



The hidden 'plant side' of insect novel foods: A DNA-based assessment

Jessica Frigerio^{a,1}, Giulia Agostinetto^{b,1}, Anna Sandionigi^b, Valerio Mezzasalma^a,
Nadia Maria Berterame^b, Maurizio Casiraghi^b, Massimo Labra^b, Andrea Galimberti^{b,*}

^a FEM2-Ambiente, Piazza della Scienza 2, I-20126 Milano, Italy

^b Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, I-20126 Milano, Italy

ARTICLE INFO

Keywords:

DNA barcoding
DNA metabarcoding
High throughput sequencing
Insect flour
Plant allergens
Processed food

ABSTRACT

In the context of novel foods, a category for which the market demand is increasing worldwide, the consumption of edible insects and related insect-based products is expected to grow in the next years. Insects represent an important source of energy for the human diet but there is a lack of scientific knowledge about their processing to ensure safe food items to the consumer. In this study we adopted a combined DNA-based approach to verify the identity of the declared species in five categories of commercial insect-based products (mt *COI* DNA barcoding) and to characterize plant declared ingredients or contaminants (nu ITS2 DNA metabarcoding) with particular attention to putative elements of allergenic concern belonging, for example to the insect rearing substrate. Moreover, the same approach has been used to assess its sensitivity to cases of contamination and counterfeits to insect flour with low cost (and potentially allergenic) vegetable flours like wheat and soybean. Results show the success of insect DNA barcoding authentication even for highly processed products. Furthermore, the DNA metabarcoding analysis revealed a high efficacy as a screening method to identify both plant ingredients and vegetal traces belonging to insect farming or possible adulteration events, also acting as an early warning strategy for the occurrence of allergens of human concern. This approach could support the development of new risk assessment procedures for novel foods by regulatory authorities to ensure their quality, safety, and acceptance which will become more required in order to face the challenge of feeding the world population in the next decades.

1. Introduction

Traditionally, edible insects are consumed in large parts of the world like Africa and Asia. In the last few years, they have increased in popularity as trendy foods in many Western countries (Sun-Waterhouse et al., 2016). Being rich in essential nutrients, they represent an important source of energy for human diets (Rumpold & Schlüter, 2013). Mean estimates show that the energy level of insects is around 400–500 kcal per 100 g of dry matter, making it comparable with other protein sources (Payne, Scarborough, Rayner, & Nonaka, 2016). Protein is probably the most significant component of edible insects, with an average value ranging from 30% to 65% of the total dry matter. However, edible insects are also rich in micronutrients such as iron, zinc, and calcium (Dobermann, Swift, & Field, 2017), and preliminary studies have shown that insect farming has a lower environmental footprint compared to other livestock animals (Oonincx & de Boer, 2012). Moreover, a recent study demonstrated that edible insects could represent a potential source of antioxidants with positive effects on

human health (Di Mattia, Battista, Sacchetti, & Serafini, 2019).

In very recent years, the use of insects in food fortification is emerging as a means of producing nutritious and acceptable food products for human consumption (Myers & Pettigrew, 2018). In many countries a lot of insect-based products have recently become commercialized (e.g. pasta, biscuits, and energy bars based on insect flour in combination with fruits, nuts, and other ingredients), and their consumption trend is expected to grow steadily.

In Europe, edible insects are placed in the category of novel foods, and since the beginning of 2018, the Regulation (EU) 2015/2283 entered into force in attempt to regulate the production and safety of novel foods in Europe. This regulation establishes the requirements that enable Food Business Operators to bring new foods into the EU market, while ensuring high levels of food safety to the consumers. However, concerning the possible risks for human health caused by insect-based products, the European Food Safety Authority has published an initial assessment (EFSA, 2015), and it concluded that attention should be placed to the possible occurrence of biological and chemical hazards

* Corresponding author at: Department of Biotechnology and Biosciences, University of Milano – Bicocca, Piazza della Scienza 2, I-20126 Milano, Italy.
E-mail address: andrea.galimberti@unimib.it (A. Galimberti).

¹ Both authors contributed equally to this work.

caused by novel foods.

Insects ingredients may cause allergic symptoms (Mazzucchelli et al., 2018; Pali-Schöll et al., 2019; Van der Fels-Klerx, Adamse, Punt, & Van Asselt, 2018), and there is a cross-reactivity/cosensitisation between edible insects and crustaceans (Ribeiro, Cunha, Sousa-Pinto, & Fonseca, 2018). Moreover, at the microbiological level, the microbiota of insects is highly complex and, apart from the body surface and the mouthparts, the maximum microorganism diversity is in the mid-gut with poor or no data about its effect on consumer health (Schlüter et al., 2017; Walia, Kapoor, & Farber, 2018).

Another element that can negatively impact the safety of insect-based novel foods is the composition and quality of feed used to raise the insects and their rearing conditions. In many cases, these feeds are composed of vegetables and very few ingredients of animal origin, such as fishmeal and egg and milk based-products (EU No. 2017/893). However, to date no analytical systems are available to control the diet of marketed insects.

In general, according to a recent review (Schlüter et al., 2017), there is a lack of scientific knowledge about insect processing to ensure safe novel foods. Moreover, most insect-based products are consumed as flour or processed items (e.g. pasta and bars); therefore the insect morphological traits cannot be used to verify the product authenticity and consequently its safety. Considering the high price of insect flours, we cannot exclude the occurrence of deliberate or intentional counterfeits (i.e. mix with low-cost flours, such as maize) as happens with other high-value food products, such as saffron (Petrakis, Cagliani, Polissiou, & Consonni, 2015).

Based on these assumptions, in this study we adopted, for the very first time, a DNA-based approach to analyse commercial novel food products to verify the identity of declared insect species and characterize trace amounts of plant material occurring in the same products in order to derive the source of the substrate, such as from insect diets and litter, with particular attention paid to putative elements of allergenic concern. Particularly, we used a region of the mtDNA *COI* marker to identify the insect species (DNA barcoding), and the nuclear ITS2 region to characterize insect diets and the vegetal composition of the tested products (High Throughput Sequencing HTS DNA metabarcoding). Moreover, the same approach can be used to verify product contamination and counterfeiting insect flour with low cost vegetable flours. We prepared six mock mixtures composed of wheat and soybean flours at different concentrations to also verify the limit of detection of our approach in the context of possible contamination of insect-based novel foods.

2. Materials and methods

2.1. Insect commercial food products

A total of 13 commercial insect-based products, namely flour ($n = 3$), pasta ($n = 3$), crackers ($n = 2$), protein bars ($n = 4$), and pet food ($n = 1$) were purchased via e-commerce from six different companies (Table 1). These categories offer an almost complete representation of the insect novel foods available in Europe. Based on the label information, these products contained only one insect species each, for a total of five species belonging to the orders Orthoptera (*Acheta domesticus* and *Gryllobius sigillatus*), Diptera (*Hermetia illucens*) and Coleoptera (*Alphitobius diaperinus* and *Tenebrio molitor*). Reference insect samples (RI) for each species were also retrieved (Supplementary Table S1) from a certified pet shop (AGRIPETGARDEN S.r.l., Conselve, Italy).

[Table 1]

2.2. Mock mixtures

To test for the efficacy of DNA metabarcoding in characterizing the composition of the insect-based products at the qualitative and semi-

quantitative levels, six mock mixtures were prepared (Supplementary Table S2). These were composed of insect flour (*T. molitor*) mixed with wheat (*Triticum aestivum*) flour (20 ppm, 200 ppm and 500 ppm of gluten) or soybean (*Glycine max*) flour (50 ppm, 200 ppm and 500 ppm of soybean proteins). Wheat and soybean flour concentrations were defined based on the alert threshold for allergens according to the EU regulation (No. 828/2014) and to Ballmer-Weber et al. (2007), respectively.

2.3. DNA extraction

For insect-based products and mock mixtures (see Table 1 and Supplementary Table S2), purified gDNA was obtained starting from 250 mg of samples by using the DNeasy PowerSoil Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions. All samples were prepared in three replicates. For reference insects (see Supplementary Table S1), purified gDNA was obtained starting from 25 mg of samples by using the DNeasy Blood & Tissue Kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions. Purified DNA was checked for concentration and purity by using a Qubit 2.0 Fluorometer and Qubit dsDNA HS Assay Kit (Invitrogen, Carlsbad, California, United States).

2.4. Insect identification by DNA barcoding

The 658 bp mtDNA *COI* region was used to authenticate the animal component in the purchased insect-based products. This region was amplified and sequenced for all 13 samples and for the reference insects by using primer pairs described by Folmer, Black, Hoeh, Lutz, and Vrijenhoek (1994) (Supplementary Table S3) and the protocol described in Bellati et al. (2014). The obtained sequences were submitted to the international GenBank through the EMBL platform (see Supplementary Table S1 for accession numbers). Each sequence was taxonomically assigned to the reference species (or to the declared one in the case of food items) by looking at the nearest matches with the BLAST algorithm using the following cut-off values/maximum identity > 99% and query coverage of 100%.

