Natural variation in *Caenorhabditis elegans* collective behavior reveals potential adaptation to novel environments

Youn Jae Kang[1,2], Antonio Carlos Costa[3], Siyu Serena Ding[1,4*]

1 Max Planck Institute of Animal Behavior, 78464 Konstanz, Germany
2 International Max Planck Research School (IMPRS) for Quantitative Behavior, Ecology & Evolution, 78464 Konstanz, Germany
3 Sorbonne University, Paris Brain Institute (ICM), Inserm U1127, CNRS UMR 7225, Paris, France
4 Centre for the Advanced Study of Collective Behaviour, 78464 Konstanz, Germany
* Corresponding author.

Correspondences to:
Siyu Serena Ding
Max Planck Institute of Animal Behavior, Universitaetsstrasse 10, 78464 Konstanz, Germany
+49 (0)7531 885904
serena.ding@ab.mpg.de

Preprint servers:
This manuscript is deposited at bioRxiv under a CC-BY-NC license.

Classification:
Biological Sciences/Evolution

Keywords:
Behavior evolution, collective behavior, *C. elegans*, colonization

Abstract

Recent advances in quantitative approaches to behavior have enabled rigorous trait measurements that allow us to understand how behavior evolves in multiple taxa(1–7). However, placing these patterns of divergence into a natural evolutionary context remains challenging without a well-resolved evolutionary history. Recent genomic analysis of wild *Caenorhabditis elegans* strains shows that this species has rapidly spread from its ancestral environments in the Hawaiian Pacific regions out to the rest of the world in the past 100-200 years(8, 9). This recent history sets the stage for addressing how *C. elegans* behavior has adapted to novel environments in this colonization process. We focus on the evolution of collective behavior, which is ecologically relevant for this boom-and-bust species(10). We selected a panel of 13 genetically diverse wild strains of *C. elegans* sampled from around the globe, categorized them into environmentally-relevant climate groups, and conducted aggregation behavior experiments and analysis to capture subtle natural variation among these populations. We found that the strains from the more recently colonized temperate regions show decreased spatial magnitude and temporal persistence in their aggregation behavior compared to strains from the ancestral sub/tropical regions. Furthermore, the observed behavioral divergence between the two climate groups closely aligns with genetically predicted phenotypic evolution, potentially reflecting heritable behavioral adaptation to temperate environments as the species spread globally. Our results showcase behavioral adaptation in the recent *C. elegans* colonization history and introduce a new perspective on using *C. elegans* as a model system for studying behavioral evolution.

Significance statement

Evolution has shaped phenotypic diversity as organisms respond to natural selection and adapt to their local environments. We examined how behavioral phenotypes have evolved as *C. elegans* spread out of their ancestral Hawaiian Pacific home to colonize the rest of the world within the last two centuries. We found that *C. elegans* aggregation, a form of collective behavior on food sources, shows genetically-mediated behavioral adaptation to novel temperate environments during this global colonization. Our results provide insights into broad-scale patterns of behavioral evolution across the globe, in a celebrated model organism species that holds much promise as a study system for behavioral evolution in a natural evolutionary context.

Introduction

How organisms evolve traits to better adapt to their environments is a central question to evolutionary biology. Even though behavior is a critical part of how organisms interface with their environment, phenotypes studied in evolutionary biology have traditionally focused on morphological traits that are relatively easy to define and measure. Moreover, in contrast to morphology which is relatively stable within an organism, behavior is much more dynamic and plastic, thus making it difficult to extract and attribute the proportion of behavioral variation to a heritable genetic basis(11). Rapid development in behavioral recording and analytical techniques in recent years are facilitating the endeavor to quantify behavior at a level of rigor commensurate with morphological analysis. Behavioral data thus systematically generated and measured have enabled sensitive phenotypic comparisons and evolutionary inferences in several taxa(1–7). However, as well-resolved recent evolutionary histories are lacking for most taxa, it remains challenging to set the variation in a natural evolutionary context to reconstruct or predict the trajectories of behavioral evolution in response to natural selection.

The recently characterized global colonization history of *Caenorhabditis elegans* provides a rare opportunity to place patterns of phenotypic divergence in a well-resolved evolutionary context to address how behavior adapts to local environments. *C. elegans* is a free-living microbivore nematode species commonly found in decomposing organic matter(12). Genomic analysis of 200 wild strains isolated worldwide shows a rapid global spread of this species from its ancestral environments in the Hawaiian Pacific region out to the rest of the world in the past 100-200 years(8, 9), thus raising the question of how *C. elegans* has adapted to novel environments in this colonization process. We focus on the evolution of *C. elegans* collective behavior to address this question. The invariant body form of this species makes behavior instead of morphology a plausible target of selection, and collective behavior is thought to evolve faster than individual behavior(13). Moreover, the boom-and-bust life cycle of *C. elegans* in nature(10) implicates rapid population expansion on food sources where many worms can be found in the same place (14–16) and even behaving together as a superorganism(17), underscoring the ecological relevance of studying the evolution of collective behavior in this species.

We focus on the natural variation of *C. elegans* aggregation behavior on food as a lens to understand how collective behavior has evolved in this species during its recent global colonization. Worms feeding on a bacterial food patch together can aggregate into clusters. Past work has examined behavioral variation in *C. elegans* aggregation and identified that variation in the *npr-1* gene underlies the respective solitary versus gregarious collective phenotypes in lab-domesticated versus wild strains(15). A body of subsequent work has focused on understanding the molecular and cellular mechanisms(18–20), the algorithmic emergence(21), and the potential fitness implication(22) of aggregation behavior by contrasting the canonically solitary lab reference strain N2 with its gregarious counterpart, the *npr-1* mutant. However, detailed behavioral variation amongst wild-collected strains has not been well-characterized, thus leaving open the question of how *C. elegans* aggregation may have diverged and evolved in natural environments. Meanwhile, a study of aggregation behavior among wild populations of another nematode species, *Pristionchus pacificus*, reveals natural behavioral variation associated with a regional altitudinal gradient in this species(23), further prompting a revisit of the evolutionary scenario in wild *C. elegans* populations.

We selected a panel of 13 genetically diverse wild *C. elegans* strains that were isolated from around the globe, and characterized their isolation sites into environmentally-relevant climate groups to reveal that the collection covers both sub/tropical and temperate environments. We conducted controlled behavior experiments in the lab, and developed concise measurements for the spatial magnitude and the temporal dynamics of aggregation behavior that captured natural variation across our populations. We found a consistent reduction in the magnitude and the temporal persistence of aggregation behavior in the temperate strains compared to those from sub/tropical regions. The variation in both behavioral traits has a large genetic contribution, and we further demonstrated that the direction of genetic covariance between the two traits predicts the observed phenotypic divergence

between the temperate and the sub/tropical groups. Taken together, our results suggest that *C. elegans* aggregation is an evolvable collective behavior that may have undergone adaptation to local environments during its recent global colonization.

Results

A panel of genetically divergent wild *C. elegans* strains represents distinct environmental groups

To widely sample natural variation in the collective behavior of wild *C. elegans* populations, we used a set of 12 highly divergent *C. elegans* strains predefined by *Caenorhabditis* Natural Diversity Resource (CaeNDR)(24), operating on the assumption that high levels of genetic divergence are more likely to confer sufficient behavioral variation that may reveal evolutionary signatures. These strains were isolated from various regions around the globe, representing both the ancestral Hawaiian Pacific regions (CB4856, DL238) as well as more recently colonized regions (Fig 1a). The lab reference strain N2 also forms part of the collection, even though it has long adapted to laboratory conditions after its original isolation in the 1950s(25). We additionally included JU1249, the only wild *C. elegans* strain reported to show solitary behavior on food(26) (see Table S1 for a list of strains).

