

Supplemental Information for:

Early life telomeres are influenced by environments acting at multiple temporal and spatial scales

David F. Westneat, Rebecca C. Young, Alexandra G. Cones, Aurelia C. Kucera, Angelo Anacleto, Britt J. Heidinger

Table of Contents:

A. Data sets	Page 1-5
B. Data analysis	Page 5-12

A. Data sets

Pedigree file ("Pedigree.csv")

The pedigree file consisted of 1629 individuals who were a mix of nestlings sampled at 10 days of age (1591) and their parents if known. Variables were

1. SampleID: Unique number given to individual, usually representing their blood sample number.
2. Dad: SampleID or band number of the male at the nest, confirmed as the genetic father or observed there.
3. Mom: SampleID or band number of the female at the nest, confirmed as genetic mother or observed there.

We used the program *nadiv* (Wolak 2019) to create the relatedness matrix for use in *brms*.

```
library(nadiv)
#Creating pedigree matrix ("Amat") from input dataframe "ped1"
ped2 <- prepPed(ped1, gender=NULL, check=TRUE)
Amat <- as.matrix(nadiv::makeA(ped2))
```

Wolak, M. E. (2019). *nadiv*: an R package to create relatedness matrices for estimating non-additive genetic variances in animal models. *Methods in Ecology and Evolution*.

We used PedigreeViewer to extract some pedigree statistics (Table S1) and visualize the pedigree (Fig. S1)

Table S1. Assorted statistics about the structure of the sparrow pedigree used for heritability analyses.

Item	Number
Individuals	1629
Fathers	211
Mothers	228
Max paternal family	28
Max maternal family	27
Cases of inbreeding	0
Tiers	7
Tier 1	175
Tier 2	290
Tier 3	754
Tier 4	276
Tier 5	106
Tier 6	24
Tier 7	4

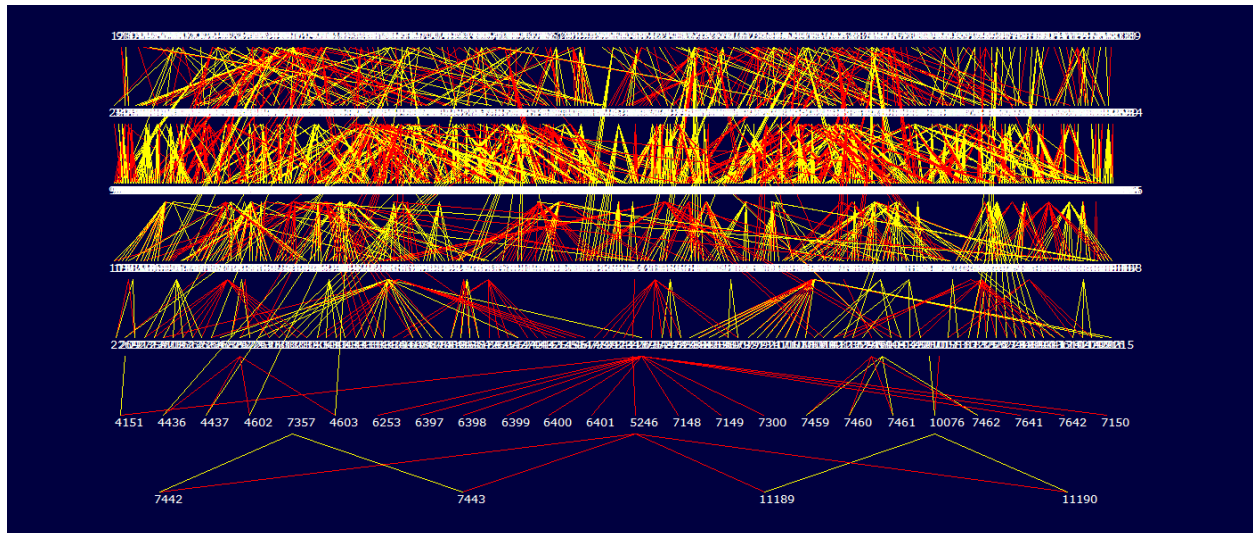


Figure S1. Map of the pedigree of 1591 individual house sparrows used in the animal model analysis.

