group size increases up to 80 individuals. This is because the coordination advantages of having a leader ($g_{\mathrm{H}} < g_{\mathrm{A}}$) mean that hierarchical individuals are able to continue generating the surplus in larger groups.

The increase in group size is driven by a positive feedback loop in which surplus production increases carrying capacity, causing an increase in group size, which then in turn allows greater amounts of surplus to be generated (eq. 2). This positive feedback loop stops when either (i) groups are too large for additional surplus to be reliably generated (eq. 2), or (ii) diminishing returns in the value of the surplus mean that the extra surplus produced by one more individual is not enough to increase carrying capacity by at least one individual (eqs. 3 and 4). When g_{H} is smaller than g_{A}, then the feedback loop can stop at a larger group size for hierarchical individuals than for acephalous individuals. Thus the ability of leaders to solve coordination problems in larger groups, combined with the effects of surplus resources on demography, means that the invasion of hierarchy produces a transition to larger-scale social groups.

The transition to a larger group size is crucial to the stability of hierarchy. This is because acephalous individuals experience density-dependent competition with hierarchical individuals on their patch, and vice versa (eq. 1). So the larger the absolute number of hierarchical individuals, the more they suppress the fitness of acephalous individuals by outcompeting them for shared resources, such as space. Conversely, when there are few hierarchical individuals, then it is relatively easy for acephalous individuals to re-invade and hierarchy to collapse. The parameter β_{k} controls the extent to which surplus production can increase group size. As Figure 2 shows, when this is low then although hierarchy can invade, it does not remain stable. As β_{k} increases, however, then the invasion of hierarchy brings about a large increase in group size that suppresses mutant acephalous individuals. The transition to larger groups thus locks individuals into hierarchy.

The degree to which group size increases when hierarchy invades also depends upon how much of the surplus the leader retains for itself. Specifically, when leaders evolve to share more surplus resources with their followers, then the group can grow to a larger size (Figs. 1a and 1b, compared to 1c and 1d).

When does cultural evolution lead to despotism?

What determines how much of the surplus the leader takes? A selection pressure exists for a leader to take more of the surplus, since this translates into an increased birth rate (eq. 5) and hence a greater number of offspring relative to the other hierarchical individuals on its patch (eq. 1). Moreover, because the leader of the next generation is chosen by random sampling of the offspring of hierarchical individuals on the patch, this increased reproduction also increases the probability that one of the current leader's offspring will remain as leader in the next generation. This continued occupancy of the leader role then increases the reproductive share of the leader's lineage even further.

However, a pressure also exists for the leader to take less surplus. This is because the total amount of surplus generated increases with increasing group size (eqn. 2), which provides an incentive for a leader to have more followers. But followers have a choice in leader since they may disperse from the group and join a different one, conditional on the amount of surplus that the leader takes (as given by their d_{f} trait). Thus, if the leader takes too much of the surplus then it will lose followers. This then means that less surplus will be generated for hierarchical individuals in the next generation, which can cause hierarchical individuals to be outcompeted by acephalous individuals on their patch.

The proportion of surplus that the leader takes is therefore a trade-off between opposing selection pressures. The balance depends upon the cost of dispersal – how easily individuals may leave one leader and follow another. If the cost of dispersal is low, then leaders are constrained in how much of the surplus they can monopolise. This is because when dispersal costs are low then followers evolve low tolerance values of d_{f}, such that they readily disperse if leaders retain a larger proportion of the surplus (Fig. 1a). Consequently, leaders evolve to share a large fraction of the surplus with their followers in order to prevent them from dispersing. On the other hand, as dispersal cost increases then followers evolve larger tolerance values of d_{f} in order to avoid paying a high dispersal cost (Fig. 1c). As a result, the strategy of leaders co-evolves to appropriate more of the surplus for their own reproduction, since their followers will not readily disperse to other groups.

Thus in an ecology where dispersal is costly, evolution leads to more despotic groups. Moreover this increased despotism is voluntarily tolerated by followers, in the sense that individuals which allow the leader to retain more surplus before dispersing outcompete both acephalous individuals, and followers that more readily disperse. Figure 3 demonstrates this co-evolution of follower dispersal preference and leader strategy for the full range of dispersal costs.

Sensitivity to parameters and model assumptions

We systematically varied the advantage in surplus production that leaders confer relative to acephalous groups (g_{A}). When leadership does not confer much advantage in surplus production, then acephalous individuals outcompete hierarchical individuals (Fig. S2). We also investigated the effect of varying the coercive power of the leader (Fig. S3), as measured by the maximal birth rate advantage it can enjoy from surplus production (β_{r}). As this increases then for a given dispersal cost leaders evolve to retain more of the surplus for themselves. Further, we investigated the effects of varying the intergenerational decay in surplus resources, ϵ, including allowing for complete decay (Appendix S2 and Fig. S4). Finally, we allowed the offspring of a leader to disperse (Appendix S3 and Fig. S5). We found that varying all of these does not qualitatively affect our main results.

Discussion

We have presented a model which captures the dynamics of the transition from small egalitarian to larger despotic groups. In line with work by Hooper et al. [15], our model demonstrates that hierarchical systems of social organisation can be voluntarily created by followers, rather than having to be imposed by a leader through coercion. This is in contrast to the current trend in archaeology that focuses on "agency", that is, on how leaders promote their own interests at the expense of others. By such accounts, leadership is seen as benefiting the leader rather than the followers [4, 6]. Yet while it is certainly the case that leaders should be expected to promote their own ends, the agency of followers must also be considered [28, 1, 36]. If leadership provides no benefit to followers, then it is hard to see why previously egalitarian individuals would accept despotic appropriation of resources, unless there were coercive institutions such as a military already in place. But such institutional coercion could not have been paid for before a leader appropriated surplus resources, making it hard to see how hierarchy could become established [16, 26].

