


## RESOURCE ARTICLE

# Evaluating next-generation sequencing (NGS) methods for routine monitoring of wild bees: Metabarcoding, mitogenomics or NGS barcoding

Morgan Gueuning<sup>1,2</sup>  | Dominik Ganser<sup>3,4</sup> | Simon Blaser<sup>1,5</sup> | Matthias Albrecht<sup>4</sup> | Eva Knop<sup>3</sup> | Christophe Praz<sup>2</sup> | Juerg E. Frey<sup>1</sup>

<sup>1</sup>Research Group Molecular Diagnostics, Genomics and Bioinformatics, Agroscope, Wädenswil, Switzerland

<sup>2</sup>Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland

<sup>3</sup>Institute of Ecology and Evolution, University of Bern, Bern, Switzerland

<sup>4</sup>Agroecology and Environment, Agroscope, Zürich, Switzerland

<sup>5</sup>Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute, University of Basel, Basel, Switzerland

## Correspondence

Morgan Gueuning, Research Group Molecular Diagnostics, Genomics and Bioinformatics, Agroscope, Wädenswil, Switzerland.

Email: [morgan.gueuning@unine.ch](mailto:morgan.gueuning@unine.ch)

## Funding information

Swiss Federal Office for Agriculture

## Abstract

Implementing cost-effective monitoring programs for wild bees remains challenging due to the high costs of sampling and specimen identification. To reduce costs, next-generation sequencing (NGS)-based methods have lately been suggested as alternatives to morphology-based identifications. To provide a comprehensive presentation of the advantages and weaknesses of different NGS-based identification methods, we assessed three of the most promising ones, namely metabarcoding, mitogenomics and NGS barcoding. Using a regular monitoring data set (723 specimens identified using morphology), we found that NGS barcoding performed best for both species presence/absence and abundance data, producing only few false positives (3.4%) and no false negatives. In contrast, the proportion of false positives and false negatives was higher using metabarcoding and mitogenomics. Although strong correlations were found between biomass and read numbers, abundance estimates significantly skewed the communities' composition in these two techniques. NGS barcoding recovered the same ecological patterns as morphology. Ecological conclusions based on metabarcoding and mitogenomics were similar to those based on morphology when using presence/absence data, but different when using abundance data. In terms of workload and cost, we show that metabarcoding and NGS barcoding can compete with morphology, but not mitogenomics which was consistently more expensive. Based on these results, we advocate that NGS barcoding is currently the seemliest NGS method for monitoring of wild bees. Furthermore, this method has the advantage of potentially linking DNA sequences with preserved voucher specimens, which enable morphological re-examination and will thus produce verifiable records which can be fed into faunistic databases.

## KEYWORDS

conservation biology, DNA barcoding, insects, molecular identification, pollinators, survey

Christophe Praz and Juerg E. Frey should be considered joint senior author.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Molecular Ecology Resources* Published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

During the last decades, insect pollinators, and especially bees, have declined in several regions of the world (Bartomeus, Stavert, Ward, & Aguado, 2019; Biesmeijer et al., 2006; Burkle, Marlin, & Knight, 2013; Ollerton, Erenler, Edwards, & Crockett, 2014; Potts et al., 2016). While these losses are extensively monitored in managed honeybees (Potts et al., 2010; vanEngelsdorp & Meixner, 2010), less is known on the status, trends and stressors of wild bee populations, as they are more difficult to survey (Goulson, Nicholls, Botias, & Rotheray, 2015; Potts, Biesmeijer, Bommarco, Kleijn, & Scheper, 2015). Due to the lack of adequate cost-effective monitoring programs, trends for the vast majority of European bee species are unknown (Goulson et al., 2015; Nieto et al., 2015; Potts et al., 2010, 2016). Therefore, there is an urgent need for developing and testing comprehensive, robust and systematic monitoring programs that deliver the information needed for policymakers to decide on the most appropriate conservation measures.

To date, most monitoring programs have relied on morphological identifications, which require a sound knowledge of taxonomy and careful analysis of each individual specimen, making it a lengthy and expensive procedure (Lebuhn et al., 2013). The recent advances of “next-generation sequencing” (NGS) techniques offer new opportunities for the assessment of biodiversity (e.g., Schnell et al., 2012; Taberlet, Bonin, Zinger, & Coissac, 2018). Molecular species identifications by DNA barcoding are particularly appealing when classical morphological identifications are not possible [e.g., eDNA, diet assessments; (Rodgers et al., 2017; Taberlet, Coissac, Pompanon, Brochmann, & Willerslev, 2012)], but DNA barcoding has also been suggested for the taxonomical assessment of morphologically identifiable taxa as a mean to reduce costs (Brunner, Fleming, & Frey, 2002; Hebert, Cywinska, Ball, & DeWaard, 2003).

Although DNA-based monitoring methods have emerged only recently, there have been numerous efforts to establish reliable molecular identification pipelines (e.g., Gibson et al., 2015; Ji et al., 2013). For the successful implementation of NGS-identification tools into monitoring programs, the approach should be reliable, reproducible, cost- and time-effective, easily applicable and, ideally, quantitative to enable assessing species abundance (Joseph, Field, Wilcox, & Possingham, 2006). To date, a variety of tools have been developed, and even though most tools have great potential, each is associated with limitations. Presently, most approaches have been assessed in terms of accuracy (species detection and abundance), but only few have been compared with regard to costs and workload (e.g., Elbrecht, Vamos, Meissner, Aroviita, & Leese, 2017). Furthermore, substantial variation in terms of species detection rates and abundance estimates can be observed among studies applying the same molecular methods (although with slightly different parameters), casting doubt on their reproducibility (e.g., see Liu et al., 2013 and Yu et al., 2012 for interstudy variation, or Brandon-Mong et al., 2015 for intrastudy variation). There is thus an urgent need for a comprehensive and reliable benchmark study assessing the strengths

and weaknesses of different methods not only in terms of species detection and abundance estimates, but also in terms of cost and workload. In this study, we assessed and compared three NGS approaches likely to be among the most suitable to be implemented in routine monitoring programs, namely metabarcoding (MB; Taberlet et al., 2012; Yu et al., 2012), mitogenomics (MG; Zhou et al., 2013) and NGS barcoding (NGSB; Shokralla et al., 2014).

As in conventional barcoding, MB relies on the amplification of a taxonomically informative gene fragment (“barcode”). However, the DNA extraction used as template in MB comes from a bulk mixture of specimens (Ji et al., 2013), rendering quantification of species abundance difficult. With NGS methods, abundance inference is generally based on the assumption that the number of output reads correlates with the initial amount of input DNA, a proxy for biomass. Thus, if the biomass of each species in the bulk mixture was known in advance, it should theoretically be possible to infer the number of specimens per operational taxonomical unit (OTU). Nevertheless, due to the very nature of the amplification steps involved in MB, this method can be subject to heavy bias, making quantifications doubtful (Dowle, Pochon, Banks, Shearer, & Wood, 2016; Elbrecht & Leese, 2015; Elbrecht et al., 2016; Piñol, Mir, Gomez-Polo, & Agustí, 2015; Tang et al., 2015; Yu et al., 2012).

To cope with the current lack of solid quantitative output from MB techniques, a PCR-free approach has been suggested (Zhou et al., 2013): MG, also referred to as mitochondrial metagenomics (Crampton-Platt et al., 2015) or mito-metagenomics (Tang et al., 2014), an ultradeep sequencing approach using mitochondrial DNA as a “super-DNA-barcode” (Tang et al., 2015). Derived from bacterial metagenomics, it has been successfully applied for mitochondrial mining on arthropod communities (Choo, Crampton-Platt, & Vogler, 2017; Crampton-Platt et al., 2015; Gillett et al., 2014; Gomez-Rodriguez, Crampton-Platt, Timmermans, Baselga, & Vogler, 2015; Linard, Crampton-Platt, Gillett, Timmermans, & Vogler, 2015; Linard et al., 2018; Liu et al., 2016; Tang et al., 2015, 2014; Wilson, Brandon-Mong, Gan, & Sing, 2019; Zhou et al., 2013). Using total DNA extraction of bulk mixtures, shotgun sequencing on high-throughput NGS platforms is performed and raw data are bioinformatically assembled either *de novo* or mapped to reference databases. MG is not subject to an amplification bias, making it more suitable for quantitative inference (Gomez-Rodriguez et al., 2015; Tang et al., 2015; Zhou et al., 2013). However, even though estimates of species abundance are approaching morphology-based results, MG is still facing methodological limitations, mostly due to the low coverage of target sequences (Crampton-Platt, Yu, Zhou, & Vogler, 2016). Although mitochondria are found in vast copy numbers in animals, mitochondrial DNA (mtDNA) only accounts for a small fraction of the total DNA compared to nuclear sequences. Consequently, the vast majority of data (e.g., 99.47%, in Zhou et al., 2013) produced with MG is not informative, making this approach hardly cost-efficient. Furthermore, as initially presented, MG relies on databases containing full mitogenomes for all investigated species. Because only few full mitogenomes are currently available, this approach is not realistic at this point. To overcome this problem, Gomez-Rodriguez et al.

(2015) compared results obtained using full mitogenomic databases with those obtained using only cytochrome oxidase I (COI) reference databases and found only a slight decrease in species detection and abundance rates in the latter.

In the third method investigated here, NGSB, each specimen is processed separately from extraction to sequencing, unlike in MB and MG (Shokralla et al., 2014). Similar to MB, this method relies on the amplification of a genetic marker, but instead of amplifying from total bulk extracts, PCR amplifications are done individually. Because each specimen is uniquely tagged, this approach is quantitative by design and therefore independent of species biomass information. An additional advantage of this method is that each specimen can be preserved for subsequent identification verification or simply to be archived in natural history collections (Wang, Srivathsan, Foo, Yamane, & Meier, 2018). However, processing all specimens individually increases cost and workload related to the library preparation, which constitutes the main limitation of this approach.

To assess the suitability of these three methods for monitoring purposes, we used a data set collected under regular monitoring conditions. The data were sampled to measure the effectiveness of three different types of flower strips (FS) in promoting wild and managed bees, and the crop pollination services they provide, in Swiss agricultural landscapes. To answer this question, we compared bee species richness and abundance (relative and absolute) found across the three different types of FS. Additionally, we evaluated the influence of plant species richness on wild bee abundance and diversity.