We used *mcmcGLMM* to duplicate the main analyses. The pedigree file had to be processed differently. To do so, we used the following script:

```
library (pedigree)
library (tidyverse)

#Rename variables
Ped <- ped1
Ped <- rename(Ped, animal = SampleID, sire = Dad, dam = Mom)

Ped <- Ped[1629:1,]
(ord <- orderPed(Ped))
Ped <- Ped[order(ord),]
```

Telomere file ("Telomere.csv")

This file contained the core data used in the animal model analysis. It consisted of 1591 entries of individuals for which we had telomere measures. Variables were:

SampleID: Unique number given to individual, usually representing their blood sample number.

NestID: Unique number assigned to each breeding attempt over the history of the study

Barn: Location of nest site, usually on a single physical structure, with one level referring to solitary nest sites not clustered on a structure.

FED: First egg date of the focal nest attempt in days with 1 Jan = 1 (Julian date).

Clutch: Number of eggs laid in focal attempt

Hatch: Number of hatchlings in focal attempt, treated as starting brood size

Band: Number of nestlings that were banded (typically at age 10 days)

Identity: Code for male (M) or female (F) if adult or for nestling (N).

Sex: M for male and F for female

Dad: ID for the sire

Mom: ID for the mother

PairID: A concatenated identity indicating the combination of the two parents

Sample10: Code for if the focal bird was sampled at day 10 (Y) or not (N).

DadAge: Count of the breeding season for the male at the breeding attempt

MomAge: Count of the breeding season for the female at the breeding attempt

Year: Year focal individual was hatched

Assay: Unique name for each assay, consisting of the plate with the samples and the reference plate.

TSratio: Measure of telomere length

TSratio10: Measure of telomere length on Day 10.

TSratioMale: Measure of telomere length for male nestlings

TSratioFem: Measure of telomere length for female nestlings

Nestling measures file ("Nmeasures.csv")

Data on the mass and tarsus length of nestlings at the time their blood was sampled. Variables:

SampleID: The unique number given to the blood sample from an individual

Date: Date of sampling in month/day/year format

JDate: Date of sampling in Julian days, where 1 Jan of the focal year = 1.

Nage: Age of nestling at time of sampling

Tarsus: Length of metatarsus in millimeters.

Weight: Mass of nestling in grams.

AvTempC: Daily mean temperature in Celsius over the previous 25 days covering the active period of the nesting attempt.

AvPrecipCM: Daily mean precipitation over the previous 25 days covering the active period of the nesting attempt.

Weather summaries file ("Weather.csv")

Summary of weather data for the years of the study. Variables:

Year: The calendar year of data collection, from 1993-2014

SumTemp: Daily mean temperature over the period April 1-August 31 in degrees Fahrenheit.

SumTC: Daily mean temperature over the period April 1-August 31 converted to Celsius

SumTMC: Variable SumTC mean centered

SumPrec: Daily mean precipitation over the period April 1-August 31 in inches

SumPM: Daily mean precipitation over the period April 1-August 31 in centimeters

SumPMC: Variable SumPM mean-centered

SpringTemp: Daily mean temperature over the period February 1-March 31 in degrees Fahrenheit.

SpringTC: Daily mean temperature over the period February 1-March 31 in degrees Celsius.

SpringTMC: Variable SpringTC mean-centered.

SpringPrec: Daily mean precipitation over the period February 1-March 31 in inches.

SpringPM: Daily mean precipitation over the period February 1-March 31 in centimeters.

SpringPMC: Variable SpringPM mean-centered.

Data processing

The telomere data file was merged with data on nestling measures and weather to produce a master analysis dataframe referred to below as "tel4".

B. Details on data analysis

We conducted the initial heritability analysis in *brms*. The samples used to obtain telomere measures were taken at roughly, but not always exactly, at 10 days of age. We ran a preliminary analysis to assess if the model ran OK with the pedigree file and used nestling age centered at day 10 as a fixed effect. That model ran successfully and revealed no impact of the age of collection on telomere length, and so we omitted that variable in most subsequent analyses.