The origin of despotism in human societies is similar to the problem addressed by reproductive skew theory [30]. In skew models, despotism is measured in terms of how much of the reproduction within a group is monopolised by a dominant individual. This is constrained by the outside options that subordinates have, either to live alone or in a different group. Skew models predict that dominants should behave more despotically as the feasibility of outside options decreases [37]. However, they do not consider the benefits leaders can provide to other group members in terms of surplus production, and so do not address how despotic leadership could evolve from an initial stable state of egalitarianism. Here, we have extended the basic logic of skew theory to incorporate the feedback between surplus production and demography that was likely to have been important during the Neolithic.

Previous work has explicitly modelled the formation of institutions to solve various collective action problems related to food production, as relevant to demographic growth in the Neolithic [31]. It was shown that groups could evolve institutionally-coordinated punishment to secure cooperation in generating surplus resources, driving demographic expansion. This paper builds upon these results by investigating the political ecology of such institutions, in terms of the opportunities that they create for despotism as group size increases.

Hooper et al. [15] showed that hierarchy can evolve if leaders help to secure cooperation in the production of large-scale public goods, using a model with complete dispersal between groups every generation. Their static analysis implied that despotism should rise as the cost for followers of switching to a different leader increases. Our model has independently confirmed that this prediction holds in a demographically realistic setting, where the cost of switching leader is given a biological basis in terms of dispersal cost. Moreover, our model incorporates dynamic group size alongside explicit co-evolution of leader despotism and follower tolerances. This framework has allowed us to demonstrate that the equilibrium of large groups with despotic leadership can actually be reached by gradual evolution, from an initial state of small egalitarian groups. Understanding the dynamics of this transition is one of the most pressing issues in Neolithic social evolution [16, 26]. But previous models have not addressed the interaction between subsistence intensification, population size, and dispersal costs. Our results demonstrate that the interaction between these factors provides a cogent explanation for the transition to large and despotic groups. We now turn to discuss the empirical evidence for this interaction during the Neolithic.

There is strong evidence that the presence of a leader conferred advantages in solving coordination problems related to food production in both complex hunter-gatherers [38, 28, 18] and agriculturalists [21, 22, 23]. Arnold [17] stresses the role of leaders in technological innovation that increased carrying capacity. For example the Chumash, a maritime culture in the north American Pacific, developed large boats made of rare materials, which required teams of specialists to construct. Consequently, only high status individuals could finance and organise their construction. The boats greatly increased productivity by allowing access to new marine resources, and by increasing the amount of resource that could be transferred simultaneously. This increased carrying capacity [17], but also led to increased stratification by providing surplus resources that boat owners could monopolise.

There is also evidence that leaders coordinated the construction of irrigation systems [21, 22, 23], even if not in the state-building sense argued by Wittfogel [20]. Spencer [22] presents archaeological evidence that the Purrón dam, an irrigation system in prehispanic Mexico, was constructed by a faction that aspired to leadership. Because canal irrigation was essential for agriculture in this area, other individuals would have benefitted from following this faction in order to gain access to water [22]. Spencer presents evidence that population growth subsequently occurred, causing the leadership faction to coordinate many followers in the construction of a larger dam. Moreover, there is evidence that this expansion of both population size and the irrigation system led to increased social stratification, with elites beginning to trade surpluses that they controlled for prestige goods [22]. This fits the feedback between demographic expansion and hierarchy formation captured by our model.

An important question is why despotic hierarchy evolved under intensive food production, but not under hunting and gathering? Our results suggest that demography plays an important role in the stability of despotism. When groups are small, then hierarchy can easily collapse if despots take too much resource. But if groups are larger, then density-dependent competition means that hierarchical individuals can outcompete acephalous individuals for shared resources, even when despots retain most of the surplus. Demographic expansion can therefore cause individuals to become locked into hierarchy, by destroying the viability of a previous non-hierarchical niche. Although human health appears to have declined with the origin of agriculture [39], and agriculture may initially have been less productive than hunter-gathering [40], cemetery data strongly implies that a demographic expansion indeed occurred during the Neolithic [8]. Other data indicates that the population density of hunter-gatherer groups is usually below 0.1 person/sq. mi., while that of early dry farmers is around 4 persons/sq. mi, and that of early irrigation farmers from 6 to 25 person/sq. mi [7]. The construction of irrigation systems, for example, could thus trigger the co-evolution of demographic expansion and despotism.

Our model predicts that despotism should increase with increasing dispersal costs, for which there is strong empirical support [41, 5, 30]. Carneiro [41] presents evidence that state formation (increased hierarchy) happens when relatively small areas of productive agricultural land are surrounded by geographical barriers. This then allows leaders to extract tribute from other individuals, whose options to leave the group are limited. For example in Peru, early states evolved where agriculture was practiced in narrow valleys, making dispersal difficult. By contrast, states did not so readily evolve in the Amazon basin where there were large expanses of agricultural land available, making dispersal relatively easy [41]. Allen [5] also stresses the role of dispersal costs in the creation of the despotic ancient Egyptian state. He argues that the deserts bordering the Nile made dispersal very costly, thus allowing the Pharaohs to extract a large surplus from agriculturalists. Similarly, technological development can increase dispersal costs. For example, irrigation farming was likely to tie agriculturalists to the irrigation system, again limiting free movement and choice of leader [20, 22].

In conclusion, our model predicts that despotic social organisation will evolve from an initial state of egalitarianism when: 1. leaders generate surplus resources leading to demographic expansion of their groups, which removes the viability of an acephalous niche in the same area; 2. high dispersal costs subsequently limit outside options for followers by restricting choice of leader. The empirical evidence reviewed here suggests that these conditions were likely to have been satisfied during the Neolithic.

Acknowledgements: We thank Mark van Vugt and two anonymous reviewers for useful comments on the manuscript. The computations were performed at the Vital-IT (http://www.vital-it.ch) Center for high-performance computing of the SIB Swiss Institute of Bioinformatics. This work was supported by Swiss NSF grant PP00P3-123344.