2.5. Libraries preparation and sequencing

To characterize the plant composition of the investigated insect-based products and mock mixtures, the obtained gDNA extracts were sequenced at the DNA barcode ITS2 region (Chen et al., 2010). Amplicons were obtained using the same approach described by Biella et al. (2019) with Illumina adapter (Supplementary Table S3) using puReTaq Ready-To-Go PCR beads (GE Healthcare Life Sciences, Italy) following the manufacturer's instructions in a 25 μ L reaction containing 1 μ L 10 μ M of each primer and up to 50 ng of gDNA. PCR cycles consisted of an initial denaturation step for 5 min at 94 °C, followed by 40 cycles of denaturation (30 s at 94 °C), annealing (30 s at 56 °C), and elongation (1 min at 72 °C), and, hence, a final elongation at 72 °C for 10 min. Amplicon DNA was checked for concentration by using a Qubit 2.0 Fluorometer and Qubit dsDNA HS Assay Kit (Invitrogen, Carlsbad, California, United States) (Supplementary Table S4) and amplicon length was measured by comparison against QX DNA Size Marker using the Qiaxcel Automatic electrophoresis system (QIAGEN, Hilden, Germany). Samples were sequenced by the Center for Translational Genomics and Bioinformatics (Milan, Italy). The sequencing was carried out on the MiSeq sequencing platform (Illumina, San Diego, CA, USA) with a paired-end approach (2 $\times$ 300 bp).

2.6. Bioinformatics analysis

Raw Illumina reads were paired and merged using the PEAR algorithm (Zhang, Kobert, Flouri, & Stamatakis, 2013). Pre-processing was performed using VSEARCH 2.0 algorithm (Rognes, Flouri, Nichols,

Table 1

List of analysed insect-based products. For each sample, information found on the label about the category, the species of insect, and the plant ingredients are reported. F (flour); FP (Food Product).

Code	Category	Label Declared insect	Label Declared Ingredients
F_001	Flour	<i>Tenebrio molitor</i>	-
F_002	Flour	<i>Grylodes sigillatus</i>	-
F_003	Flour	<i>Alphitobius diaperinus</i>	-
FP_004	Pasta	<i>Alphitobius diaperinus</i> (14%)	<i>Triticum durum</i> , <i>Ocimum basilicum</i> (1.5%); organic powdered egg whites.
FP_005	Pasta	<i>Alphitobius diaperinus</i> (14%)	<i>Triticum durum</i> ; organic powdered egg whites.
FP_006	Pasta	<i>Tenebrio molitor</i> (10%)	<i>Oryza sativa</i> (43); <i>Cicer arietinum</i> (43%); organic powdered egg whites (4%).
FP_007	Cracker	<i>Acheta domesticus</i> (14%)	<i>Triticum aestivum</i> ; <i>Sesamum indicum</i> (6%); <i>Olea europaea</i> .
FP_008	Cracker	<i>Tenebrio molitor</i> (10%)	<i>Triticum aestivum</i> ; <i>Cocos nucifera</i> ; <i>Avena sativa</i> ; <i>Sesamum indicum</i> (12%); <i>Porphyra</i> sp. (1.2%).
FP_009	Protein bar	<i>Acheta domesticus</i> (5.2%)	<i>Phoenix dactylifera</i> ; <i>Prunus dulcis</i> ; <i>Musa</i> spp. (11%); <i>Theobroma cacao</i> (9%); <i>Vaccinium macrocarpon</i> (8%); <i>Anacardium occidentale</i> ; <i>Cannabis sativa</i> .
FP_010	Protein bar	<i>Acheta domesticus</i> (5.5%)	<i>Phoenix dactylifera</i> ; <i>Prunus dulcis</i> ; <i>Prunus armeniaca</i> (22%); <i>Pisum sativum</i> ; <i>Helianthus annuus</i> ; <i>Lycium barbarum</i> (4.5%); <i>Salvia hispanica</i> (3.5%).
FP_011	Protein bar	<i>Acheta domesticus</i> (20%)	<i>Arachis hypogaea</i> (34%); <i>Cannabis sativa</i> ; <i>Theobroma cacao</i> ; <i>Agave</i> sp; <i>Beta vulgaris</i> ; <i>Cinnamomum</i> sp. (1%).
FP_012	Protein bar	<i>Acheta domesticus</i> (10%)	<i>Ananas comosus</i> (30%); <i>Phoenix dactylifera</i> ; <i>Anacardium occidentale</i> ; <i>Cocos nucifera</i> ; <i>Plantago</i> sp.; <i>Citrus limon</i> .
FP_013	Pet food	<i>Hermetia illucens</i> (25%)	<i>Ipomoea batatas</i> ; <i>Saccharomyces cerevisiae</i> ; <i>Lycium barbarum</i> (1%); <i>Rosmarinus officinalis</i> ; plant base glycerin.

Quince, & Mahé, 2016): reads were filtered out if ambiguous bases were detected and lengths were outside the bounds of 100 bp; moreover, an expected error = 1 was used as an indicator of read accuracy. Sequences were then dereplicated using `-derep_fulllength`. In order to decrease the false positive rate in the sequence population, a chimera detection analysis was performed on the obtained reference sequences. Since there is no reference database for ITS2 region for chimera detection, we used `-uchime_denovo` algorithm that carries out a de novo analysis without a reference. Plant features were obtained using `-cluster_fast` algorithm with a 100% sequence identity with at least a depth of 500x for each feature. A random sequence was chosen as the representative sequence of the cluster. Subsequently, DNA metabarcoding analysis was performed using the plugins of the QIIME2 suite (<https://docs.qiime2.org/>). The taxonomic assignment of the representative sequences was carried out using the classify-consensus-blast plugin implemented in QIIME2 (Camacho et al., 2009) against the local database, built with downloaded ITS2 sequences available in NCBI at 29th of January 2019, adopting a percent identity > 0.99 and a query coverage > 0.90. To evaluate the occurrence of contaminants (*T. aestivum* and *G. max*) in insect flour (F_001) and their relative abundance, we generated a heat map representation of the significant discriminatory features (plant species) obtained with the bioinformatic pipeline. Sample and feature axes were also organized using a clustering approach. The heat map was generated with the *feature-table* QIIME2 plugin (McDonald et al., 2012). To evaluate the sensitivity of the approach in detecting species based on feature depth, we performed a qualitative analysis considering the results of the taxonomy assignment described previously, assuming a depth of 500 ×, 100 ×, and 25 × for each OTU, respectively.

Python script (v.3; Pandas and NumPy libraries) was used to calculate sequence abundance weighted OTU and taxa overlap respectively (Wen et al., 2017) among the technical replicates. To evaluate significant differences among samples belonging to mock mixtures, a PERMANOVA test (permutation-based ANOVA, PerMANOVA) with 999 permutation-based Bray Curtis distance metrics (Faith, Minchin, & Belbin, 1987) was performed using the *diversity* QIIME2 plugin, considering both OTUs and taxa composition. PerMANOVA Pairwise contrast was performed through the beta-group-significance command of *diversity* plugin (Anderson, 2001).

2.7. ELISA assays

The three flour samples (F001, F002 and F003) and the mock samples prepared with *T. aestivum* (MT_20 - MT_500) were also analysed by RIDASCREEN® Gliadin kit (R-Biopharm AG, Darmstadt, Germany, prod. no. R 7001), a sandwich enzyme-linked

immunosorbent assay (ELISA) kit for gluten detection. The assay is based on the monoclonal antibody R5 (Méndez, Vela, Immer, & Janssen, 2005; Valdés, García, Llorente, & Méndez, 2003), which is specific for gliadin-fractions from wheat. The detection limit for gluten is 3 ppm (mg/kg). The manufacturer's instructions were followed.

3. Results

3.1. DNA barcoding authentication of insect ingredients

Good DNA yield (20–40 ng/μl) was obtained from all the replicates of the 13 collected samples and from the reference insect samples as well (see Supplementary Table S4). The mt *COI* DNA barcoding sequencing results indicate that all the tested insect-based products were correctly labelled concerning insect composition with the only exclusion of FP009 and FP010 (protein bars). In both cases, the DNA barcoding analysis did not find occurrences of *Acheta domestica* as expected, but the sequences matched with the *COI* of the food parasite species *Ectomyelois ceratoniae* (Insecta: Pyralidae).