We then ran the base animal model using the pedigree, assay identity, year, nest location

```
library(brms)
#Expanded animal model script
mod2 <- brm(TSratio ~ 1 + (1|Assay) + (1|Year) + (1|Barn) + (1|PairID) +
  (1|NestID) + (1 | gr(SampleID, cov = Amat)), data = tel4,
  data2 = list(Amat = Amat), family = gaussian(),
  chains = 4, iter = 10000, warmup = 1000, thin = 5, cores = 4
)
summary(mod2)
plot(mod2)
```

(barn) and breeding pair ID as random effects and default priors.

We duplicated the *brms* model in *mcmcGLMM* but with larger a larger set of iterations, inverse Wishart priors ($V = 0$, $\nu = 0.002$), and a single chain. The results were very similar to the *brms* output given that *brms* outputs standard deviations and *MCMCglmm* gives variances. Results are presented in the main text.

```
library(MCMCglmm)

tel4$animal <- tel4$SampleID

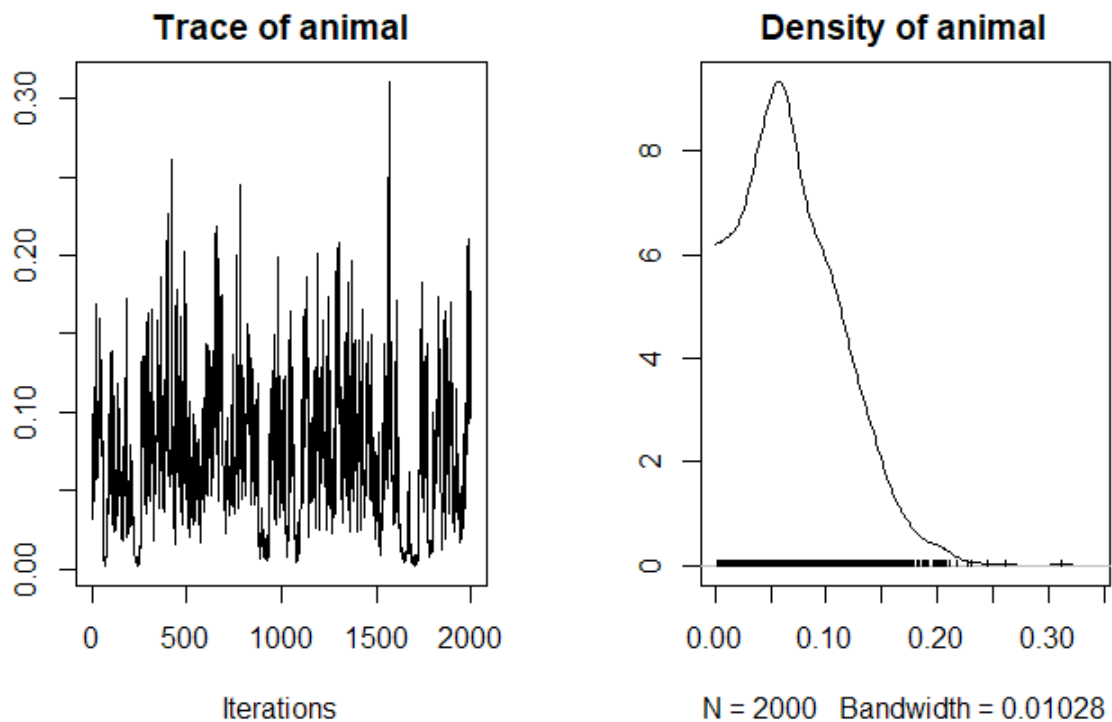
# Base animal model
prior1 <- list(G = list(G1 = list(V = 1, nu = 0.002),
  G2=list(V=1,nu=0.002),
  G3=list(V=1,nu=0.002),
  G4=list(V=1,nu=0.002),
  G5=list(V=1,nu=0.002),
  G6=list(V=1,nu=0.002)),
  R = list(V = 1, nu = 0.002))

model1 <- MCMCglmm(TSratio ~ 1, random = ~animal+Assay+Year+Barn+PairID+NestID,
  pedigree=Ped, prior=prior1,data = tel4, nitt = 110000,
  burnin = 10000, thin = 50)
summary(model1)
```