References

[1] Boehm, C., 2001 *Hierarchy in the Forest: The Evolution of Egalitarian Behavior.* Harvard University Press.

[2] Price, T. D., 1995 Social inequality at the origins of agriculture. In *Foundations of Social Inequality* (eds. T. D. Price & G. M. Feinman), pp. 129–151. New York: Plenum Press.

[3] Price, T. D. & Bar-Yosef, O., 2010 Traces of inequality at the origins of agriculture in the ancient Near East. In *Pathways to Power* (eds. T. D. Price & G. M. Feinman), Fundamental Issues in Archaeology, pp. 147–168. New York: Springer. (doi:10.1007/978-1-4419-6300-0_6).

[4] Earle, T., 1989 The evolution of chiefdoms. *Curr. Anthropol.* **30**, 84–88.

[5] Allen, R. C., 1997 Agriculture and the origins of the state in ancient Egypt. *Explor. Econ. Hist.* **34**, 135–154. (doi:10.1006/exeh.1997.0673).

[6] Hayden, B. & Villeneuve, S., 2010 Who benefits from complexity? A view from Futuna. In *Pathways to Power* (eds. T. D. Price & G. M. Feinman), Fundamental Issues in Archaeology, pp. 95–145. Springer New York. (doi:10.1007/978-1-4419-6300-0_5).

[7] Hassan, F. A. & Sengel, R. A., 1973 On mechanisms of population growth during the Neolithic. *Curr. Anthropol.* **14**, 535–542.

[8] Bocquet-Appel, J.-P., 2011 When the world's population took off: the springboard of the Neolithic Demographic Transition. *Science* **333**, 560–561. (doi: 10.1126/science.1208880).

[9] Johnson, G. A., 1982 Organizational structure and scalar stress. In *Theory and Explanation in Archaeology: the Southampton conference* (eds. C. Renfrew, M. Rowlands & B. A. Segraves-Whallon), pp. 389–420. New York: Academic Press.

[10] Olson, M., 1965 *The Logic of Collective Action: Public Goods and the Theory of Groups.* Harvard economic studies, v. 124. Harvard University Press, revised edition.

[11] Calvert, R., 1992 Leadership and its basis in problems of social coordination. *International Political Science Review* **13**, 7–24. (doi:10.1177/019251219201300102).

[12] Van Vugt, M., 2006 Evolutionary origins of leadership and followership. *Pers. Soc. Psychol. Rev.* **10**, 354–371. (doi:10.1207/s15327957pspr1004_5).

[13] King, A. J., Johnson, D. D. P. & Van Vugt, M., 2009 The origins and evolution of leadership. *Curr. Biol.* **19**, R911–R916. (doi:10.1016/j.cub.2009.07.027).

[14] O'Gorman, R., Henrich, J. & Van Vugt, M., 2009 Constraining free riding in public goods games: designated solitary punishers can sustain human cooperation. *Proc. R. Soc. B* **276**, 323–329. (doi:10.1098/rspb.2008.1082).

[15] Hooper, P. L., Kaplan, H. S. & Boone, J. L., 2010 A theory of leadership in human cooperative groups. *J. Theor. Biol.* **265**, 633–646. (doi:10.1016/j.jtbi.2010.05.034).

[16] Sterelny, K., 2013 Life in interesting times: cooperation and collective action in the holocene. In *Cooperation and Its Evolution* (eds. K. Sterelny, R. Joyce, B. Calcott & B. Fraser), pp. 89–108. Cambridge, MA: MIT Press.

[17] Arnold, J. E., 1995 Transportation innovation and social complexity among maritime hunter-gatherer societies. *Am. Anthropol.* **97**, 733–747. (doi: 10.1525/aa.1995.97.4.02a00150).

[18] Friesen, T. M., 1999 Resource structure, scalar stress, and the development of Inuit social organization. *World Archaeol.* **31**, 21–37. (doi:10.1080/00438243.1999.9980430).

[19] Alvard, M. S. & Nolin, D. A., 2002 Rousseau's whale hunt? Coordination among big–game hunters. *Curr. Anthropol.* **43**, 533–559. (doi:10.1086/341653).

[20] Wittfogel, K. A., 1957 *Oriental Despotism: a Comparative Study of Total Power*. New Haven: Yale University Press.

[21] Nichols, D. L. & Frederick, C. D., 1993 Irrigation canals and chinampas: recent research in the northern Basin of Mexico. In *Economic aspects of water management in the prehispanic New World. Research in economic anthropology, Supplement 7* (eds. V. L. Scarborough & B. L. Isaac), pp. 123–150. Greenwich: JAI Press, Inc.

[22] Spencer, C. S., 1993 Human agency, biased transmission, and the cultural evolution of chiefly authority. *J. Anthropol. Archaeol.* **12**, 41–74. (doi:10.1006/jaar.1993.1002).

[23] Carballo, D. M., Roscoe, P. & Feinman, G. M., 2012 Cooperation and collective action in the cultural evolution of complex societies. *J. Archaeol. Method. Th.* pp. 1–36. (doi: 10.1007/s10816-012-9147-2).

[24] Boone, J. L., 1992 Competition, cooperation and the development of social hierarchies. In *Ecology, Evolution and Social Behavior* (eds. E. A. Smith & B. Winterhalder), pp. 301–337. Aldine de Gruyter.

[25] Roscoe, P., 2013 War, collective action, and the "evolution" of human polities. In *Cooperation and collective action: archaeological perspectives* (ed. D. M. Carballo). Boulder, CO: University Press of Colorado.

[26] Seabright, P., 2013 The birth of hierarchy. In *Cooperation and Its Evolution* (eds.

K. Sterelny, R. Joyce, B. Calcott & B. Fraser), pp. 109–116. Cambridge, MA: MIT Press.

[27] Bingham, P. M., 1999 Human uniqueness: a general theory. *Q. Rev. Biol.* **74**, 133–169.

