Key carabid species drive spring weed seed predation of Viola arvensis
Benjamin Carbonne, David Bohan, Sandrine Petit

To cite this version:

Benjamin Carbonne, David Bohan, Sandrine Petit. Key carabid species drive spring weed seed predation of Viola arvensis. Biological Control, 2020, 141, pp.104148. 10.1016/j.biocontrol.2019.104148 . hal-02942831

HAL Id: hal-02942831
https://hal.inrae.fr/hal-02942831
Submitted on 21 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License
Title: Key carabid species drive spring weed seed predation of *Viola arvensis*.

Benjamin Carbonne¹, David A. Bohan¹, & Sandrine Petit¹

¹Agroécologie, AgroSup Dijon, INRA, Université de Bourgogne Franche-Comté, F-21000 Dijon, France

Corresponding author: Benjamin Carbonne¹

Email address: benjamin.carbonne@inra.fr

Postal address: 17 rue Sully, BP 86510, 21065 Dijon cedex, France
1. Introduction

Conservation biological control could contribute to the sustainable management of pests and weeds through the enhancement of the abundance and activity of pest natural enemies that are naturally present in agroecosystems (Begg et al., 2017; Karp et al., 2018). Understanding the complex interactions within predator communities and the level of biological control exerted by these natural enemies remains difficult, however, notably for generalist predators (Kulkarni et al., 2015; Symondson et al., 2002).

Communities of natural enemies can have disparate biological control effects because of variation in species efficacy and interspecific interactions (Letourneau et al., 2009). Negative interactions, such as intraguild predation (Finke and Denno, 2004) or interference (Lang, 2003) may disrupt prey depletion efficacy (Prasad and Snyder, 2004). Conversely, positive interactions such as complementarity, which can result from facilitation (Losey and Denno, 1998) and/or resource partitioning (Straub and Snyder, 2008), can lead to an increase in the level of biological control. In some instances, positive and negative interactions cancel out one another (Cardinale et al., 2003; Finke and Snyder, 2010), or interactions between natural enemies are weak and predators have similar and independent additive effects (Roubinet et al., 2015).

The intensity of biological control can also be strongly influenced by the occurrence of “keystone” or “key” natural enemy species within a community (Gotelli et al., 2011; Gurr et al., 2012; Letourneau, 1998; Letourneau et al., 2009; Snyder, 2019). These species exert a strong mortality on pests via characteristics such as relative greater abundance (Newman et al., 1998), high searching or competitive ability (Letourneau et al., 2009), or by exerting direct density-dependent predation (Holling, 1959; Turchin, 1995). Identifying predator assemblages that are the most effective for the suppression of specific pests is a challenge (Gotelli et al., 2011; Paredes et al., 2015). That is a prerequisite for management guidelines
promoting conservation biological control (Greenop et al., 2018; Gurr et al., 2012; Jonsson et al., 2017; Landis, 2017).

The biological control of weeds is receiving increased attention due to policy pressure to reduce herbicide use (Begg et al., 2017; Petit et al., 2018, 2015). Seed predator invertebrates, and in particular carabid beetles, have been shown to contribute to the depletion of weed seeds and weed population regulation (Westerman et al., 2003; Bohan et al., 2011; Blubaugh and Kaplan, 2016).

Carabid beetles are abundant generalist predators in agricultural fields. Certain species are specialist granivores, feeding mainly on seeds and only occasionally on animal prey, while many are omnivores feeding on both plant and animal items (Hengeveld, 1980). Trophic interactions within carabid communities have been found to impact predation (Russell et al., 2017). For example, the presence of intraguild predators can disrupt (Prasad and Snyder, 2004) or enhance (Charalabidis et al., 2017) prey suppression. Conversely, interactions between some species appear to be complementary arising from contrasting species foraging strategies for seeds (Charalabidis et al., 2019). Although several authors have demonstrated a positive effect of seed-eating carabid abundance or diversity on the level of seed predation in the field (Jonason et al., 2013; Menalled et al., 2000; Petit et al., 2017; Trichard et al., 2013), such relationship does not appear to be generic to all field situations (Saska et al., 2008). Few studies have conducted in-depth analyses of the relationships between carabid community composition and individual carabid species and effective weed seed predation (Bohan et al., 2011), or of seed consumption in the laboratory when transposed to field situations (Honek et al., 2003; Petit et al., 2014). Petit et al. (2014) underlined the importance of carabid identity in improving the prediction of weed seed predation levels using laboratory- and field-measured consumption rates in models.
As carabid beetles are mobile organisms, the spatial resolution at which the carabid community interacts with weeds, resulting in weed seed predation, should be investigated. Similarly, the resolution of study of carabid taxonomy necessary to explain predation relationships is unclear. While a species-level taxonomic approach would highlight key species for predation, higher functional group approaches, such as grouping by genus or trophic behaviour, might be a better predictor of predation (Gagic et al., 2015). Clarifying the best spatial and taxonomic resolution required to explain the predation of seeds by carabids would therefore provide advances for conservation biocontrol, and be used as a guide to better predict seed predation.

In this study, we performed an analysis of the relative contribution of selected carabid taxa to the level of weed seed depletion in order to test the expectations that key carabid species and/or genera drive this process. We compiled a dataset of 1582 field records of carabid assemblages and associated seed predation rates of the weed *Viola arvensis* gathered in 111 fields under contrasted farm management and landscape settings. Regression tree methods were applied to identify the key components (species or genera) of carabid assemblages that structure seed depletion and verify this against a null-model. These methods were applied to the dataset at taxonomic (species, genus) and spatial scales (plot, transect) in order to identify the resolution(s) at which relationships between carabid assemblages and seed depletion should be addressed.