This realistic monitoring data set allowed us to assess the performance of each NGS method with respect to variation levels found among sampling sites under realistic conditions. The number of species and specimens characterizing a data set has a large influence on the overall precision, cost and workload associated with the different NGS methods, which is why estimations of those metrics only make sense with a realistic data set. Finally, using a realistic data set allowed us to determine whether the accuracy level (presence/absence, relative and absolute abundance) of the explored methods would allow us to detect ecological patterns and reach similar conclusions, and thus validate their use in monitoring programs.

Overall, in this study we compared (a) species detection rates (presence/absence data only), (b) relative and absolute species abundances, (c) ecological patterns and finally (d) costs and workload of the three different NGS-identification methods outlined above compared to morphological identification.

## 2 | MATERIAL AND METHODS

### 2.1 | Sampling

The data set (sampling material) used in this study was collected in 2017 in agricultural landscapes of the central Swiss Midland. The sampling scheme was designed to identify the effectiveness of three types of sown FS for providing foraging resources to pollinators. In total, 20 different FS were sampled three times over the flower season (two FS were collected four times and one FS two times).

FS were sown either in April 2013 (FS type 1,  $n = 8$ ), April 2016 (FS type 2,  $n = 8$ ), and September 2016 (FS type 3,  $n = 8$ ). All three types of FS harboured unique floral mixtures, composed of species of annual (all three types) and perennial flowering plants (types 1 and 2), which were primarily selected due to their high pollen and nectar production.

To be able to obtain quantitative information on the number of pollinators present at each sampling round, a strict sweep-netting protocol was applied. During each sampling round, transects were slowly walked up while sweeping two times 25 sweeps with one-minute pause in between. After 50 sweeps, the collected material was transferred into a plastic bag and directly stored at  $-20^{\circ}\text{C}$  in a portable freezer.

Furthermore, during each sampling round, we monitored plant species richness, allowing us to additionally assess the importance of this parameter in promoting bees.

To determine the degree of variation within each FS, the exact same protocol was repeated within the same FS after five minutes (hereafter referred to as "transect I" and "transect II"). Transect II started from the end point of transect I. In total, the data set encompasses 122 sampling points [hereafter referred to as "communities":  $(17 \text{ FS} \times 3 \text{ sampling rounds} \times 2 \text{ transects}) + (2 \text{ FS} \times 4 \text{ sampling rounds} \times 2 \text{ transects}) + (1 \text{ FS} \times 2 \text{ sampling rounds} \times 2 \text{ transects})$ ].

## 2.2 | Identification methods

### 2.2.1 | Morphological identification

In the laboratory, raw sampling material was sorted to isolate wild bees from plant material, other insects, as well as honeybee workers. Each specimen ( $n = 723$ ) was then pinned, labelled, dried for at least 72 hr in a desiccator containing silica gel and identified by an expert. Most specimens were identified to species-level, but in the following cases, morphological identifications were performed to species-group level: *Bombus terrestris* group for workers belonging to *B. terrestris*, *B. lucorum* and *B. cryptarum*; *Halictus simplex* group for females of *H. simplex*, *H. langobardicus* and *H. eurygnathus*; and *Andrena ovatula* group for females of *A. ovatula* and *A. wilkella*. Morphological identification was complemented by Sanger sequencing using COI barcoding for all specimens identified to species-group level and not to species level ( $n = 29$ ) or left undetermined because of lack of intact morphological criteria ( $n = 11$ ). For clarity, we still refer to this data set as "morphological" even if for a limited number of specimens, morphological identifications have been complemented using Sanger sequencing. Details of the Sanger sequencing protocol are given in Supporting Information S1.

### 2.2.2 | Metabarcoding

Bulk DNA extractions were performed on each community using a proteinase K solution and digested overnight at  $56^{\circ}\text{C}$ . Volumes of proteinase K solutions were adapted according to the number of specimens per community so that all specimens were immersed

into the solution. To reduce costs linked to commercial kits, we purified the extracts following the Canadian Center for DNA Barcoding (CCDB) DNA extraction protocol (Ivanova, Dewaard, & Hebert, 2006). For each community, to increase species detection rates and normalize template abundance, DNA purifications were performed in triplicates and immediately pooled after extraction. To reduce workflow and limit numbers of PCRs required during the library preparation, amplification was carried out using fusion primers. In addition to the priming sequence, fusion primers have overhangs composed of Illumina indexes and a unique tag of eight base pairs (bp) designed using the software Barcode generator (Meyer & Kircher, 2010). The overhangs allow amplicons to be directly loaded onto the Illumina sequencer. To overcome the inherent limitation of Illumina platforms in sequencing low complexity libraries, we added a "heterogeneity spacer" between the labelled tag and the priming sequence, as recommended in Fadrosch et al. (2014). The PCR primer sequences of the fusion primers were those of mCOLintF and of HCO2198 (Leray et al., 2013) and targeted a 313-bp region of the COI gene. Overall, forward and reverse primers were 95 bp long ( $\pm 3$  bp). Per community, bulk amplification was performed in five different PCR replicates, each harbouring a unique combination of forward and reverse tags. Further details on MB library preparation are given in Supporting Information S2. Final library was sequenced on an Illumina MiSeq using a v3 kit ( $2 \times 300$  bp) and spiked with 20% Phix.

The majority of bioinformatics analyses (detailed in Supporting Information S3) were performed using QIIME1 (Caporaso et al., 2010). Briefly, raw data were trimmed based upon the FASTQC profile before joining paired-end reads. After demultiplexing, adaptors, spacers and primer sequences were trimmed. Chimeric sequences were identified de novo and removed using usearch61 (Edgar, 2010). Filtered sequences were then clustered using the UCLUST algorithm (Edgar, 2010) at the default similarity threshold of 97%. Taxonomical assignment of OTUs was performed using the same algorithm by fitting reads to reference sequences. To determine the impact of database quality on the species detection performance, OTUs were assigned using two separate COI databases. The first database ("uncurated") encompassed all available COI sequences of bee species (barcodes for ca. 2,000 species) available on BOLD (Barcode of Life Database) and GenBank (downloaded in June 2017). Additional verifications were made to ensure the presence of multiple barcodes ( $n \geq 3$ ) for all species present in our data set. The second database ("curated") was downloaded from BOLD and corresponds to sequences deposited by Schmidt, Schmid-Egger, Morinière, Haszprunar and Hebert (2015) in their extensive barcoding study on western European bees (dx.doi.org/10.5883/DS-GBAPI). This data set was initially missing barcodes of two species present in our data set (i.e., *Andrena flavipes* and *Chelostoma florissomne*), and barcodes for these two species were downloaded from other projects on BOLD and manually added to the database. Similarly, to determine the best similarity threshold, the MB bioinformatic pipeline was run several times using different similarities thresholds (from 90% [default] to 99%). Corresponding community

matrices were compared to the morphological community matrix, and the threshold performing best was retained for downstream analyses. The same empirical approach was applied to determine the optimal cross-validation setting among replicates (i.e., minimal occurrence of a species among replicates to be validated).

### 2.2.3 | Mitogenomics

Aliquots of the DNA extracts used for MB (prior to library preparation) were sheared using an ultrasonicator (Bioruptor). The MG library was built using a commercial Illumina 96 TruSeq DNA Nano kit following the manufacturer's recommendations. To reduce differences in sequencing depth, we homogenized sequencing depth on the number of specimens per community by applying the same correction factor as for MB (Supporting Information S2). The library was sequenced on an Illumina MiSeq using a v3 kit ( $2 \times 300$  bp) and spiked with 1% Phix.

Two different bioinformatics approaches were compared [i.e., (a) de novo assembly and (b) raw read mapping], and the approach recovering the highest number of species was retained for downstream analyses. (a) The de novo assembly approach mainly followed Crampton-Platt et al. (2015). Details are given in Supporting Information S3; briefly, libraries were quality assessed using FASTQC and residual adaptors trimmed with Trimmomatic (Bolger, Lohse, & Usadel, 2014). Then, libraries were filtered to retain only mitochondrial reads using BLASTN (Camacho et al., 2009) and a database containing all publically available (partial and full) mitogenomes of bee species (336 mitogenomes of 82 species; among which 18 present in our data set). Putative mtDNA reads were then assembled using IDBA-UD (Peng, Leung, Yiu, & Chin, 2012) with a 98% similarity threshold. Contigs were mapped at a 98% similarity against a custom database using BMap (Bushnell, 2015). Since only 18 reference mitogenomes were available for the investigated species, additional COI barcodes from the curated COI database were added to the mitogenome database. Finally, SAMtools (Li et al., 2009) was used to index and extract the number of reads that mapped reference sequences. (b) The raw read mapping approach relied on BMap (Bushnell, 2015) to map unfiltered reads against COI reference sequences. Because only a small fraction of sequences will match to the COI reference database, it is crucial for this approach that the database is not only comprehensive, but also well curated. The presence of uncurated sequences (e.g., numts) will have a major influence upon the outcome, much more than for amplicon-based approach where coverage-based filtering will in most cases obliterate errors originated from the database. Therefore, only the curated database was used in this approach. To further reduce false positives due to mapping of reads in the flanking regions of COI, sequencing spanning over the classical 658-bp COI barcoding region was filtered out of the curated database. As in Tang et al. (2015), a high similarity threshold (99%) was used to reduce false positives and reads were mapped once. Mapped reads were indexed and extracted using SAMtools (Li et al., 2009).

## 2.2.4 | NGS barcoding

Before performing bulk DNA extractions described above, a single leg of each specimen was taken for DNA extraction (one extraction per specimen) following the CCDB protocol. As for MB, fusion primers were used to amplify individually all extractions and PCRs were conducted following the same conditions as for MB. After amplification, each PCR product was examined on a 1.5% agarose gel and amplicons were pooled equimolarly as estimated based on their amplification intensity. Pooled PCR products were purified with NucleoFast 96 PCR clean-up kits (Macherey-Nagel) using 300  $\mu$ l of PCR product per well and eluted in 100  $\mu$ l ddH<sub>2</sub>O. Cleaned PCR products were sequenced on an Illumina MiSeq using a v3 kit (300 bp  $\times$  2) spiked with 20% Phix.

Data processing of the NGSB library is similar to the MB procedure. The filtered reads were clustered using UCLUST at a similarity threshold of 99%, and OTUs were taxonomically assigned using the same algorithm but with a default threshold parameter (90%). A lower taxonomical assignment threshold than for MB was used to decrease the number of unassigned OTUs since only the most abundant species assignment per specimen was retained in the final matrix. The number of false positives was therefore not affected by this lower threshold. As for MB, taxonomical assignments of OTUs were performed using the two different databases (curated and uncurated).

