Supplementary Information 2

Life history of the harlequin ladybird, *Harmonia axyridis*; a global meta-analysis

Raak-van den Berg, C.L., L. Hemerik, W. van der Werf, P.W. de Jong, and J.C. van Lenteren (2017)

Methods

Data(base)

Data search

Studies that report experimental data on life history characteristics of *H. axyridis* were gathered from previously published reviews and by carrying out a literature search in the databases CAB abstracts, ISI web of Science, Biological records, Zoological records, and Agricola. Studies from 1900 onwards were surveyed. The last search was conducted in July 2013. An additional search was performed on June 24th 2016 with Scopus and CAB abstracts. The only keyword that is used is "*Harmonia axyridis*". All resulting papers were checked by the first author for availability of data for the considered life history characteristics. Additional studies were found by searching in the bibliographies of papers. When abstracts of Chinese and Japanese studies were promising, relevant sections, tables, and figures were translated by native speakers except for studies from the period 2013-2016. When the original paper could not be obtained or translated, the data that were given in a) the abstracts, b) figures and tables in English, or, c) review papers (only when the source was clear), were included in the database (for references of selected studies see Supplementary Information 1, raw data of selected studies are deposited at: http://dx.doi.org/10.17026/dans-zt7-dhb6).

Data that were presented in figures were digitised using the Engauge Digitizer Version 4.1 and the resulting numbers were used. In other cases authors provided raw data. When ambiguity arose about definitions, details, or calculations used, clarification was asked from the authors. In case the author could not be reached and no founded assumption could be made, data were discarded.

Selection of studies

The studies had to meet some criteria to be included in the dataset. Only data that were determined under biologically relevant conditions (temperatures between 14°C and 35°C and photoperiods of 14 hours, 16 hours, or natural day light) were included. This means that extreme photoperiods (< 14 or > 18 hours of light) and temperatures (< 14°C or > 35°C) were excluded, as well as studies with artificial diet as food. Field studies are conducted under fluctuating temperatures; therefore results of field studies could not be included in the analysis.

Although under natural conditions *H. axyridis* can experience situations with limited food and although the larvae of this generalist predator can prey and survive on non-aphid food, only data generated with unlimited access to aphids or *Ephestia kuehniella* Zeller (Lepidoptera, Pyralidae) (further referred to as *Ephestia*) as prey were included. This was done for three reasons: a) toxic or unsuitable aphid species (for definition see Hodek and Honek 1996) and artificial diet often do not result in development at all; b) food limitation is difficult to quantify, and in general it directly affects development and survival; and c) results based on other non-aphid prey (e.g. ladybird eggs or larvae) are scarce and difficult to categorise and standardise. As many studies use *Ephestia* as food source, these studies were included. .

Studies that did not mention which prey was used were left out as well. When food availability was not mentioned, we assumed food was unlimited. Aphid species (Hemiptera, Aphididae) excluded because of their unsuitability for development and/or reproduction were: *Aphis craccivora* Koch, *Aphis nerii* Boyer de Fonscolombe, *Aphis pomi* De Geer, *Aphis spiraecola* Patch, *Aulacorthum magnolia* (Essig & Kuwana), *Brevicoryne brassicae* (L.), *Hyalopterus pruni* (Geoffroy), *Megoura viciae* Buckton, and *Toxoptera citricidus* (Kirkaldy). For all response variables, initial exploration of different aphid species did not show any consistent effect. Therefore, we lumped different species of aphids into one category. Studies determining life history characteristics (e.g. fecundity, survival, or longevity) of individuals that had hibernated or were hibernating were excluded. From studies with a manipulative treatment (e.g. insecticide treatment) only the data resulting from the control treatment were regarded as relevant and included in the database. Results by Thompson (1926; 1927) were excluded as well, as the species (*Leis* sp.) they worked with was not *H. axyridis*, but *Harmonia conformis* (pers. comm. Y. Jongema, Laboratory for Entomology, Wageningen).

The survival of beetles of the flightless strain of *H. axyridis* was relatively low and had a different pattern than that of beetles that were able to fly. Therefore, all data for flightless ladybirds were discarded. A total of 82 studies met the inclusion criteria.

Response variables

Twenty-two life history characteristics were defined as response variable, being a measure for immature development, immature survival, post-eclosion development, and reproduction. In the database those studies were included that reported at least one of the life history characteristics or a variable that could be converted into a life history characteristic. SE, SD, range or other, given by author, were recorded as well.