3.2. DNA metabarcoding characterization of plant composition

The High Throughput Sequencing step produced a total of 8,142,444 raw reads, with an average of 126,299 reads (SD = 108,789; range 334–482,500) per sample. After the merging, quality filtering and dereplication steps, we retrieved a total of 868,414 reads, with an average of 13,677 reads (SD = 11,453; range 5–45,066) per sample. Details on the average and standard deviation statistics about raw and filtered reads obtained for each sample, considering replicates, are provided in Supplementary Table S4. After chimera detection and 100% cluster identity, with a depth of at least 500x for each feature, we obtained 120 OTUs (Operational Taxonomic Units). Negative controls for library sequencing were not included in the analysis since the very low amount of DNA copies. OTUs and taxa diversity were analysed separately for technical replicates at each sample. Both OTUs and taxa overlap (calculated with the weight of reads per OTUs and taxa) maintained a mean of 90% for all insect-products (with a standard deviation of 0.26).

The obtained taxonomic assignment and the distribution of the assigned taxa among the sample data are depicted in Fig. 1. Overall, 120 OTUs were assigned to at least 26 plant species. Most of the assigned OTUs reached the species taxonomic level, however, in some genera, such as *Triticum* and *Brassica*, the low interspecific variability did not allow the species to be identified.

The three flour samples were largely different. Specifically, F_001 is rich in *Daucus carota* reads (93.10%) followed by *Brassica* sp. (1.80%)

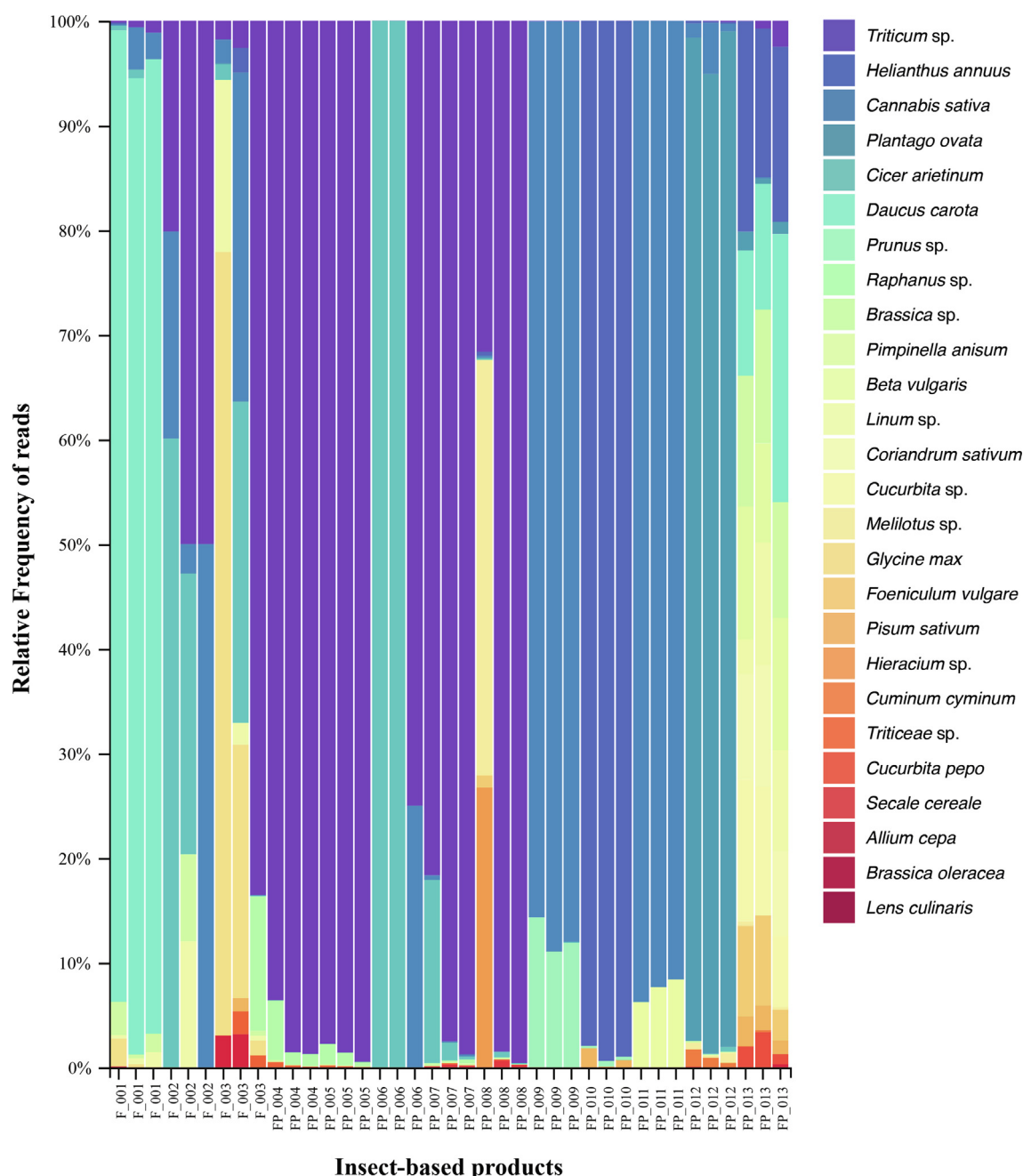


Fig. 1. Relative abundance of the plant taxa recovered in the 13 insect-based products through ITS2 metabarcoding sequencing.

and *Glycine max* (0.99%), F_002 contains mainly *Cicer arietinum* (43%), *Triticum sp.* (35%), and *Brassica sp.* (4%) reads. Finally, the F_003 flour shows many reads of *Glycine max* (33%), *Triticum sp.* (30%), and *Cicer arietinum* (11%). Moreover, all the flour samples showed a variable relative read abundance of *Cannabis sativa* (range 2.18–11%) and *Linum sp.* (range 0.76–6%).

The assigned plant taxa in the 13 samples were grouped in Expected species (E), that include taxa listed on the product label, Rearing Substrate (RS) which includes the putative plant used both as feed and litter for insect farming, and Not Expected (NE) encompassing all the remaining species. Table 2 indicates the distribution of the assigned species among the above-mentioned categories.

Concerning the processed insect-based products (from FP_004 - FP_013), the plant composition was clearly different among the tested categories. In the case of pasta, the first two samples (FP_004 and FP_005) mainly consisted of *Triticum sp.* (97% and 99% of reads

respectively) according to the label information (Table 1). Interestingly, reads of the NE *Raphanus sp.* (3% and 1% respectively) were found. Regarding the F_005 pasta sample, two out of the three replicates were largely dominated by reads of the expected species *Cicer arietinum* (67%). Strangely, no OTUs belonging to *Oryza sativa* were found.

In the two cracker products, the reads of *Triticum sp.* were mainly detected (range 74–92.68%), followed by several NE species such as *Cicer arietinum* (range 1–6.5%), *Secale cereale* (range 0.19–1%), and other taxa occurring in traces (see Table 2).

In the four protein bars (belonging to two different companies), very few OTUs were obtained, and in most cases reads of E species were not found. Finally, the composition of pet food was very complex, and OTUs belonging to > 12 species were detected (see Table 2).

Overall, the plant taxa distribution among the five insect-based product categories is schematized in Fig. 2. The contribution of plant diet is appreciable in the insect flours which are the purest and least

Table 2

List of the detected ingredients based on DNA metabarcoding assignment. For each insect-based product, the Expected species (E), the taxa belonging to the Rearing Substrate (RS), and the Not Expected (NE) species are indicated. The percentage values refer to the relative abundance of HTS ITS2 reads for each recognized ingredient.