## 2.3 | Data analyses

### 2.3.1 | Species richness

For all NGS methods, we compared species richness with morphological species richness for each community and assessed species detection rates using the Jaccard similarity index (Jaccard, 1912). To determine variation between two transects collected five minutes apart within the same FS, we also computed the Jaccard index between the samples identified based on morphology.

### 2.3.2 | Quantitative inference

In this study, species quantification (relative and absolute abundance) for both bulk methods (i.e., MB and MG) was defined as a measure of the species biomass and not numbers of specimens per species. To assess quantification accuracy for MB and MG, we correlated the number of reads per species (ln-transformed) with the corresponding species biomass measurements. For solitary bees, dry weight can be accurately estimated by the following exponential relationship (Cane, 1987):  $y = 0.77(x)^{0.405}$ , where  $y$  is the shortest linear distance between the wing plates (intertegular distance; mm) and  $x$  is the dry weight (mg). A photograph was taken of each specimen using a stereomicroscope-mounted camera (Leica M4000), and intertegular distance was measured, which enabled to measure biomass for each specimen. To compare quantitative data on the number of specimens per species among all methods, we transformed

the morphological absolute abundance (number of specimens per species) into relative abundance of biomass.

### 2.3.3 | Comparison of ecological patterns

To determine whether the detected ecological patterns would be similar across our three NGS approaches as well as the classic morphological approach, we applied the same statistical analyses on presence/absence data and on relative and absolute abundance data. First, to explore how much of the observed variance in species composition across sampling sites was explained by the identification method, we performed a nonparametric multivariate analysis of variance using distance matrices [i.e., PERMANOVA; (Anderson, 2001)]. The same test was also performed on the morphological data set to determine the biological variance found between the two transects sampled five minutes apart. These PERMANOVA tests (*adonis* function in the R cran VEGAN package) were performed using the Jaccard dissimilarity index for presence/absence data and the Bray–Curtis distance dissimilarity index for both relative and absolute abundance data. All *adonis* analyses were run with 10,000 permutations. Second, to complement the *adonis* analyses, we performed nonmetric multidimensional scaling (NMDS) to visualize and compare community compositions of FS among the identification methods. The goodness of fit between the superimposed shapes of NMDS plots was assessed by Procrustes tests computed with the *protest* function (VEGAN package). The NMDS analyses were performed with the *metaMDS* function implemented in the VEGAN package with the *noshare* function activated to use extended dissimilarities when sampling sites did not share species. “Spider” diagrams were added to connect communities sharing the same FS type. Third, to determine and compare the effectiveness of the three different types of FS in promoting wild bees, we ran linear mixed models (LMM) and generalized linear mixed models (GLMM) using the LME4 package (Bates, Mächler, Bolker, & Walker, 2015). Species richness and species abundance (relative and absolute) were used as response variables (see details of models in Supporting Information S12). Finally, to determine the importance of flower richness on promoting wild bees, we applied similar models with the predictor variable being the interaction between plant species richness and identification method. The relationship between plant species richness and bee richness or abundance was plotted using linear regressions with 95% confidence intervals.

## 2.4 | Cost and workload

Costs estimates are based upon suppliers' prices applied in 2018 in Switzerland and do not contain cost linked to workload. To compensate for the cost of wet laboratory consumables, overall costs were increased by 15%. For the morphological identifications, the workload includes mounting, labelling and databasing of the specimens and the cost corresponds to the identifications performed by the taxonomist. Regarding the workload estimate for NGS methods,

only hands-on laboratory processes were recorded, leaving out time needed for overnight digestions, PCR amplifications, electrophoresis or other incubation times.

To predict the relationship between overall cost and total number of specimens, we divided the price per specimen into fixed (i.e., independent from the number of specimens) and variable costs (dependent on the number of specimens). For the three NGS methods, we thus subtracted the cost of the sequencing kit (variable cost) to the grand total and divided the result by the number of specimens (fixed cost). Cost estimates for morphological identifications only included fixed costs.

Finally, since Illumina platforms offer the possibility to run different kits harbouring variable outputs, we estimated the overall cost and sequencing depth for all kits allowing to span our targeted read length (~ 450 bp; including tags and technical sequences) for MB and NGSB, namely the MiSeq v3 (2 × 300 bp), MiSeq v2 (2 × 250 bp) and the MiSeq v2 Nano (2 × 250 bp) kit; and for MG, the MiSeq v3 (2 × 300 bp), HiSeq 4000 (1 × 50 bp) and HiSeq 4000 (2 × 75) kit.

## 3 | RESULTS

### 3.1 | Morphological identification

Wild bees were found in 83 of the 122 sampling points. After sorting wild bees from the honeybees ( $n = 1,422$  honeybees) and other arthropods (mainly aphids, dipterans and coleopterans), we counted 723 wild bee specimens. A total of 683 specimens were identified morphologically to species level, 29 to species-group level (among which 20 were identified as workers from the *B. terrestris* group), and 11 remained unidentified. Sanger sequencing, used as complement for the identification to the species level of the species-groups and undetermined specimens, was successful for 39 of 40 specimens. The one unidentified specimen for whom Sanger sequencing failed was classified as “unidentified”.

The morphological data set, complemented with Sanger sequencing, comprised 723 specimens and 58 species, of which 382 specimens belonged to the transects I and 341 to transects II (Supporting Information S4). The median number of specimens per community was 5 and the mean ( $\pm SD$ ) number 8.71 ( $\pm 10.12$ ), with a minimum of 1 and a maximum of 55 specimens.

### 3.2 | Sequencing outputs

The MiSeq runs produced 13.8, 17.5 and 9.0 million reads, respectively, for the MB, MG and NGSB libraries (Supporting Information S5). After read merging, demultiplexing and data filtering, the MB and NGSB data sets encompassed respectively 4.5 and 3.4 million reads. Raw reads from the MG library were not filtered but directly mapped to the COI reference database. In total, 28.26%, 0.02% and 32.22% of reads mapped to the database, for MB, MG and NGSB, respectively. To estimate the average coverage per specimen and community, the number of mapped reads was divided by either the number of specimens ( $n = 723$ ) or the number of communities

( $n = 83$ ). On average, the number of reads per specimen was 5,450, 4 and 3,959 for MB, MG and NGSB, respectively, and 47,471, 38, 34,485 per community, respectively.

### 3.3 | Impact of the quality of the COI reference databases in MB and NGSB

For both MB and NGSB, species detection rates were higher while using the uncurated COI reference database (Supporting Information S6). The use of this database uncovered more true positives and decreased the number of false negatives. For NGSB, using the uncurated database, however, introduced one supplementary false positive. Based on these results, the uncurated database was used for all subsequent analyses.

### 3.4 | MB parameters

Similarity thresholds for the taxonomical assignment of OTUs considerably influenced the overall number of false positives and negatives (Supporting Information S7A). The similarity threshold providing the highest species detection rates (Jaccard similarity index) was 97% and 98%. Since species detection rates were similar for 98% and 97%, the more widely accepted threshold of 97% was favoured and used in all subsequent analyses. At this threshold, the mean percentage of unassigned OTUs per community was 18.1% (Supporting Information S8).

Cross-validation thresholds had a lesser effect and produced similar number of false positives and false negatives when validating species present in at least 1, 2, 3, 4 or 5 out of 5 replicates (Supporting Information S7B.). The less stringent thresholds (i.e., 1/5 and 2/5) introduced one additional false positive while the correlation between biomass and read numbers was slightly higher than for the more stringent thresholds. Because the higher correlation between biomass and read number did not reduce the overall difference found between the morphological and MB matrices, and because this less stringent threshold slightly increased the false-positive rate, the more conservative threshold of three out of five replicates was favoured and used for subsequent analyses.

### 3.5 | MG pipelines

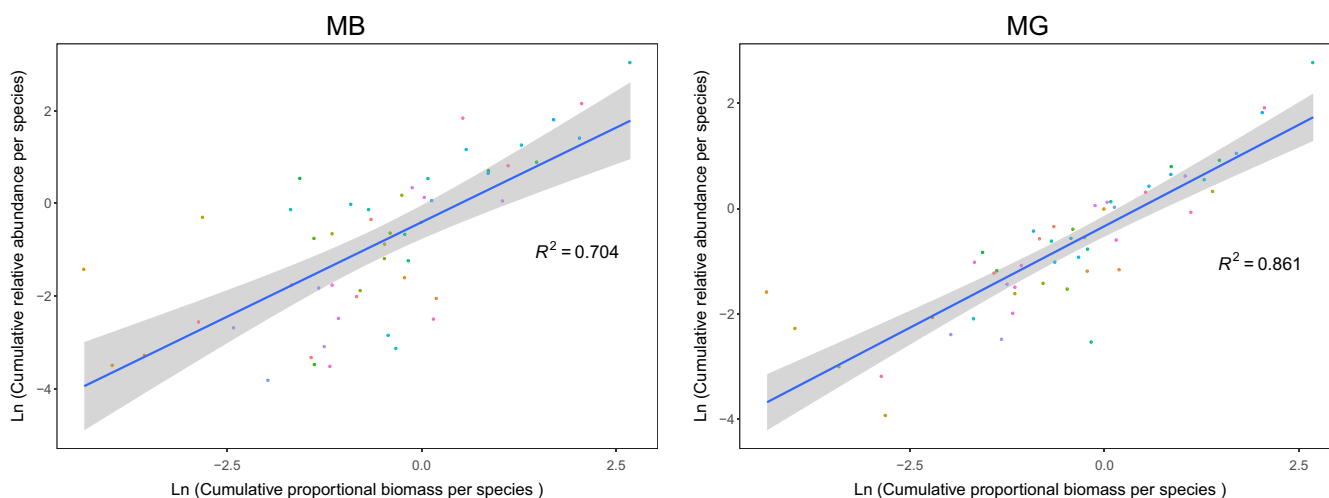
The Jaccard index for the de novo assembly pipeline was considerably lower than for the raw mapping pipeline (Supporting Information S9). The former pipeline uncovered 17 true positives whereas the latter 53 true positives. Based on these results, the raw mapping pipeline was favoured for downstream analyses.