[28] Arnold, J. E., 1996 The archaeology of complex hunter-gatherers. *J. Archaeol. Method. Th.* **3**, 77–126. (doi:10.1007/bf02228931).

[29] Betzig, L. L., 1986 *Despotism and Differential Reproduction: A Darwinian View of History.* New York: Aldine.

[30] Summers, K., 2005 The evolutionary ecology of despotism. *Evol. Hum. Behav.* **26**, 106–135. (doi:10.1016/j.evolhumbehav.2004.09.001).

[31] Powers, S. T. & Lehmann, L., 2013 The co-evolution of social institutions, demography, and large-scale human cooperation. *Ecol. Lett.* **16**, 1356–1364. (doi:10.1111/ele.12178).

[32] Lehmann, L., Perrin, N. & Rousset, F., 2006 Population demography and the evolution of helping behaviors. *Evolution* **60**, 1137–1151.

[33] Van Vugt, M., Jepson, S. F., Hart, C. M. & De Cremer, D., 2004 Autocratic leadership in social dilemmas: A threat to group stability. *J. Exp. Soc. Psychol.* **40**, 1–13. (doi: 10.1016/s0022-1031(03)00061-1).

[34] Cavalli-Sforza, L. L. & Feldman, M. W., 1981 *Cultural Transmission and Evolution: A Quantitative Approach.* Princeton, New Jersey: Princeton University Press.

[35] Chow, Y. & Hsieh, J., 2012 On multidimensional discrete-time Beverton-Holt competition models. *J. Differ. Equ. Appl.* pp. 1–16. (doi:10.1080/10236198.2012.656618).

[36] Fargher, L. F. & Blanton, R. E., 2007 Revenue, voice, and public goods in three premodern states. *Comp. Stud. Soc. Hist.* **49**, 848–882. (doi:10.1017/s0010417507000783).

[37] Vehrencamp, S. L., 1983 A model for the evolution of despotic versus egalitarian societies. *Anim. Behav.* **31**, 667–682. (doi:10.1016/s0003-3472(83)80222-x).

[38] Ames, K. M., 1994 The Northwest Coast: complex hunter-gatherers, ecology, and social evolution. *Annu. Rev. Anthropol.* **23**, 209–229. (doi: 10.1146/annurev.an.23.100194.001233).

[39] Lambert, P. M., 2009 Health versus fitness. *Curr. Anthropol.* **50**, 603–608. (doi: 10.1086/605354).

[40] Bowles, S., 2011 Cultivation of cereals by the first farmers was not more productive than foraging. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 4760–4765. (doi: 10.1073/pnas.1010733108).

[41] Carneiro, R. L., 1970 A theory of the origin of the state. *Science* **169**, 733–738. (doi: 10.1126/science.169.3947.733).

Figure legends

Figure 1: Illustration of ecological conditions under which either hierarchical (a–d) or acephalous (e–h) individuals are favoured by the co-evolution of culturally transmitted behavioural traits with demography. When the presence of a leader confers a large advantage in surplus generation (g_{H} much smaller than g_{A}), then individuals with a preference for hierarchy can invade an acephalous population (a and c). Successful generation of the surplus then drives an increase in population size (b and d). The degree of despotism, measured by the amount of surplus the leader monopolises for its own reproduction, increases with increasing dispersal cost (a and c). Conversely, if the presence of a leader does not confer a large advantage in surplus generation then hierarchy fails to invade (e–h) and groups remain acephalous. Parameters: $\beta_{\mathrm{r}} = 5$, $\beta_{\mathrm{k}} = 100$.

Figure 2: Stable hierarchy requires that surplus resources translate into demographic expansion of group size (large value of β_{k}). Demographic expansion removes the viability of the acephalous niche on a patch, locking individuals into hierarchy. Panels show the stability of hierarchy on a single patch in the metapopulation. Parameters: $g_{\mathrm{A}} = 0.15$, $\beta_{\mathrm{r}} = 2$.

Figure 3: As dispersal cost increases, followers tolerate their leader behaving more despotically (a). This in turn means that they enjoy a smaller increase in their carrying capacity, as the leader is able to direct more of the surplus into increasing its own reproductive success relative to that of its followers (b). Results show the long-run time averages over 3×10^6 generations of the stochastic simulation. Parameters: $\beta_{\mathrm{r}} = 20$, $g_{\mathrm{A}} = 0.15$, $\beta_{\mathrm{k}} = 100$.

