Metabarcoding and ecological interaction networks for selecting candidate biological control agents

Mélodie Ollivier[1], Vincent Lesieur[1,2], Johannes Tavoillot[3], Fanny Bénetière[1], Marie-Stéphane Tixier[1], Jean-François Martin[1]

[1] CBGP, Montpellier SupAgro, INRAE, CIRAD, IRD, Univ Montpellier, Montpellier, France

[2] CSIRO Health and Biosecurity, European Laboratory, Montferrier sur Lez, 34980, France

[3] CBGP, IRD, CIRAD, INRAE, Montpellier SupAgro, Univ Montpellier, Montpellier, France

Abstract

1. Classical biological control can be used to decrease the density of invasive species to below an acceptable ecological and economic threshold. Natural enemies specific to the invasive species are selected from its native range and released into the invaded range. This approach has drawbacks, despite the performance of specificity tests to ensure its safety, because the fundamental host range defined under controlled conditions does not represent the actual host range *in natura*, and these tests omit indirect interactions within community.
2. We focus on *Sonchus oleraceus* (Asteraceae), a weed species originating from Western Palearctic that is invasive worldwide and notably in Australia. We explore how analyses of interaction network within its native range can be used to 1) inventory herbivores associated to the target plant, 2) characterize their ecological host ranges, and 3) guide the selection of candidate biocontrol agents considering interactions with species from higher trophic levels. Arthropods were collected from plant community sympatric to *S. oleraceus*, in three bioclimatic regions, and interactions were inferred by a combination of molecular and morphological approaches.
3. The networks reconstructed were structured in several trophic levels from basal species (diversified plant community), to intermediate and top species (herbivorous arthropods and their natural enemies). The subnetwork centered on *S. oleraceus* related interactions contained 116 taxa and 213 interactions. We identified 47 herbivores feeding on *S. oleraceus*, 15 of which were specific to the target species (*i.e.* Generality value equal to 1). Some discrepancies with respect to published findings or conventional specificity tests suggested possible insufficient sampling effort for the recording of interactions or the

existence of cryptic species. Among potential candidate agents, 6 exhibited interactions with natural enemies.

4. *Synthesis and applications*: Adopting a network approach as prerequisite step of the CBC program can provide a rapid screening of potential agents to be tested in priority. Once ecological host range defined, we suggest that priority should be given to agent predated by a minimum species, and, when they exist, to an agent that possesses enemies from the most distant taxonomical group from those occurring in the range of introduction.

Key words: High-throughput sequencing, host range, food web, weed biocontrol, Common Sowthistle

Introduction

Biological invasions are currently threatening biodiversity to an unprecedented extent (Bellard et al., 2016; Seebens et al., 2015; Vitousek et al., 1997). When invasive species disrupt the ecological or economic balance, action is required to control their negative impact. Chemical control methods are widely used in such situations, but classical biological control (CBC) constitutes a possible alternative. CBC involves the release of natural enemies, specific to the target organism and originating from its native range, to keep the density of the invasive species below an economically and ecologically acceptable threshold (Keane and Crawley, 2002; McFadyen, 1998; Van Driesche et al., 2010). CBC is considered more sustainable than chemical control (Peterson et al., 2020), although the introduction of biocontrol agents (BCA) into a new territory may itself represents a risk for the recipient communities (Barratt et al., 2018; Hinz et al., 2019; Suckling and Sforza, 2014). Once introduced, the BCA may affect non-target species, especially if it lacks specificity (Müller-Schärer and Schaffner, 2008). Assessing the host range of a candidate BCA is, thus, crucial, to anticipate such risks. Most of the host-specificity tests performed to assess this risk are conducted under standardized conditions, through choice/no-choice experiments over a range of targets selected according to the centrifugal phylogeny approach (Briese, 2005; Wapshere et al., 1989).

Recent reviews recognized the success of such experiments for limiting the undesirable unintentional effects of the CBC of weeds (Hinz et al., 2019, 2020). However, the cumbersome nature of these tests reduces the range of species that can be screened. The candidate BCA are selected through preliminary field monitoring that may miss a species of interest. Furthermore, as the fundamental host range of a species (defined under controlled conditions) is thought to be broader than the host range actually observed in the field (known as the realized host range;

Louda et al., 2003; Schoonhoven et al., 1998; Sheppard and Harwood, 2005), these tests tend to overestimate the risk and lead to the rejection of candidate BCA based on interactions that would not occur in the field (false positives) (e.g. Groenteman et al. 2011). Most CBC programs use specificity tests under controlled conditions as proxies for field conditions due to the complexity of trophic interaction assessments in the field but this leaves room for improvement.

The characterization of ecological interactions among communities of plants and arthropods *in natura* is challenging, as it traditionally requires direct observations, the rearing of specimens and considerable taxonomic expertise, rendering the process impractical for large-scale studies. Recent advances in molecular approaches, such as the combination of DNA metabarcoding on gut content or feces and high-throughput next-generation sequencing (NGS), have opened up new opportunities to track the host range of arthropods *in natura* with both a high taxonomic resolution and high sensitivity (Derocles et al., 2018; Frei et al., 2019; Wirta et al., 2014; Zhu et al., 2019). Even interactions that are very difficult to observe, such as host-parasitoid associations, can be detected by such methods (Gariepy et al., 2014; Hrček and Godfray, 2015). This approach can be used to reconstruct networks of trophic interactions directly from studies in the field, and provides an analytical framework particularly relevant to studies of complex species assemblages. Network ecology do not only depicts species interactions, but provides elements for the understanding of recurrent patterns of antagonistic interactions between plants and herbivores, such as specialization or compartmentation (Lewinsohn et al., 2006; Thébault and Fontaine, 2010). In CBC against invasive weeds, analyses of ecological networks have been used to assess the extent to which a BCA fits into a recipient community. Such methods provided a way to quantify the direct impact of biological control on non-target plants (Memmott, 2000), and its indirect impact on other species at higher trophic levels (Carvalheiro et al., 2008; Louda et

al., 1997; Pearson and Callaway, 2003). Such studies have highlighted the usefulness of network ecology for evaluating the impact of BCA after their introduction (Memmott, 2009; Willis and Memmott, 2005), but interaction network analysis can also be used for the upstream assessment of potential candidate BCA, in a more systematic process (Ollivier et al., 2020). Adopting a network approach as prerequisite step of the CBC program, can provide a rapid screening of the ecological host range of potential agents to be tested in priority. This can also inform about species functional properties through the position and connexions the species have in the network, independently of its taxonomic assignation, which would confer a strong predictive power of the interactions possibly occurring in a novel bioclimatic region (Todd et al., 2020). Indeed, the choice of BCA should also take into account indirect effects on the recipient community due to interactions with higher trophic levels in the network, *i.e.* natural enemies (Hinz et al., 2019; Memmott, 2000). If comparable enemies than those identified in the native range are present in the range of introduction, new interactions might be created with BCA, resulting in disturbances in the ecological network through indirect interactions, *e.g.* apparent competition (Carvalheiro et al., 2008; López-Núñez et al., 2017).

The objective of this study was to determine how the analysis of interaction networks could be used to support the selection of candidate BCA for the common sowthistle, *Sonchus oleraceus* L. (Asteraceae). This plant is native to Western Europe and Northern Africa (Boulos, 1974; Hutchinson et al., 1984) and is the most widely naturalized terrestrial plant worldwide (Pyšek et al., 2017). In Australia, it has become a weed of major concern in cropping systems (Llewellyn et al., 2016; Widderick et al., 2010). Aside the development of resistance to multiple herbicides (Adkins et al., 1997; Jalaludin et al., 2018; Meulen et al., 2016), the control of this weed is complex as it is extremely prolific and seeds can germinate all year round when sufficient

113 rainfalls occurs. *Sonchus oleraceus* rapidly dominates crops, reducing yield and contaminating
114 harvested grain (Llewellyn et al., 2016). A CBC program was therefore initiated in 2017, to
115 identify candidate BCA. In this context, an analysis of ecological networks, based on direct field
116 observations and high-throughput DNA metabarcoding, was performed. Our objectives were to
117 1) establish an inventory of arthropods feeding on *S. oleraceus*, and assess the contribution of the
118 approach relatively to classical procedures, 2) delineate the ecological host range for herbivores
119 feeding on *S. oleraceus* and point out candidate BCA, and 3) identify the trophic interactions of
120 the candidate BCA with natural enemies, and consider their implications for the CBC program.
121

Materials and Methods

Sampling design

We maximized the species diversity and associated interactions, through a maximum variation design with three bioclimatic regions in France (semi-oceanic, Mediterranean and continental climates) (Ceglar et al., 2019) and three successive sampling dates (April, May and June 2018). Sampling was carried out from 10 A.M. to 4 P.M., by varying climatic conditions (wet, cloudy to sunny weathers and temperatures ranged between 10°C and 29°C). These variations did not affect our ability to capture arthropods. For each bioclimatic region and date, we employed an opportunistic sampling strategy to collect plants from several ruderal and agricultural sites, covering the diversity of habitats (open and disturbed) colonized by *S. oleraceus* (Supplementary Table 1). At each site, on each date, we sampled three quadrats (1 m^2) along a 20 m linear transect. Quadrats were placed to contain at least one *S. oleraceus* plant. Within each quadrat, arthropods were collected from plants with a forceps or brush, and stored individually in sterile 2 ml Eppendorf tubes filled with a protective buffer solution. This solution is used to prevent oxidation of polyphenols and polyamines (PCR inhibitors) (see Cruaud et al. (2018) for more details). This procedure was repeated for each plant of every plant species present in the quadrat over a period of one hour, to standardize the sampling effort. This period was deemed adapted to represent the biodiversity of the sampled unit, and to allow vagrant insects, potentially disturbed by our arrival, to settle back on their resource plant before sampling. We collected individual specimens except for colonies of aphids, thrips, and egg masses, for which at least five specimens were required to obtain sufficient DNA for analysis. We did not consider pollinators or the soil fauna in this study. Following the collection of each specimen, tools were thoroughly cleaned by successive immersions in 2.5 % bleach solution, water and 96% ethanol, to prevent cross-