### 3.6 | Species richness

The Jaccard similarity index between morphological and NGS data sets was highest for NGSB, followed by MB and MG (Table 1). For NGSB, all species present in the morphological data set were recovered and only two additional species (false positives) were identified (Supporting Information S4). The number of false negatives

**TABLE 1** Jaccard similarity index between the global diversity of morphological (Morpho) and molecular (MB, MG and NGSB) data sets. Similarity indexes per transect for the molecular methods are given in Supporting Information S10

Data sets	Transects	Species richness	# Shared species	False positives	False negatives	Jaccard index
Between transects of Morpho	I	43	30 (30/43 = 69.8%)	—	—	0.508
	II	46	30 (30/46 = 65.2%)	—	—	0.508
	I and II	58	—	—	—	—
Between MB and Morpho	I and II	57	53 (53/57 = 93.0%)	4 (4/57 = 7.0%)	5 (5/57 = 8.8%)	0.855
Between MG and Morpho	I and II	69	53 (76.8%)	16 (23.2%)	5 (7.5%)	0.716
Between NGSB and Morpho	I and II	60	58 (96.7%)	2 (3.4%)	0 (0%)	0.967

**FIGURE 1** Correlation between the ln-transformed relative read number per bee species and the ln-transformed estimate proportional biomass per species for metabarcoding and mitogenomics data sets. Grey areas represent the 95% confidence interval. Proportions were cumulated across all sampling sites. Each coloured dot represents a different species. Correlations were significant with  $p$ -values < 0.0001 [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

was similar for MB ( $n = 5$ ) and MG ( $n = 5$ ), although MG harboured substantially more false positives ( $n = 16$ ) than MB ( $n = 4$ ). There was no clear overlap in species identity between the false positives and negatives found in those two methods (Supporting Information S4). The lowest Jaccard similarity index was found among transects of the morphological identification method (Jaccard index = 0.508). Jaccard indexes between transects of each NGS method were consistently close (Supporting Information S10).

### 3.7 | Quantitative inference

Individual species biomass (as computed based on morphological identifications and measured intertegular distances) was significantly correlated with the sequencing output for both MB and MG for relative and absolute abundance (Figure 1, Supporting Information S11). For both NGS methods, correlations were higher when using relative abundance than absolute abundance. MG displayed higher correlation coefficients than MB, especially for relative abundance (Figure 1).

### 3.8 | Ecological patterns

PERMANOVA tests, performed to analyse and quantify differences in community compositions between NGS and morphological data sets, revealed significant differences in the abundance data for both MB and MG, but not for NGSB (Table 2). With presence/absence data, the differences were significant only for MG data sets. Overall, the identification method explained 0.1%, 9.0% and 10.7% of the variance found compared to the morphological data set for MPS, MB and MG, respectively.

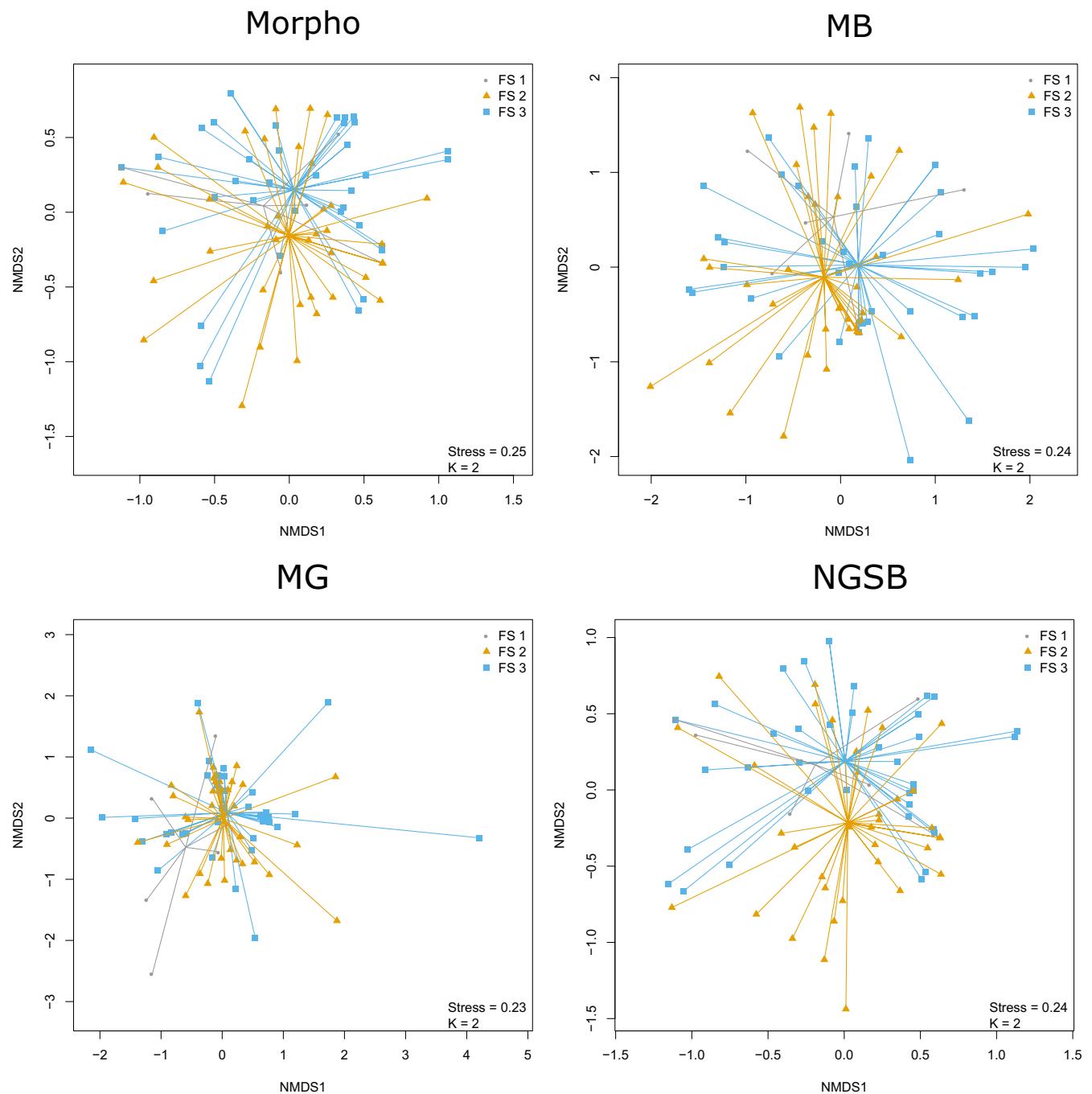
The NMDS ordinations showed similarities in community composition across the morphological and the NGS methods (Figure 2, Supporting Information S12). This was especially true for the NGSB data sets for whom the Procrustes tests revealed highly similar community compositions to the morphological one (Table 2). For the MB and MG data sets, Procrustes tests also depicted significant correlations with the morphological data set in community composition, although with lower correlation coefficients. As in PERMANOVA analyses, the lowest correlation coefficient for MB and MG was found with absolute abundance data.

**TABLE 2** Nonparametric multivariate analysis of variance on distance matrices (PERMANOVA) using the *adonis* function and Procrustes test (*protest* function) between NMDS of molecular (MB, MG and NGSB) and morphological (Morpho) identifications. Jaccard dissimilarity index was used to transform the presence/absence data sets and the Bray-Curtis index for both abundance formats. For the morphological identified data set, the PERMANOVA test was performed between transects

Method	Test	Levels	Presence/absence			Relative abundance			Absolute abundance					
			df	F model	R <sup>2</sup>	p Value	df	F model	R <sup>2</sup>	p Value	df	F model	R <sup>2</sup>	p value
Morpho	PERMANOVA	Transect	1	0.729	0.009	0.796	1	0.617	0.008	0.850	1	0.617	0.008	0.852
		Residuals	80		0.991		80		0.992		80		0.992	
		Total	81		1.000		81		1.000		81		1.000	
MB	PERMANOVA	Identification	1	0.760	0.005	0.727	1	3.614	0.022	<0.001	1	15.809	0.088	<0.001
		Residuals	164		0.995		164		0.978		164		0.912	
		Total	165		1.000		165		1.000		165		1.000	
	Procrustes			0.803	<b>0.001</b>			0.819	<b>0.001</b>			0.783	<b>0.001</b>	
MG	PERMANOVA	Identification	1	19.044	0.106	<0.001	1	10.974	0.064	<0.001	1	15.625	0.089	<0.001
		Residuals	160		0.894		160		0.936		160		0.911	
		Total	161		1.000		161		1.000		161		1.000	
	Procrustes			0.543	<b>0.001</b>			0.651	<b>0.001</b>			0.350	<b>0.001</b>	
NGSB	PERMANOVA	Identification	1	0.207	0.001	1.000	1	0.251	0.001	0.995	1	0.228	0.001	0.998
		Residuals	164		0.999		164		0.999		164		0.999	
		Total	165		1.000		165		1.000		165		1.000	
	Procrustes			0.934	<b>0.001</b>			0.854	<b>0.001</b>			0.900	<b>0.001</b>	

Note. *p*-Values under the 0.05 threshold are in bold.

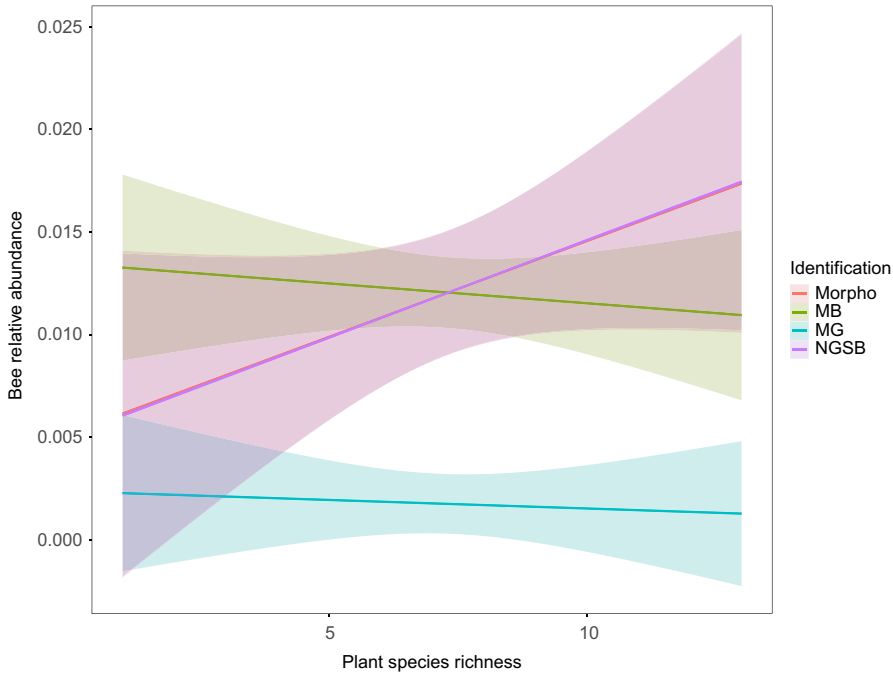




**FIGURE 2** Nonmetric multidimensional scaling (NMDS) of bees' relative abundance obtained by four different species identification methods. The NMDS analyses were performed using the Bray–Curtis index with the *metaMDS* function implemented in the *VEGAN* package. “Spider” diagrams connect communities sharing the same flower stripes (FS) type. Goodness of fit between the superimposed shapes of the molecular NMDS plots with the corresponding morphological NMDS plots was assessed using Procrustes tests, computed with the *protest* function (*VEGAN* package) (see Table 2). Note the close similarity between data sets based on morphology and NGSB [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

