Project POLBEN

The use of molecular tools in Diptera pollination ecology in Benin

coordinator Kurt Jordaens, RMCA

co-funding JEMU (call 2021), ISeBAF, SYNTHESIS

final technical report

Nele Mullens, Gontran Sonet, Lore Esselens, Jessica Ody, Carl Vangestel, Massi Virgilio

Tervuren, 3 July 2023

Summary

The project POLBEN aimed at optimizing protocols for the metabarcoding identification (ID) of pollen loads recovered from pollinating Syrphidae (Diptera) and Hymenoptera, Apoidea. After an initial literature survey, we compared the suitability of different field and lab protocols for pollen preservation and DNA extraction. We tested a range of primers and PCR conditions for the Sanger sequencing of plant DNA barcodes, with a main focus to the PCR amplification of plants DNA barcodes from the family Cucurbitaceae. Finally, we optimised and compared "in-house" protocols for preparing metabarcoding libraries, and considered cost-benefits of outsourcing library prep.

Literature survey

See a selection of references attached to this report.

Tests on optimal sample preservation and DNA extraction

In this test, we focused on pollen grains isolated from the bodies of "fresh" flower flies (Diptera, *Syrphidae*) collected by hand net in the Tervuren park, stored in 1.5ml Eppendorf tubes containing either 100% ethanol or 1ml of CTAB and subjected to:

a) Immediate pollen isolation and DNA extraction,
b) preservation for one month at room temperature (RT, ~25°C) followed by pollen isolation and DNA extraction,
c) preservation for one month in a freezer at -20°C followed by pollen isolation and DNA extraction,

Pollen DNA was extracted using either a modified CTAB extraction protocol (Supplementary data S1, elution volume 30ul) or the Qiagen DNeasy Plant Minikit (elution volume 100ul).

Table 1: Test on sample preservation and pollen DNA extraction. Experimental setup and n. of specimens processed.

		Preservation group	CTAB DNA extraction	QIAGEN DNA extraction
Control (immediate processing after collection)		a	5	5
EtOH preservation (30 days)	Room T°	b	5	5
	Freezer (-20°C)	c	5	5
CTAB preservation (30 days)	Room T°	d	5	5
	Freezer (-20°C)	e	5	5

Pollen was isolated from the insect bodies by shaking the tube twice for 5 minutes on a bead beater at 6Hz (as a compromise between vigorously shacking the pollen and avoid damaging the vouchers).

The fly was then removed from a solution of 100% EtOH and the pollen centrifuged at 13,000 rpm for 5 min. The pellet was dried in a Eppendorf® Concentrator (1400 rpm for 60 mins) and disrupted in a bead beater (VWR® Star Beater) for 2 minutes at 22,5 Hz with three 3mm stainless steel beads per tube.

DNA concentration was quantified using a fluorometer (Qubit 3, HS DNA Kit, Thermofisher Scientific, Carlsbad, CA, USA). ANOVA and the Student-Newman-Keuls (SNK) as implemented by the R package GAD (Sandrini-Neto & Camargo, 2015), were used for *a priori* and *a posteriori* hypothesis testing with Extraction Method (CTAB vs EtOH) and Preservation Group (CTAB -20°C, CTAB RT, EtOH -20°C, EtOH RT, control) as fixed, orthogonal factors. As elution volumes were different in the two extraction methods (100ul for Qiagen, 30ul for CTAB), analyses were repeated for both total DNA yields and DNA concentrations.

Total DNA concentrations measured ranged from 0.020 ng/µl to 8.52 ng/µl and DNA yields from to 0.91ng to 284 ng (a minor part of measures resulted below the instrument detection limits).

ANOVAs on DNA yields and DNA concentration showed a significant interaction between insect preservation protocol and pollen DNA extraction method. *A posteriori* comparisons revealed significantly higher DNA yields concentrations for samples preserved in CTAB and subjected to CTAB DNA extraction (Fig. 1, 2, supplementary data S2).

This pattern could be artefactual as biased by cross contamination between pollen and insect DNA (with CTAB preservation promoting somehow insect DNA extraction?). In order to further explore this hypothesis, we tentatively amplified the DNA extracted by using PCR primers for both animal and plant DNA barcoding (Folmer et al., 1994; Newmaster et al., 2006). The comparative analysis of amplification success provided a first indication that CTAB preservation followed by CTAB DNA extraction (irrespectively of preservation temperature) seems to favour cross-contamination between plant and insect DNA. In fact, primers for animal DNA barcoding generally provide positive amplification on these samples (Tab. 2).

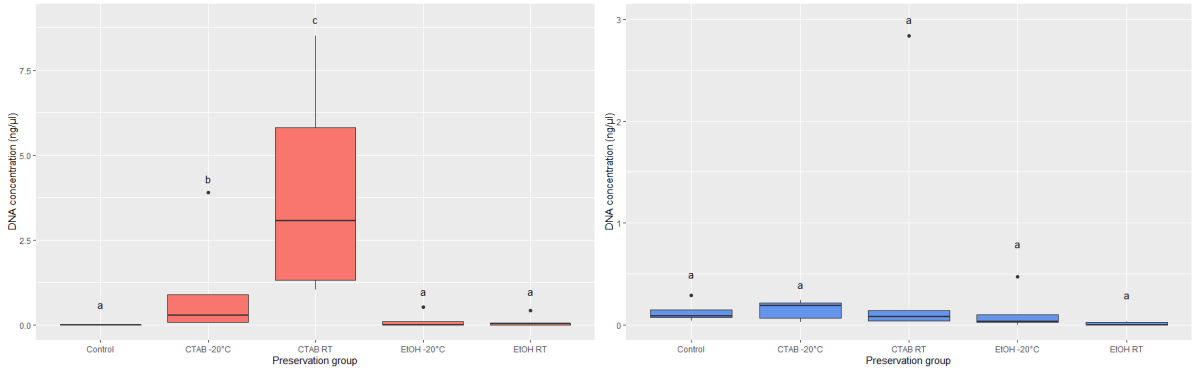

Figure 1: DNA concentrations of pollen samples as recovered by using either the Qiagen DNeasy Plant Mini Kit (left) or CTAB DNA extraction (right). Pollen was either extracted immediately after collection (control) or preserved for 30 days in ethanol 100% at room temperature (EtOH RT); at -20°C (EtOH -20°C), in CTAB at room temperature (CTAB RT) or at -20°C (CTAB -20°C).

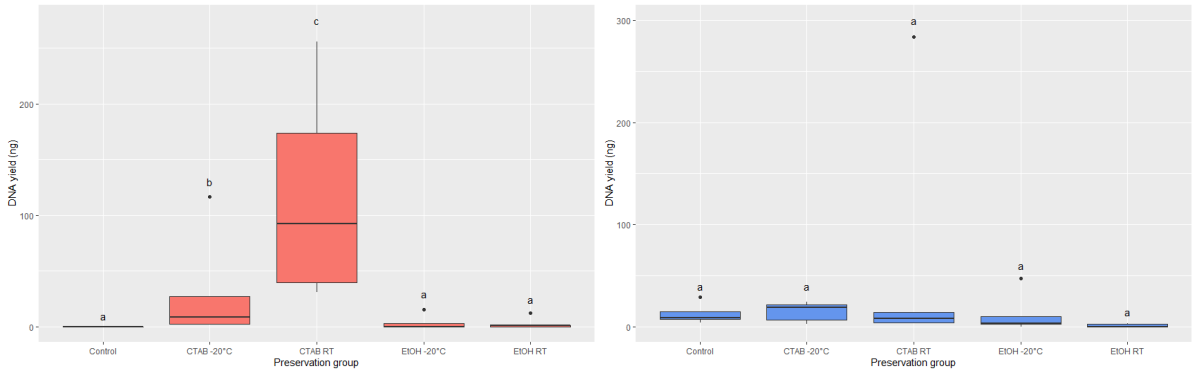

Figure 2: Total DNA yield of pollen samples as recovered by using either the Qiagen DNeasy Plant Mini Kit (left) or CTAB DNA extraction (right). Pollen was either extracted immediately after collection (control) or preserved for 30 days in ethanol 100% at room temperature (EtOH RT); at -20°C (EtOH -20°C), in CTAB at room temperature (CTAB RT) or at -20°C (CTAB -20°C).