2. **Materials and methods**

2.1. **Field data**

Data from five different experiments on the seed predation of *Viola arvensis* (Murray 1770) were compiled. These five experiments were conducted within a 50km radius of the city of Dijon in Burgundy, France (47°19’18”N, 5°02’29”E) between 2012 and 2018 and
have previously been reported in Trichard et al., (2014); Lubac, (2014), Minvielle-Debat, (2016), Petit et al., (2017); Ricci et al., (2019). Seed predation was quantified by seed cards (Westerman et al., 2003) enclosed in a mesh cage (18 x 11 x 9 cm, 1 cm² wire mesh) to exclude vertebrates. The number of seeds remaining on the cards after 4 or 7 days of exposure was used to estimate the amount of seeds removed, and to calculate a seed predation rate.

Carabids were pitfall trapped with plastic containers of 10 cm depth and 8 cm diameter, filled with 150 ml of a mixture of salt water and anti-freeze or odourless dishwashing liquid. To protect the pitfall traps from rain inundation a plastic roof was suspended 10 cm above each of the traps. Carabid names have been updated following Gargominy et al. (2018).

In total, 111 fields were sampled that were under diverse types of farm management (6 organic farming, 86 conservation agriculture, 35 conventional agriculture) and covered a wide range of landscape settings. These landscape and farm management gradients were not included in analysis, but contribute to the diversity of carabid communities in our data set. Twenty of the fields were monitored over successive years. This gave a full data set containing 171 field-year records (15 organic farming, 86 conservation agriculture, 70 conventional agriculture). These field-year records were predominantly conducted in winter cereals (69%, mostly wheat and a few barley fields), spring crop (9.3 %, mainly soybean), 9% oilseeds (oilseed rape and mustard), 6% of legume crops (alfalfa, spring peas and winter bean) and 1.75 % of grassland or a mix of grass and legumes. Several sampling dates were used in each field-year, with several locations also being sampled. The database therefore included 1582 sample observations (Table 1), which corresponded to the carabid fauna caught in a single trap, and the predation rates measured on the associated seed card, defined as the ‘plot’ level. This level corresponds to the most accurate spatial resolution tested. Plot carabid data from Trichard et al.(2014) were derived from two traps and the activity density of species were pooled and divided by two. A second database, designated the ‘transect’ level, was
created where 4 plots were spatially aggregated along transects. The transect scale covered a range of distances ranging from 48 m to 98 m, depending on the sampling design used in the original studies. The four traps was either placed along a line or formed a rectangle, depending on the study. This spatial level corresponds to the larger spatial resolution tested. Predation rates were averaged and carabid species abundances were summed. Depending to the study, the fields surveyed included 1 or several transects (see Table 1) to give 470 data records. We defined an ‘intermediate’ spatial level, with predation estimated at the plot level because predation is very local and seeds are immobile, and with carabid species abundances pooled at the transect level because carabids are mobile in the field.

2.2. Linking carabid assemblages to seed predation

Data for carabid species that were likely to affect seed predation were extracted from the carabid database. These included carabid species described as granivore or omnivore in the literature (Purtauf *et al.*, 2005a; Homburg *et al.*, 2014; Pakeman and Stockan, 2014; Pey *et al.*, 2014). We also considered potential intraguild predators, which we defined here as species described as carnivores in the literature and with body length greater than 1 cm, i.e. larger than the seed feeding carabids (average size 9.26 mm (SD=2.78)) (Woodward and Hildrew, 2002; Prasad and Snyder, 2004). We thus assumed that small carnivores (< 1 cm) do not prey on similar or larger sized carabid species (Snyder and Wise, 1999; Prasad and Snyder, 2004, 2006). Due to functional redundancy, species of the same genus or tribes generally have similar food specialization and seed preference (Goldschmidt and Toft, 1997; Honek *et al.*, 2007; Talarico *et al.*, 2016). However, seed preference can also differ between species of the same genus, especially among species belonging to *Amara* genus (Klimeš and Saska, 2010). We thus decided to conduct two distinct analyses, one at the genus level and one at the species level. In the dataset, we identified 44 species and 26 genera of seed-eaters and 19 species belonging to 16 genera as potential intraguild predators (Table 2). In our
dataset, two genera of seed-eating carabids included numerous distinct species, the genus *Harpalus* (11 species) and the genus *Amara* (15 species) (Supplementary Table 1).

### 2.3. Statistical analysis

We combined two complementary methodologies that identify key predator taxa that are the most effective for the suppression of specific pest. A linear mixed model (LMM) tree algorithm methodology (Fokkema et al., 2017) was selected because it analyses the role of the assemblage of carabid species in seed depletion. This method identifies carabid taxa and the associated activity-density thresholds that repeatedly split observations into nodes that have dissimilar seed predation rates (Fokkema et al., 2017). The regression trees can simultaneously handle many potential predictor variables and their interactions to reveal antagonisms or complementarities between species in weed seed predation. The null model approach of Gotelli et al. (2011) was also used to analyze the contribution of individual carabid species to the intensity of seed depletion, and confirm the regression tree outputs. The advantage of the regression tree analysis, in comparison to the Gotelli (2011) approach was the identification of combinations of carabid taxa that exhibited particularly high predation effects, consistent with a positive interaction e.g. complementarity, and the estimation of the abundances of the carabid taxa necessary to obtain seed depletion.