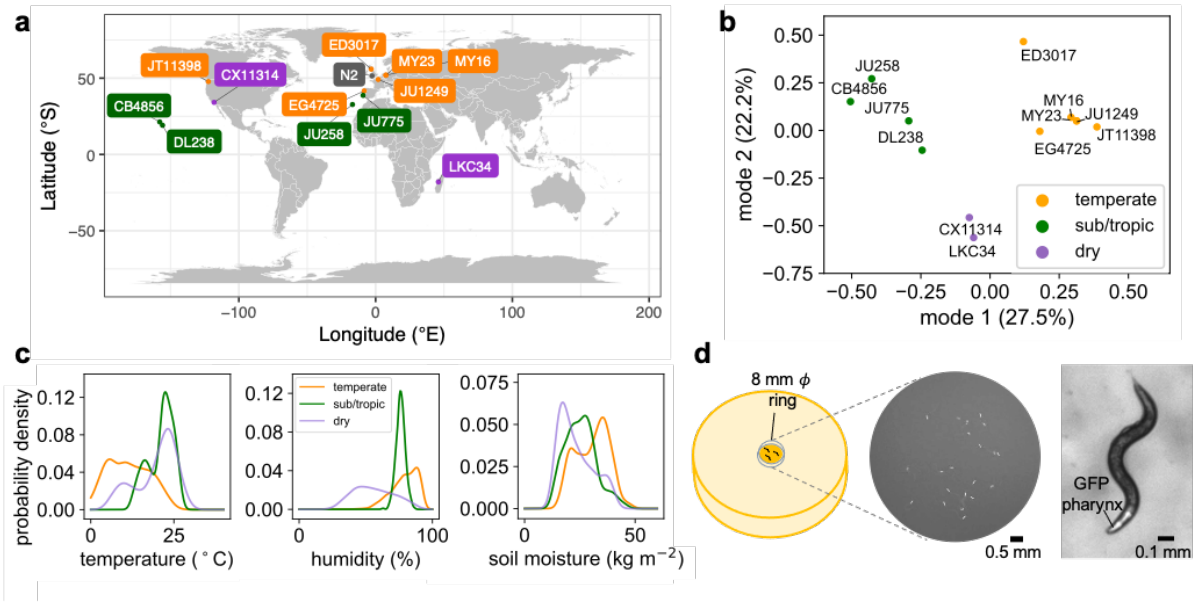


Figure 1: Climate categorization of the *C. elegans* strains. a. Geographical locations of the isolation sites. b. Constant shift embedding space of environments across the sampling locations, projected along the first two eigenvectors. Percentage denotes variance explained by the eigenvector. c. Kernel density estimation of the distribution of climate variables for each climate group, estimated with a Gaussian kernel with unit variance. d. Aggregation behavioral assay. The center of NGM agar was seeded with a patch of 10 µL of OP50 food surrounded by a polypropylene ring. Twenty worms were then placed inside the ring and recorded for 25 minutes using GFP fluorescence imaging to visualize the pharynges of the worms.

Organisms inhabiting diverse environments can exhibit different phenotypic and genetic structures that reveal possible trajectories of adaptation(27). To characterize the environments where our strains come from, we analyzed the climate data of the isolation sites and categorized the populations into climatically relevant groups (Fig 1b,c). We chose three climate variables: near-surface temperature, near-surface relative humidity, and soil moisture of the upper column, which not only capture the broad climate trends of the location but also approximate the finer microbial environment that *C. elegans* naturally inhabits(28–30). We obtained monthly data of the global climate model(31, 32) (Fig S1a) from the Climate Data Store (CDS) (see Methods), covering the 15-year period during which most strains in our collection were isolated. Excluding the lab-domesticated N2 strain from this analysis, we examined information of both the average values and the monthly fluctuations of the environmental variables and performed constant shift embedding(33) and k-means clustering (see Methods) to identify three environmental clusters with distinct climate characteristics (Fig 1b). The temperate group (JU1249, EG4725, MY16, ED3017, MY23, JT11398) is characterized by lower average temperature and high environmental fluctuations; the subtropical or tropical group (CB4856, DL238, JU775, JU258) is characterized by higher average temperature and low fluctuations of the climate variables; and the third dry group (CX11314, LKC34) mostly has low average values and

high fluctuations in humidity (Fig 1c, Fig S1b). In line with a previous work that sought to identify specific correlations between natural genetic variation and climate variables in wild *C. elegans* strains(34), here we focused on assigning broad environmental categories using longitudinal climate model data with high spatial resolution (see Methods), to assess how behavior may have adapted to these different environments.

Behavioral assay reveals natural variation in aggregation behavior

So is there natural variation in *C. elegans* aggregation behavior then? The common notion is that all wild strains of *C. elegans* aggregate on food(15, 22) except for a single solitary wild strain JU1249(26). Indeed, under routine culturing conditions where worms are grown at very high densities, aggregation tendencies of most wild strains did not show obvious differences. However, at the density of 20 worms behaving on an 8 mm (~8 times the body length) diameter uniform food patch, we observed diverse aggregation phenotypes among our strains. Therefore, we performed 25-minute recordings under this assay condition to experimentally capture the behavioral variation. To circumvent physical overlap between individuals during aggregation and enable collective behavior analysis with individual tracking resolution(21) (Fig 1d), we generated pharynx::GFP (*mIs12*) introgression lines of the wild strains (Table 1). We modified a high throughput behavioral rig(35) for fluorescence imaging, and recorded the experiments under GFP imaging conditions to visualize head positions over time (see Methods). We used age-matched clonal individuals to reduce inter-individual variation and conducted the experiments under controlled laboratory conditions, such that the behavioral variation can mainly be attributed to strain-level differences. We observed that some strains aggregate more strongly than the others, and that the dynamic formation and dissolution of aggregates also varied across our strains. Therefore, our behavioral assay accentuates previously uncharacterized spatial and temporal variation in the collective behavior of these wild strains, which we subsequently quantified for further evolutionary inferences.

Temperate strains show reduced magnitude and persistence in aggregation behavior

To quantify the observed variation from video recordings, we tracked individual head positions using TRex(36) (see Methods) and computed aggregation magnitude over time. Previous aggregation metrics such as the percentage of animals inside groups(15, 18–20, 22, 23) or the spatial distribution of individuals(21) ignore the temporal aspect of this dynamic collective behavior. To faithfully capture the rich dynamics, we defined an aggregation metric that can be calculated on a per-frame basis: the normalized inverse spatial entropy of the distribution of individual positions, $A(t)$ (Fig 2a, Fig S3) (see Methods). If worms are tightly aggregated, the probability of finding individuals in that spatial location increases, hence the total entropy decreases and the aggregation is higher (Fig 2a). To summarize the time series, we extracted one scalar value each from the spatial and the temporal axes of the data to obtain time averaged magnitude μ_A and temporal persistence τ_A (see Methods). Higher mean magnitude signifies more clustered positions, and higher temporal persistence indicates more stable behavior over time. We found varying levels of aggregation magnitude (Fig 2b, Fig S4a) and persistence (Fig 2c, Fig S4b) among our strains which corroborate our qualitative assessment from raw videos, with solitary strains such as N2 and JU1249 exhibiting low values for both behavioral traits. Our metrics thus effectively compress high dimensional behavior data into two simple and intuitive trait values whilst retaining the sensitivity to capture natural variation and key features of our dataset.

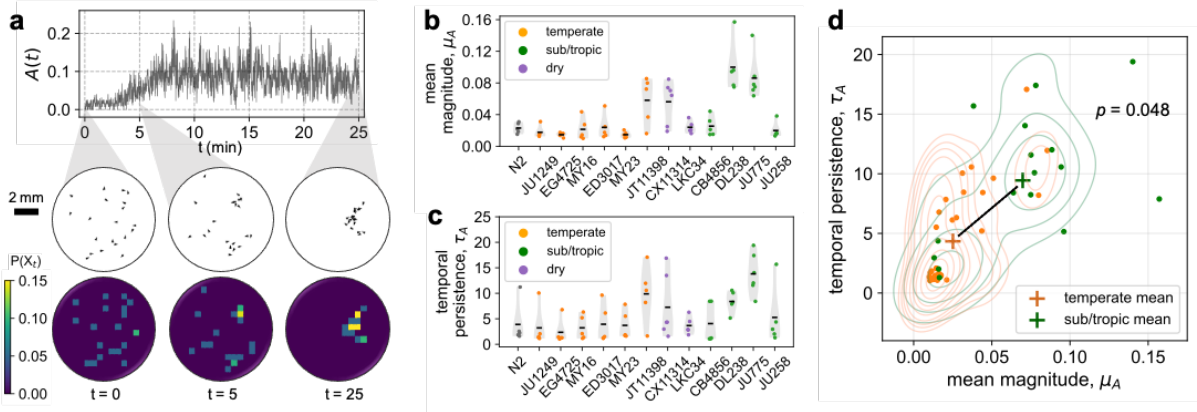


Figure 2: Natural variation in aggregation phenotypes. a. Sample aggregation time series, from a recording of the DL238 strain. The top panel shows the time series, the middle panel shows actual worm positions at a given time point, and the bottom panel depicts the corresponding spatial probability. Aggregation magnitude, *A(t),* is defined as the inverse spatial entropy. b. Time averaged magnitude, μ_A, and c. temporal persistence, τ_A, for all strains (see Methods). Each dot represents a single experimental replicate and black lines are the means of the replicates. d. Joint distribution of the two aggregation traits as the contour, estimated by a Gaussian kernel with unit variance. Each dot represents a single experiment (N=30 for temperate, N=21 for sub/tropic). Bivariate permutation test accounting for the strain subgroup reports significant divergence between the two climate groups (p = 0.048).

Next, we compared the aggregation traits between the climate groups to identify potential behavioral adaptations. Since the temperate and the sub/tropic groups contain the majority of our strains, we focused our comparative analysis on these two major climate groups. We performed rank-based permutation tests and found a significant difference in the mean aggregation magnitude (p = 0.039) but not temporal persistence (p = 0.108) between the two groups (Fig S6a) (see Methods). The bivariate comparison shows a significant difference (p = 0.048) (see Methods) in the joint distributions of these behavioral traits, with the temperate group showing overall less aggregation compared to the sub/tropic group (Fig 2d, Fig S6b). Since the global spread of this species originated from the tropical Hawaiian Pacific regions(8), we infer that as *C. elegans* populations spread from the sub/tropics to global temperate regions, they may have evolved lower aggregation phenotypes in response to novel temperate environments.