Code	Detected ingredients		
	Expected (E)	Rearing substrate (RS)	Not expected (NE)
F_001	<i>T. molitor</i>	<i>Daucus carota</i> (93.1%); <i>Cannabis sativa</i> (2.2%); <i>Brassica</i> (1.8%); <i>Glycine max</i> (1%); <i>Linum sp.</i> (0.8%); <i>Cicer arietinum</i> (0.4%); <i>Triticum sp.</i> (0.6%); <i>Helianthus annuus</i> (0.1%).	<i>Melilotus sp.</i> (0.01%).
F_002	<i>G. sigillatus</i>	<i>Cicer arietinum</i> (43%); <i>Triticum sp.</i> (35%); <i>Cannabis sativa</i> (11%); <i>Linum sp.</i> (6%); <i>Brassica</i> (4%).	–
F_003	<i>A. diaperinus</i>	<i>Glycine max</i> (33%); <i>Triticum sp.</i> (30%); <i>Cannabis sativa</i> (11%); <i>Cicer arietinum</i> (11%); <i>Linum sp.</i> (6%); <i>Helianthus annuus</i> (1%); <i>Pisum sativum</i> (1%).	<i>Raphanus sp.</i> (4%); <i>Allium cepa</i> (1%).
FP_004	<i>A. diaperinus</i> ; <i>Triticum sp.</i> (97%).	–	<i>Raphanus sp.</i> (3%).
FP_005	<i>A. diaperinus</i> ; <i>Triticum sp.</i> (99%).	–	<i>Raphanus sp.</i> (1%).
FP_006	<i>T. molitor</i> ; <i>Cicer arietinum</i> (67%).	<i>Triticum sp.</i> (25%); <i>Cannabis sativa</i> (8%).	–
FP_007	<i>A. domesticus</i> ; <i>Triticum sp.</i> (92.7%).	<i>Cicer arietinum</i> (6.5%); <i>Cannabis sativa</i> (0.2%); <i>Brassica sp.</i> (0.2%); <i>Helianthus annuus</i> (0.06%).	<i>Beta vulgaris</i> (0.01%); <i>Raphanus sp.</i> (0.2%); <i>Secale cereale</i> (0.2%).
FP_008	<i>T. molitor</i> ; <i>Triticum sp.</i> (74%).	<i>Cicer arietinum</i> (1%).	<i>Melilotus sp.</i> (13%); <i>Hieracium sp.</i> (9%); <i>Foeniculum vulgare</i> (1%); <i>Secale cereale</i> (1%).
FP_009	<i>Cannabis sativa</i> (87%); <i>Prunus</i> (12%)	–	<i>Ectomyelois ceratoniae</i>
FP_010	<i>Helianthus annuus</i> (97%); <i>Pisum sativum</i> (2%); <i>Prunus sp.</i> (1%).	–	<i>Ectomyelois ceratoniae</i>
FP_011	<i>A. domesticus</i> ; <i>Cannabis sativa</i> (93%); <i>Beta vulgaris</i> (7%).	–	–
FP_012	<i>A. domesticus</i> ; <i>Plantago ovata</i> (96%).	<i>Cannabis sativa</i> (3%).	<i>Melilotus sp.</i> (1%); <i>Cuminum cyminum</i> (1%).
FP_013	<i>H. illucens</i>	<i>Helianthus annuus</i> (17%); <i>Daucus carota</i> (17%); <i>Pimpinella anisum</i> (12%); <i>Brassica sp.</i> (12%); <i>Linum sp.</i> (8%); <i>Pisum sativum</i> (2%); <i>Triticum sp.</i> (2%).	<i>Cucurbita</i> (11%); <i>Coriandrum sativum</i> (10%); <i>Foeniculum vulgare</i> (7%); <i>Cucurbita pepo</i> (2%); <i>Plantago ovata</i> (1%).

processed products. In Pasta, Crackers, and Protein bars the HTS analysis reveals the expected ingredient as the most abundant. Finally, in the pet food we detected the highest percentage of rearing substrate plant species.

To better characterize the composition of the processed products (from FP_004 to FP_013), we tested the identification performance reducing the reads filtering parameter from 500 to 100 and 25 reads per OTU. Fig. 3 shows the distribution of plant taxa among the three categories (i.e. E, RS, and NE) at different filtering thresholds. As expected, the whole number of assigned plant taxa increased with the decreasing threshold. Furthermore, the number of NE species increases dramatically compared to the increase of the Expected species.

3.3. DNA metabarcoding characterization on flour mock mixtures

The results of the HTS DNA metabarcoding analysis performed on the six flour mock mixtures are shown in the heat map diagram of Fig. 4. Pure insect flours (F_001) cluster together and show the occurrence of several RS plants, especially *D. carota*. In these samples, the two contaminants *T. aestivum* and *G. max* are absent. Conversely, in both the types of flour mock mixture, the abundance of wheat and

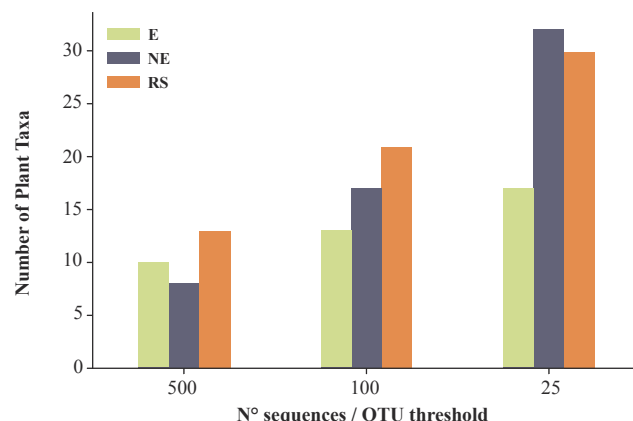


Fig. 3. Numbers of plant taxa recovered in the 13 insect-based products and mocks (E, RS, and NE) through ITS2 metabarcoding sequencing using different thresholds of numbers of sequences per OTU (500, 100, and 25, respectively).

soybean reads increases along with the amount of the admixed contaminant flours. In the case of samples MT_200 and MT_500, the

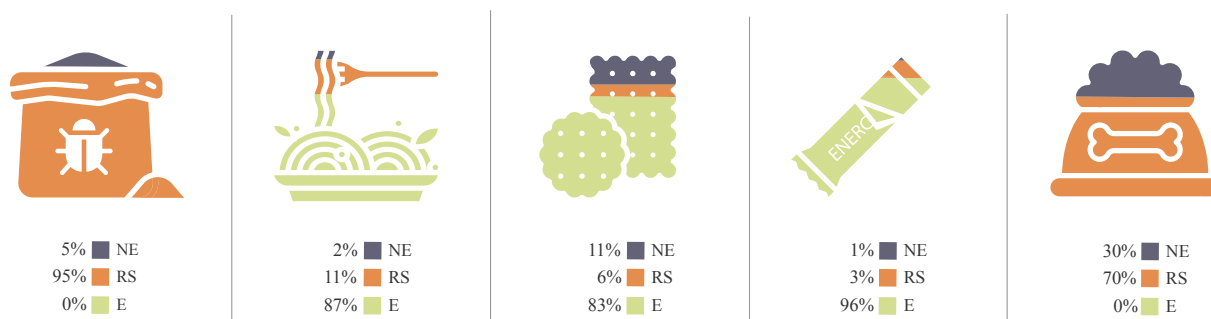


Fig. 2. Comprehensive graphic summary of the detected plant taxa in the five insect-based product categories (from left: Flour, Pasta, Crackers, Protein Bars and Pet Food). E: expected species, RS: Rearing substrate species, NE: Not expected species.