The LMM tree algorithm is a model-based recursive partitioning algorithm (MOB, Zeileis et al., 2008; Fokkema et al., 2017) of 4 distinct steps: in step (1) the model is fitted to all observations in order to estimate parameters, with the clustered structure of datasets being considered by the inclusion of random effects. (2) Assessment is used to decide whether observations should be split by performing a fluctuation test for parameter instability for this model (Zeileis et al., 2008; Zeileis and Hornik, 2007). These tests assess whether the parameter estimates are stable over each particular ordering of the partitioning variables, or whether the observations can be divided with respect to one of the partitioning variables to
reduce the instability and improve the fit of the model in each of the splitting subgroups (Zeileis et al., 2008). A test statistic is applied to assess whether there is overall instability and to calculate a p-value for every partitioning variable (Zeileis and Hornik, 2007). The observation are split only if there is significant instability associated to any of the partitioning variables. In step (3), the partitioning variable associated with the highest parameter instability, corresponding to the minimal p-value, is selected. Then, the optimal partitioning threshold associated with this partitioning variable is identified and the sample is split. Finally, in step (4), the procedure is iterated on the splitting subgroups obtained, also called nodes, until there are no further splits.

We built the regression trees using the \texttt{lmertree} function (package \texttt{glmertree}), with the predation rate as the response variable, weighted by the initial number of seeds exposed on cards. The partitioning variables were the activity density of the selected carabid genera / species. For the parameter instability tests, the significance level $\alpha$ was fixed to 0.05 to avoid overfitting and selected the most pertinent partitioning variables. We did not impose any constraints on tree size, which means that trees can have variable sizes and number of terminal nodes. For each node, we also present the three main surrogate splits, corresponding to the partitioning variables that were not selected in the tree, but had a p-value < 0.05 for the parameter instability tests (Supplementary Table 5, 6 and 7). The random effects were the \textit{Experiment name}, the \textit{Field identifier}, and the \textit{Month} in interaction with the \textit{Year} of sampling, to take into account the structure of the data set. The stability of the selected partitioning variables was assessed graphically using the \texttt{stabletree} function (package \texttt{stablelearner}) (Philipp et al., 2016), by checking the variable selection frequencies based on regression trees built on 500 bootstraps keeping 80% of the initial data set. Once the regression tree is built, a global LMM is performed to fit the predation rates against the different subgroups obtained as a fixed effect. Based on this global LMM, the estimated marginal means (EMMs) of the
predation rates of each terminal node were compared by Tukey pairwise comparison using the functions *emmeans* and *pairs* (package *emmeans*) with Bonferroni correction (Supplementary Table 2). The performance of the regression trees was obtained by evaluating $R^2$, using the ratio of the sum of squared errors (SSE) of the observations around the mean of the terminal node, to the total sum of squares deviations (SST) of the observations around the mean of the entire dataset.

In parallel, all the regression trees were also performed with all the species, including super dominant small carnivorous species, to verify the stability of trees and the importance given to dominant species.

To validate the identity of key species identified by the regression trees, we adapted the null-model approach developed by Gotelli et al. (2011). We performed LMM regression between the abundance of each species and the predation rates of *V. arvensis*. The random effects were the Experiment name, the Field identifier, and the Month in interaction with Year. The observed slopes were compared with slopes obtained with LMM performed with 1000 random permutations of the predation rate vector. A standardized effect size (SES) was calculated for each species as:

$$SES = \frac{S_{obs} - S_{sim}}{\sigma_{sim}}$$

Where $S_{obs}$ and $S_{sim}$ are the observed and the average of the slopes obtained on permuted data, respectively, and $\sigma_{sim}$ is the standardized deviation of the slopes obtained on the 1000 permuted data. SES values $> 2$ or $< -2$ indicate significant relationships (Gotelli et al., 2011) between the abundance of a given species and the predation rate of *V. arvensis*.

All statistical analyses were performed using the R software version 3.5.1 (R Development Core Team, 2005).
3. Results

The dataset included 30686 carabids belonging to 99 species and 40 genera (Supplementary Table 3 and 4). The 44 seed-eating species observed represented 32.4 % of the catch and, in decreasing order, were *Poecilus cupreus* (58.4 %), the *Harpalus* genus (18.8 %) mainly represented by *H. dimidiatus* and *H. affinis*, and *Pseudoophonus rufipes* (8 %). The 19 potential intraguild predator species represented 12 % of the total carabids catches and were dominated by *Pterostichus melanarius* (8 %) and *Carabus convexus* (1.5 %). At the plot level, there were on average 6.10 (SD = 11.22) individuals and 1.62 (SD = 1.40) species per plot of seed-eating carabids, and 2.26 (SD = 5.44) individuals and 0.79 (SD = 0.83) species per plot that were potential intraguild predators. At the transect level, there were on average 21.17 (SD = 37.46) individuals and 2.91 (SD = 2.04) species per transect of seed-eating carabids, and 8.77 (SD = 17.07) individuals and 1.58 (SD = 1.12) species per transect that were potential intraguild predators. The predation rates of *V. arvensis* ranged from 0 to 1, at both the plot and transect level, with means of 0.43 (SD = 0.40) and 0.41 (SD = 0.33), respectively at the plot and transect level.

3.1. Relating seed predation to carabid assemblages at the plot level

For carabid genera, we obtained a three terminal node tree that explained 10.6 % of the total variance in seed predation rate. The predation rate was best described by the activity density of the carabids from the genus *Harpalus* (Figure 1 A), with relatively high predation rates (\( \bar{x} = 72 \% \), \( n = 197 \)) being associated with high activity density of *Harpalus* (> 2.5). The predation rate was lower (\( \bar{x} = 34 \% \), \( n = 986 \)) at low activity density of *Harpalus* (≤ 0.5).

Carabids of the genus *Harpalus* were present in 41.5 % of the traps, with an average of 2.77 (SD = 3.62) individuals per plot where present. The genus *Harpalus* was selected in 98.2 % of the trees built on the bootstrapped data (Figure 2 A). Following *Harpalus*, the next most frequently selected splitting variables were the genera *Pseudoophonus* and *Poecilus*, which...
appear respectively in 50.2% and 28.6% of trees (Figure 2 A). They were also present as a surrogate split for node 1 (Supplementary Table 5 A). The null model approach confirmed the significance of the genera *Harpalus* (SES 8.09), *Pseudoophonus* (SES 6.70) and *Poecilus* (SES 6.11). The genera *Pterostichus* and *Scybalicus* were also identified as groups exhibiting a positive, but more moderate (SES 3.9 and 3.1 respectively), effect on *V. arvensis* predation.