The genetic basis of the behavioral traits predicts actual phenotypic divergence

To evaluate the potential genetic basis of this behavioral adaptation, we tested if the behavioral traits are heritable. Accounting for the phylogenetic relationship (Fig 3a), we estimated the genetic component of the aggregation traits in all 13 strains using linear mixed model-based heritability calculations from CaeNDR (Fig 3b) (see Methods). We estimated broad-sense heritability, revealing that 67.6% (CI: [50.2%, 72.5%]) of the variation in mean aggregation magnitude, and 33.1% (CI [0%, 42.2%]) of the variation in temporal persistence are explained by the strain-group effect (Fig 3b top). Next, we calculated narrow-sense heritability by including the additive genetic information using the covariance of the genotypes in single-nucleotide polymorphism (SNP) markers (see Methods). The results demonstrate that 48.3% (CI [33.8%, 54.9%]) of the variation in mean magnitude and 20% (CI [3.2%, 27.7%]) of the variation in temporal persistence are explained by the additive genetic effect (Fig 3b bottom). These heritability scores greatly surpass the average heritability of behavior (23.5%)(11), suggesting that *C. elegans* aggregation behavior could readily adapt via genetic mediation.

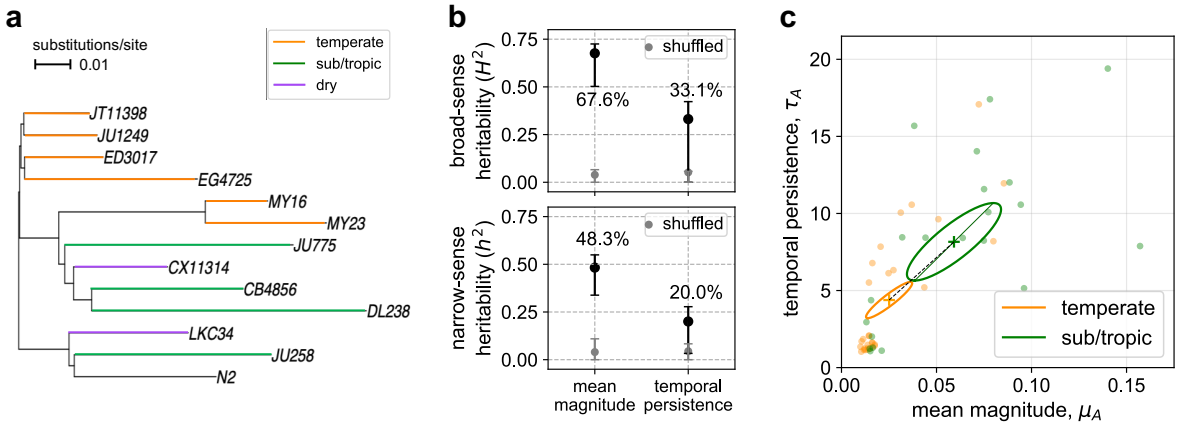


Figure 3: The genetic basis of behavioral evolution. a. Phylogenetic tree showing relatedness of the 13 divergent strains, adapted from CaeNDR. Length of the scale bar denotes the number of nucleotide substitutions per site. b. Heritability of aggregation traits. Broad-sense heritability assumes strains are independent, whereas narrow-sense considers the strain's genetic covariances. Error bars denote 95% confidence interval of the bootstrapped heritability. Grey points are control estimates for heritability generated by random shuffling across the strains. c. Bivariate phenotypic space of aggregation traits superimposed with G matrices as ellipses (see Methods). Each dot represents a single experiment (N=30 for temperate, N=21 for sub/tropic). The dashed line between the two ellipse centers represents the direction of phenotypic divergence, which aligns closely with the primary axis of the G matrix of the sub/tropic group.

To further evaluate how natural genetic variation relates to the observed phenotypic divergence, we assessed the presumed evolutionary trajectory between the sub/tropical and temperate groups. Using a bivariate linear mixed model, we estimated the genetic variance-covariance component, the *G* matrix, of the aggregation phenotypes in each climate group (see Methods). *G* matrices were represented as ellipses, with the bivariate phenotypic mean as their centers(27, 37). The primary axis of genetic covariance *G* matrix reflects the direction of the least genetic resistance, in which corresponding phenotypes are expected to evolve(38–40). Our result shows a high level of estimated genetic covariance between the mean magnitude and the temporal persistence of aggregation in both the temperate and the sub/tropic groups (Fig 4c). Importantly, the direction of genetic covariance in the sub/tropic group, represented as the diagonal primary axis of the green ellipse, aligned closely with the direction of the phenotypic change between the groups, represented as the line between the two ellipse centers (Fig 4c). Our result shows that not only do the genetic components of the two aggregation traits covary, but that the direction of predicted adaptation also highly explains where the phenotypes have actually diverged to. In other words, the phenotypic divergence of aggregation behavior between the temperate and the sub/tropic groups conforms to the expected evolutionary scenario inferred from the colonization history and the genetic structure of the traits.

Discussion

The recently resolved colonization history of *C. elegans* provides a pertinent natural history context to examine behavioral evolution in this species. We show that *C. elegans* aggregation is not only a heritable collective behavior, but also that the traits could have potentially been selected for the local environments that the worms have experienced during its recent colonization of worldwide temperate regions. Moreover, mean aggregation magnitude and temporal persistence exhibit high genetic covariance (Fig 3c), which suggests the possibility of pleiotropy(37). Hence, as *C. elegans* spread globally, aggregation behavior could have experienced selective pressure in the new temperate environments with one or both traits under selection.

The question remains what the relevant selective pressure might be and whether this behavioral evolution confers adaptive benefits. We postulate that the stability of microbial food resources could be the driver of selection, and that a reduction in aggregation behavior (Fig 2d) could reflect altered foraging behavior that potentially confers fitness benefits in highly fluctuating temperate environments (Fig S1) where microbial resources for *C. elegans* are more ephemeral. Specifically, *C.*

elegans encodes the variability of food sources in their environment to modify their internal search behavior(41, 42) and switches between local dwelling and global roaming modes(43). The latter behavior facilitates the dispersal of animals that could benefit their survival by hedging their bets on potential future resources, as opposed to depending solely on the current unpredictable food source(44). On the other hand, if food supply is predictable, animals are allowed much more liberty to stay in their local search(42, 45), which increases the probability of aggregation. Therefore, selection against local dwelling could potentially serve to benefit the survival of individuals inhabiting temperate climates. One limitation of our study is that we rely on broader climate variables as a proxy for microbial ecology(28–30). Future sample collections that include direct measurements of microbiome stability from the sites where the worms have evolved would help to strengthen the potential connection between the behavioral phenotype and this adaptive explanation. Furthermore, the speculated role of individual foraging behavior on aggregation implies that altered collective phenotypes may arise from modifications to individual behavior alone. However, interactions between individuals may also play a role in shaping the collective behavior(21).Therefore, further studies scrutinizing the individuals' internal foraging states and their interactions inside the group could help disentangle intra- and inter-individual level mechanisms in the evolution of *C. elegans* aggregation behavior.

A key development of our study is the experimental and analytical methods to effectively expose and measure natural variation in *C. elegans* aggregation behavior. Most previous studies of *C. elegans* aggregation treat it as a static and binary phenotype(15, 18–20, 22, 23); while semi-quantitative measurements were sufficient for *P. pacificus* where wild populations show a strong divergence between solitary or aggregating phenotypes(23), the case of *C. elegans* requires finer tuning of experimental parameters and rigorous behavioral quantifications to reveal more subtle variations. This allowed us to further examine the genetic covariance of the behavioral traits and make evolutionary inferences on the adaptive history of *C. elegans* collective behavior. Furthermore, as the aggregation behavior was measured in controlled laboratory settings in our study, there is no direct effect of the natural environment on the behavior. In this way, we could isolate the indirect effect of the environment acting through selection. For instance, temperature directly affects *C. elegans* locomotion speed and the wavelength of undulation(46), so that if locomotive behavior were measured in natural habitats at different temperatures, behavior variation can be attributed to the temperature effect, confounding the evolutionary effect that comes from adaptation to different environments. Our controlled experiments thus dissociate the direct environmental effect on behavior and enable a clear interpretation of genetically mediated evolutionary effect.

Our study reveals potential behavioral adaptations to novel temperate environments in the history of *C. elegans* global colonization. The study is based on a small number of genetically diverse wild strains, but hundreds more exist in public collection(24). Further characterization of aggregation behavior in more wild strains could not only strengthen our current results, but also help identify novel genetic regulators of the behavior through e.g. genome-wide association studies. Furthermore, being a major laboratory model organism, we can leverage the vast experimental feasibility available to *C. elegans* to e.g. conduct experimental evolution to validate the putative selection process on collective behavior, and to elucidate the detailed molecular and neuronal mechanisms in response to this selection. In summary, with the evidence supporting a potential adaptive history of collective behavior, a vast collection of geo-referenced wild strains showing genetic and behavioral variation, and empirical possibilities for downstream mechanistic dissection, *C. elegans* aggregation behavior serves as a powerful opportunity to exploit the synergy between proximate and ultimate explanations for future studies of behavioral evolution.