145 contamination. At the end of the one-hour insect sampling period, all the plants within the
146 quadrat were collected individually (by cutting the stem at the soil surface), for further dissection.
147 Back in the laboratory, the plants were identified morphologically, and their organs (stems,
148 leaves, flowers) were dissected to collect endophagous arthropods, which were transferred into
149 tubes as described for the arthropods collected in the field. For each arthropod specimen
150 collected, we identified the plant species from which arthropods were sampled, and recorded the
151 specimen stage and condition (degraded, parasitized), and putative identification (at least
152 taxonomic group, with identification to species level if straightforward). All arthropod samples
153 were frozen at -20°C until DNA analysis. Thus, while plants were identified morphologically,
154 arthropods were identified via molecular technologies. Each plant was transferred to a paper bag
155 and oven-dried at 70°C for 72 h, for the determination of aboveground dry biomass (g) as an
156 estimate of plant abundance per quadrat. Arthropod abundances were determined based on the
157 number of individuals collected per quadrat for each taxon. Sampling was performed for 57
158 quadrats, over the three sampling dates.

159 ## High-throughput DNA metabarcoding

160 We characterized the interaction network by directly observing plant-arthropod interactions
161 (recording only interactions for which an observation of feeding was verified); while arthropod-
162 arthropod interactions were revealed by molecular analysis. We first isolated total DNA from
163 each arthropod individual (Cruaud et al., 2018). As presented in Supplementary Figure S1, we
164 then performed metabarcoding on each arthropod sample, with a two-step DNA amplification
165 and high-throughput sequencing method adapted from the procedure described by Galan et al.
166 (2017). We sequenced three short COI fragments, with primer combinations and PCR protocols
167 developed elsewhere (HCO forward: Leray et al. 2013, HCO reverse: Folmer et al. 1994, LEP F.

and R.: Brandon-Mong et al. 2015, HEX F. and R.: Marquina et al. 2019), to overcome the problem of the lack of primer universality among arthropods. Error-proof indices for individual sample identification were developed with the high-throughput sequencing process described by Martin (2019). The libraries were sequenced with Illumina technology, using a Miseq 2x250 run for date 1 (April), and one lane of Hiseq 3000 each for dates 2 (May) and 3 (June).

The markers for each sample were demultiplexed with CutAdapt v2.3, and all paired-end reads were filtered for minimal length (280 bp), corrected for sequencing errors, and pairs of overlapping reads were merged with the Dada2 v1.12 R package (Callahan et al., 2016). A matrix was thus obtained, containing samples as variables and amplicon variant sequences (ASVS) as observations. A variant is a set of identical corrected and merged paired-end reads. We used Qiime2 (Bolyen et al., 2018) with a 2% divergence threshold, to merge ASVS, to decrease their number without the loss of taxonomic information. The summed number of reads for each merged variant for a given arthropod sample was reported as the intersection of samples and ASVS.

Each ASVS was assigned, by BLAST, to a barcoding reference database of cytochrome oxidase subunit I (COI) nucleotide sequences (658 bp) compiled from three different sources and curated by expert analysis. These reference barcodes were retrieved from BOLDSYSTEM (Ratnasingham and Hebert, 2007), the CBGP - Continental Arthropod collection (Centre de Biologie pour la Gestion des Population, 2019) and a local database specifically designed for this study. Our database contained barcodes of the most frequently encountered species during this sampling campaign (extra-specimens collected) and field surveys (2017-2020) conducted through Europe and North Africa for the search of *S. oleraceus* natural enemies (see below). In total, these three sources compiled 1 699 995 sequences from 119 299 species available for ASVS

assignation. We retained successful assignments to the ranks of species, genus and family, but not those to higher taxonomic levels, because arthropod biology is too variable at higher taxonomic ranks to be informative for our purpose. The assignments obtained for each marker were grouped together in a single table and the numbers of reads were summed by assigned taxon. The resulting file was therefore an interaction matrix in BIOM file format, in which the assigned taxa replaced ASVS. The matrix was curated and manually transformed to obtain an adjacency matrix (in which the observations are sources and the variables are consumers) usable for further network analyses. For each pair of consumer/prey species, occurrence frequencies of interaction were calculated (Supplementary Text 1 and Figure S2).

Assessment of sampling robustness and global network description

We first evaluated the completeness of sampling over the entire sampling campaign, and generated taxon accumulation curves (the 57 quadrats were added in a random order, with 1,000 permutations) for plants and arthropods, using the *specaccum* function of the R package *vegan* (Oksanen *et al*., 2019). We estimated the extrapolated taxonomic richness by calculating the Chao 1 index (Chao, 1984) with the *specpool* function. Likewise, the robustness of sampling for the characterization of interactions was assessed by generating accumulation curves for pairwise interactions. We first generated an accumulation curve including all the types of direct interactions (*e.g.* plant-herbivores, herbivores-natural enemies, etc.) present in the meta-network (*i.e.* pooling interactions from all sites). The 57 quadrats were added in a random order, with 1000 permutations. We finally generated a curve focusing on interactions involving *S. oleraceus* as a source, to evaluate the performance of the sampling design for addressing our objective of establishing an inventory of the arthropods feeding on *S. oleraceus*, corresponding to candidate

BCA. For both curves, we estimated the extrapolated interaction richness with the Chao 1 index (Chao 1984), using the *specpool* function.

Prior to interaction analyses, a global description of the meta-network (pooling interactions data from all sites) and subnetwork (centred on *S. oleraceus* related interactions) was performed. Several metrics were calculated: the number of links (L), the number of nodes (S) (connected and isolated), connectance (C) and link density (LD) (Bersier et al., 2002; Warren, 1994). Connectance is the proportion of the possible trophic links actually realized; here cannibalism is not permitted, so C = L/S(S-1). Link density is the mean number of links per taxon, calculated as LD = L/S. We also characterized the taxon assemblage by determining taxonomic richness (*i.e.* number of taxa) for each trophic level (plants, herbivores and natural enemies).

Selection of candidate biocontrol agents

The selection of candidate BCA was decided according two criteria: a restricted ecological host range and limited interactions with natural enemies. Thus, based on the interactions retrieved from the meta-network, we selected a subnetwork considering only the arthropods having *S. oleraceus* as a source plant, as well as all their complementary plant resources. We also included natural enemies associated with these herbivores (*i.e.* parasitoids and predators). We assessed and visualized the specificity of these herbivores, by plotting interactions between herbivores encountered on *S. oleraceus* and all their complementary resource plants as a grid matrix, in which plants were ordered by their degree of phylogenetic relatedness to *S. oleraceus*, as defined by the current classification of angiosperms (Chase et al., 2016). Arthropods were ordered by the increasing generality values (*i.e.* the number of resources per taxon) characterizing ecological host range. To assess and visualize the dependence of natural enemies on these herbivores, we constructed a second level grid matrix in connexion with the previous, and calculated arthropod

vulnerability values (*i.e.* the number of consumer per taxon). Multipartite network and grid matrices were constructed with *igraph* R package.

Assessing the contribution of the approach for the biocontrol program

To discuss the contribution of the method herein proposed, we used, as a point of reference, a survey performed following classical procedures (sampling, rearing and identification of specimens exclusively collected from *S. oleraceus*) in the frame of this CBC program (Lesieur et al., in prep). However, we acknowledge that this classical survey covered a longer period of sampling (2017-2020) and a much larger geographical area was prospected (10 countries through Europe and North Africa).

Results

Summary of the molecular results

In total, 2,834 arthropod specimens were collected and analyzed by metabarcoding, to reconstruct the interaction network at a global scale. We obtained DNA sequences and taxonomic assignments for 1,803 of the 2,834 arthropods initially collected (63.6 %). This proportion of exploitable information reached 71% (2,011 specimens) after manual validation of the matrix. The molecular analysis provided a total of 107,483,410 reads, 19.2% of which were retained after screening with quality filters; we obtained a final dataset of 2,014 COI variant sequences (Supplementary Table 2). Before, manual validation, we observed that a large proportion of the diversity (33% of the families and 40% of the species) was recovered by the use of all markers, the rest being recovered by a combination of two markers, or specifically found with only one marker (Supplementary Figure S3). LEP increased identification rates by 20% for families and 25% for total species, consistent with its widespread use in the research community (Brandon-Mong et al., 2015). The other two markers also provided original information, albeit to a lesser

extent, at least as far as the number of taxa recovered was concerned, as 15% of the families and 16% of the species would not have been recovered with LEP alone. After data validation, 269 taxa were identified for arthropods, with 84% identified to species level (17 orders, 90 families and 189 genera). While plant taxonomic diversity (relying on morphological identifications) accounted for 132 taxa, 80% of which were classified to species level (25 orders, 29 families and 87 genera) (Supplementary figure S4).