When implemented using carabid species, the tree explained 6.2% of the predation rate variation and the abundance of *Harpalus dimidiatus* had the highest explanatory power for predation rate (Figure 3 A). Low abundances of *H. dimidiatus* (≤ 0.5) were associated with low predation rates (\(\bar{x} = 39\%\), \(n = 1292\)), and conversely, higher abundances of *H. dimidiatus* (> 0.5) were related to higher seed predation (\(\bar{x} = 65\%\), \(n = 290\)). *Harpalus dimidiatus* was selected in slightly less than 60% of the regression trees from the bootstrapped samples. *Harpalus affinis* did not appear on the principal regression tree but was selected in 76.6% of the bootstrapped regression trees (Figure 4 A) and also appears as a surrogate split (Supplementary Table 5 B). The null model approach corroborated the significance of *H. affinis* (SES 4.68) and *H. dimidiatus* (SES 3.93), and also highlighted the role of *P. rufipes* (SES 4.63) and *P. cupreus* (SES 4.60).

### 3.2. Relating plot level seed predation to transect level carabid assemblages

When considering carabids at the genus level, a five terminal node regression tree was obtained that explained 19.7% of the total variation in seed predation (Figure 1 B). Three important carabid genera were identified: *Harpalus*, *Poecilus* and *Pseudoophonus*, which is consistent with the variable selection frequencies barplot (Figure 2 B). Predation rate was high (\(\bar{x} = 86\%\), \(n = 112\)) when *Harpalus* was abundant (> 8) and associated with high abundance of *Poecilus* (> 17). More than 55% of the observations were associated with low predation rates (\(\bar{x} = 29\%\), \(n = 898\)) and were related to low activity densities of *Harpalus* (≤ 4) and *Pseudoophonus* (≤ 1). There were three intermediate nodes, the node 4 (\(\bar{x} = 61\%\), \(n = 112\))...
223) related to presence of *Pseudoophonus* (> 1) but low abundance of *Harpalus* (≤ 4), the
node 6 (x̅ = 55 %, n = 273), associated with low AD of *Poecilus* (≤ 17) and intermediate AD
of *Harpalus* (4 < and ≤ 8), and the node 8 (x̅ = 61 %, n = 76) related with high abundance of
*Poecilus* (17 >) and intermediate AD of *Harpalus* (4 < and ≤ 8). The predation rates of these
three subgroups were not significantly different, and presented a high level of variability. The
null model approach corroborated the significance of *Harpalus* (SES 5.37), *Pseudoophonus*
(SES 4.88) and *Poecilus* (SES 4.49) found in the trees. *Pterostichus* and *Scybalicus* also had a
positive but more moderate effect on seed predation rates (SES 3.07 and 2.87 respectively).

When considering carabid species, a six terminal node regression tree was obtained that
explained 13.5 % of seed predation variation, with a key role of *H. dimidiatus*, *H. affinis*, *P.
ruipes*, *Carabus auratus* and *Amara eurynota* (Figure 3 B). There were two nodes associated
with very high predation rates, nodes 9 (x̅ = 86 %, n = 64) and 10 (x̅ = 99 %, n = 16), linked
to *H. affinis* (> 1), *P. ruipes* (> 2) and absence of *C. auratus* (≤ 0). Most of the observations
indicate low rates of predation (node 3; x̅ = 34 %, n = 1037) that were related to low
abundance of *H. affinis* (≤ 1) and *H. dimidiatus* (≤ 3). *P. cupreus* was frequently present in the
trees built on bootstrapped data (47 %), whereas it was not present on the main tree, in
contrast to *C. auratus* and *A. eurynota* that were rarely selected (Figure 4 B). The null model
approach corroborated the significance of *H. dimidiatus* (SES 7.22), *H. affinis* (SES 6.94), *P.
ruipes* (SES 6.89) and *P. cupreus* (SES 6.35). *P. melanarius*, *S. oblongiusculus* and *A.
equestris* also had a positive but moderate effect on seed predation rates (SES 4.12, 2.93 and
2.38 respectively).

### 3.3. Relating seed predation to carabid assemblages at the transect level

When considering carabid genera, we obtained a four terminal node regression tree that
explained 29 % of seed predation variation. The genera *Harpalus*, *Poecilus* and
*Pseudoophonus* explained the most variability in predation (Figure 1 C), and frequently
appeared in trees on bootstrapped data (Figure 2 C). Predation rate was high for node 7 ($\bar{x} = 74 \%$, $n = 50$) when *Poecilus* was abundant in the transect ($> 15$), and when the abundance of *Harpalus* exceeded 4 individuals. However, this situation was rather unusual in fields, occurring in only about 10% of observations. Most of the observations were associated with low predation rates ($\bar{x} = 28 \%$, $n = 301$) and related to low abundance of *Harpalus* ($\leq 4$) and *Pseudoophonus* ($\leq 2$). There were two intermediate nodes presenting high level of predation variability, nodes 4 ($\bar{x} = 65 \%$, $n = 38$) and 6 ($\bar{x} = 58 \%$, $n = 81$) related to combination of low and high abundance of the three key genera. The null model approach confirmed the overriding importance of *Harpalus* (SES 5.43), *Pseudoophonus* (SES 4.74) and *Poecilus* (SES 4.54). *Pterostichus* and *Scybalicus* also had a positive but moderate effect on seed predation rates (SES 2.82 and 2.14 respectively).