Methods

C. elegans strains

Eleven divergent wild strains of *C. elegans* were obtained from CaeNDR(24), and JU1249(26) was obtained from the lab of Marie-Anne Félix. GFP-pharynx label (*mIs12*) was introgressed(47) into these wild strain backgrounds; details of the wild strain introgression lines are in Table 1. OMG10 strain, the GFP-labelled version of N2(21), was obtained from the lab of Andre Brown. Animals were regularly cultured and maintained on nematode growth media (NGM) plates using standard protocol(48), and fed *E. coli* OP50 as worm food. *E. coli* OP50 was obtained from the *Caenorhabditis* Genetics Center (CGC) and was cultured in Luria Broth using standard liquid culture protocol(49).

Clustering of climate groups

CMIP6 global climate model(31, 32) data was obtained from Copernicus Climate Change Service (C3S) Climate Data Store (DOI: 10.24381/cds.c866074c); specific model details are in Table S2. The relevant time range was 2000-2014 to cover the isolation period of most of our *C. elegans* strains except for CB4856, which was collected in 1972. However, given the periodicity of the climate variables (Fig S1a) and the simplicity of the climate measurements we used (mean and standard deviation), we expect that these climate variables are relatively consistent beyond the years we investigated to support our overall conclusions. Each climate variable was interpolated into spatial resolution of 1.4° latitude and longitude spanning the entire globe. Isolation sites of the 13 strains were matched to geographical coordinates with less than 0.7° error in both latitude and longitude (see Data availability).

We clustered the isolation sites of 12 strains (excluding the lab-domesticated N2) based on the climate characteristics. First, we z-scored the climate variables across all strains to normalize their pooled range, and then we computed the mean and the standard deviation of each variable for each strain (Fig S1b). To account for the behavioral contribution of each climate variable, we estimated the correlation coefficients between the climate variables and the aggregation traits (Fig S2b), and used the absolute value of the correlation coefficients averaged over the two aggregation phenotypes as a weight. Each strain ends up with a unique climate vector that consists of the mean and the standard deviation of the normalized temperature, humidity, and soil moisture weighted by the corresponding correlation to the behavioral traits. We then generated the environmental distance matrix with Euclidean distances between pairs of climate vectors (Fig S2a). Finally, we performed constant shift embedding to map the environmental differences into a space in which Euclidean distances equal the environmental differences(33) (Fig 1c). We then chose the 10 highest eigenvectors which explain ≈98.76% total variance (Fig S2c), and performed k-means clustering of the environmental space (50). We estimated cluster stability by bootstrapping the points and re-estimating k-means assignments, identifying the fraction of points that consistently clustered together across bootstrapped samples(51). We tested different numbers of clusters and found that maximum cluster stability was achieved with three environmental clusters (Fig S2d).

Behavioral assay and video processing

Synchronized animals were obtained using a bleach-free L1 synchronization protocol(52) for a less invasive preparation. Synchronized L1 worms were refed onto the seeded NGM plates to develop for 65 ± 1 hours, when we conducted imaging experiments with them as freshly Day-1 adults. Imaging conditions were maintained at 19 ± 1 ºC, and humidity 70 ± 5 %. To create a uniform food patch, a fresh overnight liquid culture of *E. coli* OP50 (OD_{600} = 2.05 ± 0.05) was diluted 1:1 in LB broth (20 g LB broth mix, H_2O to 1 liter) and 10 µL of the diluted culture was freshly seeded in the middle of each well (35 mm diameter) of a 6-well plate filled with 306 mL NGM agar. The plates were left to dry at 20 ºC for 5 minutes. Manually seeding 10 µL of OP50 yielded a circle of 7 ± 1 mm diameter. A polypropylene barrier of 8 mm inner diameter (fabricated by cross-section slicing 1000 µL pipette tips) was then added to surround the food patch in each of the wells to restrict the behavioral arena to the size of the food patch. No strong bordering tendency was observed in our experiments on these uniform bacterial patches (Fig 1d, 2a).

Day-1 adults were washed twice with 1 mL M9 buffer (3 g KH_2PO_4, 6 g Na_2HPO_4, 5 g NaCl, 1 mL 1M $MgSO_4$, H_2O to 1 liter) and dispensed as a concentrated droplet onto an unseeded 3.5 cm diameter NGM plate. The droplet was left to dry at room temperature for 5 minutes, and 20 ± 1 worms were picked and placed onto the food patches in 6-well plates. Animals were allowed one minute acclimatization inside the imaging box and were recorded with GFP excitation (450 - 490 nm) for 25 minutes at 25 Hz (Megapixel camera array(35) adapted for GFP fluorescence imaging, magnification 20X, blue light 32 ± 1 Lux, exposure 39 ms). Videos obtained from the imaging assay were tracked using the automated visual tracking software TRex(36). Parameters optimized for the tracking are available on the data repository. Our videos feature only the pharynges of worms without the body (Fig 1d), therefore the tracked data are positional coordinates of the centroids of individual pharynges for each video frame. Due to overlaps between pharynges in close proximity, not all individuals were tracked in every frame. This was corrected by normalizing the aggregation measurements (see below).

Aggregation trait measurements

To estimate the magnitude of aggregation at each time t, we compute the inverse of the spatial entropy of the (*x,y*) coordinates of all individual worms,

$$\hat{A}(t) := -\frac{1}{\sum_{X_t} P(X_t) log P(X_t)}$$

where the positions of individuals at time t are $X_t = (x_t, y_t)$ and the spatial probability of the positions is $P(X_t)$ (Fig 2a). We approximate $P(X_t)$ through a spatial binning into bins of size 0.5 mm, half the body length of the worm (Fig 2a). To account for the variable number of tracked worms n_t in each frame t, we normalize the metric by subtracting the expected aggregation magnitude for a uniform distribution of n_t worms, to obtain$A(t) := \hat{A}(t) - log(n_t)$ (Fig S3b).

From the time series $A(t), t \epsilon \{\delta t, \ldots, T\delta t\}$ obtained from each trial, where δt is the inverse frame rate, we estimate mean magnitude as,

$$\mu_A =< A(t) >= \frac{1}{T}\int A(t)\, dt.$$

To capture the temporal dynamics, we measured the normalized autocorrelation function,

$$C(\tau) = \frac{< (A(t)-\mu_A)(A(t+\tau)-\mu_A) >}{< (A(t)-\mu_A)^2 >}.$$

The complex nature of the correlation functions (Fig S4b), which deviate from simple exponential decay (likely due to finite-size effects), challenges the inference of decay times for each trial (53, 54). To nonetheless capture the strength of temporal correlations, we computed the area under the curve until the autocorrelation first reaches zero, obtaining $\tau_A := max(\int_0^T C(\tau) d\tau)$ as an overall measure of the temporal persistence of the aggregation dynamics.

Statistical test of aggregation phenotypes

We performed a rank-based permutation test for a non-parametric comparison between the temperate and the sub/tropic groups, accounting for within-strain dependency across samples. We computed the null distribution by random shuffling across strains, rather than across each data point to avoid independent sampling assumption. We then estimated the p-value as the proportion of the null distribution higher than the observed differences of mean magnitude and temporal persistence ranks between the temperate and sub/tropic groups. For a two-dimensional comparison of the joint distribution of aggregation phenotypes, we estimated the Euclidean distance between climate groups with respect to the mean aggregation ranks and the mean temporal persistence ranks to represent the bivariate change in a phenotypic space:

$$\sqrt{(< rank(\mu_{temp}) > -< rank(\mu_{trop}) >)^2 + (< rank(\tau_{temp}) > -< rank(\tau_{trop}) >)^2}$$

We generated the null distribution of Euclidean distances by the same random shuffling, and estimated p-value as the proportion of the null distribution higher than the observed Euclidean distance between the two climate groups.

Heritability estimation

We estimated broad-sense heritability (H^2) and narrow-sense heritability (h^2) to assess the respective strain-group or additive genetic effect on the behavioral straits. The linear mixed model is,

$$y_{ij} = m + u_i + \varepsilon_{ij}, \text{ where}$$

$$m = < y_{ij} > \text{ for the fixed effect,}$$

$$u \sim N(0, M\sigma_u^2) \text{ for strain/genetic effect,}$$

$$\varepsilon \sim N(0, I\sigma_\varepsilon^2) \text{ for the residuals,}$$

and $y_i = [y_{i1} \cdots y_{ij}]$, $\varepsilon_i = [\varepsilon_{i1} \cdots \varepsilon_{ij}]$, for i = strain number (n), and j = replicate number. We use the R sommer package to estimate σ_u^2 and σ_ε^2 by maximizing the log likelihood function computed based on the Restricted Maximum Likelihood (REML) of the Y distribution(55). Heritability, $\frac{\sigma_u^2}{\sigma_u^2+\sigma_\varepsilon^2}$, captures the ratio of strain-group or additive genetic variance over total phenotypic variance. The model is inferred with $M = I$ (the identity matrix) for H^2 and $M = A$ for h^2. The additive genetic relations matrix A was constructed using the Van Raden(56) method, $A = \frac{MM^T}{\Sigma 2p_i q_i}$, where the covariance of the genotypes among strains (MM^T) at 58,760 different SNPs is normalized by allele frequency (p_i = frequency of allele 1 at locus i and q_i = frequency of allele 2 at locus i)(24, 57, 58).