Sampling robustness

The accumulation curve of plants seemed to approach an asymptote, but this was not the case for arthropods (Figure 1). The Chao 1 index indicated an extrapolated taxonomic richness value for plants of 164 taxa (± 12), with 132 taxa actually sampled. By contrast, for arthropods, the extrapolated taxonomic richness value was 442 taxa (± 39), but only 269 taxa were actually sampled. Sampling robustness was high over the entire sampling scheme for plants but sampling efficiency was lower for arthropods. Likewise, we assessed the completeness of pairwise interactions detected over the whole network. We observed a linear increase associated with a Chao1 index of 1245 (± 183) expected interactions, where 350 links were actually reconstructed (Figure 1). However, this is less of an issue for interactions involving *S. oleraceus*, the focus of the analysis for which this sampling was designed. The accumulation curve in question tended towards an asymptote, with a Chao 1 index of 63 (± 10) expected interactions and 47 interactions sampled. Overall, these results suggest that the sampling effort was adequate for the reconstruction of a unique interaction network maximizing of the proportion of links observed (Jordano, 2016).

Meta-network and subnetwork analyses

As presented in Table 1, the complete interaction network (meta-network) consisted of 401 nodes, 241 of which were connected to another node (60%), resulting in 350 links (Supplementary Figure S5). Linkage density and connectance calculated were 1.45 and 0.006, respectively. The meta-network included 60 plants in interaction (46% of the plants collected), 136 herbivores in interaction (74% of the herbivores collected), 35 natural enemies in interaction (49% of the natural enemies collected) comprising 19 parasitoid and 16 predator taxa, and 10 omnivores (feeding at more than one trophic level). The sub-network consisted of 116 nodes and 213 links, and resulting linkage density and connectance were 1.84 and 0.008, respectively (Figure 2). A more detailed description of taxon assemblage composing *S. oleraceus* subnetwork is provided in the following section.

Identifying candidate biocontrol agents: considering host range and regulation by enemies

Analysing *S. oleraceus* subnetwork, we found 47 herbivorous taxa feeding on the target, including 37 taxa identified to species level. They belonged to five different orders, *i.e.* Hemiptera (45%), Diptera (25%), Coleoptera (19%), Lepidoptera (0.06%) and Hymenoptera (0.04%), and were distributed in nine different trophic guilds, with the flower bud sucking-piercing guild being the most represented (23%) while the less represented guild corresponded to the chewing guild (2%) (Table 2). Fifteen taxa were collected exclusively from *S. oleraceus*, and another two taxa were collected from *S. oleraceus* and *Sonchus asper* (Figure 3). These taxa are potential BCA (host range apparently restricted to the genus *Sonchus*, subtribe Sonchinae). Six additional species were detected only on members of the tribe Chicorieae (*Aphis craccivora* Koch, *Ophiomyia cunctata* Hendel, *Phytomyza lateralis* Fallén, *Campiglosa producta* Loew, *L. punctiventris* and *T. formosa*). We identified 38 other plant species as complementary resource

plants for the herbivore species collected from *S. oleraceus*. The generality of these herbivore species ranged from 1 to 18, with *Philaenus spumarius* L. the most polyphagous of the 47 herbivores species found on *S. oleraceus* (Figure 3).

The analysis of the subnetwork (Figure 2) also indicated that the herbivores collected on *S. oleraceus* were a resource for diverse natural enemies. In particular, 19 of the 47 herbivorous taxa collected were attacked by several species of parasitoid (12 species from the family Braconidae, 1 from Figitidae, and 1 from Ichneumonidae) and predators (6 Arachnida species, 1 from Cantharidae, 2 from Coccinnellidae, 3 from Syrphidae and 1 from Orthoptera). Moreover, among the 17 arthropods identified as candidate BCA for their restricted ecological host range, we detected interactions with natural enemies for six of them, one exhibiting interactions with 8 taxa from higher trophic levels (Figure 3).

Eventually, molecular analyses revealed particular patterns of omnivory involving several species from Heteroptera. We distinguished between intermediate omnivores (species feeding on both plants and herbivores, such as members of the Tephritidae and Aphididae), and top omnivores (species feeding on herbivores and natural enemies, such as members of the Syrphidae). The list of the taxa included in the subnetwork and of all the trophic interactions (*i.e.* the edge list) used to generate Figure 3 are provided in Supplementary Tables 3 and 4, respectively.

Assessing the contribution of the approach for the biocontrol program

The analysis of trophic interactions identified 47 taxa feeding on *S. oleraceus*, 37 of which were identified to species level. Nineteen of these 37 species had already been sampled in classical field surveys, the other 18 species being newly reported as herbivores of *S. oleraceus* in this CBC program (Table 2).

Discussion

The combination of observation and molecular data performs well for the characterization of interactions

The metabarcoding approach used made it possible to target a broad diversity of taxa (90 arthropod families, with identification to species level of 84% of the variants), as expected with the use of multiple markers (Alberdi et al., 2018; Creedy et al., 2019; Marquina et al., 2019). The combination of taxonomic assignments with subsequent observational data and information available in the literature was essential: 1) to validate the trophic links (predation and parasitism) and 2) to complement the identification in cases of failed amplification or taxonomic assignment, as advocated in other contexts (Derocles et al., 2018; Wirta et al., 2014). In the meta-network, 60% of the taxa interacted, suggesting that our methods performed very well for the reconstruction of interactions. More specifically, interaction detection rates obtained for plant-herbivores and herbivores-natural showed higher values than those usually reported in comparable contexts (Braukmann et al., 2017; Clare, 2014; Erickson et al., 2017; García-Robledo et al., 2013; Roslin and Majaneva, 2016). The high rate of interaction reported here for herbivorous arthropods can be explained by our decision to focus on intensive plant dissection and morphological determination. Retaining feeding interactions only after verification reduced the risk of false positives, over-estimating species interactions, related to the use of co-occurrence data (*i.e.* tourist insects on plants rather than actual trophic links) (Zhu et al., 2019). For arthropods, the rather low rate of natural enemies positive for preys can be multifactorial; *e.g.* mismatch between the primer pairs used and the prey species, low sequencing depth given the DNA yield ratio between consumer and prey, and degradation of DNA from consumed preys

(Hosseini et al., 2008; Macías-Hernández et al., 2018; Sheppard and Harwood, 2005; Sheppard et al., 2004).

A complementary inventory of herbivores feeding on *S. oleraceus*

The 47 taxa collected from *S. oleraceus*, covered a wide range of trophic guilds. This evidenced that the method herein employed did not bias the selection towards a particular trophic guild but allows the detection of herbivores exhibited diversified feeding habits. The selection of one or several BCA from these trophic guilds could offer a good complementarity of actions (Buccellato et al., 2019). Although the guild of flower head sucker-piercer was the richest, further dedicated experiments would be necessary to assess whether those candidate agents actually provide the best regulation action of *S. oleraceus* (Morin et al., 2009).

Several elements indicated that our approach is a good complement of the classical survey. For example, we provided a more detailed description of pollen chewers (only *Brassicogethes aeneus* (Fabricius, 1775) was recorded in the classical survey). For Diptera, the Tephritidae flies, *Tephritis cometa* Loew and *T. vespertina* Loew, are newly recorded. A phylogeny of this taxa based on the CO1 barcode showed that those species are closely related to *T. formosa* Loew (Smit et al., 2013), and we cannot, therefore, rule out possible molecular misidentification due to the short CO1 barcode used or host race differentiation, as frequently observed in this group (Diegisser et al., 2006). During the classical survey, despite intensive collections, the Tephritidae species *Campiglossa producta* has been sampled only on *S. oleraceus* from the Canary Islands. The intensive sampling efforts put in the present study on *S. oleraceus* led to the collection of rare *C. producta* specimens (only 15 specimens from two sites) in the continental bioclimatic region in France. This highlights that, despite a reduced area prospected and a limited time frame, some rare species were sampled and their ecological host range described. We acknowledge we

missed some species occurring later in the season or out of the sampled area. For example, *Cystiphora sonchi* Vallot (Diptera: Cecidomyiidae) was collected in the classical surveys and passed specificity tests (Lesieur et al., 2020), but was not sampled from *S. oleraceus* in the three bioclimatic regions from April to June. This lack of detection in our sampling campaign was expected, as rates of infestation with this species peak in summer (Rizzo and Massa, 1998). Hence, this study should be regarded as a complement of usual procedures. However, with the rapid development of molecular technologies and associated drop in price (Kennedy et al., 2020), we believe that this approach will be soon applicable at larger sampling scales.

Selection of BCA

First criterion: a restricted ecological host range

Based on the present results, 15 of the 47 herbivores feeding on *S. oleraceus* seemed to have an ecological host range restricted to *S. oleraceus*, and another two taxa appeared to be restricted to the genus *Sonchus*. All these taxa recovered from the ecological network are, thus, of particular interest as candidate BCA. However, contradictions were observed between the ecological host range described by network analysis and published findings or specificity test results (Lesieur et al., in prep). These discrepancies may be due to insufficient sampling for the recording of species interactions (as shown by accumulation curve on total interactions) or to the presence of cryptic host races or cryptic species that have yet to be deciphered. Additional studies will be required to characterize the ecological host ranges of these species further. In particular, two species, *Liriomyza sonchi* Hendel and *Ensina sonchi* L., were found associated with *S. oleraceus* and *S. asper* and were, therefore, considered to be candidate BCA because these plants are both invasive weeds in Australia (Cullen et al., 2012). However, a wider range of food resources has been reported in literature for these two species (Table 2). Conversely, the promising galling insect *T.*

393 *formosa* passed specificity tests during the traditional phase of the CBC program. It was found to
394 be restricted to the genus *Sonchus*, contrary to the results reported here, as we found *T. formosa*
395 on *Crepis vesicaria* L. (belonging to Cripidinae, the same tribe but a different subtribe to *S.*
396 *oleraceus*.). This plant was not indented to be tested as a potential food plant for *T. formosa*, and
397 these results therefore highlight the complementarity of the ecological network approach for
398 clarifying herbivore host range.