While testing for difference in bee species richness or abundance among the three different types of FS, the GLMM (presence/absence) and LMM (relative and absolute abundance) analyses depicted no statistical difference among FS types for all identification methods (Figure 4, Supporting Information S13). Similarly, using the plant species richness as predictor, all identification methods showed comparable relationships between plant

species richness and bee species richness (Supporting Information S14–S15). However, the relationships between plant species richness and bee relative abundance were significantly different from the morphological data set for MB or MG (Figure 3, Supporting Information S15). Indeed, MB and MG showed a negative relationship between bee abundance and plant species richness whereas this relationship was positive for the morphological and NGSB

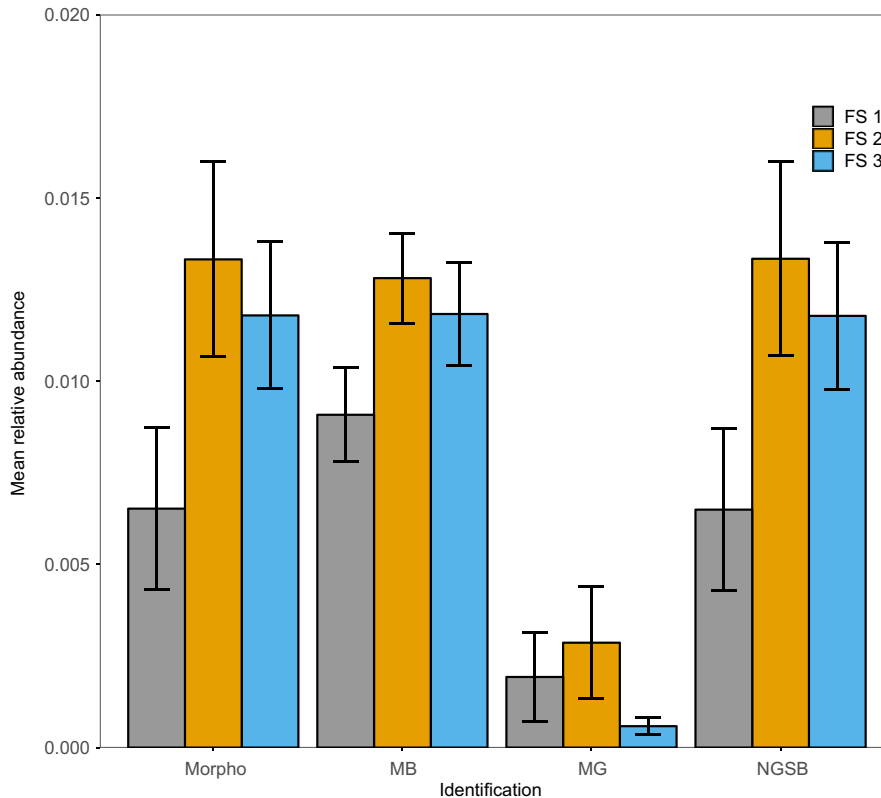


**FIGURE 3** Relationship between plant species richness and the relative abundance of bees for different identification methods. Lines were computed by linear regressions as implemented in ggplot2. Coloured areas represent the 95% confidence interval. Statistical differences in relationships of the molecular identification method compared to the morphological identification method were assessed by linear mixed models. For bee species richness, no difference in relationship was found between the morphology and NGSB (regressions overlap), while MB and MG showed significant deviation compared to the relationship based on morphological identifications (See Supporting Information S15 for LMM results) [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

data sets. Furthermore, MG overall underestimated the bee relative abundance, whereas MB overestimated it for plots low in species abundance and underestimated it for species-rich abundant plots (Figure 3).

### 3.9 | Cost and workload

With respect to cost, morphological identification was approximately half the price of the cheapest NGS-identification method



**FIGURE 4** Mean relative abundance of bees for three different types of flowering strips (FS). Means were computed per identification methods, and error bars correspond to the mean standard error. Statistical difference among means within each identification method was assessed with linear mixed models. No statistical difference among types of FS was found within method [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

(MB) and approximately three times cheaper than the priciest one (MG) (Supporting Information S16), when cost was estimated based on the number of specimens included in this study. For all investigated NGS methods, the sequencing kits used in this study represented the principal fraction of the overall cost. Since the sequencing kit cost is independent from the number of specimens sequenced (as long as the desired sequencing depth is reached), we calculated costs with increasing number of specimens. Based on this calculation, after approximately 1,795 and 4,639 specimens, MB and NGSB would become respectively more cost-efficient than morphology-based identifications for the MiSeq v3 kits (Supporting Information S17, see Supporting Information S18 for cost details). Because of sequencing depth limitations, MG stayed largely costlier than morphological identification. Alternatively, instead of increasing specimen numbers, cost could be reduced by using smaller, less expensive sequencing kits. Based on the mean sequencing depth of MB and NGSB, we estimated the coverage and overall cost for two alternative kits (Miseq v2 [ $2 \times 250$  bp] and MiSeq v2 Nano [ $2 \times 250$  bp]; Supporting Information S19). Although sequencing depth attained in this study for MB and NGSB was slightly sub-optimal (Supporting Information S4), coverage estimations based on these figures suggest sufficient sequencing depth, even for the smallest sequencing kits (Supporting Information S19).

Regarding workload, MB was the identification method requiring less workload. Morphological and MG required similar workloads and NGSB moderately more (Supporting Information S15).

## 4 | DISCUSSION

Overall, our results show that (a) NGSB provided the data set most similar to the morphological data set, both in terms of species detection and abundance. (b) As predicted, the correlation between biomass and read numbers was stronger for MG than for MB. Nevertheless, MG produced more false positives (23.2% against 7.0% for MB) and therefore considerably decreased similarities in community compositions compared to the morphological data set. (c) For both MB and MG, species abundance estimates were better when using relative abundance than absolute abundance. (d) Ecological patterns were similar across all identification methods when using presence/absence data. However, when using abundance data (both relative and absolute), the conclusions based on MB and MG identification, but not NGSB, differed from those based on morphology; (e) finally, the overall cost of all three NGS methods was higher than morphological identifications. However, MB and NGSB become more cost-effective by either using smaller sequencing kits (e.g., MiSeq v2 Nano kit) or by increasing specimen numbers. Hereafter, we summarize the advantages and weaknesses of each NGS method.

### 4.1 | Metabarcoding

Since Taberlet et al. (2012) proposed MB as a modern tool for assessing biodiversity, MB has been widely accepted when alternative

means of species identification are lacking (e.g., eDNA, diet analyses). However, for cases where morphological identification is possible (e.g., pollinators surveys), MB is still in a validation phase. To date, the vast majority of MB studies have been tested against laboratory-assembled communities of known composition (e.g., Elbrecht & Leese, 2015; Elbrecht et al., 2016; Piñol et al., 2015; Tang et al., 2015; Yu et al., 2012), and the reported detection rates are highly variable. For instance, Tang et al. (2015) compared the accuracy of MB and MG on a data set taxonomically similar to ours (33 wild bee species represented by 250 specimens) and found as many as 11 false negatives and 49 false positives, for 53 true positives. Based on these figures, the Jaccard similarity index between morphological and MB identification would be 0.47.

As illustrated in the study by Tang et al. (2015), MB detection rates are frequently obliterated by high numbers of false positives and negatives (Gentile Francesco Ficetola, Taberlet, & Coissac, 2016), a problem that strongly biases the overall interpretation of species detectability (Lahoz-Monfort, Guillera-Aroita, & Tingley, 2016). To overcome this limitation, replicates are crucial (Mata et al., 2019). Although it is possible to estimate the number of required replicates (Ficetola et al., 2015), the optimal replication level largely depends on the data set. In our study, we empirically tested different settings and observed no major differences among them. Although detection rates may vary across studies, to our knowledge, all rates of species detection were under 100%. Because a perfect match between NGS and morphological identification is illusive, Ji et al. (2013) investigated the effect of such discrepancies on policymaking and management issues. To do so, they compared MB with standard morphology-based data sets and found that both exhibited similar alpha- and beta-diversities, leading to similar policy conclusions. Although insightful and pioneering, this study was conducted on a very large data set (55,813 arthropods and bird specimens) in which small variations in species presence/absence would be unlikely to have a strong influence. Applying a similar approach to our much smaller data set resulted in similar conclusions: morphological and MB data sets exhibited similar species composition (Table 2), revealing similar ecological patterns with (a) no differences in bee species richness among the three different types of FS and (b) similar positive relationships between plant species richness and bee species richness (Supporting Information S14–S15).

Nevertheless, these conclusions are based on presence/absence data while the majority of monitoring programs rely on species abundance data, which gives a more precise picture of community composition (Joseph et al., 2006; MacKenzie, 2005). Therefore, there has been numerous efforts to foster the reliability of MB species count, and currently, there is an equal number of studies claiming or disclaiming quantification reliability (see Piñol, Senar, & Symondson, 2019). A study investigating the variability in quantification recorded the level of variance in read numbers associated with individual nematodes between PCR/library replicates and found an overall very consistent read count per individual ( $R^2 = 0.99$ ) (Porazinska, Sung, Giblin-Davis, & Thomas, 2010). However, their results also highlighted consistent variance in read numbers among species, even after correcting for their body size. In a similar attempt to uncover

variation sources in read quantification at the interspecies level, Elbrecht and Leese (2015) sequenced libraries build with the exact same biomass of different species and found substantial differences in read abundance among species (up to four times higher or lower read abundances). These results underline an inherent problem linked to PCR-based techniques, that is the primers' species-specific efficiency. PCR amplification efficiency is primarily (73%) influenced by the number of template-primer mismatches (Piñol et al., 2015), and therefore, the selection of primers will greatly influence the quantitative output (Piñol et al., 2019). While testing 15 common universal COI primer pairs, Piñol et al. (2019) found a significant relationship between DNA concentration pre- and post-PCR for the vast majority of primers (14/15) although  $R^2$  values were variable. The primer pair used in our study performed relatively well, even though other primers performed better (e.g., ArF5 & ArR5, Gibson et al., 2014). The problems outlined above likely contribute to the large differences in quantification inference reported in the literature. Furthermore, bulk-based approaches might inform on the biomass, but not necessarily on specimen numbers because of intraspecific biomass variations (e.g., sex or "cast" polymorphism in social bees).