Genetic covariance estimation

The model to estimate genetic covariance matrix G is equal to the narrow-sense heritability model with the multivariate input. In vector notation,

$$Y = m + ZU + \varepsilon$$

$$\begin{bmatrix} Y_\mu \\ Y_\tau \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} m_\mu \\ m_\tau \end{bmatrix} + \begin{bmatrix} Z_\mu & 0 \\ 0 & Z_\tau \end{bmatrix} \begin{bmatrix} U_\mu \\ U_\tau \end{bmatrix} + \begin{bmatrix} \varepsilon_\mu \\ \varepsilon_\tau \end{bmatrix}$$

$$U \sim N(0, G \otimes A) \text{ for the random effect,}$$

$$\varepsilon \sim N(0, I\sigma_\varepsilon^2) \text{ for the residuals, where}$$

$$G \otimes A = \begin{bmatrix} \sigma_\mu^2 A & \sigma_{\mu,\tau} A \\ \sigma_{\mu,\tau} A & \sigma_\tau^2 A \end{bmatrix}.$$

The genetic variance components of mean magnitude, y_μ (σ_μ^2) and temporal persistence, y_τ (σ_τ^2), and the genetic covariance between the two traits ($\sigma_{\mu,\tau}$) are stored in matrix form(27, 37) as,

$$G = \begin{bmatrix} \sigma_\mu^2 & \sigma_{\mu,\tau} \\ \sigma_{\mu,\tau} & \sigma_\tau^2 \end{bmatrix}.$$

The code for estimating the heritability and genetic covariance was adapted from the CaeNDR heritability calculation (https://github.com/AndersenLab/calc_heritability.git). Main dependencies include the sommer package(59) of R and bcftools(60) to generate the genotype matrix from the genome sequence of the strains (see Code Availability).

Data Availability

Datasets used in this study will be available at Zenodo upon publication (https://zenodo.org/uploads/14937989), which includes original tracking data from the experiments in this study, climate data obtained from Climate Data Store, information on the natural habitats of the wild *C. elegans* strains, and phylogenetic information from CaeNDR.

Code Availability

Codes for analyses and generating figures are deposited on Github (https://github.com/SerenaDingLab/Kang_et_al_AggEvol25), which includes the R code for the linear mixed model adapted from CaeNDR's heritability calculation.

Acknowledgements and Funding Information

We sincerely thank Erik Andersen and Ryan Greenway for their helpful insights and comments on this work. We thank Max Hofbauer for the help with modifying the behavior rig for fluorescence imaging, *Caenorhabditis* Natural Diversity Resources for the wild strains, Andre Brown and Marie-Anne Felix for sharing worm strains, and the *Caenorhabditis* Genetics Center (P40 OD010440) for *E. coli* OP50. We acknowledge the World Climate Research Programme and its Working Group on Coupled Modelling who coordinated and promoted CMIP6. We also thank multiple funding agencies who support CMIP6 and ESGF, and the climate modeling groups and the Earth System Grid Federation (ESGF) for producing and archiving the data. This research was supported by the Max Planck Institute of Animal Behavior and the International Max Planck Research School for Quantitative Behavior, Ecology & Evolution. We also thank Claire Wyart for the support, and acknowledge funding to A.C.C. from the Agence Nationale pour la Recherche LOCOCONNECT and RSCMAP awarded to C.W. An open access license has been selected for this work.

Author Contributions

Y. J. K., A. C. C. and S. S. D. conceptualized the study; Y. J. K. performed the experiments; Y. J. K. and A. C. C. analyzed the data; Y. J. K., A. C. C. and S. S. D. wrote the manuscript; S.S.D. acquired funding and supervised the project.

Declaration of interests

The authors declare no competing interests.

References

1. F. Mallard, B. Afonso, H. Teotónio, Selection and the direction of phenotypic evolution. *eLife* **12**, e80993 (2023).

2. D. G. Hernández, *et al.*, A framework for studying behavioral evolution by reconstructing ancestral repertoires. *eLife* **10**, e61806 (2021).

3. Y. Ding, A. Berrocal, T. Morita, K. D. Longden, D. L. Stern, Natural courtship song variation caused by an intronic retroelement in an ion channel gene. *Nature* **536**, 329–332 (2016).

4. R. A. York, R. D. Fernald, The repeated evolution of behavior. *Front. Ecol. Evol.* **4**, 143 (2017).

5. J. T. Walsh, S. Garnier, T. A. Linksvayer, Ant Collective Behavior Is Heritable and Shaped by

515 Selection. *Am. Nat.* **196**, 541–554 (2020).

516 6. J. C. Perry, *et al.*, Experimental evolution under hyper-promiscuity in Drosophila melanogaster.
517 *BMC Evol. Biol.* **16**, 131 (2016).

518 7. S. J. Helms, *et al.*, Modelling the ballistic-to-diffusive transition in nematode motility reveals
519 variation in exploratory behaviour across species. *J. R. Soc. Interface* **16**, 20190174 (2019).

520 8. E. C. Andersen, *et al.*, Chromosome-scale selective sweeps shape Caenorhabditis elegans
521 genomic diversity. *Nat. Genet.* **44**, 285–290 (2012).

522 9. D. Lee, *et al.*, Balancing selection maintains hyper-divergent haplotypes in Caenorhabditis
523 elegans. *Nat. Ecol. Evol.* **5**, 794–807 (2021).

524 10. L. Frézal, M.-A. Félix, C. elegans outside the Petri dish. *eLife* **4**, e05849 (2015).

525 11. N. A. Dochtermann, T. Schwab, M. Anderson Berdal, J. Dalos, R. Royauté, The Heritability of
526 Behavior: A Meta-analysis. *J. Hered.* **110**, 403–410 (2019).

527 12. B. S. Samuel, H. Rowedder, C. Braendle, M.-A. Félix, G. Ruvkun, Caenorhabditis elegans
528 responses to bacteria from its natural habitats. *Proc. Natl. Acad. Sci.* **113**, E3941–E3949 (2016).

529 13. G. N. Doering, *et al.*, Emergent collective behavior evolves more rapidly than individual
530 behavior among acorn ant species. *Proc. Natl. Acad. Sci.* **121**, e2420078121 (2024).

531 14. M. de Bono, D. M. Tobin, M. W. Davis, L. Avery, C. I. Bargmann, Social feeding in
532 Caenorhabditis elegans is induced by neurons that detect aversive stimuli. *Nature* **419**, 899–903
533 (2002).

534 15. M. de Bono, C. I. Bargmann, Natural Variation in a Neuropeptide Y Receptor Homolog
535 Modifies Social Behavior and Food Response in C. elegans. *Cell* **94**, 679–689 (1998).

536 16. A. D. Cutter, Nucleotide Polymorphism and Linkage Disequilibrium in Wild Populations of the
537 Partial Selfer Caenorhabditis elegans. *Genetics* **172**, 171–184 (2006).

538 17. D. M. Perez, R. Greenway, T. Stier, N. Font Massot, S. S. Ding, Nematode Towering Behavior
539 as a Powerful Experimental Model for Collective Dispersal. [Preprint] (2024). Available at:
540 https://papers.ssrn.com/abstract=4989935 [Accessed 28 April 2025].

541 18. B. H. Cheung, M. Cohen, C. Rogers, O. Albayram, M. De Bono, Experience-dependent
542 modulation of C. elegans behavior by ambient oxygen. *Curr. Biol.* **15**, 905–917 (2005).

543 19. C. Rogers, A. Persson, B. Cheung, M. De Bono, Behavioral Motifs and Neural Pathways
544 Coordinating O2 Responses and Aggregation in C. elegans. *Curr. Biol.* **16**, 649–659 (2006).

545 20. P. T. McGrath, *et al.*, Quantitative mapping of a digenic behavioral trait implicates globin
546 variation in C. elegans sensory behaviors. *Neuron* **61**, 692–699 (2009).

547 21. S. S. Ding, L. J. Schumacher, A. E. Javer, R. G. Endres, A. E. Brown, Shared behavioral
548 mechanisms underlie C. elegans aggregation and swarming. *eLife* **8**, e43318 (2019).

549 22. Y. Zhao, *et al.*, Changes to social feeding behaviors are not sufficient for fitness gains of the
550 Caenorhabditis elegans N2 reference strain. *Elife* **7**, e38675 (2018).

551 23. E. Moreno, A. McGaughran, C. Rödelsperger, M. Zimmer, R. J. Sommer, Oxygen-induced
552 social behaviours in Pristionchus pacificus have a distinct evolutionary history and genetic

regulation from Caenorhabditis elegans. *Proc. Biol. Sci.* **283**, 20152263 (2016).

24. T. A. Crombie, *et al.*, CaeNDR, the Caenorhabditis natural diversity resource. *Nucleic Acids Res.* **52**, D850–D858 (2024).

25. M. G. Sterken, L. B. Snoek, J. E. Kammenga, E. C. Andersen, The laboratory domestication of Caenorhabditis elegans. *Trends Genet. TIG* **31**, 224–231 (2015).