Table 2: Exploratory, semi-quantitative test for contamination of pollen samples. Comparative PCR amplification (+ = amplification, - = no amplification) of pollen samples from different preservation groups (see Tab. 1) via universal primers for plant (rbcl) and animal DNA (COI) barcoding. Positive COI amplification suggests contamination from insect DNA.

DNA extraction	preservation group			rbcl (plant)	COI (animal)
Qiagen	b	EtOH	RT	+	+
Qiagen	b	EtOH	RT	+	-
Qiagen	b	EtOH	RT	+	-
Qiagen	b	EtOH	RT	+	-
CTAB	d	CTAB	RT	+	+
CTAB	d	CTAB	RT	+	+
CTAB	d	CTAB	RT	+	+
CTAB	d	CTAB	RT	+	-
CTAB	e	CTAB	-20°C	+	+
CTAB	e	CTAB	-20°C	+	+
CTAB	e	CTAB	-20°C	+	+
CTAB	e	CTAB	-20°C	+	-

Optimization of primers and PCR conditions for Sanger sequencing

Wet-lab pipelines for the amplification of 4 markers generally used in plant DNA barcoding ID were developed: internal transcribed spacer 1 and 2 (ITS1 and ITS2), ribulose 1,5-biphosphate carboxylase (rbcL), and maturase K (matK).

Fourteen primer pairs were tested. Four of them were designed *ex novo* using Primer3 (Untergasser *et al.,* 2012) on alignments including 1500-2100 publicly available plant sequences (focus on *Cucurbitaceae*, minimum sequence length 500bp). Gradient PCRs (55°C < T < 65°C) were used to test optimal annealing temperature (Ta, supplementary data S3) on DNA extracts from three cucurbits (*Cucumis sativus* L.), pumpkins (*Cucurbita maxima* Duchesne), watermelon (*Citrullus lanatus* (Thunb.) Matsumura & Nakai)).

PCR were performed in a final volume of 25µl using the Platinum™ Taq DNA Polymerase (Invitrogen™). The PCR reaction mixture contained 2,50µl of PCR Buffer 10x, 0,75µl MgCl Platinum™ 50mM, 2,50µl of dNTP 2mM, 0,5µl of the forward primer (20µM) and 0,5µl of the reverse primer (20µM), 0,15µl Taq Platinum™ (5U/µl). PCR cycles included an initial heat activation for 5 min at 94°C; followed by 40 cycles of 30 s at 94°C, 30 s at Ta (see Tab. 3), and 1 min at 72°C; followed by a final extension of 10 min at 72°C.

Table 3: Primer list for ITS1, ITS2, rbcL and matk DNA barcodes. The expected amplicon size was inferred using a selection of plant DNA sequences downloaded from the NCBI reference database.

gene fragment	Primer pair ID	Primer Forward	Primer Reverse	Expected amplicon size (bp)	Reference
ITS1	ITS1-390	AGTCGTAACAAGGTTTCCGT	GGGATTCTGCAATTCACACC	390	J. Ody – RMCA, unpublished
	ITS1-380	AGTCGTAACAAGGTTTCCGT	AACTTGCGTTCAAAGACTCG	380	J. Ody – RMCA, unpublished
ITS2	ITS2-23	ATGCGATACTTGGTGTGAAT	GACGCTTCTCCAGACTACAAT	460	Chen *et al*., 2010
	ITS2-34	GCATCGATGAAGAACGCAGC	TCCTCCGCTTATTGATATGC	350	White *et al*., 1990
	ITS2-54	CCTTATCATTTAGAGGAAGGAG	TCCTCCGCTTATTGATATGC	750	Chen *et al*., 2010
	ITS2-Uni	TGTGAATTGCARRATYCMG	CCCGHYTGAYYTGRGGTCDC	310	Moorhouse-Gann *et al*., 2018
rbcL	rbcL-506	ATGTCACCACAAACAGAGACT	AGGGGACGACCATACTTGTTCA	506	Modified from De Vere *et al*., 2012
	rbcL-375	ATGTCACCACAAACAGAGACT	ACCCACAATGGAAGTAAACATGT	375	J. Ody – RMCA, unpublished
	rbcL-320	ATGTCACCACAAACAGAGACT	GCAAATCCTCCAGACGTAGA	320	J. Ody – RMCA, unpublished
	rbcL-23506	CTTACCAGYCTTGATCGTTACAAAGG	AGGGGACGACCATACTTGTTCA	275	García-Robledo *et al*., 2013 ; De Vere *et al*., 2012
	rbcL-T	ATGTCACCACAAACAGAGACT	GAAACGGTCTCTCCAACGCAT	660	Modified from Gous *et al*., 2019
	rbcL-2623	CCTTTGTAACGATCAAGRCTGGTAAG	CTTACCAGYCTTGATCGTTACAAAGG	380	*García-Robledo et al.*, 2013
	rbcL-A	ATGTCACCACAAACAGAGACTAAAGC	CTTCTGCTACAAATAAGAATCGATCTC	600	Kress & Erickson, 2007
matK	KIM	CGTACAGTACTTTTGTGTTTACGAG	ACCCAGTCCATCTGGAAATCTTGGTTC	890	Laha *et al*., 2017

The performance of primers at suboptimal DNA concentrations were explored by diluting the three plant DNA extracts (one for each of the target cucurbits) to 0,1 - 0,05 - 0,01 - 0,001 - 0,0001 ng/µl and by verifying their amplification success. It's important to notice that concentrations below 0,05 ng/µl are below the detection limits of most fluorometers. Part of the 210 PCR products obtained (approx. 6%) were sequenced (Macrogen) to verify the amplification of the target gene fragment and exclude potential amplification / sequencing issues.

Five primer pairs (ITS1-390, ITS1-380, rbcL-506, rbcL-320, rbcL-23506) worked also at the lowest concentrations, rbcl-A did not provide any PCR product and was discarded from further consideration, the other 8 primer pairs generally worked at higher DNA concentrations (Fig. 3).

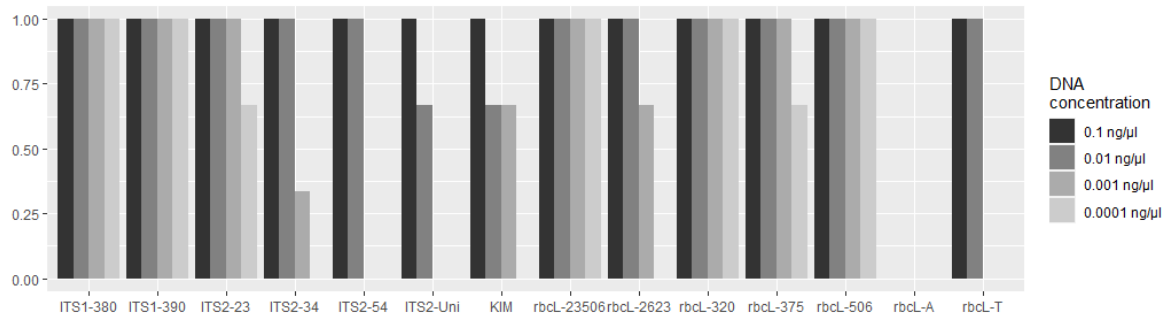

Figure 3: Exploratory analysis of amplification success at suboptimal DNA concentrations based on plant DNA extracts from 3 cucurbits (proportion of successful PCRs, n = 3).

The amplification success of each primer was then tested on 6 pollen DNA extracts from *Apis mellifera* collected from cucurbit crops in Tanzania (preserved in EtOH, Qiagen DNA extraction, DNA concentration range = 0.52-0.024 ng/ul). The primer pairs ITS2-23 and rbcL-320 worked with all the 6 pollen DNA extracts tested. The other twelve primer pairs generally worked at a lower success rate (Fig.4).

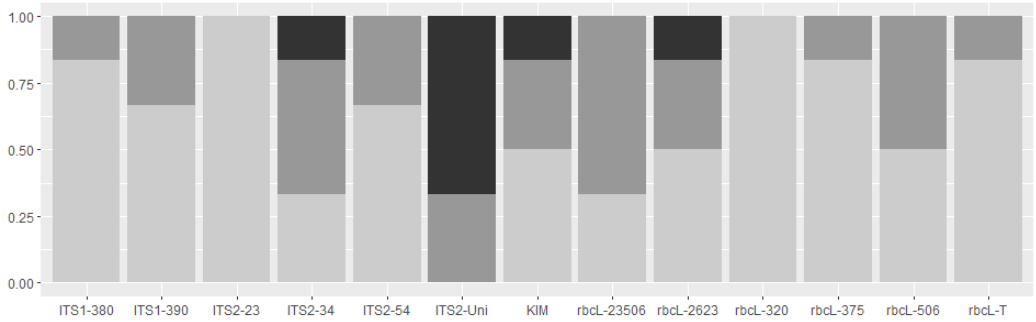

Figure 4: Proportions of successful (light grey), uncertain (grey) or failed PCRs (black) as obtained using different primers on fresh pollen DNA extracts (n=6).