Heritability was then calculated from the output of this model as follows:

```
v_all1<-as.data.frame(apply(model1$VCV,2,median))
v_animal1 <- v_all1["animal",]
v_assay1 <- v_all1["Assay",]
v_year1 <- v_all1["Year",]
v_pair1 <- v_all1["PairID",]
v_barn1 <- v_all1["Barn",]
v_nest1 <- v_all1["NestID",]
v_r1 <- v_all1["units",]
her1 <- as.mcmc(v_animal1 / (v_animal1 + v_assay1 + v_year1 + v_barn1 + v_pair1 +
v_nest1 + v_r1))
summary(her1)
```

Figure S2. Plots of (a) the trace pattern of estimates from the base model and (b) the posterior distribution of heritability.



MOLECULAR ECOLOGY

We reran the models above using only adults that were also sampled at 10 days of age. All results were quite similar, and those results are summarized in main figure 1.

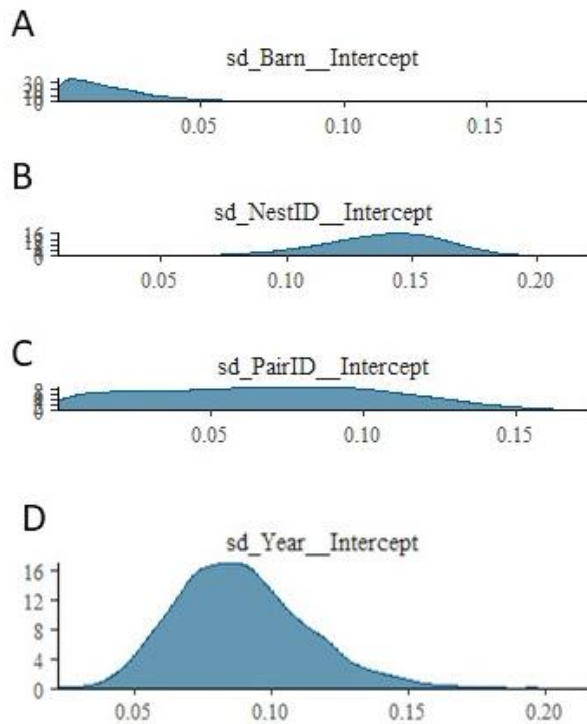


Figure S3. Plots of posterior distributions for estimates of variance components in telomere length from an animal model. Shown are posteriors of A) nest location (Barn), B) breeding attempt (NestID), C) pair identity (PairID), and D) cohort (Year).

Contingent heritability

We explored if heritability depended on the age of the parents. A constraint was the large proportion of samples in which adults paired to only one partner for all attempts, so we did not have the replicates necessary to estimate contingent heritabilities on father (range 1-9 years) and mother (range 1-6 years) age simultaneously. For this analysis we used *brms*, dropped BarnID and replaced PairID with either DamID (“Mom”) or SireID (“Dad”). Default priors and Cauchy priors gave similar results. Code for mother’s age is below.

```
prior_brms <- set_prior(prior = 'cauchy(0,2)', class='sd')
mod2.1 <- brm(TSratio ~ 1 + MomAge + (1|Assay) + (1|Year) + (1|Mom) + (1|NestID) +
  (1 + MomAge | gr(SampleID, cov = Amat)),
  data = tel4,
  data2 = list(Amat = Amat),
  family = gaussian(),
  prior = prior_brms,
  chains = 2, iter = 10000, warmup = 1000, thin = 5, cores = 4,
  control = list(adapt_delta = 0.9)
)

summary(mod2.1)
```