26. I. Beets, *et al.*, Natural Variation in a Dendritic Scaffold Protein Remodels Experience-Dependent Plasticity by Altering Neuropeptide Expression. *Neuron* **105**, 106-121.e10 (2020).

27. G. M. Walter, *et al.*, Environmental effects on genetic variance are likely to constrain adaptation in novel environments. *Evol. Lett.* **8**, 374–386 (2024).

28. X. Li, *et al.*, Temperature and moisture gradients drive the shifts of the bacterial microbiomes in 1000-year-old Mausoleums. *Atmosphere* **14**, 14 (2022).

29. J. Wirsching, *et al.*, Temperature and soil moisture change microbial allocation of pesticide-derived carbon. *Eur. J. Soil Sci.* **74**, e13417 (2023).

30. A. Borowik, J. Wyszkowska, Soil moisture as a factor affecting the microbiological and biochemical activity of soil. (2016).

31. V. Eyring, *et al.*, Overview of the Coupled Model Intercomparison Project Phase 6 (CMIP6) experimental design and organization. *Geosci. Model Dev.* **9**, 1937–1958 (2016).

32. R. Petrie, *et al.*, Coordinating an operational data distribution network for CMIP6 data. *Geosci. Model Dev.* **14**, 629–644 (2021).

33. V. Roth, J. Laub, M. Kawanabe, J. M. Buhmann, Optimal cluster preserving embedding of nonmetric proximity data. *IEEE Trans. Pattern Anal. Mach. Intell.* **25**, 1540–1551 (2003).

34. K. S. Evans, *et al.*, Correlations of Genotype with Climate Parameters Suggest Caenorhabditis elegans Niche Adaptations. *G3 Bethesda Md* **7**, 289–298 (2017).

35. I. L. Barlow, *et al.*, Megapixel camera arrays enable high-resolution animal tracking in multiwell plates. *Commun. Biol.* **5**, 253 (2022).

36. T. Walter, I. D. Couzin, TRex, a fast multi-animal tracking system with markerless identification, and 2D estimation of posture and visual fields. *Elife* **10**, e64000 (2021).

37. S. J. Arnold, R. Bürger, P. A. Hohenlohe, B. C. Ajie, A. G. Jones, Understanding the evolution and stability of the G-matrix. *Evolution* **62**, 2451–2461 (2008).

38. D. Schluter, Adaptive radiation along genetic lines of least resistance. *Evolution* **50**, 1766–1774 (1996).

39. G. M. Walter, Experimental Evidence That Phenotypic Evolution but Not Plasticity Occurs along Genetic Lines of Least Resistance in Homogeneous Environments. *Am. Nat.* **201**, E70–E89 (2023).

40. M. I. Lind, K. Yarlett, J. Reger, M. J. Carter, A. P. Beckerman, The alignment between phenotypic plasticity, the major axis of genetic variation and the response to selection. *Proc. R. Soc. B Biol. Sci.* **282**, 20151651 (2015).

41. A. J. Calhoun, *et al.*, Neural Mechanisms for Evaluating Environmental Variability in

Caenorhabditis elegans. *Neuron* **86**, 428–441 (2015).

42. M. Hendricks, Neuroecology: Tuning foraging strategies to environmental variability. *Curr. Biol.* **25**, R498–R500 (2015).

43. S. W. Flavell, *et al.*, Serotonin and the Neuropeptide PDF Initiate and Extend Opposing Behavioral States in C. elegans. *Cell* **154**, 1023–1035 (2013).

44. D. W. Stephens, The logic of risk-sensitive foraging preferences. *Anim. Behav.* **29**, 628–629 (1981).

45. K. R. Hopper, RISK-SPREADING AND BET-HEDGING IN INSECT POPULATION BIOLOGY1. *Annu. Rev. Entomol.* **44**, 535–560 (1999).

46. K. MacMillan, D. Mahdaviani, D. Matuszewski, Effect of temperature on caenorhabditis elegans locomotion. *The Expedition* **1** (2011).

47. A. Doroszuk, L. B. Snoek, E. Fradin, J. Riksen, J. Kammenga, A genome-wide library of CB4856/N2 introgression lines of Caenorhabditis elegans. *Nucleic Acids Res.* **37**, e110–e110 (2009).

48. S. Brenner, The genetics of Caenorhabditis elegans. *Genetics* **77**, 71–94 (1974).

49. L. Byerly, R. C. Cassada, R. L. Russell, The life cycle of the nematode Caenorhabditis elegans: I. Wild-type growth and reproduction. *Dev. Biol.* **51**, 23–33 (1976).

50. S. Lloyd, Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **28**, 129–137 (1982).

51. T. Lange, M. Braun, V. Roth, J. Buhmann, Stability-based model selection. *Adv. Neural Inf. Process. Syst.* **15** (2002).

52. R. A. Ellwood, *et al.*, Sulfur amino acid supplementation displays therapeutic potential in a C. elegans model of Duchenne muscular dystrophy. *Commun. Biol.* **5**, 1255 (2022).

53. J. Desponds, *et al.*, Precision of Readout at the hunchback Gene: Analyzing Short Transcription Time Traces in Living Fly Embryos. *PLOS Comput. Biol.* **12**, e1005256 (2016).

54. A. C. Costa, G. Sridhar, C. Wyart, M. Vergassola, Fluctuating Landscapes and Heavy Tails in Animal Behavior. *PRX Life* **2**, 023001 (2024).

55. G. Covarrubias-Pazaran, Quick start for the sommer package.

56. P. M. VanRaden, Efficient Methods to Compute Genomic Predictions. *J. Dairy Sci.* **91**, 4414–4423 (2008).

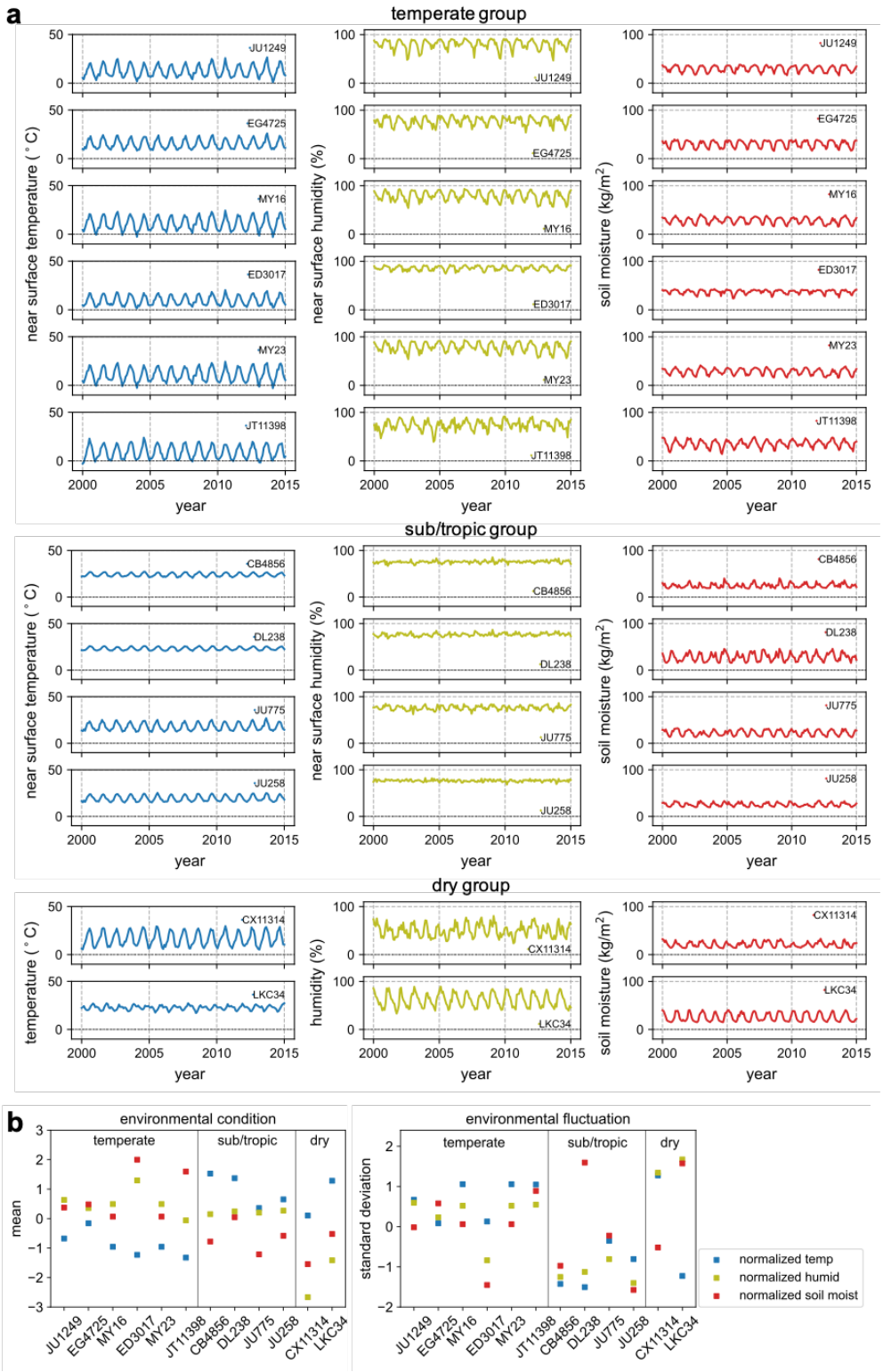
57. J. B. Endelman, J.-L. Jannink, Shrinkage Estimation of the Realized Relationship Matrix. *G3 GenesGenomesGenetics* **2**, 1405–1413 (2012).