Performance of “in -house” metabarcoding:
commercial kits vs custom pipelines for library prep

With the objective of defining cost- and time effective standards for research on pollen and microbial profiling, we preliminarily and qualitatively considered expected performance and costs of common wet lab pipelines for DNA metabarcoding (results not reported here).

We eventually designed two custom wet lab pipelines (S5) complementary or alternative to the popular Nextera XT pipeline (S6) for DNA metabarcoding. The designed pipelines aimed at using different reagents (semi-custom library prep) or different reagents and indexes (fully custom library prep) to be purchased in bulk and used on batches of samples of different sizes (from only a few to a few hundreds). The main rationale of this approach was to achieve a relatively low and uniform cost/sample and to increase scalability compared to a quite expensive commercial kit which only allows processing either 24 or 96 samples.

Due to the relatively high costs of HTS technologies, we adopted a step-by-step approach with the objective of developing the fully custom pipeline only in case the semi-custom protocol would have provided effective advantages in terms of performance or time/cost effectiveness.

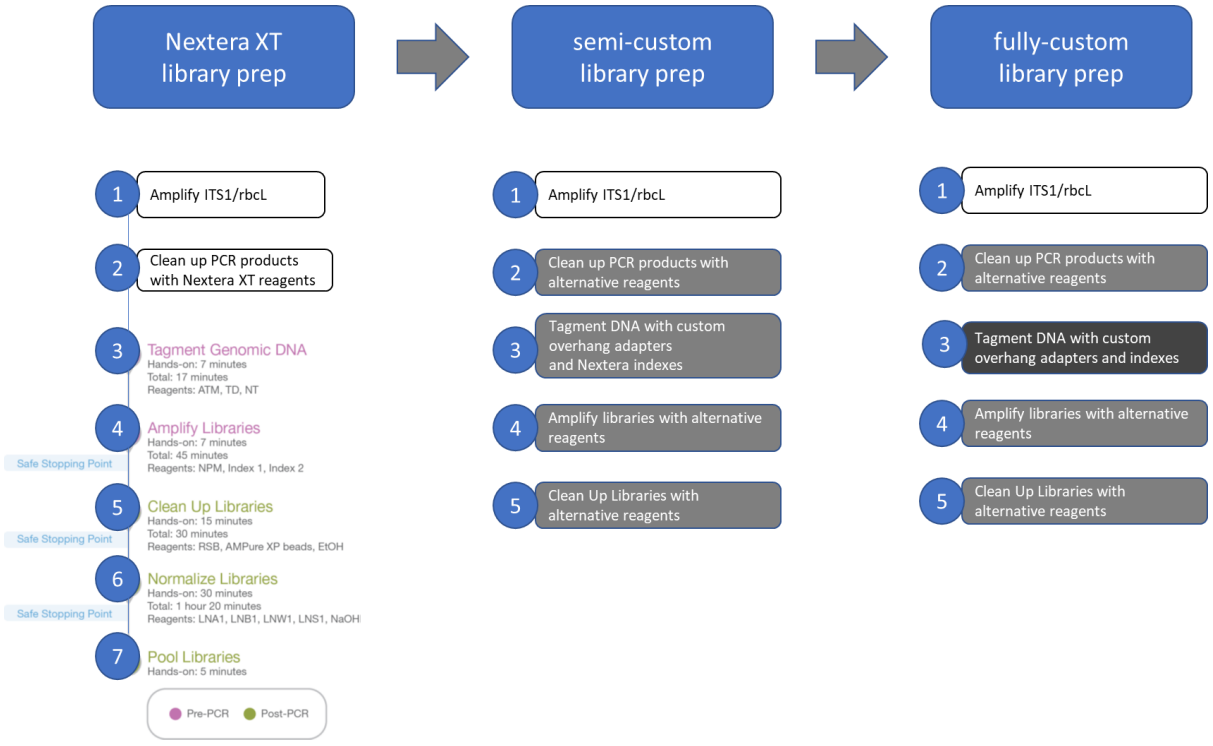

Fig. 5: Complementary (semi-custom library prep) or alternative pipelines (fully-custom library prep) to Nextera XT DNA metabarcoding (left).

The semi-custom pipeline included reagents routinely used at RMCA such as the DNA polymerases kit (Invitrogen 10966-050) for 1st amplicon PCR and 2nd indexing PCR (steps 1 and 3 in Fig. 5) and the AMPure XP beads (Beckman Coulter A63881) for DNA cleanup and size selection (steps 2 and 5). The fully-custom pipeline also included the use of custom-made dual indexes (including P adaptors) which could have been synthesized and purchased from specialized companies (such as Macrogen or Eurogentec). For a detailed overview of the pipeline, see S5.

The performance of the semi-custom and of the Nextera XT library prep were compared during a lab test organized at RMCA in 2022. DNA was extracted from 24 pollen loads from flower flies (Diptera, Tephritidae) and bees (Hymenoptera Apoidea) as per manufacturer's instructions of the DNeasy Plant Mini Kit (Qiagen cat. 69106). Following DNA quality check, each DNA extract was aliquoted in 4 samples, which were subjected either to

1. Nextera XT library prep following ITS2 (primerpair ITS2-34, see S3) amplification or
2. Nextera XT library prep following rbcL (primerpair rbcL-320, see S3) amplification or
3. Semi-custom library prep following ITS2 amplification or
4. Semi-custom library prep following rbcL amplification

The 96 metagenomic libraries obtained (metadata available in S4) were pooled and, after standardising their DNA concentrations, submitted to Macrogen for High Throughput Sequencing on a single MiSeq flowcell (300 PE, 8Gb output).

The performance of the semi-custom library preparation pipeline was generally lower compared to Nextera XT (Fig. 6 and 7), with 79.2% of semi-custom libraries (n=48) showing lower yields in terms of raw reads, 83.3% in terms of n. of Amplicon Sequence Variants (ASVs)and 52.1% in terms of cumulative n. ASVs. The Nextera XT libraries outperforming semi-custom library prep showed an average gain of raw reads of 29.8% (SD=20.8%), while the average gain of raw and filtered reads in semi-custom library prep libraries outperforming Nextera XT was 8.2% (SD=6.6%). The Nextera XT libraries outperforming semi-custom library prep with respect to the ratio between n. filtered / n. raw reads showed an average gain 44.2% (SD=34.1%), while the average gain of in semi-custom library prep libraries outperforming Nextera XT was 20.7% (SD=17.1%). The average gain in terms of n. of ASVs and cumulative n. of ASVs of Nextera XT libraries outperforming semi-custom library prep was 54.8% (SD=26.6%) and 40.9% (SD=35.8%), while the average gain of in semi-custom library prep libraries outperforming Nextera XT was 20.2% (SD=17.8%) and 37.8% (30.1%).

Fig. 6: Comparative output of DNA extracts subjected to semi-custom or Nextera XT library prep following amplification of ITS2. Gain / loss % in n. filtered / n. reads and in n. of ASVs and cumulative n. of ASVs are shown.