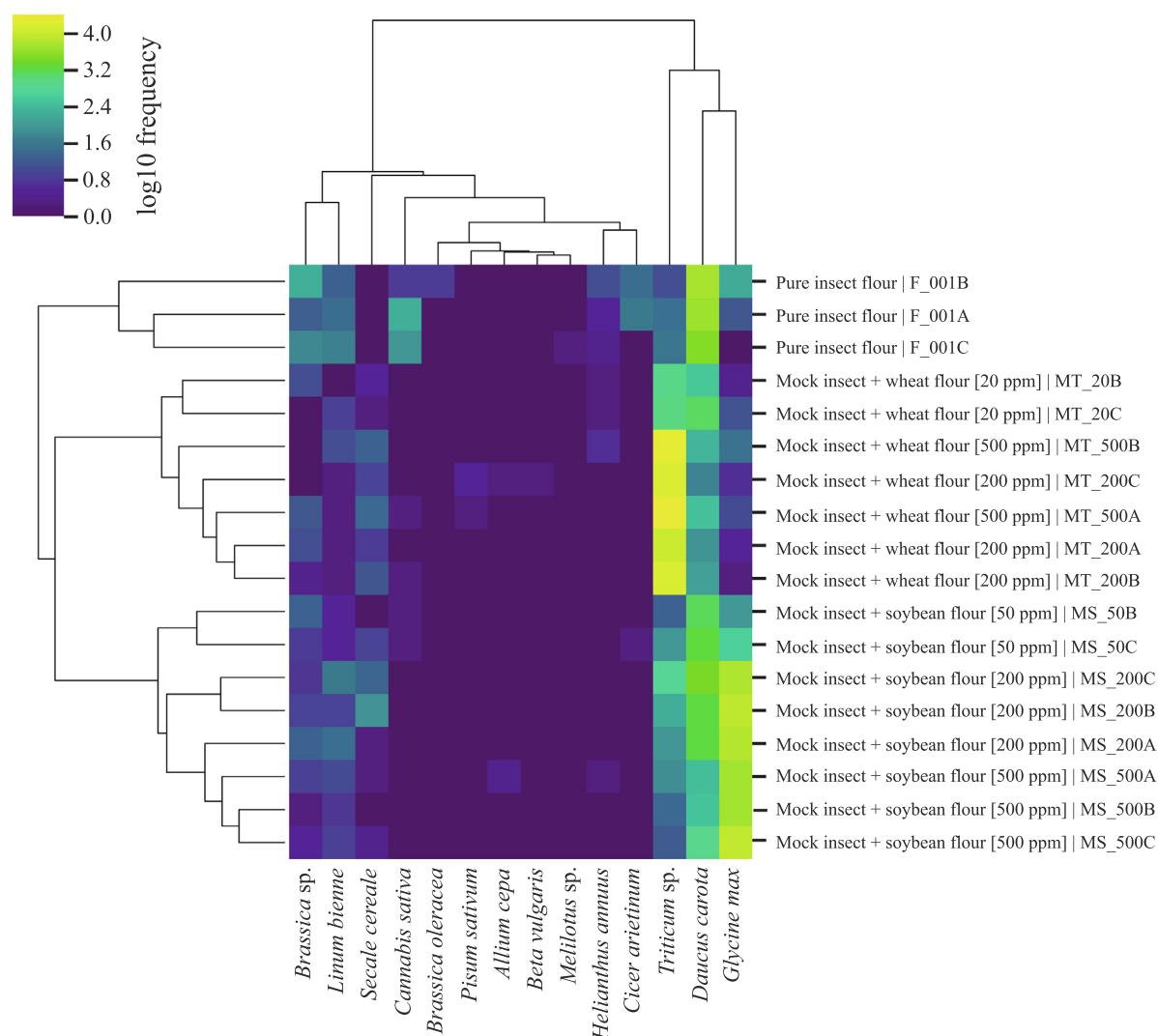


Fig. 4. Heat map diagram showing the abundance of plant taxa in the pure insect flours and mock mixtures. All the replicates, with the only exception of MT_20A and MS_50A, are shown. Color shading in the heat map indicates the abundance, expressed as log₁₀ frequency, of each species in the sample.

frequency of wheat reads reach the maximum level.

The PERMANOVA analysis shows that there is no statistical difference between samples belonging to the two categories of mock mixtures, both considering OTUs and taxa composition.

3.4. ELISA assays

Flour samples (i.e., F_001, F_002 and F_003) and the mocks prepared with *T. aestivum* flour (i.e., MT_20, MT_200 and MT_500) were analysed for gluten detection. Sample F_002, the mock MT_200, and MT_500 were positive to gluten content. (Details are provided in Table 3).

Table 3
Results of ELISA analyses on insect flours and wheat mock mixtures.

Sample	ELISA results (mg/kg)	Gluten-free claim
F_001	< 3	Yes
F_002	> 20	No
F_003	11	Yes
MT_20	15	Yes
MT_200	> 20	No
MT_500	> 20	No

4. Discussion

The characterization of novel food composition is essential to ensure safety (Patel, Suleria, & Rauf, 2019) and our results suggest that the mtDNA *COI* region is useful for identifying the declared insect species in flours and in almost all the analysed insect-based products. However, this approach showed some limitations, since in two protein bars, the *COI* DNA barcoding analysis identified the parasite *Ectomyelois ceratoniae* instead of the declared *Acheta domestica*. The occurrence of *E. ceratoniae* is not unexpected as its larvae typically parasitize raw food material, especially when almond occurs as an ingredient (Mortazavi, Samih, Ghajarieh, & Jafari, 2015). We cannot exclude that the industrial processing steps fragmented or degraded the *A. domestica* DNA in some way, making it not targetable by the used primer pair. The use of species-specific DNA probes (Tramuta et al., 2018) or another DNA metabarcoding approach targeting at the insect ingredients using short genetic regions (< 400 bp), would have detected the occurrence of both the moth and the insect species declared on product label (Frigerio et al., 2019).

Concerning the plant components, the ITS2 region was efficient in providing information on species composition including those taxa probably belonging to the insect rearing substrate. There are several open questions about the composition of the feeds used for insects farming and its influences on the quality and safety of the final products

(Magara et al., 2019; Van Broekhoven, Oonincx, Van Huis, & Van Loon, 2015). The selection of suitable feeding substrates is very important to enhance the nutritional characteristics of the insects (Magara et al., 2019; Oonincx & de Boer, 2012; Oonincx, Laurent, Veenenbos, & van Loon, 2019; Van Broekhoven et al., 2015) and can also affect the total farming yield (Ganda, Zannou-Boukari, Kenis, Chrysostome, & Mensah, 2019). To date, the EU regulation (EC no 1069/2009) clarifies that the insect rearing substrate has to contain only products of non-animal origin. The circular economy strategies support the adoption of bio-waste and by-products of different agricultural and industrial origins claiming for potential benefits in terms of sustainability (EFSA, 2015). Therefore, the insect feeds are usually composed of vegetable ingredients derived from different agricultural supply chains which are likely impossible to be identified using morphological parameters. This was confirmed by our data which highlighted that the insect feed was almost completely characterized by horticultural plant sources, such as carrots, cabbages, and chickpea (Magara et al., 2019). Therefore, the proposed molecular approach offers a universal diagnostic system to identify the composition of the rearing substrate and to verify compliance with the current and future regulations.

In agreement with Schlüter and co-workers (Schlüter et al., 2017), insects must be reared under a defined substrate to avoid contamination and possible food borne outbreak for the consumer (e.g., due to pathogenic microorganisms, toxins, and antinutrients). Our data suggest that the ITS2 barcode region was also able to identify the plant-based substrate used for insect rearing. For example, in the analysed flour samples, we detected DNA from hemp (*Cannabis sativa*) and linen (*Linum* sp.), which are commonly used as litter material (data confirmed by interviewed companies). We underline that *Cannabis sativa* is also used as an ingredient in some of the tested protein bars. Unfortunately, our method is not able to distinguish between the two sources.

In many processed products, we obtained most reads belonging to the expected highly abundant species, such as wheat and chickpeas, which likely hide the less represented OTUs. For example, the species that possibly constitute the rearing substrate are not very evident when compared to the analysed pure flours. Therefore, the main limitation of our analysis certainly resides in the sensitivity to detect the less abundant species due to primer bias. This issue has already been discussed by Bista and colleagues (Bista et al., 2018) and Krehenwinkel and colleagues (Krehenwinkel et al., 2017). Both studies agreed that a PCR-free whole genome sequencing could permit to avoid this effect. Furthermore, we demonstrated that different thresholds of OTUs size (in terms of number of reads) dramatically affect the final list of plant taxa recovered in the analysed samples. It is important to underline that in a context of food authentication and traceability, the adopted conservative criteria (i.e., n° sequences/OTU > 500) is essential to preserve information on the most representative species. Conversely, a deeper data exploration, using fewer conservative parameters (i.e., n° sequences/OTU (500), retrieves information on trace species but increases the risk of including false positives. Probably, a multi-marker approach, coupled with a dedicated reference molecular database encompassing the expected plants, as well as the most common contaminants (or species prohibited by law) could improve the plant characterization of insect-based products and could be useful to exclude the false positives. Other authors (e.g., Zhang, Chain, Abbott, & Cristescu, 2018) demonstrated that the combined use of at least two barcode markers improves species detection. Another possible limitation of DNA metabarcoding resides in the completeness and reliability of the reference dataset that could lead to incorrect reads assignment (Murali, Bhargava, & Wright, 2018). In our study, we chose the ITS2 barcode due to its higher capability of distinguishing congeneric plant species due to a higher mutation rate (Al-Juhani, 2019; Yu et al., 2017). Moreover, in recent years, the ITS2 database is growing exponentially and this improves the suitability of this locus to taxonomically assigning DNA metabarcoding data.

4.1. DNA metabarcoding to identify and quantify allergens

In our study, we tested the ability of DNA metabarcoding to find plant contaminants in edible insect flours. The obtained results suggested that we are able to identify DNA of *T. aestivum* or *G. max* in the tested mock mixtures, starting from 20 ppm and 50 ppm of allergenic proteins, respectively. According to European regulations (EU No. 828/2014), these concentrations are the maximum limits for commercial gluten and soybean-protein free products.

Therefore, the HTS DNA metabarcoding analysis detects low amounts of contaminant products and allergens, with a limit of detection even lower than the ELISA analysis (i.e., 3 ppm). We underline that the three tested insect-based flours are declared as gluten-free products, but only F_001 and F_003 comply with the limit established by the European Commission.

Concerning the ability of DNA metabarcoding to quantify the 'putative plant contaminants', our study seems to indicate a weak relationship between the dry weight and the number of reads. However, there is a fervent debate about the effectiveness of providing quantitative inferences using HTS data. Some recent studies reported their findings in a quantitative manner where the relative read abundance is interpreted as the relative abundance of biomass (Lamb et al., 2019). Others use a frequency of occurrence approach, also referred to as weighted occurrence (Deagle et al., 2018), where the proportion of samples in which a given sequence was detected is used to infer a different sort of quantitative measure (De Barba et al., 2014).