When considering carabid species, the tree explains 30% of predation variation and three species emerge as important: *H. affinis*, *H. dimidiatus* and *P. rufipes*. The highest level of predation (node 9; $\bar{x} = 84 \%$, $n = 26$) was associated with *H. affinis* ($> 1$) and *P. rufipes* ($> 2$) (Figure 3 C). Although it did not appear on the main tree, *P. cupreus* was selected in 29% of the trees based on bootstrapped data (Figure 4 C), and was a surrogate split (Supplementary Table 7 B). The null model approach corroborated the significance of *H. dimidiatus* (SES 4.30), *H. affinis* (SES 5.36) and *P. rufipes* (SES 4.70). It also highlighted the role of *P. cupreus* (SES 4.45), *P. melanarius* (SES 2.81) and *S. oblongiusculus* (SES 2.14).

### 4. Discussion

Our study provides some methodological advances and guidelines to explore the relative contribution of key taxa to pest regulation in field situations. We investigated this relationship at different spatial and taxonomic resolutions, with the objective of providing guidelines for future research. Our work consistently identifies four seed-eating carabid species, genera and their assemblages as key drivers of the depletion of *V. arvensis* seeds, and suggests that intra-
guild predators within carabid community and their trophic interactions only played a minor role in seed predation. Finally, our analysis shows that in many of the field situations explored in this paper, seed predation was limited because key carabid species driving seed predation were not present at high enough abundance.

4.1. The key taxa driving seed depletion

Among the 44 species and 26 genera of seed-eating carabids present in our dataset, four key carabid taxa were consistently estimated to have a significant effect on the predation of *V. arvensis* seeds. The granivore genus *Harpalus*, appears as a key group that explains the predation of *V. arvensis*. Two species of *Harpalus* were found to be important, *H. affinis* and *H. dimidiatus*. These are two medium sized spring breeding granivores (Homburg et al., 2014) that made up approximately 6 % and 8.7 % of the seed feeders caught, respectively. After spatial aggregation of carabids, we detected effects of two other medium size species, *P. cupreus* and *P. rufipes*, which are spring and autumn breeder species respectively (Cole et al., 2002; Homburg et al., 2014; Pey et al., 2014), and are the dominant species of the genera *Poecilus* and *Pseudoophonus*, respectively, in our data. The importance of these key genera corroborates the findings of previous studies that have shown that genera from the tribe *Harpalini*, such as *Harpalus* or *Pseudoophonus*, consumed significant quantities of *V. arvensis* seeds in laboratory cafeteria tests (Honek et al., 2007; Petit et al., 2014; Deroulers and Bretagnolle, 2019). Interestingly, *P. cupreus* does not show a preference for *V. arvensis* seeds in cafeteria tests, eating some 0.40 seeds / day in comparison to 4.65 and 15.37 seeds / day for *H. affinis* and *P. rufipes*, respectively (Petit et al., 2014). *P. cupreus* is dominant, however, being present in 69.6 % of the sampled transects, with on average 12.11 (SD = 28.79) individuals per trap. Thus, *Poecilus* appeared as a key species due to mass effect in the field, compensating for its low consumption of *V. arvensis* seeds, while *Pseudoophonus* and *Harpalus* are key species due to their high per capita consumption of *V. arvensis* seeds. The
genus *Amara*, known to consume substantial amounts of weed seeds (Honek et al., 2007; Petit et al., 2014) was rarely selected in the regression trees but this could be related to its low acceptance of *V. arvensis* seeds, as previously shown in laboratory cafeteria tests (Honek et al., 2007; Petit et al., 2014). The highest predation rates were associated with the combination of relatively high abundances of *Harpalus* and *Poecilus*, at the genera level, and *H. affinis* and *P. rufipes*, at the species level. The combination of relatively high activity density of *Harpalus* and *Poecilus* better explained high predation rates than the two genera separately, which may be an indication of additivity or complementarity effects, such as resource partitioning or facilitation when they co-occur (Gurr et al., 2012; Letourneau et al., 2009). Charalabidis et al. (2019) highlighted the differences and complementarities of foraging strategies between *H. affinis* and *P. cupreus* under laboratory conditions, which may provide a mechanism for a positive association of these two species. *Poecilus cupreus* has a low latency time before consumption, but is limited in the range of weed seed species consumed, whereas *H. affinis* start consuming after a longer delay, but consume a wide range of weed species (Charalabidis et al., 2019). The per capita consumption of *V. arvensis* seeds seems to be unaffected by carabid conspecific density, suggesting that seed predation is proportional to the abundance of conspecific for both *H. affinis* and *P. cupreus* (Carbonne et al., 2019). In addition, seed consumption by *H. affinis* increases under risk of intraguild predation, in contrast to consumption by *P. cupreus* that remained unchanged (Charalabidis et al., 2019, 2017). In addition to presenting complementary foraging strategies, these two species present temporal complementarity, with *P. cupreus* mostly described as a day active carabid, and *H. affinis* as a night active species (Ribera et al., 1999; Barbaro and Van Halder, 2009). Finally, although we expected that intraguild predators would be important drivers of the patterns of seed consumption in our field situations, our analysis did not support this
convention. While *Pterostichus melanarius* was identified as being important in our analysis, it appears in only 5 to 10% of the bootstrapped trees and is actually identified as a positive contributor to seed predation, contrary to expectation. The mechanisms that would explain this impact are, however, unclear. Charalabidis et al. (2019) showed that the risk of predation exercised by *P. melanarius* can increase the weed seed consumption by *H. affinis*. It is also not excluded that *P. melanarius* consumes seeds. Other large intraguild predators are sometimes selected in the main or bootstrapped regression trees. For example, when *H. affinis* (>1) and *P. rufipes* (>2) are in combination, the presence of *Carabus auratus* tends to increase the variability in seed predation rates, and it cannot be excluded that *C. auratus* reduces weed seed predation by preying upon seed-feeding carabids (Prasad and Snyder, 2004). The presence of complex habitats and heterogeneous vegetation can be a solution to dampen antagonistic interactions by reducing the chances of encounters among natural enemies (Gontijo, 2019; Snyder, 2019). Here, only indirect effects of intraguild predation within carabid communities are assessed, as information on other non-carabid predator groups was not available. The presence of other seed predators, like micro mammals (Blubaugh et al., 2017), might provide a better understanding of the effect of intraguild predation on the weed seed predation, and increase our explanatory power.