399 Moreover, in the interaction network, *C. producta* was identified on three different plant species
400 from the Chicorieae tribe. This species, found only on *S. oleraceus* in the Canary Islands during
401 the classical surveys, was considered a promising BCA for testing. The results presented here
402 indicate that its host range would not be compatible with its use as a BCA, potentially leading to
403 its exclusion from the list of candidate BCA. This example shows how the network developed
404 here is complementary to classical procedures, making it possible to narrow down the list of
405 candidate BCA to be tested.

406 The same applies to *Cheilosia latifrons* Zetterstedt, a species collected in the classical survey. In
407 our study, we did not sample this species on *S. oleraceus*, but the meta-network indicated it was
408 collected from *S. asper* and *Picris echioides* (L.), revealing its oligophagous dietary behavior.
409 However, little is known about the biology and the host plants of *C. latifrons* (Schmid and
410 Grossmann, 1996) and its taxonomy seems to be unsolved, calling into question the existence of a
411 species complex defined on the basis of host plant use (Speight, 2014). We also observed
412 discrepancies for specimens from Cynipidae that appeared to be generalist herbivores in the
413 interaction network (associated with both *S. oleraceus* and *Carduus pycnocephalus* Spreng.),
414 whereas subsequent analysis of the variants assigned to Cynipidae indicated a genetic structure
415 more consistent with multiple cryptic species potentially specializing on the host plants from

which they were collected. One species from Cynipidae is a known stem galler of *S. oleraceus*: *Aulacidea follioti* Barbotin (Bladmineerders Online database, 2020). However, this species is not yet present in any of the barcoding databases used here and could therefore only be assigned to family level. Further prospections to collect other Cynipidae specimens and rear them to adulthood would be required to confirm this identification.

Second criterion: limited interactions with natural enemies

By using metabarcoding to reconstruct interactions between arthropods, we were able to detect a wide range of parasitoids from their herbivore hosts, and some predators. We detected omnivorous dietary behavior in several groups from Heteroptera. Opportunistic predation through carnivory is common in Lygaeidae (Burdfield-Steel and Shuker, 2014) to supplement the low levels of protein supplied by plants. Carnivory has also been reported in Miridae (Wheeler, 2001) and sometimes leads to intraguild predation interactions. We found that both Syrphidae (Diptera) and Miridae (Heteroptera) fed on aphid species, but we also revealed that mirids could prey upon syrphids, as already demonstrated in arena experiments (Fréchette et al., 2006). Members of the Lygaeidae and Miridae were also found to be able to access and feed on larval stages of several Tephritidae species whilst inside the flower heads of *S. oleraceus*. This interaction does not seem to have been observed before and provides insight useful not only for the CBC program against *S. oleraceus*, but also with direct implications for other biological control programs, particularly those involving the conservation biological control of insect pests.

More specifically, among candidate BCA exhibiting a restricted ecological host range, some were associated to an important diversity of natural enemies, and should be considered of lower priority for testing (*i.e. Ensina sonchi*). We suggest that priority should be given to agent predated by a minimum parasitoid and predator species, and, when they exist, to an agent that

possesses enemies from the most distant taxonomical group from those occurring in the range of introduction (Ollivier et al., 2020). It has been shown that newly created interactions between hosts and parasitoids in the introduced range are predictable based on the realised interactions in native range (Paynter et al., 2017; Veldtman et al., 2011). Further steps in this program would consist in investigating the diversity of natural enemies occurring in the range of invasion to anticipate new potential interactions and refine BCA choice.

Conclusion

We demonstrate here the potential of network ecology for characterizing candidate BCA and their ecological interactions in the field. This characterization clearly benefited from the use of complementary approaches (morphological and molecular analyses) to identify plant/arthropod and arthropod/arthropod interactions and provided a solid framework for the establishment of an inventory of herbivores feeding on the target weed, their realized host range and interactions with natural enemies. Avenues for further investigation have been identified and in-depth studies are now required. The strength of this approach also lies in its capacity to screen field host ranges for multiple herbivore species simultaneously, without the need for as many tests as species. This potential to narrow down the list of candidate BCA for testing should help to save both time and money. Finally, in addition to the potential value of ecological network analysis to the CBC targeting Common Sowthistle in Australia, the data reported here are potentially useful for other future programs.

Authors' contributions

Conceptualization: M. Ollivier, V. Lesieur, M. S. Tixier, J.-F. Martin

Methodology: M. Ollivier, J. Tavoillot, F. Bénetière, V. Lesieur, J.-F. Martin

461 Data acquisition: M. Ollivier, J. Tavoillot, F. Bénetière, V. Lesieur, J.-F. Martin

462 Data analysis: M. Ollivier

463 Supervision: M. S. Tixier, J.-F. Martin,

464 Writing original draft: M. Ollivier

465 Writing review and editing: M. Ollivier, J. Tavoillot, V. Lesieur, M. S. Tixier, J.-F. Martin

466 Acknowledgments

467 We are particularly grateful to A Coeur d'acier, G Delvare, B Derepas, J Harran, B Michel, A
468 Migeon, E Pierre, JM Ramel and JC Streito for providing their taxonomic expertise for arthropod
469 identification for construction of the local database. We also thank G Fried for assistance with the
470 identification of some plant specimens. We also warmly thank all our coworkers involved in field
471 sampling: P Audiot, M Corbin, R Guilhot, A Loiseau, L Olazcuaga, T Thomann, and N Vieira.
472 This project is supported by funding from the Australian Government Department of Agriculture
473 and Water Resources, as part of its Rural R&D for Profit programme, through AgriFutures
474 Australia (Rural Industries Research and Development Corporation) (PRJ--010527).

475

Literature cited

Adkins, S.W., Wills, D., Boersma, M., Walker, S.R., Robinson, G., Mcleod, R.J., and Einam, J.P. (1997). Weeds resistant to chlorsulfuron and atrazine from the north-east grain region of Australia. Weed Research *37*, 343–349.

Alberdi, A., Aizpurua, O., Gilbert, M.T.P., and Bohmann, K. (2018). Scrutinizing key steps for reliable metabarcoding of environmental samples. Methods in Ecology and Evolution *9*, 134–147.

Aphids on the wolrd's plants Database (2020). Aphids on the wolrd's plants Database - An online identification and information guide.

Barratt, B.I.P., Moran, V.C., Bigler, F., and van Lenteren, J.C. (2018). The status of biological control and recommendations for improving uptake for the future. BioControl *63*, 155–167.

Bellard, C., Genovesi, P., and Craine, J.M. (2016). Global patterns in threats to vertebrates by biological invasions. Proceedings of the Royal Society B: Biological Sciences *283*, 20152454.

Benavent-Corai, J. (2005). Catalogue of the Hosts-plants of the World Agromyzidae (Diptera) (Università degli Studi di Milano).

Bersier, L.-F., Banašek-Richter, C., and Cattin, M.-F. (2002). Quantitative Descriptors of Food-Web Matrices. Ecology *83*, 2394–2407.

Bladmineerders Online database (2020). Plant Parasites of Europe – leafminers, galls and fungi.

Bolyen, E., Rideout, J.R., Dillon, M.R., Bokulich, N.A., Abnet, C., Al-Ghalith, G.A., Alexander, H., Alm, E.J., Arumugam, M., Asnicar, F., et al. (2018). QIIME 2: Reproducible, interactive, scalable, and extensible microbiome data science (PeerJ Inc.).

Boulos, L. (1974). Revision systematique du genre Sonchus L. s.l. VI. Sous-genre. 3. origosonchus. genres Embergeria, Babcockia et -Taeckholmia. species exclusae et dubiae. index. Botaniska notiser.

Brandon-Mong, G.-J., Gan, H.-M., Sing, K.-W., Lee, P.-S., Lim, P.-E., and Wilson, J.-J. (2015). DNA metabarcoding of insects and allies: an evaluation of primers and pipelines. Bull. Entomol. Res. *105*, 717–727.

Braukmann, T.W.A., Kuzmina, M.L., Sills, J., Zakharov, E.V., and Hebert, P.D.N. (2017). Testing the Efficacy of DNA Barcodes for Identifying the Vascular Plants of Canada. PLoS One *12*.

Briese, D.T. (2005). Translating host-specificity test results into the real world: The need to harmonize the yin and yang of current testing procedures. Biological Control *35*, 208–214.

Buccellato, L., Byrne, M.J., Fisher, J.T., and Witkowski, E.T.F. (2019). Post-release evaluation of a combination of biocontrol agents on Crofton weed: testing extrapolation of greenhouse results to field conditions.

509 Burdfield-Steel, E.R., and Shuker, D.M. (2014). The evolutionary ecology of the Lygaeidae. Ecol
510 Evol *4*, 2278–2301.