58. G. Su, O. F. Christensen, T. Ostersen, M. Henryon, M. S. Lund, Estimating Additive and Non-Additive Genetic Variances and Predicting Genetic Merits Using Genome-Wide Dense Single Nucleotide Polymorphism Markers. *PLoS ONE* **7**, e45293 (2012).

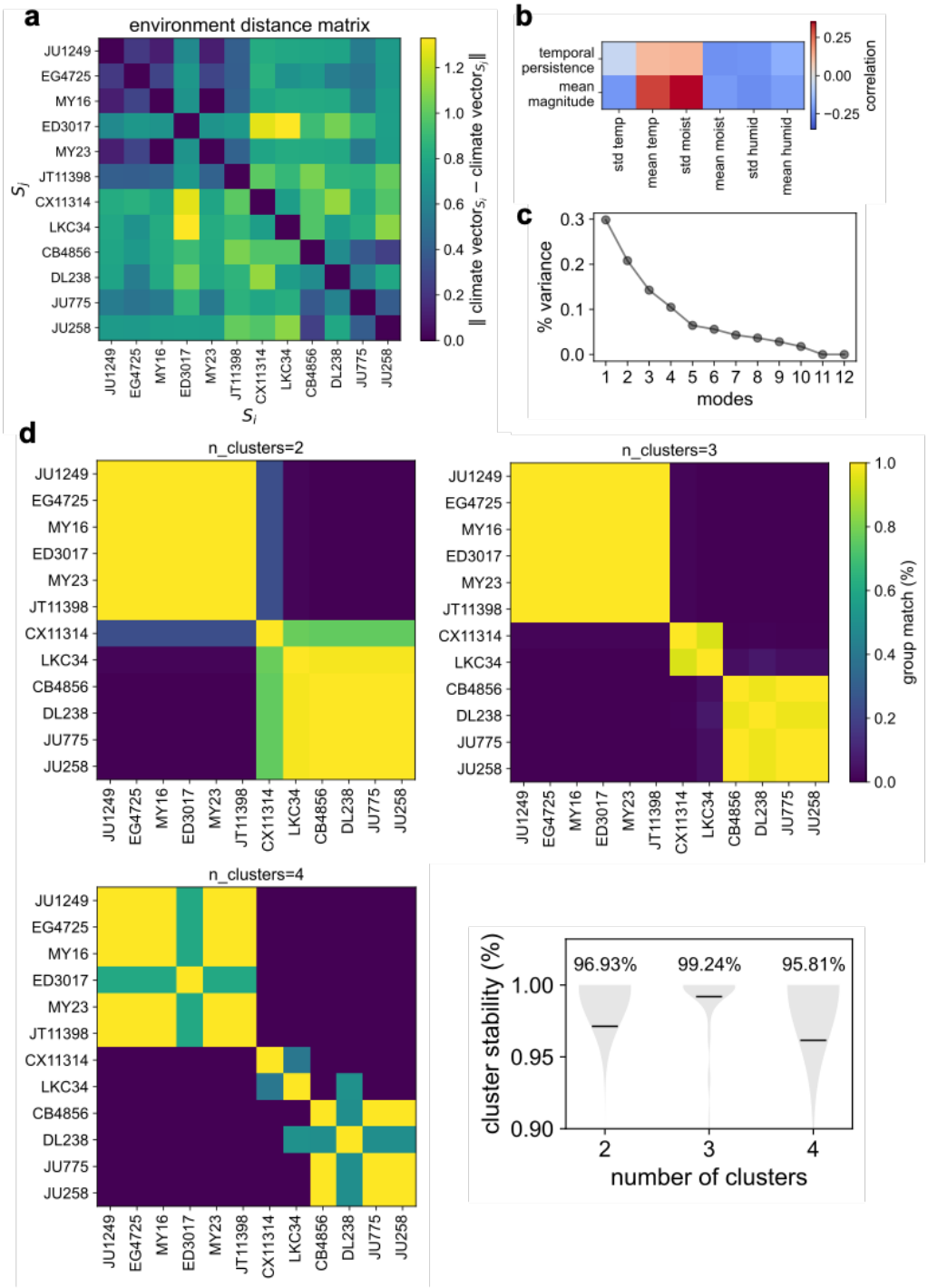
59. G. Covarrubias-Pazaran, Genome-Assisted Prediction of Quantitative Traits Using the R Package sommer. *PLOS ONE* **11**, e0156744 (2016).

60. P. Danecek, *et al.*, Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008 (2021).

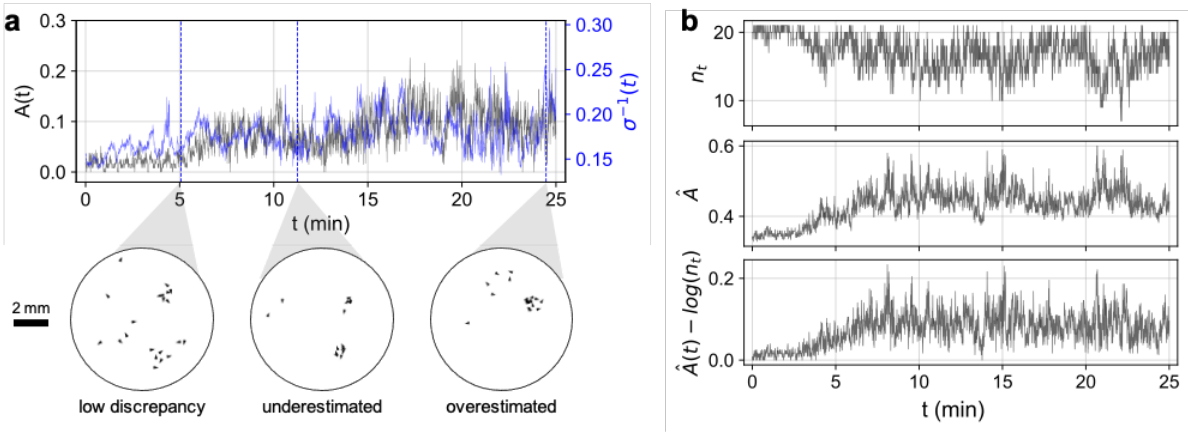
Supporting Information



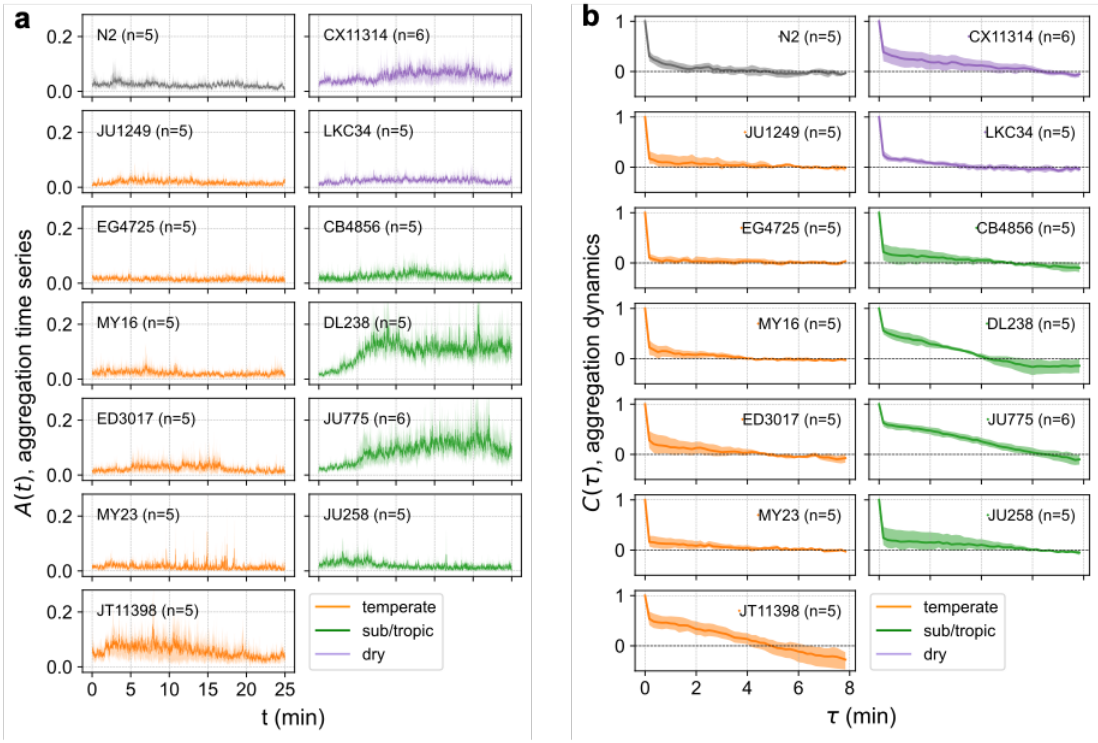
[Fig S1] Climate information of the strain isolation sites. a. Monthly records of near-surface temperature, near-surface relative humidity, and upper column soil moisture of the isolation sites from the period 2000-2014 (see Methods). b. Mean and standard deviation for each of the z-scored climatic variables across the isolation sites. The lab-domesticated N2 strain is excluded from this analysis.