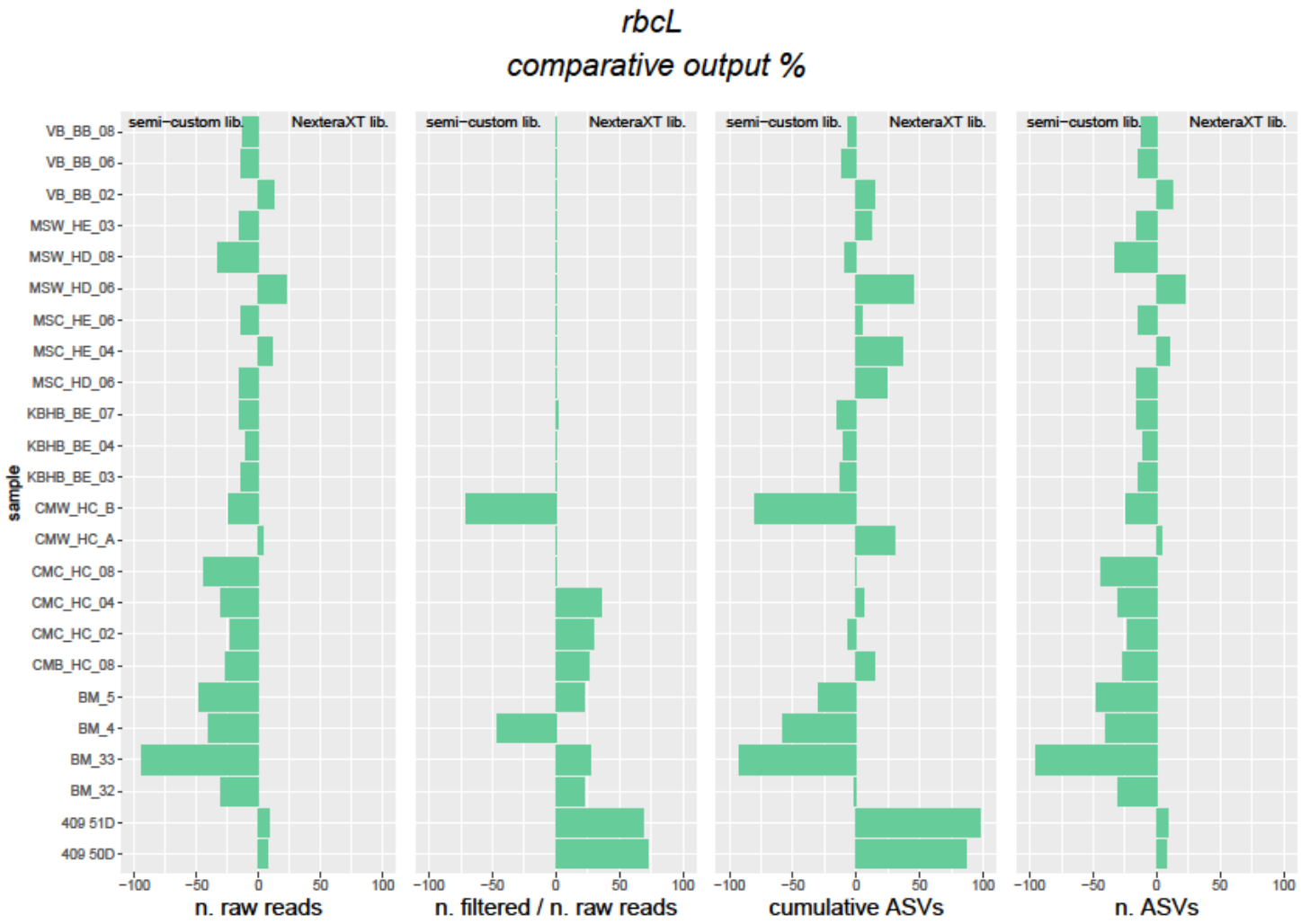

Fig. 7: Comparative output of DNA extracts subjected to semi-custom or Nextera XT library prep following amplification of rbcL. Gain / loss % in n. filtered / n. raw reads and in n. of ASVs and cumulative n. of ASVs are shown.

ANOVA (Tab. 4) detected significantly higher n. of raw reads in Nextera XT compared to semi-custom library prep, and higher n. of ASVs in ITS2 Nextera XT libraries (while no differences were detected between rbcL Nextera XT and semi-custom libraries). Additionally, and irrespectively of library prep, rbcL amplification yielded a higher ratio n. filtered / n. raw reads and a higher n. of cumulative ASVs compared to ITS amplification.

Tab. 4: ANOVAs testing the effects of marker amplification (ITS2, rbcl) and library prep (semi-custom, Nextera XT) on (a) n. of raw reads, (b) ratio n. filtered / n. raw reads, (c) n. of ASVs and (d) cumulative n. ASVs in 96 DNA metabarcoding libraries. *: P<0.05, **: P<0.01, ns: not significant. Significant effects are highlighted in yellow.

		df	MS	F	P	
n. raw reads	Marker: MAR	1	0.075	0.025	0.875	ns
	Library prep: LIB	1	38.425	12.711	0.001	*
	MAR:LIB	1	1.780	0.589	0.445	ns
	Residual	92	3.023			
	C = 0.434, p = 0.016, transformation: fourth root					
n. filtered / n. raw reads	Marker: MAR	1	0.892	23.240	0.000	*
	Library prep: LIB	1	0.085	2.221	0.140	ns
	MAR:LIB	1	0.075	1.948	0.166	ns
	Residual	92	0.038			
	C = 0.378, p = 0.114, transformation: none					
n. ASVs	Marker: MAR	1	36.046	19.577	0.000	*
	Library prep: LIB	1	47.077	25.569	0.000	*
	MAR:LIB	1	9.937	5.397	0.022	*
	Residual	92	1.841			
	C = 0.343, p = 0.315, transformation: square root					
	SNK test MAR x LIB:	ITS2	custom lib. < NexteraXT lib.			
		ITS2	custom lib. = NexteraXT lib.			
cumulative n. ASVs	Marker: MAR	1	13824.765	7.427	0.008	*
	Library prep: LIB	1	587.413	0.316	0.576	ns
	MAR:LIB	1	1615.921	0.868	0.354	ns
	Residual	92	1861.516			
	C = 0.336, p = 0.376, transformation: square root					

As the performance of semi-custom library prep was significantly lower than Nextera XT In terms of raw reads and n. of ASVs, and as semi-custom library prep did not prove to be cost or time effective compared to outsourcing (see below), we did not proceed any further with the setting up of a second experiment to test the performance of fully-custom library prep.

Conversely, we proceeded exploring the results obtained from pollen loads from different groups of bees and flower flies. These analyses were implemented on the data obtained using the better performing Nextera XT library prep. In particular, we tested differences between:

- groups of pollen loads (each including 6 replicated libraries) from *Syrphidae* recently collected (2022, fresh flower flies), *Apis mellifera* recently collected (2022, fresh honeybees) and collection honeybees (Museum honeybees) collected in 1922 (n=2), 1947 (n=2), 1963 (n=1), 1993 (n=1).

- groups of pollen loads (each including 3 replicated libraries) from three species of recently collected *Syrphidae* (*Betasyrphus adligatus* (Wiedemann), *Ischiodon aegyptius* (Wiedemann), *Toxomerus floralis*(Fabricius))

Surprisingly, the first test did not show significant output differences from the DNA metabarcoding of fresh flower flies, fresh honeybees and Museum honeybees (Fig. 8 and 9, Tab. 4 and 5) either in terms of n. of raw reads or of n. of ASVs and cumulative ASVs. This was unexpected due to (a) the larger pollen loads which were isolated from fresh honeybees compared to fresh flower flies (as qualitatively observed during pollen isolation) and (b) to the expected lower quality of pollen DNA isolated from honeybees dating back up to 1922. Yet, significant differences were found in the performance of the two markers, with rbcL yielding a higher n. filtered / n. raw reads ratio in fresh bees compared to ITS2.

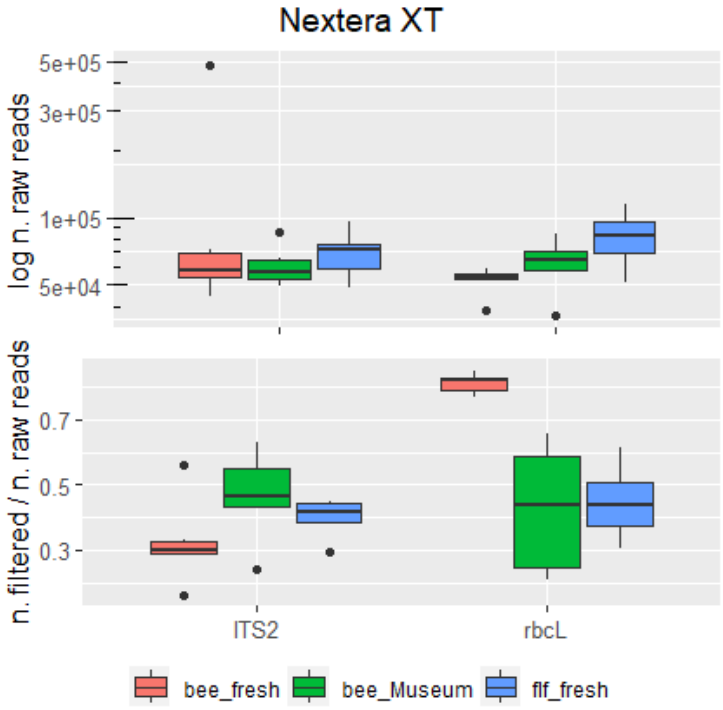

Fig. 8: N. of raw and filtered reads obtained from 36 DNA Nextera XT metabarcoding libraries following ITS2 and rbcl amplification in three sample groups including pollen loads from fresh flower flies, fresh honeybees, Museum honeybees.