Environmental effects

We merged the telomere data file with the weather summaries file. We then employed two approaches to exploring different environmental effects further. Our first approach was to ask if unknown environmental factors might be interacting to affect telomere lengths. There are several ways to achieve this, but the handiest was using SAS Proc Mixed to split the among-unit variance at lower levels among the categories of the higher level and test this against a simpler model using the likelihood ratio test. This presumably can be done in *brms* or *MCMCglmm* using a random “slope” term that is the higher level, but assessing if this really provides a better fit to the data is not so straightforward. Here we provide the code used in SAS for splitting residual variance by year, followed by breeding attempt variance by year. The improved fit was tested with a likelihood ratio test against a model without the split, which had a -2log-likelihood of 1233.4.

```
Proc Mixed Data = Hosp.telomere Method = REML Covtest;
Class Year NestID;
Model TSRatio10 = / s;
Random Intercept/Subject = Year;
Random Intercept/Subject = NestID;
Repeated/ Group = Year;
Run;

Proc Mixed Data = HOSP.telomere Method = REML Covtest;
Class Year NestID;
Model TSRatio10 = / s;
Random Intercept/Subject = Year;
Random Intercept/Subject = NestID Group = Year;
Run;
```

The likelihood ratio test of these models is presented in the main text, but individual yearly estimates are in Table S2.

Table S2. Partitioning of among-attempt variance and residual variance in house sparrow telomere lengths by year, obtained from a frequentist mixed model in SAS Proc Mixed.

	Among attempts Variance ± SE	Residual (among nestlings) Variance ± SE
Intercept (1994)	0.004 ± 0.003	0.004 ± 0.003
1993	-	0.147 ± 0.076
1995	0.650 ± 0.313	0.885 ± 0.266
1996	0.021 ± 0.022	0.139 ± 0.033
1997	0.014 ± 0.021	0.051 ± 0.016
1998	0.009 ± 0.018	0.093 ± 0.023

1999	0.013 ± 0.018	0.087 ± 0.019
2000	-	0.123 ± 0.023
2001	0.027 ± 0.018	0.091 ± 0.015
2002	0.030 ± 0.018	0.093 ± 0.016
2003	0.033 ± 0.017	0.089 ± 0.015
2004	0.010 ± 0.013	0.099 ± 0.018
2005	-	0.071 ± 0.020
2006	0.043 ± 0.027	0.134 ± 0.029
2007	0.015 ± 0.015	0.115 ± 0.020
2008	0.003 ± 0.015	0.103 ± 0.021
2009	0.005 ± 0.012	0.076 ± 0.014
2010	0.036 ± 0.017	0.149 ± 0.023
2011	0.331 ± 0.121	0.302 ± 0.053
2012	0.023 ± 0.015	0.095 ± 0.016
2013	0.006 ± 0.009	0.072 ± 0.012
2014	0.061 ± 0.145	0.116 ± 0.114
-2Loglikelihood	1132.6	1096.1

We also assessed if sire or dam age might have similar interactive effects with variables acting either at the among-year level or the among-breeding attempt level. For this, we lumped all ages of 5 years and above into one group. The code for dam age is provided below and that for male age is nearly identical.

```
Proc Mixed Data = HOSP.telomere Method = REML Covtest;
Class Year NestID MomAgeC;
Model TSRatio = / s;
Random Intercept/Subject = Year;
Random Intercept/Subject = NestID Group = MomAgeC;
Run;
```

Table S3. Estimates of the variance in telomere lengths among breeding attempts or among nestlings within attempts by sire or dam age categories and the likelihood ratio compared to a model that estimated a single variance among breeding attempts or for residuals.

Among breeding attempts		
Age category	Dam estimate SE	Sire estimate SE
1	0.013 ± 0.006	0.029 ± 0.012
2	0.021 ± 0.009	0.009 ± 0.007
3	0.027 ± 0.016	0.012 ± 0.009
4	-	0.016 ± 0.013
5+	0.033 ± 0.095	0.011 ± 0.018

Likelihood ratio test	600.6-597.1 = 3.5	600.6-597.1 = 3.5
Among nestlings (residual)		
1	0.10 ± 0.008	0.11 ± 0.012
2	0.10 ± 0.010	0.088 ± 0.008
3	0.12 ± 0.018	0.10 ± 0.012
4	0.07 ± 0.017	0.11 ± 0.016
5+	0.04 ± 0.049	0.09 ± 0.023
Likelihood ratio test	600.6 – 596.1 = 4.5	600.6-596.9 = 3.7