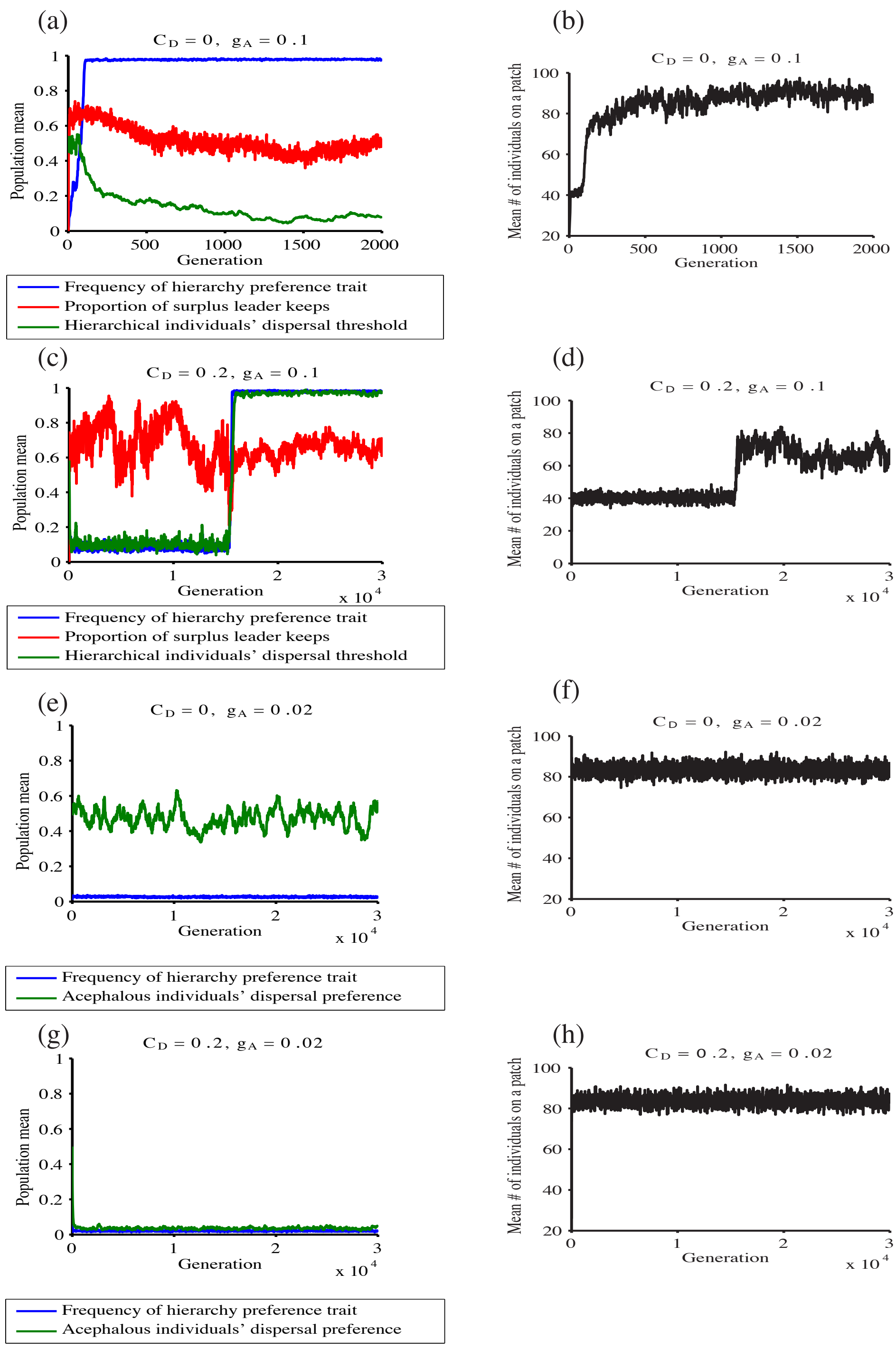
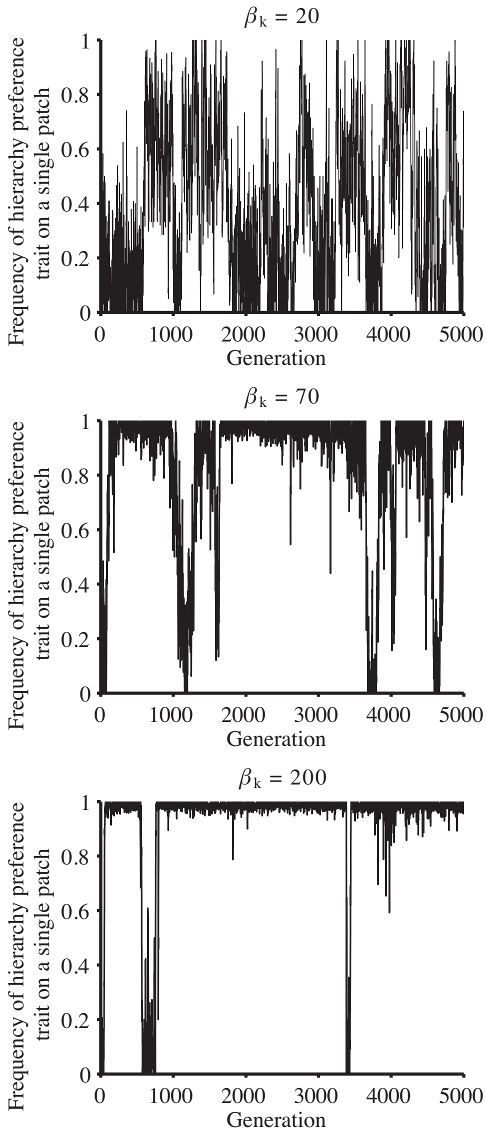

(a)
$C_D = 0,\ g_A = 0.1$
Population mean
Generation
Frequency of hierarchy preference trait
Proportion of surplus leader keeps
Hierarchical individuals' dispersal threshold
(b)
$C_D = 0,\ g_A = 0.1$
Mean # of individuals on a patch
Generation
(c)
$C_D = 0.2,\ g_A = 0.1$
Population mean
Generation
x 10^4
Frequency of hierarchy preference trait
Proportion of surplus leader keeps
Hierarchical individuals' dispersal threshold
(d)
$C_D = 0.2,\ g_A = 0.1$
Mean # of individuals on a patch
Generation
x 10^4
(e)
$C_D = 0,\ g_A = 0.02$
Population mean
Generation
x 10^4
Frequency of hierarchy preference trait
Acephalous individuals' dispersal preference
(f)
$C_D = 0,\ g_A = 0.02$
Mean # of individuals on a patch
Generation
x 10^4
(g)
$C_D = 0.2,\ g_A = 0.02$
Population mean
Generation
x 10^4
Frequency of hierarchy preference trait
Acephalous individuals' dispersal preference
(h)
$C_D = 0.2,\ g_A = 0.02$
Mean # of individuals on a patch
Generation
x 10^4