Finally, when regression trees were built on all genera or species without any selection, only the two trees with species at the plot and intermediate scale were slightly different (Supplementary Figure 1 and 2). Simply one new species, *Brachinus crepitans* (Linnaeus, 1758), appears in the tree built with species at the plot scale. *B. crepitans* is an ectoparasitoid carabid which develop on pupae of spring breeding species, and can be highly correlated with some granivorous species, in particular with the genus *Amara* (Saska and Honek, 2008). Trees are therefore relatively stable, and did not seem to over-represent the dominant species.

4.2. The spatial and taxonomic resolution
Although the identification of key taxa was consistent across the three spatial resolutions explored, our results indicate that the relationships between carabid assemblages and seed predation were much better explained at the transect scale than at the plot scale. The apparent mismatch between carabid abundance and predation rates at fine spatial scales has already been shown in field studies (Saska, 2008; Trichard et al., 2014). The higher explanatory power achieved at the transect scale may be explained by the aggregation of data, and particularly the pooling of the predation rate over several plots that averages away some of the variability of predation. It may also reflect the heterogeneous spatial distribution of carabids in arable fields (Thomas et al., 2002; Trichard et al., 2014) and that carabids caught in a specific plot location may have foraged at distances of up to 50 meters away (Wallin and Ekbom, 1988). Finally, aggregating the carabid data increased the pool of species considered with the inclusion of new taxa as partitioning variables in regression trees and to a more precise detection of species preying upon V. arvensis seeds. For future experiments, spatial aggregation of plots, by transect or zone, would provide a beneficial increase in prediction of seed predation by taking into account carabid mobility.

Finally, we assumed that analysing the relationship between carabid taxa and seed predation at the species level would provide more accurate results than an analysis at the genus level due to the potential difference in seed acceptance between species belonging to the same genus. Our results show, however, that there is an increase in explanation in considering predation at the genus level rather than the species level at the plot and intermediate scale. By grouping species at the genus level we formed functional groups including functionally redundant species that if considered separately would have little weight, and would not be included in the tree. This result is supported by previous studies showing that functional groups are more informative for predicting prey suppression than species (Gagic et al., 2015; Russell et al., 2017).
Overall, the 30% of variation in seed depletion explained by assemblages of carabid taxa, at the transect level, is fairly high. The use of regression trees permit the explanation of predation rates in situations where key species are absent or present. Nodes with component parts of the group of four key species identified or intermediate levels of activity-density have highly variable levels of predation (see for example Figure 1 A node 4 and B node 4, 6 and 8). The 70% of the unexplained variation in seed predation is likely due to the occurrence of other invertebrate seed predators such as crickets or ants (Baraibar et al., 2009; Brust, 1994) not accounted for here. Variability in the availability of alternative prey for carabids, but also the presence of competitor or predator, might also affect the predation of the target prey (Koss and Snyder, 2005; Symondson et al., 2006) and account for unexplained variation in the depletion of V. arvensis seeds in the sampled transects. Finally, a part of the seed loss can be due to environmental conditions such as rain, wind or loss of adhesive power of the cards (Westerman et al., 2003).

4.3. Enhancing weed seed predation in arable landscapes

Beside the identification of key taxa delivering seed predation, our study also emphasizes the fact that in the vast majority of our field situations, the predation of V. arvensis was limited by the absence or low abundances of key carabid species. Our study was based on data collected in 111 farmer’s fields conducted under contrasting farming managements and located in very different landscape settings. Yet, in more than 60% of the transects surveyed, key carabid taxa were absent or were poorly represented, and associated with low seed predation. Nodes of the regression trees exhibiting favourable key carabid assemblages delivering high seed predation gathered only a limited number of observations. This suggests that seed predation could be enhanced in arable landscapes, and that currently only a small proportion of the transects surveyed were favourable to the key carabid species associated with high seed predation rates. It is interesting to note that fields under conservation agriculture and organic fields tended to
be more represented in these nodes than conventional fields. Indeed, for the regression tree performed on genera at the transect scale, 17% and 18% of the conservation agriculture and organic field transects are associated to high seed predation (node 7), respectively, compared to only 2% of the conventional agriculture transects. This is in line with studies that have demonstrated the positive effect of conservation agriculture on seed-eating carabid species (Menalled et al., 2007), including *P. rufipes* (Trichard et al., 2014). Tillage is known to strongly reduced adult activity, oviposition rate and larval capture in fields (Blubaugh and Kaplan, 2015). Similarly, organic farming has been shown to enhance the abundance of seed-eating carabids such as *H. affinis* (Diekötter et al., 2016) and *P. rufipes* (Döring and Kromp, 2003). These elements could explain the much higher proportion of fields exhibiting high predation rates in conservation and organic agriculture compared to conventional fields. Yet, our results also suggest that adopting organic or conservation farming practices do not seem to ensure high predation rates in all situations and there may be several reasons for this. First, conversion to organic or conservation agriculture may be too recent to result in significant changes in carabid communities and subsequent increases in weed seed predation rates. For example, increases in predation rates in conservation agriculture were shown to be apparent at least four years after (Petit et al., 2017). In addition, pesticide use can be high in conservation agriculture, with detrimental effects on key carabid species and weed seed predation (Ricci et al., 2019; Trichard et al., 2013). Second, the landscape context of arable fields can greatly affect the composition of carabid communities and modulate the effect of the local farming management in the focal field. For example, the presence of grassy field boundaries can positively affect the in-field abundance of *P. rufipes* (Labruyere et al., 2017, 2016b; Rouabah et al., 2015). At the landscape scale, the cover of grassland (Labruyere et al., 2016a; Purtauf et al., 2005b; Trichard et al., 2013) and the proportion of organic agriculture (Diekötter et al., 2016) enhance the key carabid taxa identified in this study. These elements suggest that
different combinations of local field and landscape scale management could increase the very low abundance of key species and the level of seed predation observed here in conventional arable fields.