511 Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016).
512 DADA2: High resolution sample inference from Illumina amplicon data. Nat Methods *13*, 581–
513 583.

514 Carvalheiro, L.G., Buckley, Y.M., Ventim, R., Fowler, S.V., and Memmott, J. (2008). Apparent
515 competition can compromise the safety of highly specific biocontrol agents. Ecology Letters *11*,
516 690–700.

517 Ceglar, A., Zampieri, M., Toreti, A., and Dentener, F. (2019). Observed Northward Migration of
518 Agro-Climate Zones in Europe Will Further Accelerate Under Climate Change. Earth's Future *7*,
519 1088–1101.

520 Centre de Biologie pour la Gestion des Population (2019). CBGP - Continental Arthropod
521 Collection (Portail Data INRAE).

522 Chao, A. (1984). Nonparametric Estimation of the Number of Classes in a Population.
523 Scandinavian Journal of Statistics *11*, 265–270.

524 Chase, M.W., Christenhusz, M.J.M., Fay, M.F., Byng, J.W., Judd, W.S., Soltis, D.E., Mabberley,
525 D.J., Sennikov, A.N., Soltis, P.S., and Stevens, P.F. (2016). An update of the Angiosperm
526 Phylogeny Group classification for the orders and families of flowering plants: APG IV. Bot J Linn
527 Soc *181*, 1–20.

528 Clare, E.L. (2014). Molecular detection of trophic interactions: emerging trends, distinct
529 advantages, significant considerations and conservation applications. Evolutionary Applications *7*,
530 1144–1157.

531 Creedy, T.J., Ng, W.S., and Vogler, A.P. (2019). Toward accurate species-level metabarcoding of
532 arthropod communities from the tropical forest canopy. Ecology and Evolution *9*, 3105–3116.

533 Cruaud, A., Gonzalez, A.-A., Godefroid, M., Nidelet, S., Streito, J.-C., Thuillier, J.-M., Rossi, J.-
534 P., Santoni, S., and Rasplus, J.-Y. (2018). Using insects to detect, monitor and predict the
535 distribution of Xylella fastidiosa : a case study in Corsica. Sci Rep *8*, 1–13.

536 Cullen, J., Julien, M., and McFadyen, R. (2012). Biological Control of Weeds in Australia.
537 (Collingwood: CSIRO Publishing).

538 Derocles, S.A.P., Bohan, D.A., Dumbrell, A.J., Kitson, J.J.N., Massol, F., Pauvert, C.,
539 Plantegenest, M., Vacher, C., and Evans, D.M. (2018). Chapter One - Biomonitoring for the 21st
540 Century: Integrating Next-Generation Sequencing Into Ecological Network Analysis. In Advances
541 in Ecological Research, D.A. Bohan, A.J. Dumbrell, G. Woodward, and M. Jackson, eds.
542 (Academic Press), pp. 1–62.

543 Diegisser, T., Seitz, A., and Johannesen, J. (2006). Phylogeographic patterns of host-race evolution
544 in Tephritis conura (Diptera: Tephritidae). Molecular Ecology *15*, 681–694.

Diputacio Barcelona (2019). Trobada d'Estudiosos de Sant Llorenç del Munt i l'Obac - Comunicacions presentades el 21 i 22 de novembre de 2017 a Castellar del Vallès.

Erickson, D.L., Reed, E., Ramachandran, P., Bourg, N.A., McShea, W.J., and Ottesen, A. (2017). Reconstructing a herbivore's diet using a novel rbcL DNA mini-barcode for plants. AoB PLANTS *9*.

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol. Marine Biol. Biotechnol. *3*, 294–299.

Fréchette, B., Rojo, S., Alomar, O., and Lucas, É. (2006). Intraguild predation between syrphids and mirids: who is the prey? Who is the predator? BioControl *52*, 175.

Frei, B., Guenay, Y., Bohan, D.A., Traugott, M., and Wallinger, C. (2019). Molecular analysis indicates high levels of carabid weed seed consumption in cereal fields across Central Europe. J Pest Sci *92*, 935–942.

Galan, M., Pons, J.-B., Tournayre, O., Pierre, É., Leuchtmann, M., Pontier, D., and Charbonnel, N. (2017). Metabarcoding for the parallel identification of several hundred predators and their prey: Application to bat species diet analysis. Molecular Ecology Resources *18*, 474–489.

García-Robledo, C., Erickson, D.L., Staines, C.L., Erwin, T.L., and Kress, W.J. (2013). Tropical plant-herbivore networks: reconstructing species interactions using DNA barcodes. PLoS One *8*, e52967.

Gariepy, T.D., Haye, T., and Zhang, J. (2014). A molecular diagnostic tool for the preliminary assessment of host-parasitoid associations in biological control programmes for a new invasive pest. Mol. Ecol. *23*, 3912–3924.

Groenteman, R., Fowler, S.V., and Sullivan, J.J. (2011). St. John's wort beetles would not have been introduced to New Zealand now: A retrospective host range test of New Zealand's most successful weed biocontrol agents. Biological Control *57*, 50–58.

Haye, T., Achterberg, C. van, Goulet, H., Barratt, B.I.P., and Kuhlmann, U. (2006). Potential for classical biological control of the potato bug Closterotomus norwegicus (Hemiptera: Miridae): description, parasitism and host specificity of Peristenus closterotomae sp. n. (Hymenoptera: Braconidae). Bulletin of Entomological Research *96*, 421–431.

Hinz, H.L., Winston, R.L., and Schwarzländer, M. (2019). How Safe Is Weed Biological Control? A Global Review of Direct Nontarget Attack. The Quarterly Review of Biology *94*, 1–27.

Hinz, H.L., Winston, R.L., and Schwarzländer, M. (2020). A global review of target impact and direct nontarget effects of classical weed biological control. Current Opinion in Insect Science *38*, 48–54.

579 Hosseini, R., Schmidt, O., and Keller, M.A. (2008). Factors affecting detectability of prey DNA in
580 the gut contents of invertebrate predators: a polymerase chain reaction-based method. Entomologia
581 Experimentalis et Applicata *126*, 194–202.

582 Hrček, J., and Godfray, H.C.J. (2015). What do molecular methods bring to host–parasitoid food
583 webs? Trends in Parasitology *31*, 30–35.

584 Hutchinson, I., Colosi, J., and Lewin, R.A. (1984). THE BIOLOGY OF CANADIAN WEEDS.:
585 63. Sonchus asper (L.) Hill and S. oleraceus L. Can. J. Plant Sci. *64*, 731–744.

586 Jalaludin, A., Widderick, M.J., Broster, J., and Walsh, M.J. (2018). Glyphosate and 2,4-D amine
587 resistance in common sowthistle (Sonchus oleraceus) and fleabane (Conyza bonariensis) in the
588 northern grain growing region of Australia. p.

589 Jordano, P. (2016). Sampling networks of ecological interactions. Functional Ecology *30*, 1883–
590 1893.

591 Keane, R.M., and Crawley, M.J. (2002). Exotic plant invasions and the enemy release hypothesis.
592 Trends in Ecology & Evolution *17*, 164–170.

593 Kennedy, S.R., Prost, S., Overcast, I., Rominger, A.J., Gillespie, R.G., and Krehenwinkel, H.
594 (2020). High-throughput sequencing for community analysis: the promise of DNA barcoding to
595 uncover diversity, relatedness, abundances and interactions in spider communities. Dev Genes Evol
596 *230*, 185–201.

597 Landolt, P.J., Worth, R.A., and Zack, R.S. (2010). First Report of *Hecatera dysodea* (Denis and
598 Schiffermüller) (Noctuidae) in the Pacific Northwest of the United States. Journal of the
599 Lepidopterists' Society *64*, 192–196.

600 Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V., Boehm, J.T., and
601 Machida, R.J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial
602 COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut
603 contents. Frontiers in Zoology *10*, 34.

604 Lesieur, V., Thomann, T., Ollivier, M., and Raghu, S. (2020). Making host specificity testing more
605 efficient: Exploring the use of abridged test plant lists. Journal of Applied Entomology *144*, 546–
606 551.

607 Lesieur et al. in prep. Feasibility of classical biological control of Sonchus oleraceus in Australia,
608 Biocontrol Science and Technology.

609 Lewinsohn, T.M., Prado, P.I., Jordano, P., Bascompte, J., and Olesen, J.M. (2006). Structure in
610 plant–animal interaction assemblages. Oikos *113*, 174–184.

611 Llewellyn, R., Ronning, D., Ouzman, J., Walker, S., Mayfield, A., and Clarke, M. (2016). Impact
612 of weeds on Australian grain production : the cost of weeds to Australian grain growers and the
613 adoption of weed management and tillage practices (GRDC).

López-Núñez, F.A., Heleno, R.H., Ribeiro, S., Marchante, H., and Marchante, E. (2017). Four-trophic level food webs reveal the cascading impacts of an invasive plant targeted for biocontrol. Ecology *98*, 782–793.

Louda, S.M., Kendall, D., Connor, J., and Simberloff, D. (1997). Ecological Effects of an Insect Introduced for the Biological Control of Weeds. Science *277*, 1088–1090.

Louda, S.M., Pemberton, R.W., Johnson, M.T., and Follett, P.A. (2003). Nontarget effects--the Achilles' heel of biological control? Retrospective analyses to reduce risk associated with biocontrol introductions. Annu. Rev. Entomol. *48*, 365–396.