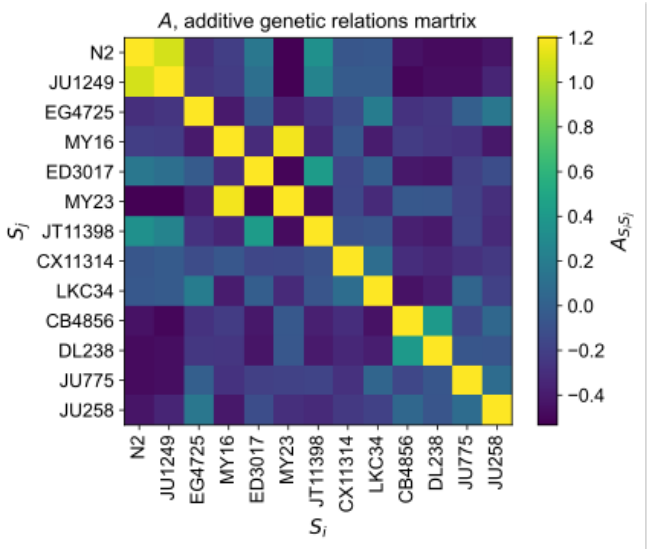
[Fig S2] Clustering of the strains into distinct climate groups. a. Environment distance matrix depicting the covariance of climate vectors between strains, weighted by correlation coefficients of the aggregation phenotypes (see Methods). b. Correlation coefficients between aggregation traits and climate statistics. c. Percentage of variance explained by the eigenvectors from the constant shift embedding of the environment distance matrix. d. Testing k-means grouping stability via bootstrapping. Climate variables were bootstrapped and clustered with varying cluster numbers. The highest stability was with three groups.



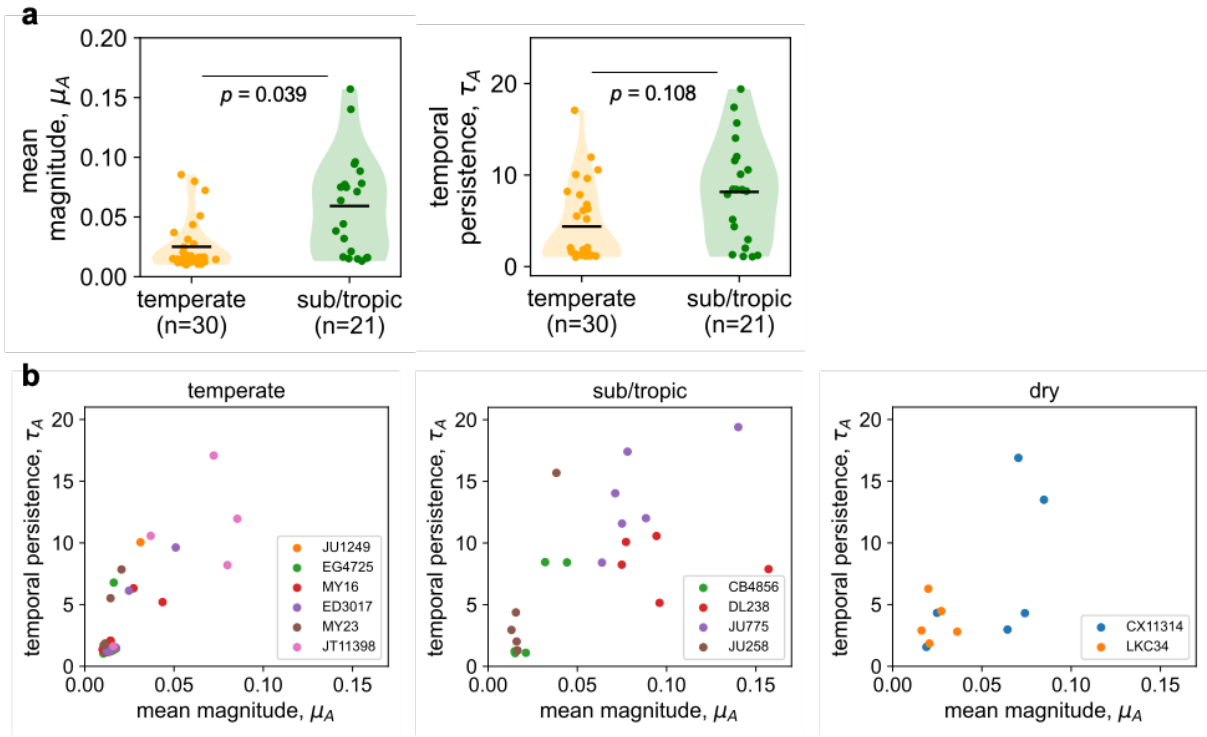
[Fig S3] Robustness of inverse entropy as a measure for aggregation magnitude. a. Comparison between the inverse entropy, $A(t)$ (black) and the inverse standard deviation, σ^{-1} (blue)(21) as measures of aggregation magnitude, using the same sample recording of DL238 as Fig 2. The general increasing trends are similar, with occasionally high discrepancies between the two measures. The bottom plots show the actual spatial distribution of the 20 worms. Inverse entropy measured aggregation more faithfully, especially for occasions with multiple aggregates (underestimated by $\sigma^{-1}(t)$), or mixed states of aggregates and dispersed individuals (overestimated by $\sigma^{-1}(t)$). b. Number of tracked individuals over time (top), aggregation time series before normalization (middle), and aggregation after normalization (bottom) (see Methods).



[Fig S4] Natural variation in aggregation behavior using time series data. a. Aggregation of the 13 strains over the 25 minutes recordings. Line depicts a strain's average aggregation magnitude over different experimental replicates, and experimental variation is captured in the 95% bootstrapped confidence interval as shading. b. Autocorrelation function of the 13 strains as the measure of aggregation dynamics. Line depicts a strain's average autocorrelation function over different experimental replicates, and experimental variation is captured in the 95% bootstrapped confidence interval as shading.



[Fig S5] Additive genetic relations matrix of the 13 strains. The matrix depicts the covariance of the SNP genotypes between strains s_i and s_j at different loci (see Methods).



[Fig S6] Phenotypic comparisons between climate groups. a. Independent comparison of aggregation phenotypes between the temperate and the sub/tropic groups. Black lines denote the average of each group. A rank-based permutation test accounting for the strain subgroup (see Methods) found significant differences between the group means (p = 0.039) for mean magnitude but not temporal persistence (p = 0.108). b. Bivariate phenotypic space of the three climate groups. Each dot represents a single experiment.

[Table S1] *C. elegans* strains

Original strain	Isolation site of the original strain	GFP-tagged strain	Genotype of the GFP-tagged strain
N2	England, UK	OMG10	*mIs12*[myo-2p::GFP+pes-10p::GFP+F22B7.9p::GFP]II; in N2 background (backcrossed 6x)
JU1249	Ile-de-France, France	SSD13	*mIs12II*; in JU1249 (backcrossed 6x)
EG4725	Braga, Portugal	SSD21	*mIs12II*; in EG4725 (backcrossed 8x)
MY16	Nordrhein-Westfalen, Germany	SSD22	*mIs12II*; in MY16 (backcrossed 8x)
ED3017	Scotland, UK	SSD20	*mIs12II*; in ED3017 (backcrossed 8x)
MY23	Nordrhein-Westfalen, Germany	OMG40	*mIs12II*; in MY23 (backcrossed 6x)
JT11398	Washington state, USA	SSD6	*mIs12II*; in JT11398 (backcrossed 6x)
CX11314	California state, USA	SSD19	*mIs12II*; in CX11314 (backcrossed 8x)
CB4856	Hawaii state, USA	OMG8	*mIs12II*; in CB4856 (backcrossed 10x)
LKC34	Antananarivo Province, Madagascar	SSD18	*mIs12II*; in LKC34 (backcrossed 8x)
DL238	Hawaii state, USA	SSD17	*mIs12II*; in DL238 (backcrossed 8x)
JU775	Lisbon, Portugal	SSD8	*mIs12II*; in JU775 (backcrossed 6x)
JU258	Madeira, Portugal	SSD10	*mIs12II*; in JU258 (backcrossed 6x)

[Table S2] Source information of the CMIP6 climate model

source ID	institution ID	citation
CNRM-ESM2-1	CNRM-CERFACS	Seferian, Roland (2018). *CNRM-CERFACS CNRM-ESM2-1 model output prepared for CMIP6 CMIP historical*. Version *20241215*. Earth System Grid Federation. https://doi.org/10.22033/ESGF/CMIP6.4068