In our study, we found strong correlations between read numbers and estimated biomass, especially when using relative abundance data (up to  $R^2 = 0.704$ ; Figure 1). The beta-diversity of MB and morphological data sets was also highly similar for relative abundance with only 2.2% variance explained by the identification methods alone (Table 2). Furthermore, the Procrustes test depicted a relatively high correlation between the NMDS shapes of the MB and morphological data sets ( $R^2 = 0.819$ ,  $p$ -value < 0.001; Figure 2). Although these results are promising, we still found evidence of a bias introduced because of quantitative inference. Indeed, the LMM analysis depicted contrasting relationships between plant species richness and bee relative abundance depending on the identification method (Figure 3, Supporting Information S15); while the relationship between these variables was positive for the morphological data set, it was slightly negative for the MB data set (Figure 3). These results show that regardless of high correlations between estimated biomass and inferred abundance in morphology and MB, the overall ecological patterns are skewed by a biased estimate of species abundance, ultimately leading to incorrect ecological conclusions.

## 4.2 | Mitogenomics

As initially suggested by Zhou et al. (2013) and several follow-up studies (Gomez-Rodriguez et al., 2015; Tang et al., 2015), we corroborate that quantitative inference based on biomass is less biased with a PCR-free approach: regardless of the quantitative community format, Pearson's correlations were higher for MG (relative abundance:  $R^2 = 0.861$ ; absolute abundance:  $R^2 = 0.623$ ) than MB (relative abundance:  $R^2 = 0.704$ ; absolute abundance:  $R^2 = 0.549$ ) (Figure 1, Supporting Information S11). Interestingly, the correlation coefficients found in our study are similar to those found in other MG analyses (Zhou et al., 2013:  $R^2 = 0.64$ ; Gomez-Rodriguez et al., 2015:  $R^2 = 0.69$ ; but see Tang et al., 2015:  $R^2 = 0.25$ ).

While estimates of biomass appear to be more reliable and precise when using MG, the higher number of false positives and negatives (Table 1) skewed the overall species composition and introduced greater variance than with MB (Table 2). Although often claimed as less prone to false positives and negatives than PCR-based methods (Tang et al., 2015; Zhou et al., 2013), we nevertheless found in our study substantially more false positives (23.2%) than with MB (7.0%). We argue that these high rates could mainly be attributed to two factors: the reference database and the low coverage. First, the database used in our study featured sequences for considerably more species (>450 species) than present in our data set (58 species). This approach was favoured to mimic monitoring conditions with limited a priori knowledge on species richness. To date, previous studies often opted for a more conservative approach and used the same DNA extracts for building the reference databases and the NGS library (e.g., Gomez-Rodriguez et al., 2015), which most likely increases the mapping success. Additionally, using a full mitogenomes reference database has been shown to slightly decrease the false negatives and positives rates (Gomez-Rodriguez et al., 2015), but is presently illusive for monitoring purposes due to the lack of published and annotated mitogenomes. In our study, the reduced number of false positives found with the de novo assembly approach (Supporting Information S9) also indicates that an exhaustive database can considerably improve the outcome of MG. Second, higher coverage rates could help reducing false discovery rates by filtering out all mappings under a certain threshold or by adding replicates to cross-validate species presence/absence as we did here on the MB data set. In general, sequencing depth is a major limitation for MG as the vast majority of sequences produced with MG do not correspond to mitochondrial sequences and are therefore currently uninformative (although see Linard et al., 2015). In our study, approximately 0.02% of all reads mapped to the COI reference database for the raw read mapping pipeline (Supporting Information S5). For the de novo assembly pipeline, approximately 5% of the reads were mapped to the mtDNA reference database. Using full-mitogenome databases unsurprisingly increases the overall percentage of mapped reads, but in most cases, the mitochondrial fraction will nevertheless plateau around 1% (see review on MG by Crampton-Platt et al., 2016).

Despite these limitations, this PCR-free method has the advantage of not relying on taxon-specific primers and is therefore universally applicable to any group of animal, or even to plants, fungi or bacteria if other organelles or genes are targeted.

## 4.3 | NGS barcoding

In terms of species detection and abundance, NGSB performed best by far. Indeed, we found highly similar community compositions compared to the morphological identification data (Tables 2, Figures 2–4, Supporting Information S4–S10). Noteworthy, in transect II, two specimens belonging two *H. simplex* (as determined by Sanger sequencing) were most probably miss-identified

as *H. langobardicus* by NGSB, a species for which barcoding is often challenged due to the co-amplification of nuclear copies of mitochondrial genes (i.e., numts; unpublished data C. Praz). For most western European bee fauna, COI barcoding is reliable and provides enough resolution to discriminate at the species level; however, there are some known cases of barcode sharing. In our data set, only one problematic case of barcoding sharing species was sampled (i.e., *Andrena dorsata*, which shares barcodes with *A. propinqua*). After verification, this species was correctly identified for two out of the three methods (i.e., MB and NGSB). For MG, *A. dorsata* was not identified however neither was its sister species (i.e., *A. propinqua*). Therefore, potential biases due to barcoding sharing can be excluded in our study.

The PERMANOVA and Procrustes tests on relative or absolute abundance data also indicate high similarity between this method and morphology in terms of species abundance and ecological patterns (Figures 2,3). The level of accuracy found in this study is in the range of previous studies. For instance, Shokralla et al. (2015) applied NGSB to a diverse data set of arthropods (11 orders) and obtained an overall recovery rate of 97.3% ( $n = 1,010$ ), and 96.5% for Hymenoptera alone ( $n = 226$ ). Likewise, Wang et al. (2018) sequenced over 4,000 ants using NGSB and obtained 95% of correspondence between taxonomy and morphology.

Besides high accuracy, NGSB holds several other advantages over bulk-based approaches (i.e., MB and MG). First, individual DNA extractions and the preservation of associated specimens provide the possibility of verifying unexpected records (e.g., rare species or species outside their known range) through morphology since exoskeletons remain mostly unaltered after proteinase K digestions. Alternatively, DNA extractions can be performed on single legs as done in our study, and reference specimens could be kept nearly intact, although at the cost of additional workload. The preservation of reference specimens provides a valuable back-up, and therefore, NGSB data are more likely to be considered for national or international databases, which can be used for purposes other than monitoring (e.g., compiling red lists or more generally for conservation biology). Second, DNA barcodes generated using NGSB can be fed into existing DNA databases since a link to the specimen is maintained. Third, DNA extractions can further be used for population genetic or phylogenetic studies. Finally, contrary to MB, NGSB does not require PCR replicates. Thereby, the sequencing runs of NGSB can encompass larger data sets and provide higher coverages and thus further reduce costs.

Dealing at the specimen instead of community level has, however, a major drawback. Individual extractions and PCRs considerably increase cost and workload linked to the library preparation. This additional workload and cost difference with bulk-based approaches will, however, largely depend on the number of specimens sampled per community.

#### 4.4 | Cost and workload effectiveness

One of the main arguments brought forward for promoting NGS-identification tools in monitoring programs is the potential cost

reduction in identifications. Although often stated as more cost-efficient than morphological identification, only few studies have systematically assessed the financial advantages of NGS tools over morphology using "real" monitoring data sets. Overall, we found that all investigated NGS-identification methods were costlier than morphological identification (Supporting Information S16). For MB and NGSB, sequencing kits constituted the largest fraction of the total cost (Supporting information S18). To reduce the overall cost for both methods, it is possible to either use smaller sequencing kits or increase the number of specimens by sequencing run. Based upon estimations, the smallest MiSeq sequencing kit able to span our targeted fragment would considerably decrease costs without compromising sequencing depths (Supporting Information S19). Although the output of a MiSeq v2 Nano kit ( $2 \times 250$  bp) corresponds to approximately 1/30 of a MiSeq v3 kit, the estimated coverage will remain high, with over 100 mapped reads per specimens. Higher coverages can be expected if clustering optima during sequencing runs are reached. Using the MiSeq v2 Nano sequencing kit, the overall cost of MB and NGSB is largely reduced and drop in the range of morphological identification (Supporting Information S19). Alternatively, with the same sequencing kit used in this study, we estimated that MB and NGSB become more cost-efficient than morphology after 1,675 and 4,434 specimens, respectively. Noteworthy, several steps of our pipeline could be optimized to even further reduce cost and labour time. For instance, one could reduce hands-on time required for DNA extraction to only a few minutes by using quick DNA extraction kits such as QuickExtract DNA Extraction kit (Lucigen; see Kranzfelder, Ekrem, & Stur, 2016). Studies reducing as much as possible laboratory costs report that sequencing can be performed for approximately 0.50\$ per specimen (Wang et al., 2018). Nevertheless, such cost reduction often implies fine-tuning protocols for the targeted taxon, mainly because DNA is amplified through direct PCR (Wong, Tay, Puniamorthy, Balke, Cranston & Meier, 2014). Additionally, such price optimization requires running libraries on partial kits/lanes, which is not always possible or proposed by sequencing suppliers. For MG, our cost estimations on the Illumina MiSeq platform show that this method will hardly overpass morphology in terms of cost-efficiency. Indeed, sequencing depth is a main bottleneck for this method since only a minor fraction of the data is informative. Therefore, we would recommend sequencing MG libraries on more appropriated platforms, such as HiSeq 4000, HiSeq X or even NovaSeq 6000. Although there are indications that the ability to sequence shorter fragments negatively affects the overall mitochondrial proportion, and therefore, the fraction of reads corresponding to mitochondrial DNA may be reduced on a HiSeq sequencer (Crampton-Platt et al., 2016; Maddock et al., 2016), using larger scale sequencing platforms will drastically reduce costs and increase species detection rates.