Our study suggests that for the conservation biological control of *V. arvensis* it is necessary to promote the activity density of *Harpalus, Poecilus* and *Pseudoophonus*. The statistical modelling approach we develop could be a useful method to identify key natural enemies that exert high pest mortality, and can be a useful guide to select field management that specifically supports these species to improving the efficacy of biocontrol.

5. Acknowledgements

The authors thank the staff at INRA Dijon experimental farm for their technical support. We also would like to thank Jean Villerd for his help with regression tree, and Achim Zeileis for his help with analyzing data with the *glmertree* function. We thank Adam Vanbergen for his critical review of the manuscript before submission.

Funding: This work was supported by the projects FRB EcoPhyto SEBIOPAG-PHYTO, ERA-NET C-IPM BioAWARE and INRA MP SMaCH.
6. References


Hengeveld, R., 1980. Polyphagy, oligophagy and food specialization in ground beetles


Russell, M.C., Lambrinos, J., Records, E., Ellen, G., 2017. Seasonal shifts in ground beetle


**Table 1:** Summary of the information on the five datasets contained in the database. For each dataset we specify year of data collection, number of session, number of fields, number of transect per field, and the number of records at the plot and transect level. More information about the protocols can be found on the corresponding publication (Trichard et al., (2013); Petit et al., (2017); Ricci et al., (2019)).
Table 2: Summary of all genus and species of seed eaters and intraguild predators integrated in the analysis with their corresponding activity-density (AD) and occurrence (%).
Figure 1: Regression trees displaying interaction among activity-density of seed eaters and potential intraguild predator carabid beetles genera on weed seed predation of *V. arvensis* at different spatial scale: (A) Plot scale ($R^2 = 10.60\%$), (B) intermediate scale with predation measures at the plot level and carabid beetles at the transect level ($R^2 = 19.7\%$), and (C) transect scale ($R^2 = 29\%$). Box plots represent medians and upper and lower quartiles for each sub group of the regression trees. Difference between subgroup obtained are indicated by different letter on the top of each boxplot.
Figure 2: Relative frequency of selection of the most frequent carabid genera in the regression trees, built on 500 bootstraps of the data set, for explaining V. arvensis predation rates. The variables selected in the regression tree based on the initial data set are indicated in dark grey, and the associated genera names are underlined. Letters indicate the spatial aggregation tested: (A) plot scale, (B) intermediate scale (predation measure at the plot level, and carabid beetles at the transect level), and (C) transect scale.
Figure 3: Regression trees displaying interaction among activity-density of seed eaters and potential intraguild predator carabid beetles species on weed seed predation of *V. arvensis* at different spatial scale: (A) Plot scale (R² = 6.20 %), (B) intermediate scale with predation measures at the plot level, and carabid beetles at the transect level (R² = 13.5 %), and (C) transect scale (R² = 29.7 %). Box plots represent medians and upper and lower quartiles for each sub group of the regression trees. Difference between subgroup obtained are indicated by different letter on the top of each boxplot.
Figure 4: Relative frequency of selection of each carabid species in the regression trees built on 500 bootstraps of the data set, for explaining *V. arvensis* predation rates. The variables selected in the regression tree based on the initial data set are indicated in dark grey, and the associated genera names are underlined. Letters indicate the spatial aggregation tested: (A) plot scale, (B) intermediate scale (predation measure at the plot level, and carabid beetles at the transect level), and (C) transect scale.
(A) Node 3 (n = 986) Node 4 (n = 400) Node 5 (n = 196)

(B) Node 3 (n = 898) Node 4 (n = 223) Node 6 (n = 273) Node 8 (n = 80) Node 9 (n = 108)

(C) Node 3 (n = 301) Node 4 (n = 38) Node 6 (n = 81) Node 7 (n = 50)

Seed removal rates
<table>
<thead>
<tr>
<th>Project</th>
<th>Main objective</th>
<th>Year</th>
<th>No. sessions</th>
<th>No. of fields</th>
<th>No. of transects per field</th>
<th>Records at plot level</th>
<th>Records at transect level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trichard et al., (2013)</td>
<td>Investigate spatio-temporal association between weed seed predation and carabids in two type of soil management: conventional and no tillage.</td>
<td>2011</td>
<td>5</td>
<td>2</td>
<td>14</td>
<td>280</td>
<td>140</td>
</tr>
<tr>
<td>Petit et al., (2017)</td>
<td>Investigate the effect of landscape and conservation agricultural practices, in interaction, on the weed seed predation in-field.</td>
<td>2012</td>
<td>1</td>
<td>69</td>
<td>2</td>
<td>546</td>
<td>138</td>
</tr>
<tr>
<td>Lubac (2014)</td>
<td>Investigate the effect of crop type on weed, carabid communities and weed seed predation.</td>
<td>2014</td>
<td>1</td>
<td>8</td>
<td>1</td>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>Ricci et al., (2019)</td>
<td>Investigate the effect of landscape and local pesticide use intensity, in interaction, on weed seed, aphids and Lepidoptera eggs predation.</td>
<td>2015 to 2018</td>
<td>2</td>
<td>20</td>
<td>1</td>
<td>640</td>
<td>160</td>
</tr>
<tr>
<td>Minvielle-Debat (2016)</td>
<td>Characterize carabid communities at different spatial scales.</td>
<td>2016</td>
<td>2</td>
<td>12</td>
<td>1</td>
<td>95</td>
<td>24</td>
</tr>
</tbody>
</table>
## Seed Eaters (44 species)