Macías-Hernández, N., Athey, K., Tonzo, V., Wangensteen, O.S., Arnedo, M., and Harwood, J.D. (2018). Molecular gut content analysis of different spider body parts. PLoS One *13*, e0196589.

Marquina, D., Andersson, A.F., and Ronquist, F. (2019). New mitochondrial primers for metabarcoding of insects, designed and evaluated using in silico methods. Molecular Ecology Resources *19*, 90–104.

Martin, J.-F. (2019). error-proof_indexes v1.0 (Zenodo). http://doi.org/10.5281/zenodo.3350207

McFadyen, R.E. (1998). Biological control of weeds. Annu. Rev. Entomol. *43*, 369–393.

Memmott, J. (2000). Food webs as a tool for studying nontarget effects in biological control. In Nontarget Effects of Biological Control, P.A. Follett, and J.J. Duan, eds. (Boston, MA: Springer US), pp. 147–163.

Memmott, J. (2009). Food webs: a ladder for picking strawberries or a practical tool for practical problems? Philos Trans R Soc Lond B Biol Sci *364*, 1693–1699.

Meulen, A.W. van der, Widderick, M., Cook, T., Chauhan, B.S., and Bell, K. (2016). Survey of glyphosate resistance in common sowthistle (Sonchus oleraceus) across the Australian Northern Grains Region. In 20th Australasian Weeds Conference, (Perth, Western Australia), p. 108.

Morin, L., Reid, A.M., Sims-Chilton, N.M., Buckley, Y.M., Dhileepan, K., Hastwell, G.T., Nordblom, T.L., and Raghu, S. (2009). Review of approaches to evaluate the effectiveness of weed biological control agents. Biological Control *51*, 1–15.

Müller-Schärer, H., and Schaffner, U. (2008). Classical biological control: exploiting enemy escape to manage plant invasions. Biol Invasions *10*, 859–874.

Ollivier, M., Lesieur, V., Raghu, S., and Martin, J.-F. (2020). Characterizing ecological interaction networks to support risk assessment in classical biological control of weeds. Current Opinion in Insect Science *38*, 40–47.

Oksanen J., Blanchet F. G., Friendly M., Kindt R., Legendre P., McGlinn D., Minchin P R., O'Hara R. B., Simpson G. L., Solymos P., Stevens M. H. H., Szoecs E. and Wagner H. (2019). vegan: Community Ecology Package. R package version 2.5-6. https://CRAN.R-project.org/package=vegan

649 Paynter, Q., Fowler, S.V., and Groenteman, R. (2017). Making weed biological control predictable,
650 safer and more effective: perspectives from New Zealand. BioControl 1–10.

651 Pearson, D.E., and Callaway, R.M. (2003). Indirect effects of host-specific biological control
652 agents. Trends in Ecology & Evolution *18*, 456–461.

653 Peterson, P.G., Merrett, M.F., Fowler, S.V., Barrett, D.P., and Paynter, Q. (2020). Comparing
654 biocontrol and herbicide for managing an invasive non-native plant species: Efficacy, non-target
655 effects and secondary invasion. Journal of Applied Ecology *57*, 1876–1884.

656 Pyšek, P., Pergl, J., Essl, F., Lenzner, B., Dawson, W., Kreft, H., Weigelt, P., Winter, M., Kartesz,
657 J., Nishino, M., et al. (2017). Naturalized alien flora of the world: species diversity, taxonomic and
658 phylogenetic patterns, geographic distribution and global hotspots of plant invasion. Preslia *89*,
659 203–274.

660 Ratnasingham, S., and Hebert, P.D.N. (2007). bold: The Barcode of Life Data System
661 (http://www.barcodinglife.org). Mol Ecol Notes *7*, 355–364.

662 Rizzo, M.C., and Massa, B. (1998). On two Italian gall midges (Diptera Cecidomyiidae) and their
663 parasitoids. Entomologica *32*, 121–131.

664 Roslin, T., and Majaneva, S. (2016). The use of DNA barcodes in food web construction-terrestrial
665 and aquatic ecologists unite! Genome *59*, 603–628.

666 Schmid, U., and Grossmann, A. (1996). Eiablage von Cheilosia latifrons (Zetterstedt,
667 1843)(Díptera, Syrphidae) an Leontodón autumnalis L. Volucella *2*, 86–87.

668 Schoonhoven, L.M., Jermy, T., and Loon, J.J.A. van (1998). Insect-plant biology: from physiology
669 to evolution. (London: Chapman and Hall).

670 Seebens, H., Essl, F., Dawson, W., Fuentes, N., Moser, D., Pergl, J., Pyšek, P., Kleunen, M. van,
671 Weber, E., Winter, M., et al. (2015). Global trade will accelerate plant invasions in emerging
672 economies under climate change. Global Change Biology *21*, 4128–4140.

673 Sheppard, S.K., and Harwood, J.D. (2005). Advances in molecular ecology: tracking trophic links
674 through predator–prey food-webs. Functional Ecology *19*, 751–762.

675 Sheppard, S.K., Henneman, M.L., Memmott, J., and Symondson, W.O.C. (2004). Infiltration by
676 alien predators into invertebrate food webs in Hawaii: a molecular approach. Mol Ecol *13*, 2077–
677 2088.

678 Smit, J., Reijnen, B., and Stokvis, F. (2013). Half of the European fruit fly species barcoded
679 (Diptera, Tephritidae); a feasibility test for molecular identification. Zookeys 279–305.

680 Speight, M.C.D. (2014). Species accounts of European Syrphidae (Diptera), 2014. 319.

681 Suckling, D.M., and Sforza, R.F.H. (2014). What Magnitude Are Observed Non-Target Impacts
682 from Weed Biocontrol? PLOS ONE *9*, e84847.

683 Tay, E.B. (1972). Population Ecology of Cicadella Viridis (L.) and Bionomics of Graphocephala
684 Coccinea (Forster)(Homoptera, Cicadellidae). (University of London (Imperial College of Science
685 and Technology)).

686 Thébault, E., and Fontaine, C. (2010). Stability of Ecological Communities and the Architecture
687 of Mutualistic and Trophic Networks. Science *329*, 853–856.

688 Todd, J.H., Pearce, B.M., and Barratt, B.I.P. (2020). Using qualitative food webs to predict species
689 at risk of indirect effects from a proposed biological control agent. BioControl.

690 Van Driesche, R.G., Carruthers, R.I., Center, T., Hoddle, M.S., Hough-Goldstein, J., Morin, L.,
691 Smith, L., Wagner, D.L., Blossey, B., Brancatini, V., et al. (2010). Classical biological control for
692 the protection of natural ecosystems. Biological Control *54*, S2–S33.

693 Vázquez Albalate, X. (2002). European fauna of Oedemeridae: Coleoptera (Barcelona: Argania
694 Ed).

695 Veldtman, R., Lado, T.F., Botes, A., Procheş, Ş., Timm, A.E., Geertsema, H., and Chown, S.L.
696 (2011). Creating novel food webs on introduced Australian acacias: indirect effects of galling
697 biological control agents. Diversity and Distributions *17*, 958–967.

698 Vitousek, P.M., Mooney, H.A., Lubchenco, J., and Melillo, J.M. (1997). Human Domination of
699 Earth's Ecosystems. Science *277*, 494–499.

700 Wang, Q. (2017). Cerambycidae of the World: Biology and Pest Management (CRC Press).

701 Wapshere, A.J., Delfosse, E.S., and Cullen, J.M. (1989). Recent developments in biological control
702 of weeds. Crop Protection *8*, 227–250.

703 Warren, P.H. (1994). Making connections in food webs. Trends in Ecology & Evolution *9*, 136–
704 141.

705 Wheeler, A.G. (2001). Biology of the plant bugs (Hemiptera: Miridae): pests, predators,
706 opportunists. Biology of the Plant Bugs (Hemiptera: Miridae): Pests, Predators, Opportunists.

707 White, I.M. (1988). Tephritid flies: Diptera: Tephritidae. Tephritid Flies: Diptera: Tephritidae. *10*.

708 Widderick, M.J., Walker, S.R., Sindel, B.M., and Bell, K.L. (2010). Germination, emergence, and
709 persistence of Sonchus oleraceus, a major crop weed in subtropical Australia. Weed Biology and
710 Management *10*, 102–112.

711 Willis, A.J., and Memmott, J. (2005). The potential for indirect effects between a weed, one of its
712 biocontrol agents and native herbivores: A food web approach. Biological Control *35*, 299–306.

713 Wirta, H.K., Hebert, P.D.N., Kaartinen, R., Prosser, S.W., Várkonyi, G., and Roslin, T. (2014).
714 Complementary molecular information changes our perception of food web structure. PNAS *111*,
715 1885–1890.

716 Zhu, C., Gravel, D., and He, F. (2019). Seeing is believing? Comparing plant–herbivore networks
717 constructed by field co-occurrence and DNA barcoding methods for gaining insights into network
718 structures. Ecology and Evolution *9*, 1764–1776.