Tab. 4: ANOVAs testing the effects of Nextera XT marker amplification (ITS2, rbcl) and sample group (fresh flower flies, fresh honeybees, Museum honeybees) on n. of raw reads and on the ratio n. filtered / n. raw reads in 36 DNA metabarcoding libraries. ns: not significant.

.		df	MS	F	P	
n. raw reads	Marker: MAR	1	1.996	0.480	0.494	ns
	Group: GRO	2	1.961	0.472	0.629	ns
	MAR:GRO	2	6.638	1.596	0.219	ns
	Residual	30	4.158			
	C = 0.811, p = 1.548e-07, transformation: fourth root					
n. filtered / n. raw reads	Marker: MAR	1	0.038	9.450	0.004	**
	Group: GRO	2	0.007	1.782	0.186	
	MAR:GRO	2	0.045	11.374	0.000	***
	Residual	30	0.004			
	C = 0.442, p = 0.051, transformation: fourth root					

SNK test MAR x GRO:	bee fresh	rbcl > ITS2
	bee Museum	rbcl = ITS2
	flf fresh	rbcl = ITS2

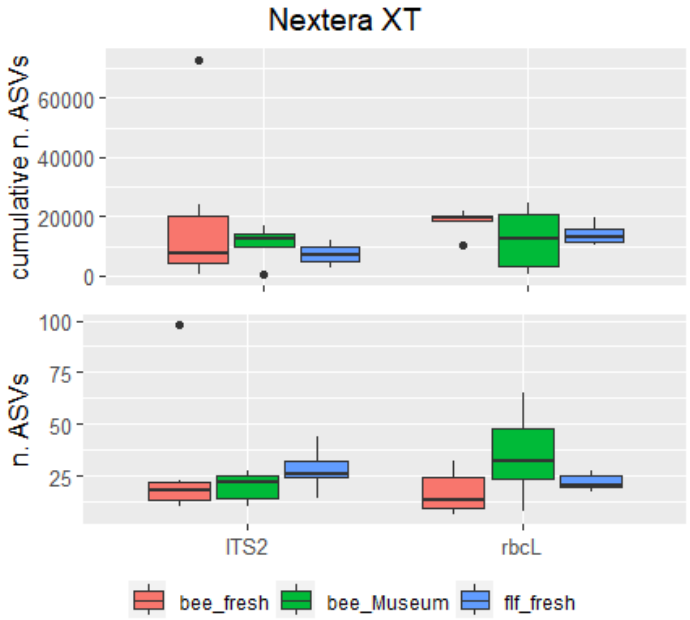

Fig. 9: N. of ASVs and cumulative ASVs obtained from 36 DNA Nextera XT metabarcoding libraries following ITS2 and rbcl amplification in three sample groups including pollen loads from fresh flower flies, fresh honeybees, Museum honeybees.

Tab. 4: ANOVAs testing the effects of Nextera XT marker amplification (ITS2, rbcl) and sample group (fresh flower flies, fresh honeybees, Museum honeybees) on n. of ASVs and cumulative n. ASVs in 36 DNA metabarcoding libraries. ns: not significant.

		df	MS	F	P	
cumulative n. ASVs	Marker: MAR	1	10.045	1.625	0.212	ns
	Group: GRO	2	5.618	0.909	0.414	ns
	MAR:GRO	2	3.160	0.511	0.605	ns
	Residual	30	6.182			
	C = 0.402, p = 0.11, transformation: fourth root					
n. ASVs	Marker: MAR	1	0.105	0.048	0.828	ns
	Group: GRO	2	1.132	0.517	0.602	ns
	MAR:GRO	2	4.509	2.058	0.145	ns
	Residual	30	2.191			
	C = 0.465, p = 0.03, transformation: square root					

Similarly, we did not observe significant interspecific differences from the DNA metabarcoding of pollen loads from three flower fly species (Fig. 10 and 11, Tab. 5 and 6), while as already observed for the main analysis reported in the beginning of this section, the DNA metabarcoding of rbcL provided significantly higher output in terms of ratio n. filtered / n. raw reads and cumulative n. of ASVs.

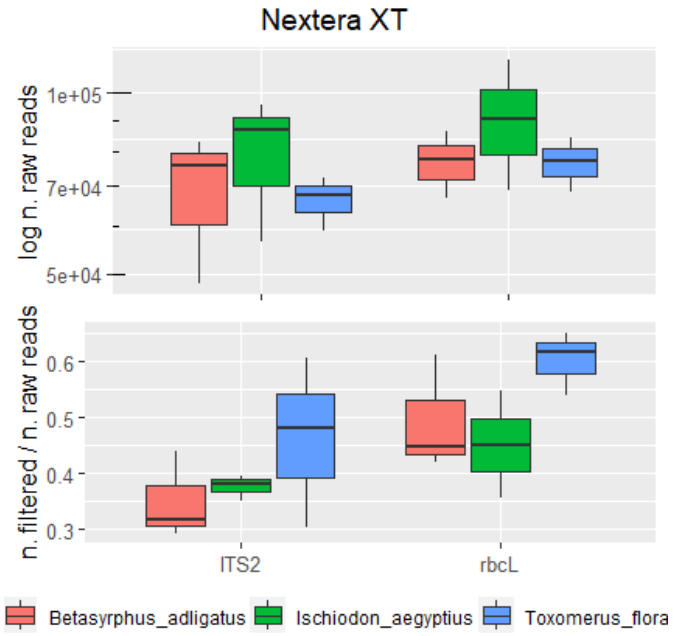

Fig. 10: N. of raw reads and on the ratio n. filtered / n. raw reads obtained from 18 DNA Nextera XT metabarcoding libraries following ITS2 and rbcl amplification in three sample groups including pollen loads from recently collected (2022) flower flies from three species (*Betasyrphus adligatus, Ischiodon aegyptius, Toxomerus floralis*).

Tab. 5: ANOVAs testing the effects of Nextera XT marker amplification (ITS2, rbcl) and flower fly species (*Betasyrphus adligatus, Ischiodon aegyptius, Toxomerus floralis*) on n. of raw reads and on the ratio n. filtered / n. raw reads in 18 DNA metabarcoding libraries. ***: p<0.001, ns: not significant.

		df	MS	F	P	
n. raw reads	MAR	1	4.3E+08	1.776	0.207	ns
	SPE	2	3.5E+08	1.419	0.280	ns
	MAR:SPE	2	4710460	0.019	0.981	ns
	Residual	12	2.4E+08			
	C = 0.353, p = 0.678, transformation: none					
n. filtered / n. raw reads	MAR	1	0.064	7.188	0.020	*
	SPE	2	0.026	2.955	0.090	ns
	MAR:SPE	2	0.002	0.255	0.779	ns
	Residual	12	0.009			
	C = 0.419, p= 0.394, transformation: none					

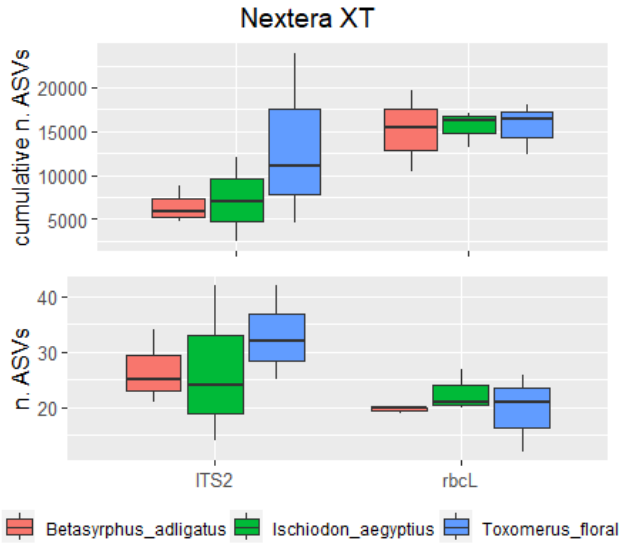

Fig. 11: N. of ASVs and cumulative n. of ASVs obtained from 18 DNA Nextera XT metabarcoding libraries following ITS2 and rbcl amplification in three sample groups including pollen loads from recently collected (2022) flower flies from three species (*Betasyrphus adligatus, Ischiodon aegyptius, Toxomerus floralis*).