<table>
<thead>
<tr>
<th>Genus/Species</th>
<th>Total AD</th>
<th>Occurrence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Poecilus cupreus</td>
<td>5638</td>
<td>55.94</td>
</tr>
<tr>
<td>Harpalus</td>
<td>1815</td>
<td>41.53</td>
</tr>
<tr>
<td>H. dimidius</td>
<td>713</td>
<td>21.68</td>
</tr>
<tr>
<td>H. affinis</td>
<td>584</td>
<td>22.06</td>
</tr>
<tr>
<td>H. distinguendus</td>
<td>241</td>
<td>5.25</td>
</tr>
<tr>
<td>H. rubripes</td>
<td>132</td>
<td>4.80</td>
</tr>
<tr>
<td>H. serripes</td>
<td>70</td>
<td>2.47</td>
</tr>
<tr>
<td>H. tardus</td>
<td>50</td>
<td>2.47</td>
</tr>
<tr>
<td>H. sp</td>
<td>14</td>
<td>0.25</td>
</tr>
<tr>
<td>H. atratus, H. latus, H. lateicornis, H. smaragdinus</td>
<td>13</td>
<td>0.13-0.25</td>
</tr>
<tr>
<td>Pseudoophonus rufipes</td>
<td>788</td>
<td>20.42</td>
</tr>
<tr>
<td>Amara</td>
<td>465</td>
<td>12.20</td>
</tr>
<tr>
<td>A. similata</td>
<td>236</td>
<td>5.63</td>
</tr>
<tr>
<td>A. eurynota</td>
<td>83</td>
<td>2.97</td>
</tr>
<tr>
<td>A. aenea</td>
<td>45</td>
<td>1.14</td>
</tr>
<tr>
<td>A. consularis</td>
<td>40</td>
<td>1.45</td>
</tr>
<tr>
<td>A. ovata</td>
<td>34</td>
<td>1.20</td>
</tr>
</tbody>
</table>

## Intraguild Predators (19 species)

<table>
<thead>
<tr>
<th>Genus/Species</th>
<th>Total AD</th>
<th>Occurrence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pterostichus</td>
<td>2403</td>
<td>37.93</td>
</tr>
<tr>
<td>P. melanarius</td>
<td>2394</td>
<td>37.80</td>
</tr>
<tr>
<td>P. anthracinus</td>
<td>9</td>
<td>0.51</td>
</tr>
<tr>
<td>Abax</td>
<td>21</td>
<td>1.39</td>
</tr>
<tr>
<td>A. parallelepipedus</td>
<td>20</td>
<td>1.33</td>
</tr>
<tr>
<td>A. parallelus</td>
<td>1</td>
<td>0.06</td>
</tr>
<tr>
<td>Carabus</td>
<td>864</td>
<td>22.57</td>
</tr>
<tr>
<td>C. convexus</td>
<td>462</td>
<td>13.02</td>
</tr>
<tr>
<td>C. auratus</td>
<td>296</td>
<td>5.31</td>
</tr>
<tr>
<td>C. violaceus</td>
<td>50</td>
<td>3.48</td>
</tr>
<tr>
<td>C. monilis</td>
<td>49</td>
<td>2.53</td>
</tr>
<tr>
<td>C. coriaceus</td>
<td>7</td>
<td>0.76</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genus/Species</th>
<th>Total AD</th>
<th>Occurrence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antisodactylus</td>
<td>287</td>
<td>2.09</td>
</tr>
<tr>
<td>A. signatus</td>
<td>274</td>
<td>1.77</td>
</tr>
<tr>
<td>A. binotatus</td>
<td>13</td>
<td>0.63</td>
</tr>
<tr>
<td>Ophonus</td>
<td>265</td>
<td>4.74</td>
</tr>
<tr>
<td>O. puncticeps</td>
<td>224</td>
<td>2.34</td>
</tr>
<tr>
<td>O. azureus</td>
<td>39</td>
<td>2.53</td>
</tr>
<tr>
<td>O. sabulicola</td>
<td>2</td>
<td>0.13</td>
</tr>
<tr>
<td>Scybalicus oblongiusculus</td>
<td>230</td>
<td>1.33</td>
</tr>
<tr>
<td>Diachromus germanus</td>
<td>116</td>
<td>1.14</td>
</tr>
<tr>
<td>Pterostichus madidus</td>
<td>18</td>
<td>1.01</td>
</tr>
<tr>
<td>Stenolophus teutonus, Acupalpus meridianus, Semiophonus signaticornis, Zabius tenebriones, Cryptophonus tenebrinosus, Paraphonus mendax, Polistichus connexus, Acupalpus dubius</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nebria</td>
<td>245</td>
<td>10.18</td>
</tr>
<tr>
<td>N. salina</td>
<td>203</td>
<td>8.15</td>
</tr>
<tr>
<td>N. brevicollis</td>
<td>42</td>
<td>2.47</td>
</tr>
<tr>
<td>Calathus fascipes</td>
<td>32</td>
<td>2.15</td>
</tr>
<tr>
<td>Molops piceus, Chlaenius nigricornis, Cylindera germanica, Poecilus versicolor, Bembidion lunukatum, Cicindela campestris, Platynus assimilis</td>
<td>19</td>
<td>0.06-0.70</td>
</tr>
</tbody>
</table>
Classification of weed seed predation of *Viola arvensis* according to the composition of carabid key genera