The major difficulty was that different authors used different definitions, measurements, and methods for the life history characteristics. Therefore, recalculations had to be made. When studies reported values for response variables that, after recalculation, did not fit the definitions of the response variables used in this paper, it was not possible to include the results in this meta-analysis.

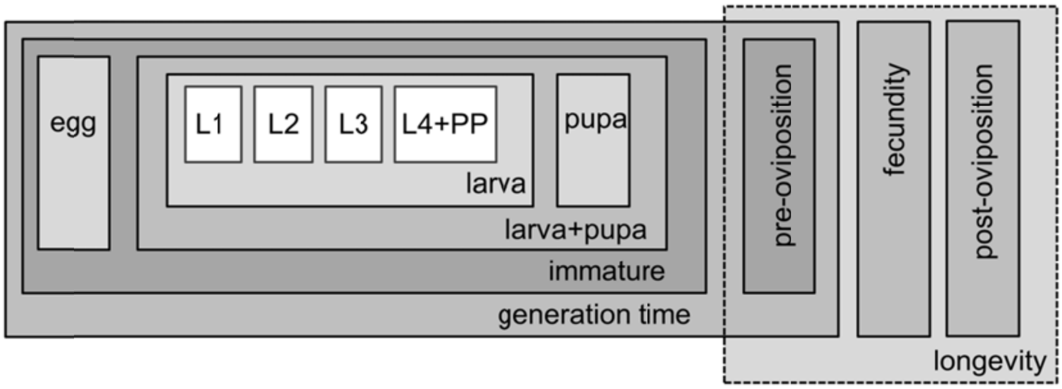

Figure SI 2.1 A schematic overview of the life stages of *H. axyridis* used to define the life history characteristics and their relationships. L1, L2, L3, and L4 refer to four larval instars. The fourth instar includes the pre-pupal stage (PP)

Figure S1.1 gives a schematic overview of the life cycle of *H. axyridis* consisting of four larval instars and a pupal stage, separated by moults. The moult into the next stage is the cut-off for the earlier stage. The fourth instar has a mobile and an immobile (prepupal) phase, not clearly separated by a moult. In this analysis the prepupal stage is considered as part of the L4 stage, therefore data are combined when authors report separate data for L4 and prepupa. When authors did not clearly define the stages, it was assumed that the authors included the prepupa in the fourth instar.

Both development rate and fraction survival were defined as response variables for the following nine developmental stages: egg, L1, L2, L3, L4, larva (from egg hatch to moult into pupa), pupa (from moult into pupa to adult emergence), larva+pupa (from egg hatch to adult emergence), and immature stages (from egg to adult emergence) (figure 1). When studies only gave values for individual instars, data were lumped together and combined to cover a longer period, only if those values were obtained from the same group of insects (e.g. separate development values for the instars L1+L2+L3+L4 are combined into one value for larval development). Although developmental time is often determined for larvae of unknown sex, while pre-oviposition period is determined for females, the values of immature development and pre-oviposition are combined to obtain more data for the response variable generation time.

Development time is the length of time the larva needs to develop from one stage to the next. Development rate is the reciprocal of the development time. Survival is the fraction of larvae that develop into the next developmental stage. When mortality was reported it was recalculated into survival (=1 - fraction mortality). The fraction of egg hatching is similar to egg survival.

After adult emergence the adult lives a certain lifespan, called longevity. This period can be split in a pre-oviposition period (from adult emergence to first oviposition), an oviposition period (from first to last oviposition), and a post-oviposition period (from last oviposition until death). Generation time is the time that elapses from egg deposition until the first oviposition by the resulting adult. Longevity, the rate (reciprocal) of pre-oviposition period, and the rate (reciprocal) of generation time are three response variables for the adult stage. The fourth adult response variable is fecundity, measured as the number of eggs laid. When fecundity was given as a total number over a certain period, it was recalculated into a daily fecundity (i.e. the average number of eggs oviposited per day over the studied period). When total fecundity was determined over the total period of longevity, which also includes the pre-oviposition period and the post-oviposition period (figure 1), dividing by longevity led to an underestimation of daily fecundity compared with daily fecundity that was measured over a short period starting with the first oviposition. Therefore, daily fecundity was recalculated using the actual oviposition period when available (6 data points). When no data on post-oviposition period were given (4 data points), longevity was reduced with pre-oviposition period only. For five data points only longevity was available.