Tab. 6: ANOVAs testing the effects of Nextera XT marker amplification (ITS2, rbcl) and flower fly species (*Betasyrphus adligatus, Ischiodon aegyptius, Toxomerus floralis*) on n. of ASVs and cumulative n. ASVs in 18 DNA metabarcoding libraries. *: p<0.05, ns: not significant.

		df	MS	F	P	
n. ASVs	MAR	1	296.056	4.62989	0.052	ns
	SPE	2	15.0556	0.23545	0.793	ns
	MAR:SPE	2	34.0556	0.53258	0.600	ns
	Residual	12	63.9444			
	C = 0.526, p = 0.142, transformation: none					

cumulative n. ASVs	MAR	1	1.9E+08	7.174	0.020	*
	SPE	2	2.3E+07	0.854	0.449	ns
	MAR:SPE	2	1.9E+07	0.726	0.503	ns
	Residual	12	2.6E+07			
	C = 0.594, p= 0.065, transformation: none					

These results suggest that DNA metabarcoding could be profitably implemented also from reduced amount of pollen, as is the case for pollen loads isolated from small-sized flower flies or wild bees. Pollen metabarcoding from the insect historical collections of RMCA/RBINS should be technically feasible by using standard and routinely used wet-lab procedures. Its technical feasibility however, would not exclude major problems with sample cross contamination (e.g. across specimens preserved in the same box).

A more in detail analysis on pollen compositional differences across the sample groups targeted by this project as well as on the performance of DNA metabarcoding IDs provided by different markers is currently ongoing in the the framework of the project ISeBAF (project partners M. Virgilio, JEMU RMCA and C. Vangestel, JEMU RBINS). These results will be communicated in the framework of the ongoing collaborative research between RMCA, , the Sokoine University of Agriculture, RBINS and the Botanical garden of Meise.

Time and cost-effectiveness of “in-house” vs. outsourced metabarcoding library prep

Besides the technical performance of pipelines for DNA metabarcoding, we also considered the time and cost-effectiveness of outsourcing all or part of the wet-lab pipelines to specialized companies. The main rationale behind this analysis was to reduce as much as possible the working costs for DNA metabarcoding. The main assumption was that external companies would provide the same (or higher) quality standards and output than those achieved with “in-house” library prep. Working costs were calculated both in terms of lab consumables and of personnel costs for wet-lab time, which were estimated by considering 96 libraries / working week / person, and a gross year salary / person of 60k €.

Below (Fig. 12) a schematic representation of cost/sample as calculated for a batch of 96 samples (as this is the sample size commonly loaded on a Miseq lane for HTS) and one marker (e.g. ITS2) is shown. The lowest cost / samples are reported in orange and to an offer for a batch of minimum 227 samples.

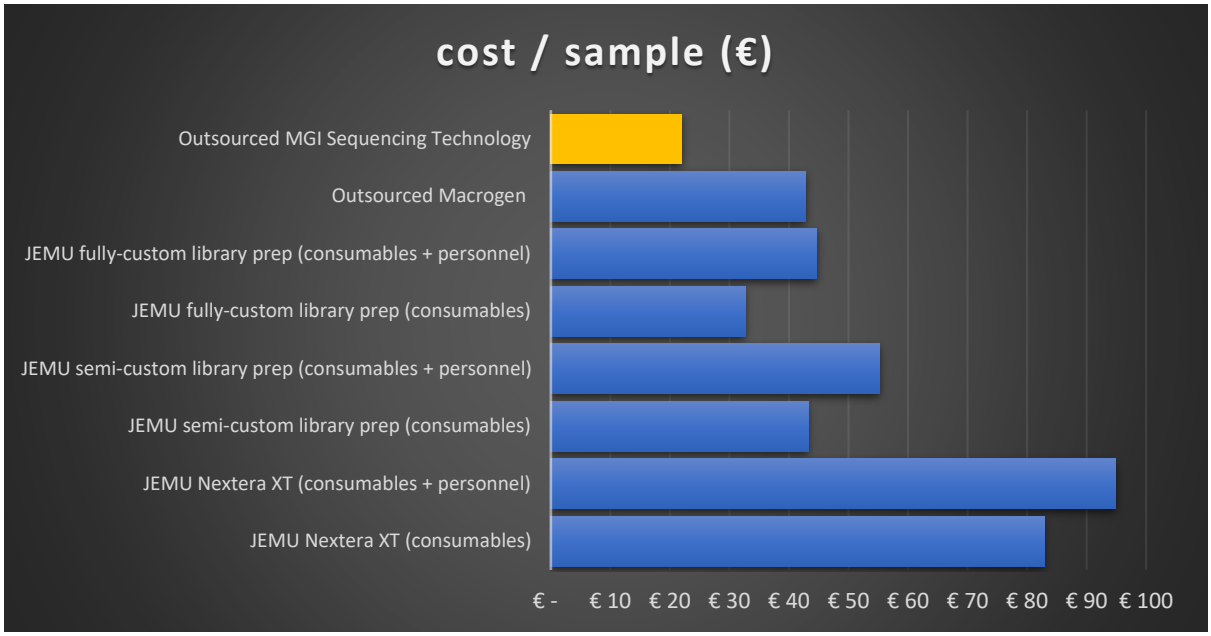

Fig. 12: Total cost / sample including library prep and HTS, as calculated for a typical run of 96 samples on a Miseq lane (in blue). The lowest cost/sample is represented in orange and refers to a batch of minimum 227 samples (in yellow). Costs for “in-house” library prep (labelled as “JEMU”) are provided for consumables only and for consumables and personnel costs.

Irrespectively of performance (which was generally lower in the semi-custom pipeline, compared to the pipeline based on a commercial kit and regardless we acknowledge that in-house library prep might still present some advantage in terms of scalability, ad hoc optimization, last minute substitution of a few samples based on QC, etc., it is very clear how outsourcing provides the most convenient option for the DNA metabarcoding of relatively large batches of samples (at least 96 samples, with the lowest costs obtained with an investment of 5k € for 227 samples (see supplementary data S7).

Conclusions and recommendations

These tests were conceived to verify the suitability of field and lab procedures to process pollen loads recovered from flower flies preserved in Ethanol (as this is the preservative we generally use to transport insect samples from Africa).

Based on the results obtained, EtOH sample preservation, followed by Qiagen DNeasy Plant Mini Kit DNA extraction seems to be a suitable (possibly the most suitable) combination for pollen DNA barcoding in flower flies.

Conversely, CTAB preservation and CTAB DNA extraction provided inconsistent results, probably due to cross contamination between the insect voucher and the pollen recovered from its body (in fact, the high DNA concentration measured in CTAB seems to be artefactual).

Furthermore, compared to CTAB, the Qiagen protocol is faster (2-3 hours vs 4-6 hours), highly standardised and safer for the health of the operator (as not using β-mercaptoethanol). But in comparison, far more expensive.

The comparisons implemented between "in-house" metabarcoding library prep using a popular commercial kit and a semi-custom made pipeline showed that this latter seems to have poorer performances compared to library prep based on commercial kit. However, even if we observed a significantly lower output in n. of raw reads and cumulative n. of ASVs, both the semi-custom pipeline and the commercial kit have comparable output in terms of n. of ASVs recovered and of ratio n. filtered / n. raw reads. Regardless of these differences, outsourcing seems to provide the most cost- and time-effective approach to DNA metabarcoding, and should be currently considered as the best option particularly for the routine processing of large batches of samples.

Acknowledgements

This JEMU project was co-financed via ISeBAF (BRAIN-be 2.0) and SYNTHESYS+,

Selected References

Bafeel, S. O., Arif, I. A., Bakir, M. A., Khan, H. A., Al Farhan, A. H., Al Homaidan, A. A., Ahamed, A., & Thomas, J. (2011). Comparative evaluation of PCR success with universal primers of maturase K (matK) and ribulose-1, 5-bisphosphate carboxylase oxygenase large subunit (rbcL) for barcoding of some arid plants. Plant OMICS, 4(4), 195–198.