719

Tables

Table 1: Global description of the meta-network and subnetwork centred on S. oleraceus. *Metrics measured are the total number of nodes (in brackets is the number of nodes connected in the network), the number of links, the linkage density and the connectance of the network. For each network is also presented the number of species per trophic level (in brackets is the number of species in interaction with another species). Omnivorous species are regarded as natural enemies.*

		Meta-network	Subnetwork*
Global metrics	Nb of nodes (connected)	401 (241)	116
	Nb of links	350	213
	Linkage density	1,45	1,84
	Connectance	0,006	0,008
Taxon assemblage	Nb of plant species (in interaction)	132 (60)	39
	Nb of herbivore species (in interaction)	185 (136)	47
	Nb of natural enemies species (in interaction)	79 (35)	30

***** All nodes are connected

728 Table 2: Herbivores from *Sonchus oleraceus* (SO) collected and identified through intense field sampling in three bioclimatic regions
729 (Semi-oceanic, Mediterranean and Continental) in France in spring 2018. The field host range of these herbivores was defined by
730 network analysis (S - SO: Specific to SO, S – S: Specific to the genus Sonchus, S – C: Specific to the tribe Cichorieae, G: Generalist,
731 ?: Unknown host range). The column 3-years surveys refers to the species collected following classical procedures of biocontrol
732 program (see Materials and Methods section).

733

Order / Family	***Species***	**Trophic guild**	**3-years survey**	**Field host range**	**Literature host range (Reference)**
Coleoptera					
Cerambicydae	*Agapanthia cardui* (Linnaeus, 1767)	Stem boring	No	G	
Cerambicydae	*Agapanthia pannonica* Kratochvil, 1985	Stem boring	No	S - SO	Polyphagous (Wang, 2017)
Curculionidae	*sp.*	Crown/stem boring	_	G	
Curculionidae	*Lixus punctiventris* (Boheman, 1835)	Stem boring	Yes	S - C	
Dasytidae	*Dasytes tristiculus* Mulsant & Rey, 1868	Flower/pollen chewing	No	S - SO	Polyphagous (Diputacio Barcelona, 2019)
Dasytidae	*Psilothrix viridicoeruleus* (Geoffroy, 1785)	Flower/pollen chewing	No	G	
Oedemeridae	*Oedemera crassipes* Ganglbauer, 1881)	Flower/pollen chewing	No	S - SO	Polyphagous (Vázquez Albalate, 2002)
Oedemeridae	*Oedemera flavipes* (Fabricius, 1792)	Flower/pollen chewing	No	G	
Nitidulidae	*Brassicogethes aeneus* (Fabricius, 1775)	Flower/pollen chewing	Yes	G	
Diptera					
Agromyzidae	*sp.*	Leaf mining	_	G	
Agromyzidae	*Liriomyza sonchi* Hendel, 1931	Leaf mining	Yes	S - S	Oligophagous (Benavent-Corai, 2005)
Agromyzidae	*Ophiomyia cunctata* (Hendel, 1920)	Leaf mining	Yes	S - C	
Agromyzidae	*Phytomyza horticola* Goureau, 1851	Leaf mining	Yes	G	
Agromyzidae	*Phytomyza lateralis* Fallén, 1823	Leaf mining	Yes	S - C	
Tephritidae	*sp.*	Flower bud galling/seed feeding	_	S - SO	Oligophagous (White, 1988)
Tephritidae	*Campiglossa producta* (Loew, 1844)	Flower bud galling/seed feeding	Yes	S - C	
Tephritidae	*Ensina sonchi* (Linnaeus, 1767)	Flower bud galling/seed feeding	Yes	S - S	Oligophagous (White, 1988)
Tephritidae	*Tephritis sp.*	Flower bud galling/seed feeding	_	G	
Tephritidae	*Tephritis cometa* (Loew, 1840)	Flower bud galling/seed feeding	No	S - SO	Oligophagous (Bladmineerders Online database, 2020)
Tephritidae	*Tephritis formosa* (Loew, 1844)	Flower bud galling/seed feeding	Yes	S - C	

Tephritidae	*Tephritis vespertina* (Loew, 1844)	Flower bud galling/seed feeding	No	S - SO	Oligophagous (Bladmineerders Online database, 2020)
Hemiptera					
Aphididae	*sp.*	Systemic sucking piercing	_	S - SO	Polyphagous (Aphids on the wolrd's plants Database, 2020)
Aphididae	*Aphis craccivora* Koch, 1854	Systemic sucking piercing	No	S - C	
Aphididae	*Aphis fabae* Scopoli, 1763	Systemic sucking piercing	No	G	
Aphididae	*Hyalopterus pruni* (Geoffroy, 1762)	Systemic sucking piercing	No	S - SO	Oligophagous (Aphids on the wolrd's plants Database, 2020)
Aphididae	*Hyperomyzus lactucae* (Linnaeus, 1758)	Systemic sucking piercing	Yes	G	
Aphididae	*Macrosiphum rosae* (Linnaeus, 1758)	Systemic sucking piercing	No	S - SO	Polyphagous (Aphids on the wolrd's plants Database, 2020)
Aphididae	*Uroleucon sonchi* (Linnaeus, 1767)	Systemic sucking piercing	Yes	G	
Anthocoridae	*Orius sp.*	Flower head sucking piercing	No	G	
Aphrophoridae	*Philaenus sp.*	Leaf sucking piercing	No	G	
Aphrophoridae	*Philaenus spumarius* (Linnaeus, 1758)	Leaf sucking piercing	Yes	G	
Cicadellidae	*Cicadella viridis* (Linnaeus, 1758)	Leaf sucking piercing	No	S - SO	Polyphagous (Tay, 1972)
Coreidae	*Coreus marginatus* (Linnaeus, 1758)	Flower head sucking piercing	Yes	G	
Lygaeidae	*Lygaeus equestris* (Linnaeus, 1758)	Flower head sucking piercing	Yes	G	
Miridae	*sp.*	Flower head sucking piercing	_	G	
Miridae	*Closterotomus norvegicus* (Gmelin, 1790)	Flower head sucking piercing	No	S - SO	Polyphagous (Haye et al., 2006)
Miridae	*Lepidargyrus ancorifer* (Fieber, 1858)	Flower head sucking piercing	No	G	
Orsillidae	*Nysius cymoides* (Spinola, 1837)	Flower head sucking piercing	Yes	G	
Pentatomidae	*Dolycoris baccarum* (Linnaeus, 1758)	Flower head sucking piercing	Yes	G	
Pentatomidae	*Nezara viridula* (Linnaeus, 1758)	Flower head sucking piercing	No	G	
Rhopalidae	*Liorhyssus hyalinus* (Fabricius, 1794)	Flower head sucking piercing	Yes	G	
Rhopalidae	*Stictopleurus punctatonervosus* (Goeze, 1778)	Flower head sucking piercing	Yes	G	
Hymenoptera					
Cynipidae	*sp.*	Stem boring	Yes	G	
Tenthredinidae	*Cephaledo bifasciata* (Müller, 1766)	Leaf chewing	No	S - SO	
Lepidoptera					
Noctuidae	*Hecatera dysodea* (Denis & Schiffermüller, 1775)	Flower bud chewing	Yes	S - SO	Oliphagous (Landolt et al., 2010)
Tortricidae	*sp.*	Leaf mining/chewing	_	S - SO	Polyphagous (Bladmineerders Online database, 2020)
Tortricidae	*Cnephasia stephensiana* (Doubleday, 1849)	Leaf mining/chewing	No	S - SO	Polyphagous (Bladmineerders Online database, 2020)

734

735 # Figures

736

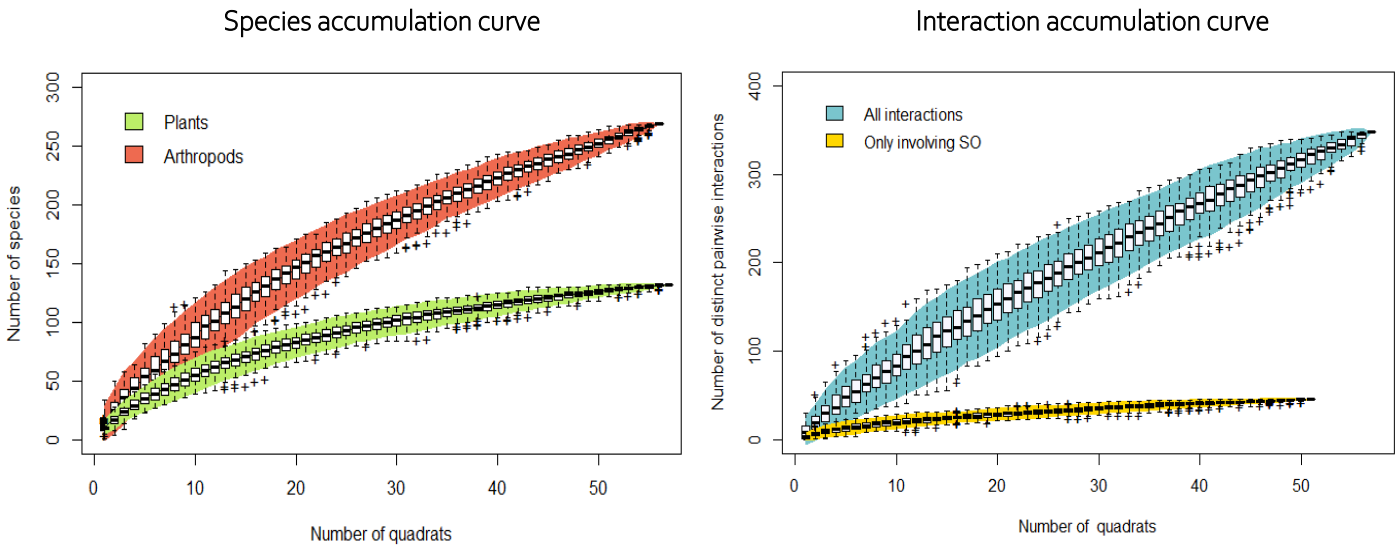

Figure 1: Accumulation curves representing a) species richness in plants and arthropods and b) pairwise interactions from the meta-network and focusing on interactions involving S. oleraceus. *Curves were constructed with 1000 random resampling events over the 57 quadrats analyzed along the sampling campaign (Spring 2018).*