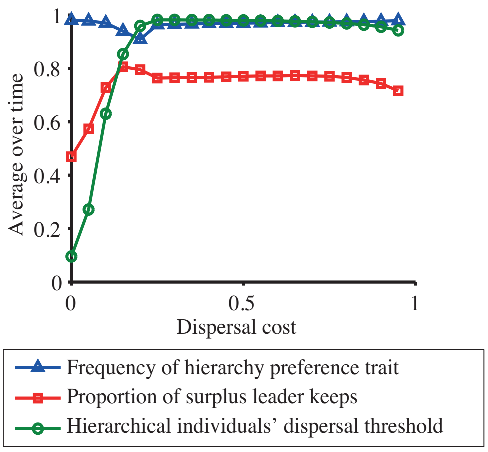
Figure 1



β_k = 20
Frequency of hierarchy preference trait on a single patch
0
0.2
0.4
0.6
0.8
1
0
1000
2000
3000
4000
5000
Generation
β_k = 70
Frequency of hierarchy preference trait on a single patch
0
0.2
0.4
0.6
0.8
1
0
1000
2000
3000
4000
5000
Generation
β_k = 200
Frequency of hierarchy preference trait on a single patch
0
0.2
0.4
0.6
0.8
1
0
1000
2000
3000
4000
5000
Generation

Figure 2

(a)


Average over time
1
0.8
0.6
0.4
0.2
0
0
0.5
1
Dispersal cost
Frequency of hierarchy preference trait
Proportion of surplus leader keeps
Hierarchical individuals' dispersal threshold

(b)

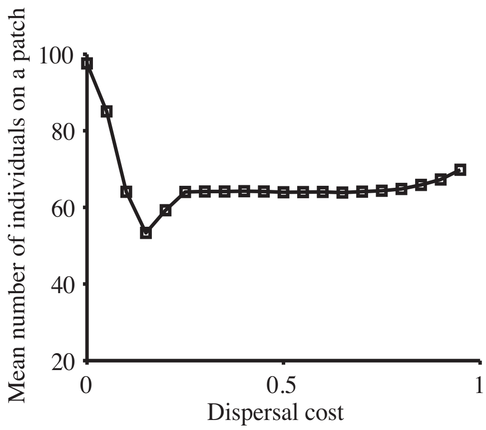

Mean number of individuals on a patch
100
80
60
40
20
0
0.5
1
Dispersal cost

Figure 3

Electronic Supplementary Material for

The transition from leadership to despotism in Neolithic human groups

Simon T.Powers*

Laurent Lehmann†

July 10, 2014

This file includes:

SI Appendix S1-S3

Figures S1-S5

*Department of Ecology & Evolution, University of Lausanne, Switzerland, Simon.Powers@Unil.ch

†Department of Ecology & Evolution, University of Lausanne, Switzerland, Laurent.Lehmann@Unil.ch

Appendix S1: Evolution of acephalous individuals' dispersal probability trait, d_{a}

In order to better understand the evolution of dispersal in acephalous individuals, we consider in this section a version of the model without hierarchical individuals present.

The dispersal probability of acephalous individuals decreases under natural selection as dispersal costs increase, in agreement with classic theory on the evolution of dispersal [1, 2]. Figure S1a shows a baseline case where there is no possibility for surplus resources to decrease local resource competition ($\beta_{\mathrm{k}} = 0$), and hence to benefit individuals. In this case, acephalous individuals evolve high dispersal rates if dispersal is costless (Fig. S1a). Intuitively this is because by dispersing, individuals reduce local density-dependent competition for resources with their relatives.

However, where surplus resources can reduce local resource competition ($\beta_{\mathrm{k}} = 100$), as in the *Main Text*, then acephalous individuals do not evolve such high dispersal rates even when dispersal cost is zero (Fig. S1b). This is because the absolute amount of surplus generated is proportional to group size, so that larger groups can generate a larger amount of surplus. In particular, the presence of one extra group member can produce a marginal surplus that increases carrying capacity by more than one. In that case, by remaining philopatric and helping to generate a larger surplus, an individual can reduce local competition for its relatives by a greater amount than if it dispersed. This explains why acephalous individuals do not evolve very high dispersal rates in Fig. 1e of the *Main Text*, even though dispersal is costless. The effect becomes greater as the probability of acephalous individuals generating the surplus decreases (Fig. S1b). Intuitively, this is because as the probability of surplus generation decreases then it pays individuals to make sure that when it is generated, it is as large as possible. However, in all cases we find that dispersal does not evolve to zero, since individuals can on average reduce local kin competition more by sometimes dispersing, even when surplus production favours larger groups.

Appendix S2: Effect of varying the surplus decay rate, ϵ

Here we show that our main conclusions are robust to varying the carrying capacity surplus decay rate, ϵ. There is still an increase in the proportion of the surplus retained by the leader as dispersal cost increases from zero, across the full range of ϵ (Fig. S4a). This holds even when there is complete decay in the surplus between generations ($\epsilon = 1$). However, larger decay rates cause large fluctuations in carrying capacity, and hence group size, when the surplus is not generated. Consequently, the mean group size over time is smaller for larger decay rates (Fig. S4b).

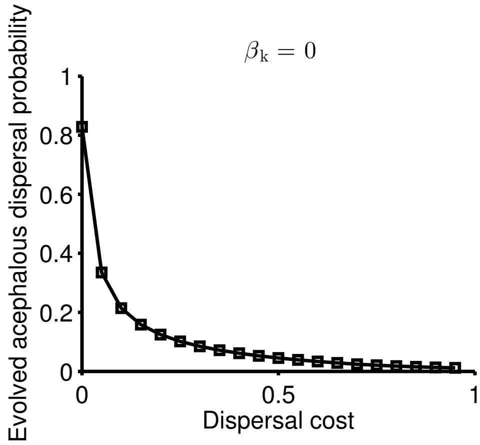
Appendix S3: Effect of allowing the offspring of a leader to disperse

Finally, we have investigated the effect of allowing the offspring of a leader to disperse, conditional on the proportion of the surplus that their parent took in the previous generation (Fig. S5). In that case, we find that leaders evolve to retain less of the surplus for themselves when dispersal costs are high. In particular, the proportion of surplus that the leader retains reaches its maximal value for intermediate dispersal costs (around 0.2 in Fig. S5). As dispersal costs increase beyond this, leaders slowly begin to reduce the proportion of surplus that they retain. This is in order to prevent their offspring from conditionally dispersing, and hence suffering a high dispersal cost.