Bell, K. L., De Vere, N., Keller, A., Richardson, R. T., Gous, A., Burgess, K. S., & Brosi, B. J. (2016). Pollen DNA barcoding: Current applications and future prospects. *Genome*, *59*(9), 629–640. https://doi.org/10.1139/gen-2015-0200

Bell, K. L., Loeffler, V. M., & Brosi, B. J. (2017). An rbcL Reference Library to Aid in the Identification of Plant Species Mixtures by DNA Metabarcoding . *Applications in Plant Sciences*, *5*(3), 1600110. https://doi.org/10.3732/apps.1600110

Bruns, T. D., Lee, S. B., & Taylor, J. W. (1990). *Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics*. *May 2014*.

Chen, S., Yao, H., Han, J., Liu, C., Song, J., Shi, L., Zhu, Y., Ma, X., Gao, T., Pang, X., Luo, K., Li, Y., Li, X., Jia, X., Lin, Y., & Leon, C. (2010). Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE*, *5*(1), 1–8. https://doi.org/10.1371/journal.pone.0008613

Cheng, T., Xu, C., Lei, L., Li, C., Zhang, Y., & Zhou, S. (2016). Barcoding the kingdom Plantae: New PCR primers for ITS regions of plants with improved universality and specificity. *Molecular Ecology Resources*, *16*(1), 138–149. https://doi.org/10.1111/1755-0998.12438

Coghlan, S. A., Shafer, A. B. A., & Freeland, J. R. (2021). *Development of an environmental DNA metabarcoding assay for aquatic vascular plant communities*. *June 2020*, 372–387. https://doi.org/10.1002/edn3.120

De Vere, N., Rich, T. C. G., Ford, C. R., Trinder, S. A., Long, C., Moore, C. W., Satterthwaite, D., Davies, H., Allainguillaume, J., Ronca, S., Tatarinova, T., Garbett, H., Walker, K., & Wilkinson, M. J. (2012). *DNA Barcoding the Native Flowering Plants and Conifers of Wales*. *7*(6), 1–12. https://doi.org/10.1371/journal.pone.0037945

De Vere, N., Jones, L. E., Gilmore, T., Moscrop, J., Lowe, A., Smith, D., Hegarty, M. J., Creer, S., & Ford, C. R. (2017). Using DNA metabarcoding to investigate honey bee foraging reveals limited flower use despite high floral availability. *Scientific Reports*, *7*(January), 1–10. https://doi.org/10.1038/srep42838

Doyle, J. J., & Doyle, J. L. (1987). Doyle_plantDNAextractCTAB_1987.pdf. In *Phytochemical Bulletin* (Vol. 19, Issue 1, pp. 11–15). https://webpages.uncc.edu/~jweller2/pages/BINF8350f2011/BINF8350_Readings/Doyle_plantDNAextractCTAB_1987.pdf

Erickson, D. L., Reed, E., Ramachandran, P., Bourg, N. A., McShea, W. J., & Ottesen, A. (2017). Reconstructing a herbivore's diet using a novel rbcL DNA mini-barcode for plants. *AoB PLANTS*, *9*(3). https://doi.org/10.1093/aobpla/plx015

Fazekas, A. J., Burgess, K. S., Kesanakurti, P. R., Graham, S. W., Newmaster, S. G., Husband, B. C., Percy, D. M., Hajibabaei, M., & Barrett, S. C. H. (2008). Multiple multilocus DNA barcodes from the plastid genome discriminate plant species equally well. *PLoS ONE*, *3*(7). https://doi.org/10.1371/journal.pone.0002802

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/10.1071/ZO9660275

García-Robledo, C., Erickson, D. L., Staines, C. L., Erwin, T. L., & Kress, W. J. (2013). Tropical Plant-Herbivore Networks: Reconstructing Species Interactions Using DNA Barcodes. *PLoS ONE*, *8*(1). https://doi.org/10.1371/journal.pone.0052967

Gous, A., Swanevelder, D. Z. H., Eardley, C. D., & Willows-Munro, S. (2019). Plant–pollinator interactions over time: Pollen metabarcoding from bees in a historic collection. *Evolutionary Applications*, *12*(2), 187–197. https://doi.org/10.1111/eva.12707

Henry, R. J. (2009). Plant DNA extraction. *Plant Genotyping: The DNA Fingerprinting of Plants*, 239–249. https://doi.org/10.1079/9780851995151.0239

Kress, W. J., & Erickson, D. L. (2007). A Two-Locus Global DNA Barcode for Land Plants: The Coding rbcL Gene Complements the Non-Coding trnH-psbA Spacer Region. *PLoS ONE*, *2*(6). https://doi.org/10.1371/journal.pone.0000508

Laha, R. C., Mandal, S. De, Ralte, L., Ralte, L., & Kumar, N. S. (2017). Meta - barcoding in combination with palynological inference is a potent diagnostic marker for honey floral composition. *AMB Express*. https://doi.org/10.1186/s13568-017-0429-7

Lalhmangaihi, R., Ghatak, S., Laha, R., Gurusubramanian, G., & Senthil Kumar, N. (2014). Protocol for optimal quality and quantity pollen DNA isolation from Honey samples. *Journal of Biomolecular Techniques*, *25*(4), 92–95. https://doi.org/10.7171/jbt.14-2504-001

Lucas, A., Bodger, O., Brosi, B. J., Ford, C. R., Forman, D. W., Greig, C., Hegarty, M., Neyland, P. J., & de Vere, N. (2018). Generalisation and specialisation in hoverfly (Syrphidae) grassland pollen transport networks revealed by DNA metabarcoding. *Journal of Animal Ecology*, *87*(4), 1008–1021. https://doi.org/10.1111/1365-2656.12828

Moorhouse-Gann, R. J., Dunn, J. C., Vere, N. de, Goder, M., Cole, N., Hipperson, H., & Symondson, W. O. C. (2018). New universal ITS2 primers for high-resolution herbivory analyses using DNA metabarcoding in both tropical and temperate zones. *Scientific Reports*, *November 2017*, 1–15. https://doi.org/10.1038/s41598-018-26648-2

Newmaster, S. G., Fazekas, A. J., & Ragupathy, S. (2006). DNA barcoding in land plants: Evaluation of rbcL in a multigene tiered approach. *Canadian Journal of Botany*, *84*(3), 335–341. https://doi.org/10.1139/B06-047

Peel, N., Dicks, L. V., Clark, M. D., Heavens, D., Percival-Alwyn, L., Cooper, C., Davies, R. G., Leggett, R. M., & Yu, D. W. (2018). *Semi-quantitative characterisation of mixed pollen samples using MinION sequencing and Reverse Metagenomics*. 2–41.

Prosser, S. W. J., & Hebert, P. D. N. (2017). Rapid identification of the botanical and entomological sources of honey using DNA metabarcoding. *Food Chemistry*, *214*, 183–191. https://doi.org/https://doi.org/10.1016/j.foodchem.2016.07.077

Richardson, R. T., Lin, C.-H., Sponsler, D. B., Quijia, J. O., Goodell, K., & Johnson, R. M. (2015). Application of ITS2 Metabarcoding to Determine the Provenance of Pollen Collected by Honey Bees in an Agroecosystem. *Applications in Plant Sciences*, *3*(1), 1400066. https://doi.org/10.3732/apps.1400066

Sandrini-Neto, L., & Camargo, M. G. (2015). GAD: an R package for ANOVA designs from general principles. *Available on CRAN*Vere, N. De, Rich, T. C. G., Ford, C. R., Trinder, S. A., Long, C., Moore, C. W., Satterthwaite, D., Davies, H., Allainguillaume, J., Ronca, S., Tatarinova, T., Garbett, H., Walker, K., & Wilkinson, M. J. (2012). *DNA Barcoding the Native Flowering Plants and Conifers of Wales*. *7*(6), 1–12. https://doi.org/10.1371/journal.pone.0037945

White, T. J., Bruns, T. D., Lee, S. B., & Taylor, J. W. (1990). *Amplification and direct sequencing of fungal ribosomal RNA Genes for phylogenetics*.