As already noticed in our previous paper on processed food (Bruno et al., 2019), the amplicon DNA metabarcoding efficacy could be biased by the PCR amplification step using "universal" markers. The occurrence of bias during PCR amplification may cause the inaccurate estimation of quantities, and this was at least partially demonstrated for metazoans and plants (Balech et al., 2018; Thudi, Li, Jackson, May, & Varshney, 2012). This bias generates a variable number of template-primer mismatches across species, resulting in a final amplified DNA mixture that does not always reflect the original proportion of each species, limiting the quantitative potential of DNA metabarcoding (Bista et al., 2018; Piñol, Senar, & Symondson, 2019). Nevertheless, our analysis suggests that DNA metabarcoding has a relative quantitative ability, as already demonstrated by Lamb and colleagues (Lamb et al., 2019), and this methodology can be intended as an early warning method for allergen detection in food products.

4.2. DNA metabarcoding of insect-based novel food: An overview

DNA metabarcoding is currently used for food authentication (Galvin-King, Haughey, & Elliott, 2018; Prosser & Hebert, 2017; Utzeri et al., 2018). Moreover, Haynes and colleagues (Haynes, Jimenez, Pardo, & Helyar, 2019) recommend this approach to enhance the quality control along the food supply chain. In order to present an overview of DNA metabarcoding applied to insect food authentication and safety, we developed a SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis (Fig. 5). The strengths are related to the high efficiency of the technique. DNA metabarcoding can detect traces of ingredients due to its high sensitivity and allows to obtain simultaneously different information about food safety and quality. The main weaknesses are the higher cost compared to the current available analytical approaches like ELISA tests or target PCR assays. However, we cannot exclude that in the very next future, the panel of targets required to assess authenticity and safety of insect-based products will be so wide to make the HTS DNA metabarcoding approach much more convenient than the use of multiple single target tests. DNA tests indeed are related to DNA quality and quantity, so highly processed insect products like protein bars can be challenging to analyse. Finally, the results depend on the database which may be incomplete with a subsequent incorrect assignment. The opportunities are related to the novelty of insect food products. It is possible to create new economic

S TRENGTHS	O PPORTUNITIES
• Higher sensitivity compared to EU-required analyses (e.g., ELISA) • More accurate food quality and traceability assessments than conventional methods • Increased food safety (multitarget allergen detection)	• Quickly support the compliance of future regulations • Create new economic opportunities in the analytical field • Stimulate the insect base novel food market
• Depends on DNA quality and quantity • Depends on completeness and accuracy of references databases • More expensive compared to EU-required analyses (e.g., ELISA)	• DNA authentication control of food products is not mandatory (e.g., in Europe) • Insect feed compositions and farming details are not required on the label • Lack of DNA metabarcoding references applied to insect-based novel food
W EAKNESSES	T HREATS

Fig. 5. Overview of the main strengths, weaknesses, opportunities and threats (SWOT analysis) related to the use of DNA metabarcoding as tool for insect-based novel food products.

opportunities in the analytical field with the goal to guarantee a safe product, stimulating the insect-based novel food market. HTS techniques, due to the completeness of the results, could also quickly support the compliance to forthcoming regulations. Finally, the first threat is that information on insect feed composition is not reported on the product label and this impedes to compare the detected composition with the declared one. Furthermore, there is currently no scientific reference to DNA metabarcoding applied to the quality and safety control of the new insect-based foods. In addition to that, DNA analyses are currently not mandatory, so this can interfere with the spread of DNA metabarcoding as a routine analysis.

5. Conclusions

Novel foods demand is increasing, and their consumption is expected to grow in the next years. This condition encourages an increase in the number of ingredient species and enhances the risk of misidentification, contamination, and counterfeiting. This is well documented in the case of fish, where more new species are available on the market, and in many countries the DNA barcoding approach is considered an essential tool to avoid fish frauds (Fox, Mitchell, Dean,

Elliott, & Campbell, 2018).

Considering that many food products contain a mixture of species, we strongly encourage the adoption of DNA metabarcoding to better elucidate not only the food composition but also to assess trace elements belonging to different steps of the food supply chains. Unfortunately, this approach does not accurately estimate the biomass of the ingredient taxa, and although in our case the HTS DNA metabarcoding approach highlighted the occurrence of allergenic species (even at limit concentration values), we cannot use this method as an alternative to the standard ELISA test. This is also because the presence of DNA of a species is not necessarily correlated with the occurrence of allergens. However, we proposed the DNA metabarcoding analysis as a preliminary screening, especially for novel foods, because this method offers the ability to identify, at the semiquantitative level, several potential allergenic plants with a single analysis. Therefore, DNA-based analysis can be used to select which ELISA tests, and in general which of the more reliable toxicological assays for the detection of plant contaminants to use. In this sense, the DNA metabarcoding approach offers an opportunity to enhance the food safety of novel food products, such as those based on insect ingredients.

Data statement

The dataset generated for this study will be submitted to the EBI metagenomics portal (<https://www.ebi.ac.uk/metagenomics/>) upon paper acceptance. In any case, we will provide the dataset if required.

Acknowledgements

The authors are indebted with Paola Re for support in preparing the graphical part of the manuscript and with F. Braciola and M. Barnaby for advice and support during manuscript preparation.

Funding

This work was supported by Regione Lombardia in the framework of the Program 'Accordi per la ricerca e l'innovazione' under the project 'Food Social Sensor Network Food NET', grant number: E47F17000020009. The funder had no role in conducting the research and/or during the preparation of the article.

Declaration of Competing Interest

FEM2-Ambiente s.r.l., provided support in the form of a salary for authors J.F. and V.M. The company only provided financial support in the form of research materials.

Appendix A. Supplementary material

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodres.2019.108751>.