737

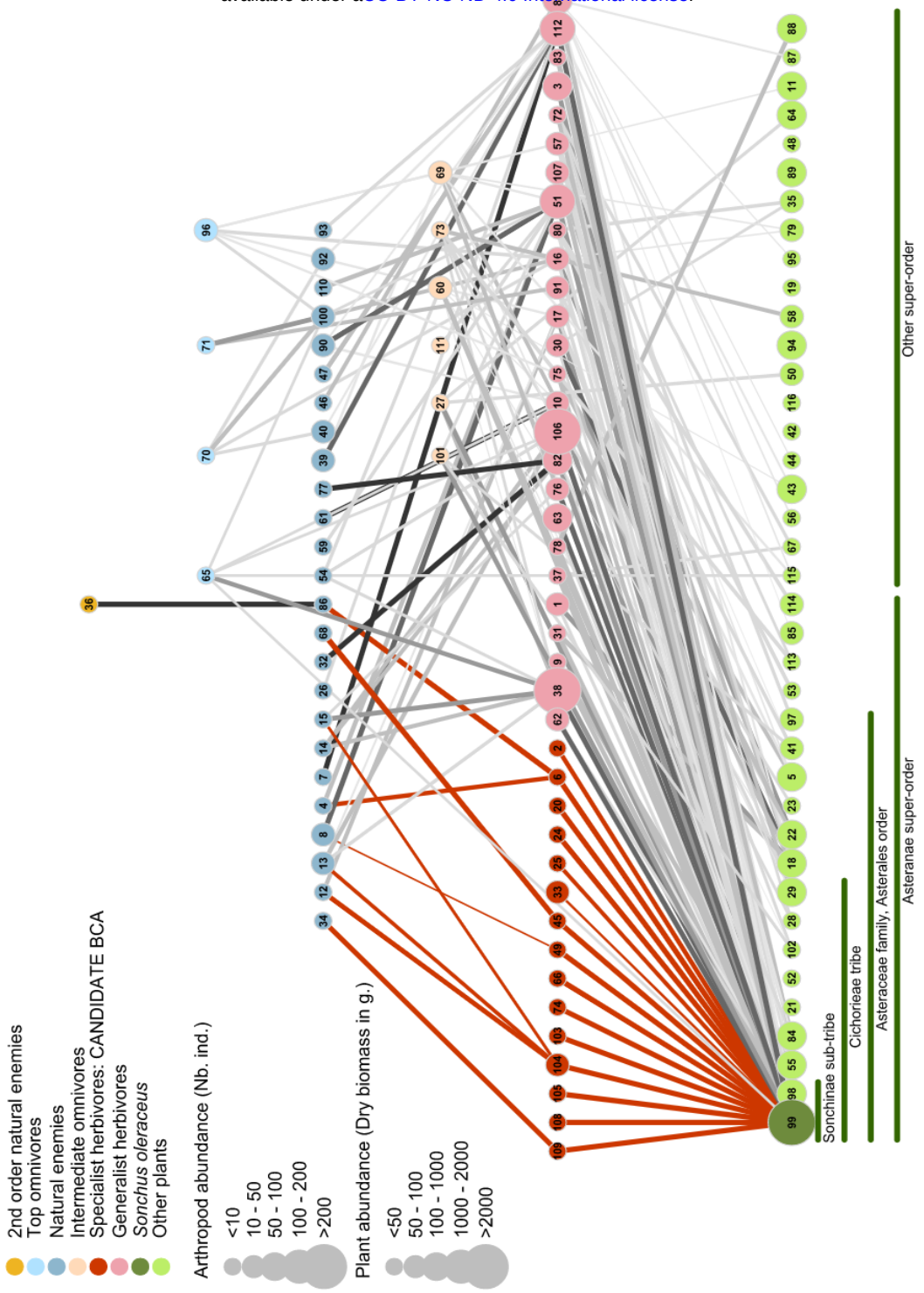

Figure 2: Multitrophic network reconstructed from S. oleraceus *(dark green node) and 38 other plants (light green nodes) used by* S. oleraceus *herbivores (red nodes represent putative specialist herbivores, and pink nodes correspond to herbivores feeding on other species in addition to* S. oleraceus*). Plants are ordered by their phylogenetic relatedness to* S. oleraceus*. Natural enemies of herbivore species are represented by dark blue nodes. Nodes at intermediate levels (beige and light blue) correspond to omnivorous species identified by molecular analyses. The width of edges reflects frequencies of interactions between pairs of species and edges coloured in red emphasize interactions involving potential candidate BCA. The list of taxa corresponding to each node and the edge list are provided in Supplementary Tables 3 and 4, respectively.*

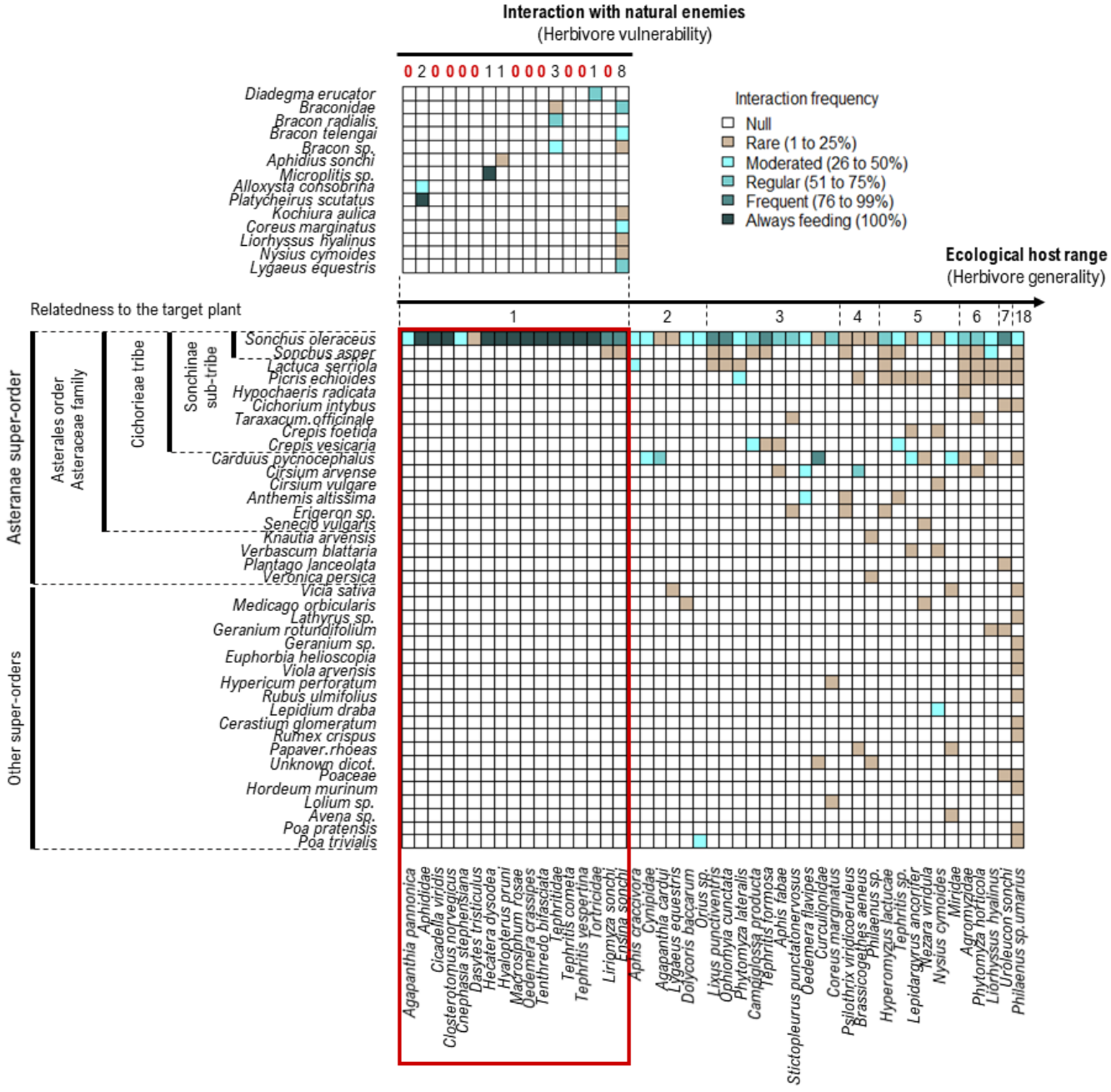

Figure 3: Interaction matrix between herbivores sampled on Sonchus oleraceus *and their resource plants, indicating the ecological host range of the herbivores, as defined by intense field sampling in France, during spring 2018. Plants are ordered by their phylogenetic relatedness to* S. oleraceus *and arthropods are ordered by increasing generality values (i.e. the number of resources per species). Red rectangle highlights the 17 species considered as candidate BCA for their restricted ecological host range. The second level matrix diplays interactions between candidate BCA and their natural enemies. Interaction are represented as semi-quantitative information, via occurences frequencies of interactions culculated for each species pair.*