In terms of workload, MB was the least labour-intensive method with approximately 27% less hands-on work than morphological identification. NGSB is unsurprisingly the method requiring most workload, although it is in a close range to MG and morphological identification. Compared to MB, NGSB relies on individual DNA extraction,

which is a time demanding procedure, especially since extractions were performed on single legs. With a well-organized protocol, the sorting and DNA extractions required for NGSB may considerably be reduced, potentially to a similar level than MB. Indeed, bulk-based approaches such as MB or MG also require presorting of raw sampling material to isolate bees from plant material, from numerous honeybees ( $n = 1,422$ , thus nearly twice as many wild bees in our data set) and other insects. If none-targeted taxa, and especially honeybees, are not removed, the sequencing depth, and therefore detection rates and biomass estimations would largely be affected.

## 4.5 | Conclusions

For routine monitoring of wild bees using molecular identification methods, we recommend NGSB. The reliability and accuracy levels of this method are hardly attainable with bulk-based approaches, especially for species abundance estimation. Furthermore, this approach provides a valuable supplementary security since specimens can be re-examined morphologically if required. NGSB is thus more likely to yield occurrence data that can be validated and integrated into national faunistic databases and thus used by bee experts and by conservation practitioners. Feeding national faunistic databases is an important by-product of monitoring programs (e.g., in Switzerland: <http://www.biodiversitymonitoring.ch/en/home.html>).

## ACKNOWLEDGEMENTS

We would like to gratefully acknowledge Beatrice Frey and Daniel Frei for their valuable support and advice during the library preparations, as well as Melanie Schirrmann and Ernest Hennig for the fruitful discussions and advices on figure layouts and statistics. We are very thankful to the taxonomist Andreas Müller for having accepted to identify our data set. We also would like to thank all farmers involved in this study for allowing us to sample on their properties. Finally, we are grateful to Marc Matter and Jeannette Regan for having edited several sections of this manuscript. This study was funded by the Swiss Federal Office for Agriculture (FOAG) ('Wild bee metabarcoding project').

## AUTHORS CONTRIBUTION

Sampling scheme was designed and conducted by D.G., M.A. and E.K. Laboratory protocols were designed and performed by M.G. and S.B. M.G. executed the bioinformatics steps and M.G., D.G., M.A. and E.K. performed the statistical analyses. A first draft of the manuscript was written by M.G., C.P. and J.F. All authors contributed to the writing of the final version of this paper.

## DATA ACCESSIBILITY

Absolute abundance matrix of all identification methods is available at Dryad (<https://doi.org/10.5061/dryad.gh830j7>). This repertory also contains raw sequencing files and associated metadata.

## ORCID

Morgan Gueuning  <https://orcid.org/0000-0001-8574-9640>

## REFERENCES

- Anderson, M. J. (2001). Permutation tests for univariate or multivariate analysis of variance and regression. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(3), 626–639. <https://doi.org/10.1139/cjfas-58-3-626>
- Bartomeus, I., Stavert, J. R., Ward, D., & Aguado, O. (2019). Historical collections as a tool for assessing the global pollination crisis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 374(1763), 20170389. <https://doi.org/10.1098/rstb.2017.0389>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), <https://doi.org/10.18637/jss.v067.i01>
- Biesmeijer, J. C., Roberts, S. P. M., Reemer, M., Ohlemüller, R., Edwards, M., Peeters, T., ... Kunin, W. E. (2006). Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands. *Science*, 313(5785), 351. <https://doi.org/10.1126/science.1127863>
- Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15), 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>
- Brandon-Mong, G.-J., Gan, H.-M., Sing, K.-W., Lee, P.-S., Lim, P.-E., & Wilson, J.-J. (2015). DNA metabarcoding of insects and allies: An evaluation of primers and pipelines. *Bulletin of Entomological Research*, 1–11, <https://doi.org/10.1017/S0007485315000681>
- Brunner, P. C., Fleming, C., & Frey, J. E. (2002). A molecular identification key for economically important thrips species (Thysanoptera: Thripidae) using direct sequencing and a PCR-RFLP-based approach. *Agricultural and Forest Entomology*, 4(2), 127–136. <https://doi.org/10.1046/j.1461-9563.2002.00132.x>
- Burkle, L. A., Marlin, J. C., & Knight, T. M. (2013). Plant-Pollinator Interactions over 120 years: loss of species, co-occurrence, and function. *Science*, 339(6127), 1611–1615. <https://doi.org/10.1126/science.1232728>
- Bushnell, B. (2015). BBMap (version 35.14) [Software]. Retrieved from <https://Sourceforge.Net/Projects/Bbmap/>.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, 10(1), 421. <https://doi.org/10.1186/1471-2105-10-421>
- Cane, J. H. (1987). Estimation of bee size using intertegular span (Apoidea). *Journal of the Kansas Entomological Society*, 60(1), 145–147. Retrieved from <http://www.jstor.org/stable/25084877>.
- Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., ... Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*, 7(5), 335–336. <https://doi.org/10.1038/nmeth.f.303>
- Choo, L. Q., Crampton-Platt, A., & Vogler, A. P. (2017). Shotgun mitogenomics across body size classes in a local assemblage of tropical Diptera: Phylogeny, species diversity and mitochondrial abundance spectrum. *Molecular Ecology*, 26(19), 5086–5098. <https://doi.org/10.1111/mec.14258>
- Crampton-Platt, A., Timmermans, M. J. T. N., Gimmel, M. L., Kutty, S. N., Cockerill, T. D., Khen, C. V., & Vogler, A. P. (2015). Soup to tree: The phylogeny of beetles inferred by mitochondrial metagenomics of a bornean rainforest sample. *Molecular Biology and Evolution*, 32(9), 2302–2316. <https://doi.org/10.1093/molbev/msv111>
- Crampton-Platt, A., Yu, D. W., Zhou, X., & Vogler, A. P. (2016). Mitochondrial metagenomics: Letting the genes out of the bottle. *GigaScience*, 5(1), 15. <https://doi.org/10.1186/s13742-016-0120-y>

- Dowle, E. J., Pochon, X., Banks, J. C., Shearer, K., & Wood, S. A. (2016). Targeted gene enrichment and high-throughput sequencing for environmental biomonitoring: A case study using freshwater macroinvertebrates. *Molecular Ecology Resources*, 16, 1240–1254. <https://doi.org/10.1111/1755-0998.12488>
- Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19), 2460–2461. <https://doi.org/10.1093/bioinformatics/btq461>
- Elbrecht, V., & Leese, F. (2015). Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS ONE*, 10(7), 1–16. <https://doi.org/10.1371/journal.pone.0130324>
- Elbrecht, V., Taberlet, P., Dejean, T., Valentini, A., Usseglio-Polatera, P., Beisel, J.-N., ... Leese, F. (2016). Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. *PeerJ*, 4, 1–12. <https://doi.org/10.7717/peerj.1966>
- Elbrecht, V., Vamos, E. E., Meissner, K., Aroviita, J., & Leese, F. (2017). Assessing strengths and weaknesses of DNA metabarcoding-based macroinvertebrate identification for routine stream monitoring. *Methods in Ecology and Evolution*, 8(10), 1265–1275. <https://doi.org/10.1111/2041-210X.12789>
- Fadrosh, D. W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R. M., & Ravel, J. (2014). An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*, 2(1), 6. <https://doi.org/10.1186/2049-2618-2-6>
- Ficetola, G. F., Pansu, J., Bonin, A., Coissac, E., Giguet-Covex, C., De Barba, M., ... Taberlet, P. (2015). Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, 5(3), 543–556. <https://doi.org/10.1111/1755-0998.12338>
- Ficetola, G. F., Taberlet, P., & Coissac, E. (2016). How to limit false positives in environmental DNA and metabarcoding? *Molecular Ecology Resources*, 16(3), 604–607. <https://doi.org/10.1111/1755-0998.12508>
- Gibson, J., Shokralla, S., Porter, T. M., King, I., van Konynenburg, S., Janzen, D. H., ... Hajibabaei, M. (2014). Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasytematics. *Proceedings of the National Academy of Sciences of the United States of America*, 111(22), 8007–8012. <https://doi.org/10.1073/pnas.1406468111>
- Gibson, J. F., Stein, E. D., Baird, D. J., Max, F. C., Zhang, X., & Hajibabaei, M. (2015). Wetland ecogenomics – The next generation of wetland biodiversity and functional assessment. *Wetland Science and Practice*. [https://doi.org/10.1007/978-3-540-70962-6\\_5](https://doi.org/10.1007/978-3-540-70962-6_5)
- Gillett, C. P. D. T., Crampton-Platt, A., Timmermans, M. J. T. N., Jordal, B. H., Emerson, B. C., & Vogler, A. P. (2014). Bulk de novo mitogenome assembly from pooled total DNA elucidates the phylogeny of weevils (Coleoptera: Curculionidae). *Molecular Biology and Evolution*, 31(8), 2223–2237. <https://doi.org/10.1093/molbev/msu154>
- Gomez-Rodriguez, C., Crampton-Platt, A., Timmermans, M. J. T. N., Baselga, A., & Vogler, A. P. (2015). Validating the power of mitochondrial metagenomics for community ecology and phylogenetics of complex assemblages. *Methods in Ecology and Evolution*, 8, 883–894. <https://doi.org/10.1111/2041-210X.12376>
- Goulson, D., Nicholls, E., Botías, C., & Rotheray, E. L. (2015). Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. *Science*, 347(6229). <https://doi.org/10.1126/science.1255957>
- Hebert, P. D. N., Cywinska, A., Ball, S. L., & DeWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 270(1512), 313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Ivanova, N. V., Dewaard, J. R., & Hebert, P. D. N. (2006). An inexpensive, automation-friendly protocol for recovering high-quality DNA. *Molecular Ecology Notes*, 6(4), 998–1002. <https://doi.org/10.1111/j.1471-8286.2006.01428.x>
- Jaccard, P. (1912). The distribution of the flora in the alpine zone. *New Phytologist*, 11(2), 37–50. <https://doi.org/10.1111/j.1469-8137.1912.tb05611.x>
- Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., ... Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, 16(10), 1245–1257. <https://doi.org/10.1111/ele.12162>
- Joseph, L. N., Field, S. A., Wilcox, C., & Possingham, H. P. (2006). Presence-absence versus abundance data for monitoring threatened species. *Conservation Biology*, 20(6), 1679–1687. <https://doi.org/10.1111/j.1523-1739.2006.00529.x>
- Kranzfelder, P., Ekrem, T., & Stur, E. (2016). Trace DNA from insect skins: A comparison of five extraction protocols and direct PCR on chironomid pupal exuviae. *Molecular Ecology Resources*, 16(1), 353–363. <https://doi.org/10.1111/1755-0998.12446>
- Lahoz-Monfort, J. J., Guillera-Aroita, G., & Tingley, R. (2016). Statistical approaches to account for false-positive errors in environmental DNA samples. *Molecular Ecology Resources*, 16(3), 673–685. <https://doi.org/10.1111/1755-0998.12486>
- Lebuhn, G., Droege, S., Connor, E. F., Gemmill-Herren, B., Potts, S. G., Minckley, R. L., ... Parker, F. (2013). Detecting insect pollinator declines on regional and global scales. *Conservation Biology*, 27(1), 113–120. <https://doi.org/10.1111/j.1523-1739.2012.01962.x>
- Leray, M., Yang, J. Y., Meyer, C. P., Mills, S. C., Agudelo, N., Ranwez, V., ... Machida, R. J. (2013). A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Frontiers in Zoology*, 10(1), 34. <https://doi.org/10.1186/1742-9994-10-34>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., ... Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Linard, B., Crampton-Platt, A., Gillett, C. P. D. T., Timmermans, M. J. T. N., & Vogler, A. P. (2015). Metagenome skimming of insect specimen pools: Potential for comparative genomics. *Genome Biology and Evolution*, 7(6), 1474–1489. <https://doi.org/10.1093/gbe/evv086>
- Linard, B., Crampton-Platt, A., Moriniere, J., Timmermans, M. J. T. N., Andújar, C., Arribas, P., ... Vogler, A. P. (2018). The contribution of mitochondrial metagenomics to large-scale data mining and phylogenetic analysis of Coleoptera. *Molecular Phylogenetics and Evolution*, 128, 1–11. <https://doi.org/10.1016/j.ympev.2018.07.008>
- Liu, S., Li, Y., Lu, J., Su, X., Tang, M., Zhang, R., ... Zhou, X. (2013). SOAPBarcode: Revealing arthropod biodiversity through assembly of Illumina shotgun sequences of PCR amplicons. *Methods in Ecology and Evolution*, 4(12), 1142–1150. <https://doi.org/10.1111/2041-210X.12120>
- Liu, S., Wang, X., Xie, L., Tan, M., Li, Z., Su, X. u., ... Zhou, X. (2016). Mitochondrial capture enriches mito-DNA 100 fold, enabling PCR-free mitogenomics biodiversity analysis. *Molecular Ecology Resources*, 16(2), 470–479. <https://doi.org/10.1111/1755-0998.12472>
- MacKenzie, D. I. (2005). What are the issues with presence-absence data for wildlife managers? *The Journal of Wildlife Management*, 69(3), 849–860. Retrieved from <http://www.jstor.org/stable/3803327>
- Maddock, S. T., Briscoe, A. G., Wilkinson, M., Waeschenbach, A., San Mauro, D., Day, J. J., ... Gower, D. J. (2016). Next-generation mitogenomics: A comparison of approaches applied to caecilian amphibian phylogeny. *PLoS ONE*, 11(6), e0156757. <https://doi.org/10.1371/journal.pone.0156757>
- Mata, V. A., Rebelo, H., Amorim, F., McCracken, G. F., Jarman, S., & Beja, P. (2019). How much is enough? Effects of technical and biological replication on metabarcoding dietary analysis. *Molecular Ecology*, 28(2), 165–175. <https://doi.org/10.1111/mec.14779>
- Meyer, M., & Kircher, M. (2010). Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring*