References

- Al-Juhani, W. S. (2019). Evaluation of the capacity of the DNA barcode ITS2 for identifying and discriminating dryland plants. *Genetic Molecular Research*, 18(1), GMR18133. <https://doi.org/10.4238/gmr18133>.
- Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26(1), 32–46. <https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x>.
- Balech, B., Sandionigi, A., Manzari, C., Trucchi, E., Tullo, A., Licciulli, F., ... Vicario, S. (2018). Tackling critical parameters in metazoan meta-barcoding experiments: A preliminary study based on cox1 DNA barcode. *PeerJ*, 6, e4845. <https://doi.org/10.7717/peerj.4845>.
- Ballmer-Weber, B. K., Holzhauser, T., Scibilia, J., Mittag, D., Zisa, G., Ortolani, C., ... Bindslev-Jensen, C. (2007). Clinical characteristics of soybean allergy in Europe: A double-blind, placebo-controlled food challenge study. *The Journal of Allergy and Clinical Immunology*, 119(6), 1489–1496. <https://doi.org/10.1016/j.jaci.2007.01.049>.
- Bellati, A., Tiberti, R., Cocca, W., Galimberti, A., Casiraghi, M., Bogliani, G., & Galeotti, P. (2014). A dark shell hiding great variability: A molecular insight into the evolution and conservation of melanin *Daphnia* populations in the Alps. *Zoological Journal of the Linnean Society*, 171(4), 697–715. <https://doi.org/10.1111/zoj.12151>.
- Biella, P., Tommasi, N., Akter, A., Guzzetti, L., Klecka, J., Sandionigi, A., & Galimberti, A. (2019). Foraging strategies are maintained despite workforce reduction: A multi-disciplinary survey on the pollen collected by a social pollinator. *PLoS One*, 14(11), e0224037. <https://doi.org/10.1101/622910>.
- Bista, I., Carvalho, G. R., Tang, M., Walsh, K., Zhou, X., Hajibabaei, M., ... Creer, S. (2018). Performance of amplicon and shotgun sequencing for accurate biomass estimation in invertebrate community samples. *Molecular Ecology Resources*, 18(5), 1020–1034. <https://doi.org/10.1111/1755-0998.12888>.
- Bruno, A., Sandionigi, A., Agostinetto, G., Bernabovi, L., Frigerio, J., Casiraghi, M., & Labra, M. (2019). food tracking perspective: DNA metabarcoding to identify plant composition in complex and processed food products. *Genes*, 10(3), 248. <https://doi.org/10.3390/genes10030248>.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, 10, 421. <https://doi.org/10.1186/1471-2105-10-421>.
- Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., ... Luo, K. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One*, 5(1), e8613. <https://doi.org/10.1371/journal.pone.0008613>.
- De Barba, M., Miquel, C., Boyer, F., Mercier, C., Rioux, D., Coissac, E., & Taberlet, P. (2014). DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: Application to omnivorous diet. *Molecular Ecology Resources*, 14(2), 306–323. <https://doi.org/10.1111/1755-0998.12188>.
- Deagle, B. E., Thomas, A. C., McInnes, J. C., Clark, L. J., Vesterinen, E. J., Clare, E. L., ... Eveson, J. P. (2018). Counting with DNA in metabarcoding studies: How should we convert sequence reads to dietary data? *Molecular Ecology*, 28(2), 391–406. <https://doi.org/10.1111/mec.14734>.
- Di Mattia, C., Battista, N., Sacchetti, G., & Serafini, M. (2019). Antioxidant Activities in vitro of Water and Liposoluble Extracts Obtained by Different Species of Edible Insects and Invertebrates. *Frontiers in Nutrition*, 6, 106. <https://doi.org/10.3389/fnut.2019.00106>.
- Dobermann, D., Swift, J. A., & Field, L. M. (2017). Opportunities and hurdles of edible insects for food and feed. *Nutrition Bulletin*, 42(4), 293–308. <https://doi.org/10.1111/mbu.12291>.
- EFSA (2015). Risk profile related to production and consumption of insects as food and feed. *EFSA Journal*, 13(10), 4257. <https://doi.org/10.2903/j.efsa.2015.4257>.
- Faith, D. P., Minchin, P. R., & Belbin, L. (1987). Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio*, 69(1–3), 57–68. https://doi.org/10.1007/978-94-009-4061-1_6.
- Folmer, O., Black, M., Hoeh, W., Lutz, R., & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, 3(5), 294–299. <https://doi.org/10.1371/journal.pone.0013102>.
- Fox, M., Mitchell, M., Dean, M., Elliott, C., & Campbell, K. (2018). The seafood supply chain from a fraudulent perspective. *Food Security*, 10(4), 939–963. <https://doi.org/10.1007/s12571-018-0826-z>.
- Frigerio, J., Gorini, T., Galimberti, A., Bruni, I., Tommasi, N., Mezzasalma, V., & Labra, M. (2019). DNA barcoding to trace Medicinal and Aromatic Plants from the field to the food supplement. *Journal of Applied Botany and Food Quality*, 92, 33–38. <https://doi.org/10.5073/JABFQ.2019.092.005>.
- Galvin-King, P., Haughey, S. A., & Elliott, C. T. (2018). Herb and spice fraud: the drivers, challenges and detection. *Food Control*, 88, 85–97. <https://doi.org/10.1016/j.foodcont.2017.12.031>.
- Ganda, H., Zannou-Boukari, E. T., Kenis, M., Chrysostome, C. A. A. M., & Mensah, G. A. (2019). Potentials of animal, crop and agri-food wastes for the production of fly larvae. *Journal of Insects as Food and Feed*, 5(2), 59–67. <https://doi.org/10.3920/JIFF2017.0064>.
- Haynes, E., Jimenez, E., Pardo, M. A., & Helyar, S. J. (2019). The future of NGS (Next Generation Sequencing) analysis in testing food authenticity. *Food Control*, 101, 134–143. <https://doi.org/10.1016/j.foodcont.2019.02.010>.
- Krehenwinkel, H., Wolf, M., Lim, J. Y., Rominger, A. J., Simison, W. B., & Gillespie, R. G. (2017). Estimating and mitigating amplification bias in qualitative and quantitative arthropod metabarcoding. *Scientific Reports*, 7(1), 17668. <https://doi.org/10.1038/s41598-017-17333-x>.
- Lamb, P. D., Hunter, E., Pinnegar, J. K., Creer, S., Davies, R. G., & Taylor, M. I. (2019). How quantitative is metabarcoding: A meta-analytical approach. *Molecular Ecology*, 28(2), 420–430. <https://doi.org/10.1111/mec.14734>.
- Magara, H. J., Tanga, C. M., Ayieko, M. A., Hugel, S., Mohamed, S. A., Khamis, F. M., ... Roos, N. (2019). performance of newly described native edible cricket scapsipedus icipe (Orthoptera: Gryllidae) on various diets of relevance for farming. *Journal of Economic Entomology*, 112(2), 653–664. <https://doi.org/10.1093/jeet/toy397>.
- Mazzuchelli, G., Holzhauser, T., Cirkovic Velickovic, T., Diaz-Perales, A., Molina, E., Roncada, P., ... Hoffmann-Sommergruber, K. (2018). Current (Food) allergenic risk assessment: Is it fit for novel foods? Status quo and identification of gaps. *Molecular Nutrition & Food Research*, 62(1), 1700278. <https://doi.org/10.1002/mnfr.201700278>.
- McDonald, D., Clemente, J. C., Kuczynski, J., Rideout, J. R., Stombaugh, J., Wendel, D., ... Caporaso, J. G. (2012). The biological observation matrix (BIOM) format or: How I learned to stop worrying and love the ome-ome. *GigaScience*, 1, 7. <https://doi.org/10.1186/2047-217X-1-7>.
- Méndez, E., Vela, C., Immer, U., & Janssen, F. W. (2005). Report of a collaborative trial to investigate the performance of the R5 enzyme linked immunoassay to determine gliadin in gluten-free food. *European Journal of Gastroenterology & Hepatology*, 17(10), 1053–1063. <https://doi.org/10.1097/00042737-200510000-00008>.
- Mortazavi, S., Samih, M. A., Ghajarieh, H., & Jafari, A. (2015). Effect of some diets on demographic parameters of *Ectomyelois ceratoniae* (Zeller) (Lepidoptera: Pyralidae) in vitro. *Journal of Plant Protection Research*, 55(2), 212–219. <https://doi.org/10.1515/jppr-2015-0028>.
- Murali, A., Bhargava, A., & Wright, E. S. (2018). IDTAXA: A novel approach for accurate taxonomic classification of microbiome sequences. *Microbiome*, 6(1), 140. <https://doi.org/10.1186/s40168-018-0521-5>.
- Myers, G., & Pettigrew, S. (2018). A qualitative exploration of the factors underlying seniors' receptiveness to entomophagy. *Food Research International*, 103, 163–169. <https://doi.org/10.1016/j.foodres.2017.10.032>.
- Oonincx, D. G. A. B., & de Boer, I. J. M. (2012). Environmental Impact of the Production of Mealworms as a Protein Source for Humans – A Life Cycle Assessment. *PLoS One*, 7(12), e51145. <https://doi.org/10.1371/journal.pone.0051145>.
- Oonincx, D. G. A. B., Laurent, S., Veenbos, M. E., & van Loon, J. J. A. (2019). Dietary enrichment of edible insects with omega 3 fatty acids. *Insect Science*, 1–10. <https://doi.org/10.1111/1744-7917.12669>.
- Pali-Schöll, I., Verhoeckx, K., Mafra, I., Bavaro, S. L., Mills, E. C., & Monaci, L. (2019). Allergenic and novel food proteins: State of the art and challenges in the allergenicity assessment. *Trends in Food Science & Technology*, 84, 45–48. <https://doi.org/10.1016/j.tifs.2018.03.007>.
- Patel, S., Suleria, H. A. R., & Rauf, A. (2019). Edible insects as innovative foods: Nutritional and functional assessments. *Trends in Food Science & Technology*, 86, 352–359. <https://doi.org/10.1016/j.tifs.2019.02.033>.
- Payne, C. L. R., Scarborough, P., Rayner, M., & Nonaka, K. (2016). Are edible insects more or less healthy than commonly consumed meats? A comparison using two nutrient profiling models developed to combat over- and undernutrition. *European Journal of*