Outliers

Preliminary analysis of scatter plots with data points revealed some outliers in the data. When a data point had a distance of more than two standard deviations from the mean at a certain temperature it was omitted from the analysis. In some cases the development rate at 35°C was a clear outlier (for L2, L3, and L (Wang et al. 2009), which was excluded as well. The study by Lundgren (2009) showed poor performance on soybean leaves (with *Ephestia* food); therefore only the no-plant control was included in the dataset and outliers were excluded. Nedved and Kalushkov (2012) reported low survival for the larval and pupal stages studied (L2, L3, L4, pupa) and attributed it to the larvae being grouped during the experiment; therefore their data were also excluded from the dataset.

Explanatory variables

To explain variation in life history characteristics reported in the different studies, several variables were selected as factors in the analysis. Unfortunately, not all studies provided an exhaustive list of parameters to describe the experimental setup. Moreover, for the final analysis it was required to lump the factor levels into classes, despite some loss of detail. Below, a description of each factor follows, including explanation of choices made in case of missing values or lumping of data. A summary is given in table SI 2.1.

1. Temperature: The temperature, at which the experiment has been conducted, is used in the analysis as a continuous variable in the range of 14°C to 35°C. When temperature is not or cannot be specified (e.g. in case of field data), the data are excluded.

2. Origin: The origin of the beetles used in the experiment. When the paper does not explicitly state the origin of the beetles, the location of the university has been used as origin. For analysis, data were lumped into two categories: native (Asian range) and invasive (non-Asian range: North America, Europe, South America, and Africa). Due to selection on biological relevance (see above) the second category contained studies from North America and Europe and only one from Africa (Egypt).

3. Photoperiod: Studies conducted at extreme photoperiods were excluded from the dataset. In case of an unknown photoperiod, it was assumed that the experiment was performed under natural light conditions although some studies were conducted in Asia, where researchers often used extreme photoperiods. Photoperiods were lumped into three categories: short (14–15.5 hours light), long (16–18 hours light), and natural (also including data with unknown photoperiod). Under natural day conditions less than two studies for survival response variable were performed (except egg survival). Therefore, for all survival response variables data under natural day were excluded from the dataset, and models were fitted with two classes for photoperiod (short and long). For egg survival even the complete factor photoperiod was excluded, because only nine of the 24 different classes based on four factors had more than three data points.

4. Food: In the analysis no distinction was made between aphid species, one of the reasons being that for many aphid species only one data point was available, which hampers the analysis. A classification into aphid and *Ephestia* was used for the analysis. Data based on combined diets of aphids and *Ephestia*, or *Ephestia* and pollen were excluded. For the following response variables very few data with *Ephestia* as food were available. Therefore, models were fitted that did not contain the factor food: survival L1, L2, L3, L4, and pupa.

5. Strain: The populations used could either be a reared population or a (recently) field-collected population (incl. subsequent short-term rearing of max. 4 generations). When the

strain type was not clearly stated in the paper, a strain was assigned to the data based on the other information given in the paper.

6. Sex: For the response variable longevity the factor sex was added, as many studies report differences in longevity between sexes. When sex was unknown data were excluded.

7. Origin-photoperiod: For survival of L1, L2, L3, L4, larva, and pupa the data for factors origin and photoperiod were very confounded. Therefore, these factors were combined into one factor (origin-photoperiod) with three classes: native&short, native&long, and invasive&long.

In some case more than one data point was retrieved from one study. As conditions within one study presumably are more uniform than between studies, study is modelled as a random effect.

Table SI 2.1 Description of explanatory variables and factors (with their shorthand notation) used for model fit. Column ‘added’ indicates the order in which the explanatory variable were included in model when generating the most complex linear mixed model.