The second approach was the conventional method of identifying potential target variables and adding them to the model as fixed effects. We were interested in interactions between environmental factors, so we included a modest set of likely ones. To help the model run faster, we omitted the pedigree random effect, PairID, and BarnID. We standardized (value minus the mean then divided by standard deviation) all variables in the model as the scale of predictor variables varied considerably. Default priors were used.

```
mod4.2a <- brm(TSratioSTD ~ Sex +
              JdateSTD*HatchSTD +
              AvPrecipSTD*HatchSTD +
              SpringTSTD*HatchSTD +
              SpringPSTD*HatchSTD +
              SumTSTD*HatchSTD +
              SumPSTD*HatchSTD +
              (1|Assay) + (1|Year) + (1|NestID),
              data = tel4,
              family = gaussian(),
              chains = 4, iter = 10000, warmup = 500, thin = 5, cores = 4
              )
summary(mod4.2a)
```

The output from this model is presented in main text as Table 2. We tested if any of the results were sensitive to other terms in the model, especially the large number of interaction terms involving brood size. We reran the model above with a restricted set of interactions.

```

mod4.2b <- brm(TSratioSTD ~ AvPrecipSTD +
  Sex +
  JdateSTD*HatchSTD +
  SumTSTD*HatchSTD +
  SumPSTD*HatchSTD +
  (1|Assay) + (1|Year) + (1|NestID),
data = tel4,
family = gaussian(),
chains = 4, iter = 10000, warmup = 500, thin = 5, cores = 4
)

summary(mod4.2b)

```

The results reveal little change in the main conclusions (Table S4).

Table S4. Posterior means \pm SE of fixed effect predictors and the remaining unexplained levels of environmental variance (in standard deviations of standardized scores) in telomere length among house sparrow nestlings using a mixed model with a reduced set of random effects and standardized response and predictor variables.

Level/Factor	Effect \pm SE	Credible interval
Among year	0.21 \pm 0.07	0.09, 0.37
Summer temperature	0.09 \pm 0.06	-0.04, 0.21
Summer precipitation	0.17 \pm 0.07	0.04, 0.30
Among nests	0.43 \pm 0.04	0.35, 0.51
Date	-0.03 \pm 0.03	-0.10, 0.03
Brood size	-0.04 \pm 0.03	-0.10, 0.02
Mean precipitation	-0.05 \pm 0.04	-0.12, 0.02
Brood size:Date	-0.05 \pm 0.03	-0.11, 0.01
Brood size:Summer temperature	0.02 \pm 0.03	-0.05, 0.09
Brood size:Summer precipitation	0.05 \pm 0.03	-0.01, 0.11
Residual	0.79 \pm 0.02	0.75, 0.83
Offspring sex (female)	-0.38 \pm 0.86	-2.03, 1.32
Offspring sex (male)	-0.26 \pm 0.86	-1.93, 1.44

Bivariate animal model analyses of telomere and nestling size or mass.

We analyzed potential environmental and genetic correlations between telomere length and nestling size (tarsus length) or mass using Bayesian bivariate mixed models. These were run in brms with a reduced set of random effects and the fixed effect of nestling age centered at Day 10 in the equations for size or mass. We ran these models using both the default priors and those recommended by the brms help page for set_prior. The code for the latter is shown below, but results were similar between the two. Results using default priors are presented in the main text.

```
library(brms)

prior_brms2 <- set_prior("lkj(2)", class = "cor")

bf_tel <- bf(TSratio ~ 1 + (1|Assay) + (1|a|Year) + (1|d|NestID) +
  (1|p| gr(SampleID, cov = Amat)))
bf_tar <- bf(Weight ~ 1 + Nage10 + (1|a|Year) + (1|d|NestID) +
  (1|p| gr(SampleID, cov = Amat)))

bmod2.1 <- brm(bf_tel + bf_tar + set_rescor(TRUE),
  data = tel4,
  data2 = list(Amat = Amat),
  family = gaussian(),
  chains = 2, iter = 10000, warmup = 1000, thin = 5, cores = 4
)

summary(bmod2.1)
```