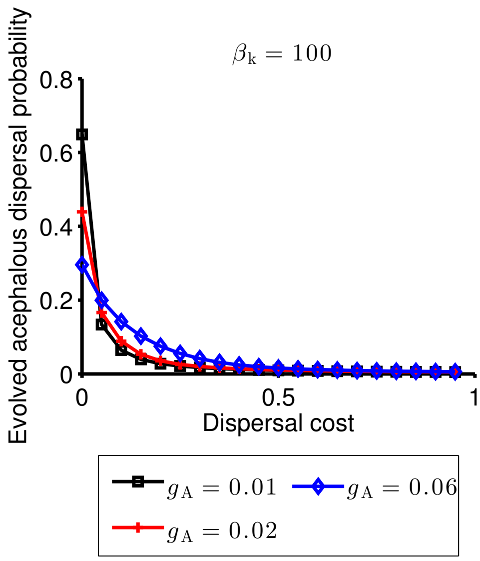
References

[1] Hamilton, W. D & May, R. M. (1977) Dispersal in stable habitats. *Nature* **269**, 578–581.

[2] Rousset, F. (2004) *Genetic Structure and Selection in Subdivided Populations*, Monographs in Population Biology. (Princeton University Press, Princeton, NJ).



(a)



(b)

Figure S1: The dispersal probability trait of acephalous individuals, d_a, decreases with increasing dispersal cost. If surplus resources cannot decrease local competition ($\beta_k = 0$) then high dispersal rates evolve if dispersal is costless (a). This is because dispersal reduces local competition with relatives. However, if the surplus can decrease local competition ($\beta_k = 100$, as in the *Main Text*) then there is a selection pressure for reduced dispersal, since larger groups can generate a larger surplus (b). Results show the long-run time averages over 3×10^6 generations of the stochastic simulation.

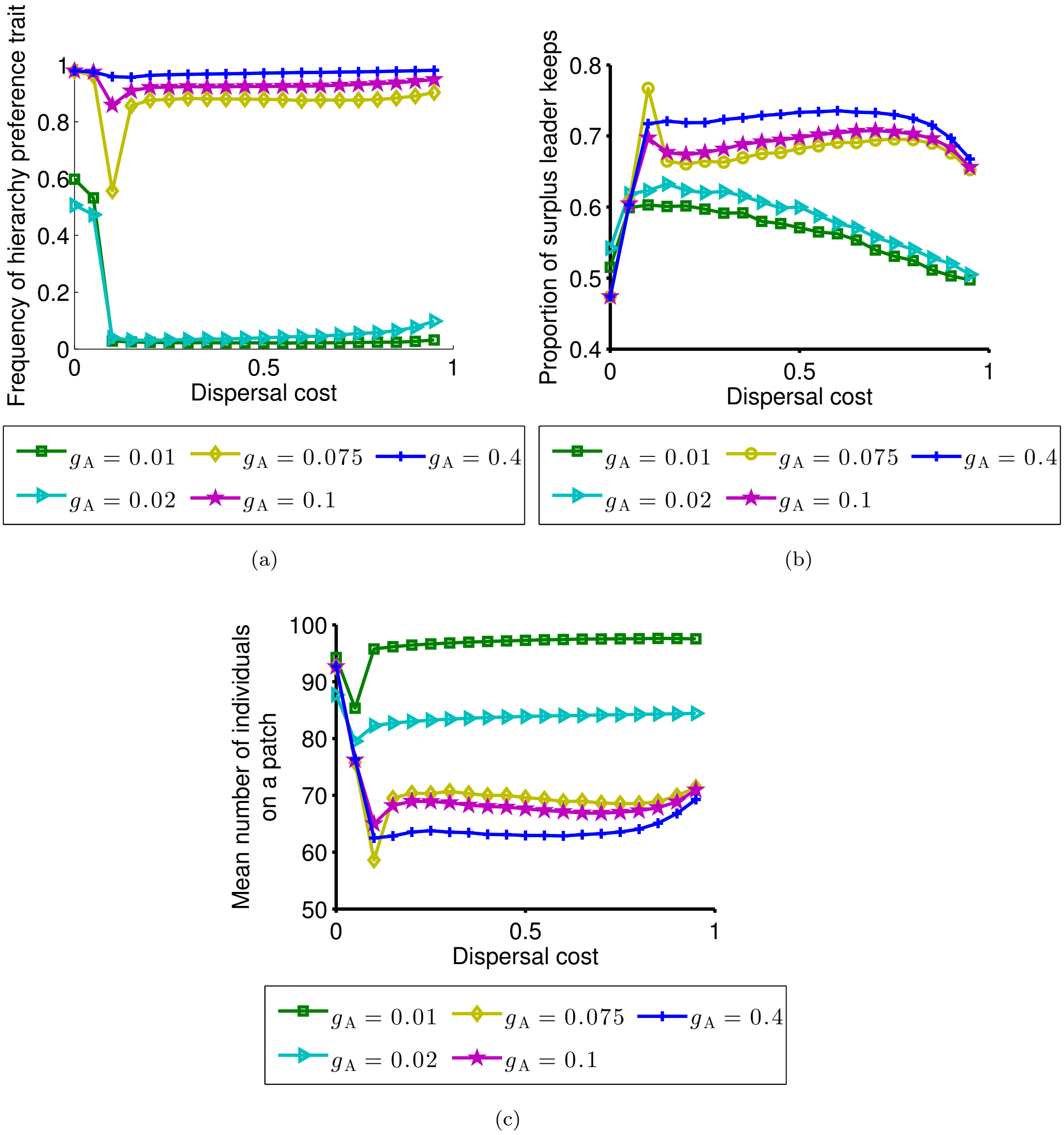


Figure S2: Effect of varying the relative advantage that a leader confers on surplus generation on (a) the frequency of individuals with a preference for hierarchy, (b) how despotically the leader behaves, and (c) population size. Results show the long-run time averages over 3×10^6 generations of the stochastic simulation. Parameters: $\beta_r = 10$.