Factor	Added	Classes	Description
Temperature (T)	1	continuous scale	14 to 35°C
origin (O)	2	native; invasive	area of origin of ladybirds
photoperiod (P)	3	short; long; natural	photoperiod during experiment: light during 14–15.5 hours, 16–18 hours, or naturally occurring (including unknown photoperiod)
Food (F)	4	Aphid; *Ephestia*	food of ladybird[c] used during experiment: only aphids or only *Ephestia kuehniella* eggs (without aphids or pollen)

Strain (S)	5	field; reared	source of population: field-collected population (incl. subsequent short-term rearing of max. 4 generations) or laboratory reared population
sex[a] (G)	6	female; male	
origin-photoperiod[b] (PO)	A[d]	native&short; native&long; invasive&long	origin and photoperiod combined, because factors were confounded

[a]for response variable longevity

[b]for response variables survival L1, L2, L3, L4, and pupa

[c]of mother in case of eggs, of larva in case of pupa

[d]added after temperature and before food if origin and photoperiod were confounded

Meta-analysis — model selection

For each response variable, first, the most complex linear mixed model was formulated, i.e. all defined factors that were independent from each other were included as long as the number of estimated effects was within the limits the number of data points allowed for. The factors used were origin, photoperiod, food, and strain. In case of small data sets, confounded factors were coded as combination origin-photoperiod. Moreover, temperature was included as continuous explanatory variable. Each factor could be part of the model as a fixed effect, or additionally in interaction with one of the other factors or with temperature. 'Study' is modelled as a random effect. Subsequently, model selection was performed; all possible less complex models that can be assembled with the factors of the most complex model, were fitted automatically to the data using the R function dredge from the R package MuMIn (version 3.0.2).

In the fitted models, temperature was included as a covariate while there were four categorical variables: geographic origin (O, 2 levels), photoperiod (P, 3 levels), food (F, 2 levels) and strain (S, 2 levels). For longevity the factor sex (G, 2 levels) was also included. Development rates (including the rates of pre-reproductive development of the adult and generation time) were fitted as linear functions of temperature and covariates. Daily fecundity was fitted with a parabolic function of temperature. In the full models (with most factors included) for the eleven development rates, longevity and fecundity, each factor and its interaction with temperature was at least included. The first-order interactions were included in the order OxP, OxF, OxS, (OxG), PxF, PxS, (PxG,) FxS, (FxG, SxG), as long as the model could be fitted (i.e. no singular matrix resulted in the fitting procedure). The models for survival included higher powers of the temperature up to the fourth power to account for non-linear relationships. For the survival of the single stages (here first, second, third and fourth larval stage and pupa) the factor strain and the combined factor origin-photoperiod were included along with their interaction with temperature. Survival of the egg, all larval stages together, “larva plus pupa” and all immature stages had a most complex model with all or almost all combinations of origin, photoperiod, food and strain and their interaction with temperature. Here, the first-order interactions were included in the order OxP, OxF, OxS, PxF, PxS, FxS, as long as the model could be fitted (i.e. no singular matrix resulted in the fitting procedure).

Thereafter, the best models were selected with Akaike’s Information Criterion (AIC), which is based on the maximum likelihood. Because of small sample sizes, we used the corrected AIC (AICc) (Motulsky and Christopoulos 2004), which is a measure of the relative goodness of fit of the model.

$$AICc = 2p - 2Log(LL) + 2p\frac{(1+p)}{(n-p-1)},$$

in which Log is the natural logarithm, *n* = number of data points, *p* = parameters estimated, AICc = corrected Akaike's Information Criterion, and *LL* is the maximum likelihood for the estimated model.

The best model or, equivalently, most supported model is characterized by the lowest AICc value among the set of competing models. Model fits were evaluated with respect to the AICc-difference (ΔAICc) between the considered model and the best model. Models within the interval ΔAICc < 2 are considered equivalent (Bolker 2008) and are referred to as the set of most supported models (Hilborn and Mangel 1997).

Within the set of most supported models Ockham's razor has been used, which is the principle of parsimony. This principle states that among competing hypotheses/models (with the same explanatory power) the one that makes the fewest assumptions should be used. The most parsimonious model i.e. the model with the least number of parameters within the set of most supported models is further referred to as the most parsimonious model. Model selection with AICc is not based on the significance of the individual factors in the model, but it selects the overall best model. As a consequence, the models are presented with the AICc value for the full model, and their relative support (=AICc weight) within the set of 20 models with lowest AICc (Burnham and Anderson 2002; Lukacs et al. 2010). An overview of the 20 best models for each life history is available at http://dx.doi.org/10.17026/dans-zt7-dhb6.

$$w_i = \frac{\exp(\frac{-\Delta\mathrm{AICc}_i}{2})}{\sum_{i=1}^{m}\exp(\frac{-\Delta\mathrm{AICc}_i}{2})} \qquad \text{(eqn. 2)}$$

Where $\Delta\mathrm{AICc}_i$ is the difference in AICc between model *i* and the best model (lowest AICc) for the same response variable and data, i.e. $\Delta\mathrm{AICc}_i = \mathrm{AICc}_i - \min_j\left(\mathrm{AICc}_j\right)$.