- Clinical Nutrition*, 70(3), 285–291. <https://doi.org/10.1038/ejcn.2015>.
- Petrakis, E. A., Cagliani, L. R., Polissiou, M. G., & Consonni, R. (2015). Evaluation of saffron (*Crocus sativus* L.) adulteration with plant adulterants by 1H NMR metabolite fingerprinting. *Food Chemistry*, 173, 890–896. <https://doi.org/10.1016/j.foodchem.2014>.
- Piñol, J., Senar, M. A., & Symondson, W. O. C. (2019). The choice of universal primers and the characteristics of the species mixture determine when DNA metabarcoding can be quantitative. *Molecular Ecology*, 28(2), 407–419. <https://doi.org/10.1111/mec.14776>.
- Prosser, S. W. J., & Hebert, P. D. N. (2017). Rapid identification of the botanical and entomological sources of honey using DNA metabarcoding. *Food Chemistry*, 1(214), 183–191. <https://doi.org/10.1016/j.foodchem.2016>.
- Regulation (Ec) No 1069/2009 of the European Parliament and of the council of 11 December 2015. Available online <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L:2015:327:FULL&from=EN> (accessed on 19 April 2019).
- Regulation (EU) 2017/893 of the European Parliament and of the council of 24 May 2017. Available online <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0893&from=en> (accessed on 19 April 2019).
- Regulation (EU) No 828/2014 of 30 July 2014. Available online <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0828&from=IT> (accessed on 19 April 2019).
- Ribeiro, J. C., Cunha, L. M., Sousa-Pinto, B., & Fonseca, J. (2018). Allergic risks of consuming edible insects: A systematic review. *Molecular Nutrition & Food Research*, 62(1), 1700030. <https://doi.org/10.1002/mnfr.201700030>.
- Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, 4, e2584. <https://doi.org/10.7717/peerj.2584>.
- Rumpold, B. A., & Schlüter, O. K. (2013). Nutritional composition and safety aspects of edible insects. *Molecular Nutrition & Food Research*, 57(5), 802–823. <https://doi.org/10.1002/mnfr.201200735>.
- Schlüter, O., Rumpold, B., Holzhauser, T., Roth, A., Vogel, R. F., Quasigroch, W., ... Kulling, S. (2017). Safety aspects of the production of foods and food ingredients from insects. *Molecular nutrition & Food Research*, 61(6), 1600520. <https://doi.org/10.1002/mnfr.201600520>.
- Sun-Waterhouse, D., Waterhouse, G. I. N., You, L., Zhang, J., Liu, Y., Ma, L., ... Dong, Y. (2016). Transforming insect biomass into consumer wellness foods: A review. *Food Research International*, 89, 129–151. <https://doi.org/10.1016/j.foodres.2016.10.001>.
- Tramuta, C., Gallina, S., Bellio, A., Bianchi, D. M., Chiesa, F., Rubiola, S., ... Decastelli, L. (2018). A set of multiplex polymerase chain reactions for genomic detection of nine edible insect species in foods. *Journal of Insect Science*, 18(5), <https://doi.org/10.1093/jisesa/iey087>.
- Thudi, M., Li, Y., Jackson, S. A., May, G. D., & Varshney, R. K. (2012). Current state-of-art of sequencing technologies for plant genomics research. *Briefings in Functional Genomics*, 11(1), 3–11. <https://doi.org/10.1093/bfpg/elfr045>.
- Utzeri, V. J., Ribani, A., Schiavo, G., Bertolini, F., Bovo, S., & Fontanesi, L. (2018). Application of next generation semiconductor based sequencing to detect the botanical composition of monofloral, polyfloral and honeydew honey. *Food Control*, 86, 342–349. <https://doi.org/10.1016/j.foodcont.2017.11.033>.
- Valdés, I., García, E., Llorente, M., & Méndez, E. (2003). Innovative approach to low-level gluten determination in foods using a novel sandwich enzyme-linked immunosorbent assay protocol. *European Journal of Gastroenterology & Hepatology*, 15(5), 465–474. <https://doi.org/10.1097/01.meg.0000059119.41030.df>.
- Van Broekhoven, S., Oonincx, D. G., Van Huis, A., & Van Loon, J. J. (2015). Growth performance and feed conversion efficiency of three edible mealworm species (Coleoptera: Tenebrionidae) on diets composed of organic by-products. *Journal of Insect Physiology*, 73, 1–10. <https://doi.org/10.1016/j.jinsphys.2014.12.005>.
- Van der Fels-Klerx, H. J. (Ine), Adamse, P., Punt, A., & Van Asselt, E. D. (2018). Data analyses and modelling for risk based monitoring of mycotoxins in animal feed. *Toxins*, 10(2), 54. <https://doi.org/10.3390/toxins10020054>.
- Walia, K., Kapoor, A., & Farber, J. M. (2018). Qualitative risk assessment of cricket powder to be used to treat undernutrition in infants and children in Cambodia. *Food Control*, 92, 169–182. <https://doi.org/10.1016/j.foodcont.2018.04.047>.
- Wen, C., Wu, L., Qin, Y., Van Nostrand, J. D., Ning, D., Sun, B., ... Zhou, J. (2017). Evaluation of the reproducibility of amplicon sequencing with Illumina MiSeq platform. *PLoS One*, 12(4), e0176716. <https://doi.org/10.1371/journal.pone.0176716>.
- Yu, N., Wei, Y. L., Zhang, X., Zhu, N., Wang, Y. L., Zhu, Y., ... Sun, A. D. (2017). Barcode ITS2: A useful tool for identifying *Trachelospermum jasminoides* and a good monitor for medicine market. *Scientific Reports*, 7(1), 5037. <https://doi.org/10.1038/s41598-017-04674-w>.
- Zhang, G. K., Chain, F. J. J., Abbott, C. Lp., & Cristescu, M. E. (2018). Metabarcoding using multiplexed markers increases species detection in complex zooplankton communities. *Evolutionary Applications*, 11(10), 1901–1914. <https://doi.org/10.1111/eva.12694>.
- Zhang, J., Kobert, K., Flouri, T., & Stamatakis, A. (2013). PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*, 30(5), 614–620. <https://doi.org/10.1093/bioinformatics/btt593>.

The hidden ‘plant side’ of insect novel foods: a DNA-based assessment

Jessica Frigerio^{a^}, Giulia Agostinetto^{b^}, Anna Sandionigi^b, Valerio Mezzasalma^a, Nadia Maria Berterame^b, Maurizio Casiraghi^b, Massimo Labra^b, Andrea Galimberti^{b*}

^a FEM2-Ambiente, Piazza della Scienza 2, I-20126 Milano, Italy.

^b Department of Biotechnology and Biosciences, University of Milano-Bicocca, Piazza della Scienza 2, I-20126 Milano, Italy.

APPENDIX A: Supplementary Material

Table S.1. List of insect samples used as reference for DNA barcoding authentication.

Specimen	Species	Accession number
RI_001	<i>Acheta domesticus</i>	LR585071
RI_002	<i>Alphitobius diaperinus</i>	LR585072
RI_003	<i>Gryllodes sigillatus</i>	LR585073
RI_004	<i>Hermetia illucens</i>	LR585074
RI_005	<i>Tenebrio molitor</i>	LR585075

Table S.2. List of mock mixtures composed of insect flour (*T. molitor*) admixed with wheat and soybean contaminant flours.

Code	Contaminant species	Contaminant concentration (mg/kg)
MT_20	<i>Triticum aestivum</i>	20 ppm
MT_200	<i>Triticum aestivum</i>	200 ppm
MT_500	<i>Triticum aestivum</i>	500 ppm
MS_50	<i>Glycine max</i>	50 ppm
MS_200	<i>Glycine max</i>	200 ppm
MS_500	<i>Glycine max</i>	500 ppm

Table S.3. List of primer pairs used for DNA barcoding and metabarcoding analyses.

Primer name	5'-3'	Barcode locus	Reference
LCO1490	GGTCAACAAATCATAAAGATATTGG	COI	Folmer et al., 1994
HC02198	TAAACTTCAGGGTGACCAAAAAATCA		
S2F	ATGCGATACTTGGTGTGAAT	ITS	Chen et al., 2010
S3R	GACGCTTCTCCAGACTACAAT		

Table S.4. List of analysed specimens with DNA concentration (ng/μl) evaluated by using a Qubit 2.0 Fluorometer and Qubit dsDNA HS Assay Kit and details about the average number of reads for each specimen, considering raw and filtered reads (after merging, quality filtering and dereplication steps) and including replicates.

Specimen	DNA concentration (ng/μl)	Average number of raw reads	Average number of filtered reads
F__001A	53	38,186	5,252
F__001B	51		
F__001C	37.9		
F__002A	34.2	6,944	1,340
F__002B	43.4		
F__002C	32		
F__003A	31	83,306	14,056
F__003B	30		
F__003C	60		
FP_004A	100	97,395	9,979
FP_004B	130		
FP_004C	350		
FP_005A	112	304,191	25,231
FP_005B	40		
FP_005C	200		
FP_006A	200	103,443	8,055
FP_006B	300		
FP_006C	276		
FP_007A	42.8	157,591	18,127
FP_007B	38.3		
FP_007C	46.3		
FP_008A	43.4	203,698	21,307
FP_008B	36.6		
FP_008C	27.8		

FP_009A	87	158,204	22,390
FP_009B	76		
FP_009C	81		
FP_010A	95	312,664	26,646
FP_010B	155		
FP_010C	94		
FP_011A	160	108,473	9,270
FP_011B	125		
FP_011C	231		
FP_012A	39	162,537	14,163
FP_012B	34		
FP_012C	34		
FP_013A	30	144,571	25,150
FP_013B	33.4		
FP_013C	33.3		
MS_50A	47.6	40,883	5,940
MS_50B	38		
MS_50C	28.3		
MS_200A	42.5	91,712	11,939
MS_200B	26.2		
MS_200C	59		
MS_500A	36.8	54,081	5,753
MS_500B	37.8		
MS_500C	38		
MT_20A	40.8	108,620	13,827
MT_20B	36		
MT_20C	34.2		
MT_200A	50	89,601	7,888
MT_200B	60		

MT_200C	58		
MT_500A	60	133,585	13,554
MT_500B	43.7		
MT_500C	60		