- Harbor Protocols, 2010(6), pdb.prot5448–pdb.prot5448. <https://doi.org/10.1101/pdb.prot5448>
- Nieto, A., Roberts, S. P. M., Kemp, J., Rasmont, P., Kuhlmann, M., Criado, M. G., ... Michez, D. (2015). European Red List of Bees. Luxembourg: Publication Office of the European Union. <https://doi.org/10.2779/77003>
- Ollerton, J., Erenler, H., Edwards, M., & Crockett, R. (2014). Extinctions of aculeate pollinators in Britain and the role of large-scale agricultural changes. *Science*, 346(6215), 1360–1362. <https://doi.org/10.1126/science.1257259>
- Peng, Y., Leung, H. C. M., Yiu, S. M., & Chin, F. Y. L. (2012). IDBA-UD: A de novo assembler for single-cell and metagenomic sequencing data with highly uneven depth. *Bioinformatics*, 28(11), 1420–1428. <https://doi.org/10.1093/bioinformatics/bts174>
- Piñol, J., Mir, G., Gomez-Polo, P., & Agustí, N. (2015). Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, 15, 819–830. <https://doi.org/10.1111/1755-0998.12355>
- 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>
- Porazinska, D. L., Sung, W., Giblin-Davis, R. M., & Thomas, W. K. (2010). Reproducibility of read numbers in high-throughput sequencing analysis of nematode community composition and structure. *Molecular Ecology Resources*, 10(4), 666–676. <https://doi.org/10.1111/j.1755-0998.2009.02819.x>
- Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). Global pollinator declines: Trends, impacts and drivers. *Trends in Ecology and Evolution*, 25(6), 345–353. <https://doi.org/10.1016/j.tree.2010.01.007>
- Potts, S., Biesmeijer, K., Bommarco, R., Kleijn, D., & Scheper, J. A. (2015). Status and trends of European pollinators. Key findings of the STEP Project. 70. Pensoft Publishers. Retrieved from <http://edepot.wur.nl/389377>.
- Potts, S. G., Imperatriz-fonseca, V., Ngo, H. T., Biesmeijer, J. C., Breeze, T. D., Dicks, L. V., ... Viana, B. F. (2016). Summary for policymakers of the thematic assessment on pollinators, pollination and food production Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Key messages Values of pollinators and pollination (pp. 1–36). Bonn, Germany: Secretariat of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services.
- Rodgers, T. W., Xu, C. C. Y., Giacalone, J., Kapheim, K. M., Saltonstall, K., Vargas, M., ... Jansen, P. A. (2017). Carrion fly-derived DNA metabarcoding is an effective tool for mammal surveys: Evidence from a known tropical mammal community. *Molecular Ecology Resources*, 17(6), 133–145. <https://doi.org/10.1111/1755-0998.12701>
- Schmidt, S., Schmid-Egger, C., Morinière, J., Haszprunar, G., & Hebert, P. D. N. (2015). DNA barcoding largely supports 250 years of classical taxonomy: identifications for Central European bees (Hymenoptera, Apoidea). *Molecular Ecology Resources*, 15(4), 985–1000.
- Schnell, I. B., Thomsen, P. F., Wilkinson, N., Rasmussen, M., Jensen, L. R. D., Willerslev, E., ... Gilbert, M. T. P. (2012). Screening mammal biodiversity using DNA from leeches. *Current Biology*, 22(8), 262–263. <https://doi.org/10.1016/j.cub.2012.02.058>
- Shokralla, S., Gibson, J. F., Nikbakht, H., Janzen, D. H., Hallwachs, W., & Hajibabaei, M. (2014). Next-generation DNA barcoding: Using next-generation sequencing to enhance and accelerate DNA barcode capture from single specimens. *Molecular Ecology Resources*, 14(5), 892–901. <https://doi.org/10.1111/1755-0998.12236>
- Shokralla, S., Porter, T. M., Gibson, J. F., Dobosz, R., Janzen, D. H., Hallwachs, W., ... Hajibabaei, M. (2015). Massively parallel multiplex DNA sequencing for specimen identification using an Illumina MiSeq platform. *Scientific Reports* 5(1).
- Taberlet, P., Bonin, A., Zinger, L., & Coissac, E. (2018). *Environmental DNA: For biodiversity research and monitoring*. Oxford, UK: Oxford University Press. <https://doi.org/10.1093/oso/9780198767220.001.0001>
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C., & Willerslev, E. (2012). Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, 21, 2045–2050. <https://doi.org/10.1111/j.1365-294X.2012.05470.x>
- Tang, M., Hardman, C. J., Ji, Y., Meng, G., Liu, S., Tan, M., ... Yu, D. W. (2015). High-throughput monitoring of wild bee diversity and abundance via mitogenomics. *Methods in Ecology and Evolution*, 6(9), 1034–1043. <https://doi.org/10.1111/2041-210X.12416>
- Tang, M., Tan, M., Meng, G., Yang, S., Su, X. u., Liu, S., ... Zhou, X. (2014). Multiplex sequencing of pooled mitochondrial genomes - A crucial step toward biodiversity analysis using mito-metagenomics. *Nucleic Acids Research*, 42(22), 1–13. <https://doi.org/10.1093/nar/gku917>
- vanEngelsdorp, D., & Meixner, M. D. (2010). A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them. *Journal of Invertebrate Pathology*, 103, 80–95. <https://doi.org/10.1016/J.JIP.2009.06.011>
- Wang, W. Y., Srivathsan, A., Foo, M., Yamane, S. K., & Meier, R. (2018). Sorting specimen-rich invertebrate samples with cost-effective NGS barcodes: Validating a reverse workflow for specimen processing. *Molecular Ecology Resources*, 18(3), 490–501. <https://doi.org/10.1111/1755-0998.12751>
- Wilson, J.-J., Brandon-Mong, G.-J., Gan, H.-M., & Sing, K.-W. (2019). High-throughput terrestrial biodiversity assessments: Mitochondrial metabarcoding, metagenomics or metatranscriptomics? *Mitochondrial DNA Part A*, 30(1), 60–67. <https://doi.org/10.1080/24701394.2018.1455189>
- Wong, W. H., Tay, Y. C., Puniamoorthy, J., Balke, M., Cranston, P. S., & Meier, R. (2014). 'Direct PCR' optimization yields a rapid, cost-effective, nondestructive and efficient method for obtaining DNA barcodes without DNA extraction. *Molecular Ecology Resources*, 14 (6), 1271–1280.
- Yu, D. W., Ji, Y., Emerson, B. C., Wang, X., Ye, C., Yang, C., & Ding, Z. (2012). Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods in Ecology and Evolution*, 3(4), 613–623. <https://doi.org/10.1111/j.2041-210X.2012.00198.x>
- Zhou, X., Li, Y., Liu, S., Yang, Q., Su, X. u., Zhou, L., ... Huang, Q. (2013). Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification. *GigaScience*, 2(1), 4. <https://doi.org/10.1186/2047-217X-2-4>

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Gueuning M, Ganser D, Blaser S, et al. Evaluating next-generation sequencing (NGS) methods for routine monitoring of wild bees: Metabarcoding, mitogenomics or NGS barcoding. *Mol Ecol Resour*. 2019;19:847–862. <https://doi.org/10.1111/1755-0998.13013>