Based on the overall shape of the plot of the selected points, the following types of models were fitted: for rates of development, pre-oviposition period, and generation time and for longevity linear mixed models with and without temperature were fitted. Fecundity models included temperature effect in polynomials of highest order equal to 2, and survival models included temperature effect in polynomials up to order 4.

For the response variables survival L1, L2, L3, L4, larva, and pupa data sets were small and the two factors day length and origin were confounded; thus, no reliable models could be fitted. Therefore, the two factors were combined into one factor with three classes: native&short, native&long, and invasive&long, reducing the number of explanatory variables, which is preferable with small data sets. The few studies that did not fit in these three classes were excluded from further analysis.

In the analysis, data were weighted with the square root of the number of replicates used to obtain the value for the life history characteristic. He et al. (1994) did not give a number of replicates in the group that they followed. Therefore, the following values were used: N = 65 for egg development, which is the median number of replicates for egg development; N = 33 for development of larvae, pupa, larva plus pupa, and all immature stages, which is the average of the median numbers of replicates of each of these response variables as data were determined on the same group of individuals; N = 17 for pre-oviposition period, generation time, fecundity, and longevity, which is equal to half of the number of replicates for development, as these variables were determined on the females only.

When in the course of the experiment the number of replicates decreased due to larval mortality, N was not adjusted accordingly, unless the authors gave separate values for N for each developmental stage. When generation time was calculated as the sum of immature development time and pre-oviposition period, N of the latter response variable was used, as pre-oviposition period is usually based on a lower number of individuals. For egg survival the number of replicates was converted into the number of eggs by either (number of females*total fecundity), or (number of clutches*clutch size), or (number of larvae*sex ratio).

Life history characteristic values for native and invasive *H. axyridis* populations

To enable a comparison between native and invasive populations, we constructed a summary overview of different life history parameters as a function of temperature. Life history parameters were predicted for 15°C, 20°C, 25°C and 30°C, using for predicting either: 1) the most parsimonious model , or 2) the best model (with lowest AICc) that contained origin as a fixed factor: further referred to as 'best origin model'. The latter was done if the most parsimonious model did not contain origin and a model with origin as a factor was within the set of most supported models. In this case, two predictions of the life history trait are given: for the native populations and for the invasive populations. A single prediction was given for both geographic origins if the set of most supported models did not include a model with origin as an explanatory variable. When other co-variables were included in the model, the prediction was made for wild and aphid-fed populations under long-day conditions.

References

Bolker BM (2008) Ecological models and data in R. Princeton University Press, Princeton, New Jersey, USA

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information-theoretic approach. Springer, New York

He JL, Ma EP, Shen YC, Chen WL, Sun XQ (1994) Observations of the biological characteristics of *Harmonia axyridis* (Pallas) (Coleoptera: Coccinellidae). Journal of Shanghai Agricultural College 12:119-124

Hilborn R, Mangel M (1997) The ecological detective: confronting models with data Princeton University Press, Princeton

Hodek I, Honek A (1996) Ecology of Coccinellidae vol 54. Series entomologica. Kluwer, Dordrecht

Lukacs P, Burnham K, Anderson D (2010) Model selection bias and Freedman's paradox Annals of the Institute of Statistical Mathematics, 62:117-125

Lundgren JG (2009) Nutritional aspects of non-prey foods in the life histories of predaceous Coccinellidae. Biol Control 51:294-305

Motulsky H, Christopoulos A (2004) Fitting models to biological data using linear and nonlinear regression : a practical guide to curve fitting. Oxford University Press, Oxford [etc.]

Nedved O, Kalushkov PK (2012) Effect of air humidity on sex ratio and development of ladybird *Harmonia axyridis* (Coleoptera: Coccinellidae). Psyche 2012

Thompson WL (1926) A life history study of important lady-beetle predators of the citrus aphid. The Florida Entomologist 10:40-46

Thompson WL (1927) A life history study of important lady-beetle predators of the citrus aphid (continued). The Florida Entomologist 10:57-59

Wang H, Ji S, Zhai W (2009) Effect of temperature on survival, development and fecundity of *Harmonia axyridis*. Chin Bull Entomol 46:449-452