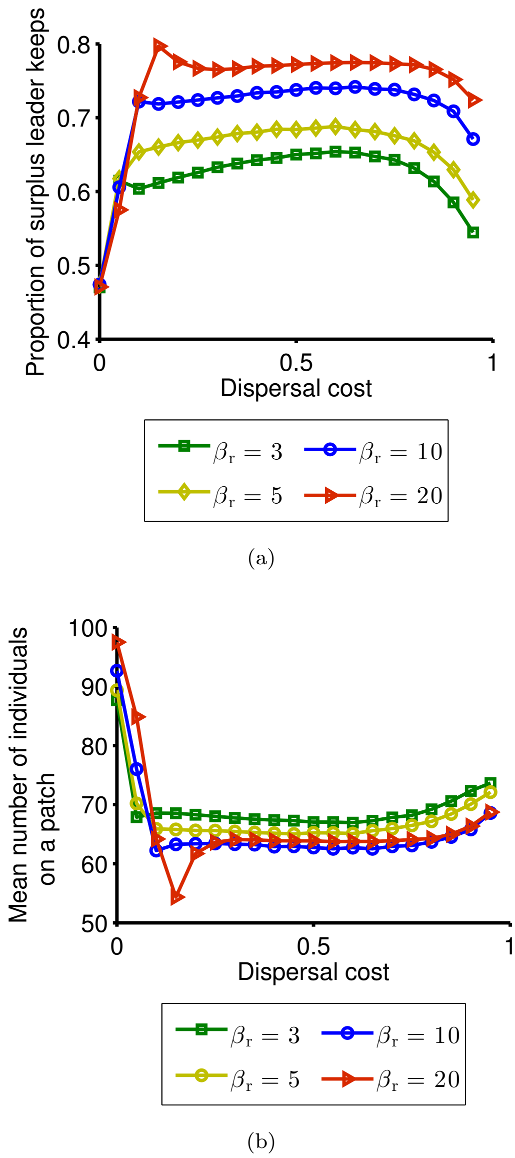
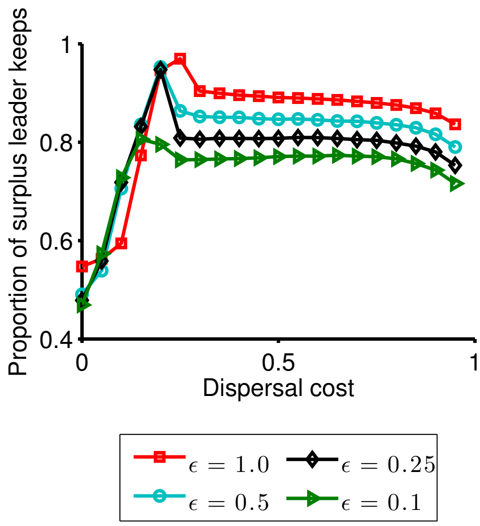
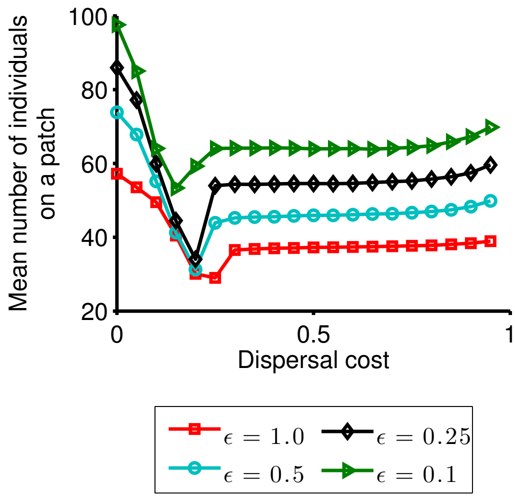


Figure S3: Effect of varying the maximal increase in birth rate that the surplus can confer on a leader, which corresponds to the degree of coercive power that the leader can exert. (a) The proportion of surplus that the leader retains increases with increasing power. (b) Population size correspondingly decreases due to less surplus being shared with followers. Results show the long-run time averages over 3×10^6 generations of the stochastic simulation. Parameters: $g_A = 0.15$.



(a)



(b)

Figure S4: The main conclusions of the model hold across the range of ϵ (a). However, a larger value of ϵ causes larger fluctuations in group size between generations, resulting in a smaller mean group size across time (b). Results show the long-run time averages over 3×10^6 generations of the stochastic simulation.

Parameters: $\beta_{\rm r} = 20$, $g_{\rm A} = 0.15$.

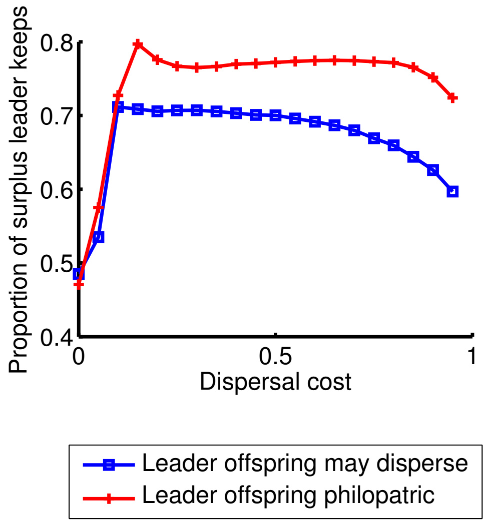


Figure S5: When the offspring of a leader can disperse based on the proportion of surplus that their parent took, then leaders evolve to take less surplus, particularly for higher dispersal costs. Results show the long-run time averages over 3×10^6 generations of the stochastic simulation.

Parameters: $\beta_{\mathrm{r}} = 20$, $g_{\mathrm{A}} = 0.15$.