Wilson, E. E., Sidhu, C. S., Levan, K. E., & Holway, D. A. (2010). Pollen foraging behaviour of solitary Hawaiian bees revealed through molecular pollen analysis. *Molecular Ecology*, *19*(21), 4823–4829. https://doi.org/10.1111/j.1365-294X.2010.04849.x

Zhang, G., Liu, J., Gao, M., Kong, W., Zhao, Q., Shi, L., & Wang, Q. (2020). Tracing the Edible and Medicinal Plant Pueraria montana and Its Products in the Marketplace Yields Subspecies Level Distinction Using DNA Barcoding and DNA Metabarcoding. In *Frontiers in Pharmacology* (Vol. 11). https://www.frontiersin.org/article/10.3389/fphar.2020.00336

Supplementary data

S1: modified CTAB DNA extraction protocol (courtesy of S. Janssens – RBG Meise)

Products

- CTAB 2x
  - 100 mM Tris-HCl
  - 1,4 M NaCl
  - 20 mM EDTA
  - 2% (w/v) CTAB
  - 1% (w/v) PVP-40
- Chloroform/isoamylalcohol 24:1
- Isopropanol (4°C)
- 70% ethanol (-20°c)
- TE buffer (pH 8,0) (4°C)
  - 10 mM Tris-HCl
  - 1 mM EDTA
- RNase A (1/10) (4°C)
  - 1 mg/mL RNase A

Protocol

1) The fly was immersed in 1 ml of 100% ethanol or CTAB solution in an Eppendorf tube
2) Pre-heat the water bath/thermoshaker 60°C
3) Shake the tube vigorously by hand for 2 min or by shaker at 5Hz for 2 min.
4) Remove the fly using forceps to a clean 1.5 ml SCREWCAP tube and add 500µl of CTAB or 100% ethanol to the new tube.
5) Shake the tube vigorously by hand for 2 min or by shaker at 5Hz for 2 min.
6) Remove the fly using forceps into a new 1.5ml clean tube and frozen at −20°C for subsequent species identification.
7) Transfer the entire volume of the first tube into the screw cap tube.
8) The tube containing the detergent and pollen was centrifuged at 13,000 rpm for 5 min
9) Remove the supernatant and dry the samples in the SPEEDVAC for 1h. (30min for ethanol).
10) Put three 3-mm stainless steel beads into each tube and disrupt the sample for 2 min at 22.5 Hz.
11) Resuspend the samples in 400 µl CTAB 2x extraction buffer and 0.5 µl (0,1%) of β-Mercaptoethanol (or 1,5µl (0,3%) might be tested) UNDER the hood
12) Vortex for 10 s
13) Put the sample in 60° C for 1h (or overnight)
14) Let the samples cool down.
15) Prepare tubes with 300 µl of Isopropanol and store in the fridge at -20°C for 30min.
16) Centrifuge the tubes at 13,000 rpm for 5 min. The ≈400 µl of supernatant was extracted to a clean 1.5 ml tube. KEEP SUPERNATANT and UNDER the hood
17) Add an equal volume (400µl) of Chloroform:Isoamylalcohol (24:1) UNDER the hood

18) Vortex vigorously for at least 5s and shake the sample for 2 min at 6Hz TWO Times.
19) Centrifuge for 15 min at 11000 rmp at 22°C.
20) Carefully transfer 400 µL of the upper aqueous phase to the isopropanol prepared tubes (0,8 v/v). CAREFUL not to disrupt the middle and bottom phase (by gently transferring the upper phase in multiple steps of 200 ul, aim for a total collected volume of ≈ 350 µl) With pollen we will only do this phase once. UNDER the hood
21) Shake by up and down (+- 50 times)

22) Store 20 min at -20°C for maximal precipitation
23) Centrifuge 10 minutes at 14000 rpm (preferably at 4°C).
24) Remove the supernatant by pouring the liquid in the isopropanol waste and place the tubes inverted for a few minutes.
25) Add 400 µL 70% cold ethanol and loosen the pellet.
26) Store 20 min at -20°C for maximal precipitation
27) Centrifuge for 10 min at 20000 rcf at 4°C.

28) Remove the supernatant by pouring the liquid in the ethanol waste and place the tubes inverted for a few minutes.
29) Put the tubes horizontal and let them air dry for around 1h. or use the speedvac for 30min.
30) Dissolve the pellet by adding 30ul TE buffer to the dried tubes.
31) Tap the pellet to dissolve it in the TE buffer
32) Add 3 µL RNase A (1/10) per tube, shake and spin down.
33) Incubate maximum 2 minutes at room temperature.
34) Store the DNA in the fridge or freezer.

S2: ANOVA on DNA concentrations and yields as obtained from pollen isolated from 50 flower flies. Test for the effects of pollen DNA extraction protocol (Qiagen vs CTAB) and preservation group (a, …, e, see Tab. 1). n.s.: not significant; ∗: P < 0.05, ∗∗: P < 0.01, *** P < 0.001. C: Cochran's test.

DNA concentrations

ANOVA

	Df	MS	F	
DNA extraction (Extr)	1	0.05642	0.6986	n.s.
Preservation (Pres-	4	1.04425	12.9309	***
Extr x Pres	4	0.53808	6.6630	***
Residual	40	0.08076		

Transformation = Fourth root

C = 0.854 n.s.

A posteriori comparisons: Extr x Pres

Qiagen extraction: EtOH RT = Et -20°C = control = CTAB-20°C = CTAB RT

CTAB extraction: control = Et -20°C = EtOH RT < CTAB-20°C < CTAB RT

DNA yield

ANOVA

	Df	MS	F	
DNA extraction (Extr)	1	0.7421	1.2837	ns
Preservation (Pres)	4	6.6469	11.4985	***
Extr x Pres	4	2.9009	5.0182	**
Residual	40	0.5781		

Transformation = Fourth root

C = 0.530 n.s.

A posteriori comparisons: Extr x Pres

Qiagen extraction: EtOH RT = Et -20°C = control = CTAB-20°C = CTAB RT

CTAB extraction: control = Et -20°C = EtOH RT < CTAB-20°C < CTAB RT

S3: Optimal annealing temperatures for each of the primer pairs tested

Marker	Forward primer	Reverse primer	Primer pair ID	Annealing Temperature	Primer sequences
matK	KIM3	KIM1	KIM	57°C	CGTACAGTACTTTTGTGTTTACGAG - ACCCAGTCCATCTGGAAATCTTGGTTC+
rbcL	rbcL-Tag-F	rbcL506	rbcL-506	58°C	ATG TCA CCA CAA ACA GAG ACT - AGGGGACGACCATACTTGTTCA
	rbcL-Tag-F	rbcL375	rbcL-375	58°C	ATG TCA CCA CAA ACA GAG ACT - GCAAATCCTCCAGACGTAGA
	rbcL-Tag-F	rbcL320	rbcL-320	61°C	ATG TCA CCA CAA ACA GAG ACT - ACCCACAATGGAAGTAAACATGT
	rbcL_230	rbcL506	rbcL-23506	56°C	CTTACCAGYCTTGATCGTTACAAAGG - AGGGGACGACCATACTTGTTCA
	rbcLF-Tag_IL	rbcLR-Tag_IL	rbcL-T	62°C	ATG TCA CCA CAA ACA GAG ACT - GA AAC GGT CTC TCC AAC GCA T
	rbcL_260	rbcL_230	rbcL-2623	58°C	GTAAAATCAAGTCCACCRCG - CTTACCAGYCTTGATCGTTACAAAGG
	rbcL-af	rbcL-ar	rbcL-A	/	ATGTCACCACAAACAGAGACTAAAGC - CTTCTGCTACAAATAAGAATCGATCTC
ITS1	ITS1-18S	ITS1-390	ITS1-390	58°C	AGTCGTAACAAGGTTTCCGT - GGGATTCTGCAATTCACACC
	ITS1-18S	ITS1-380	ITS1-380	60°C	AGTCGTAACAAGGTTTCCGT - AACTTGCGTTCAAAGACTCG
ITS2	ITS2	ITS3a	ITS2-23	53°C	ATGCGATACTTGGTGTGAAT - GACGCTTCTCCAGACTACAAT
	ITS3	ITS4	ITS2-34	56°C	GCA TCG ATG AAG AAC GCA GC - TC CTC CGC TTA TTG ATA TGC
	ITS5a	ITS4	ITS2-54	54°C	CCTTATCATTTAGAGGAAGGAG - TCCTCCGCTTATTGATATGC
	UniPlantF	UniplantR	ITS2-Uni	55°C	TGTGAATTGCARRATYCMG - CCCGHYTGAYYTGRGGTCDC

S4: Metadata sequenced samples

Specimen	Library Prep	Marker	sample group	Collection Year	taxon group	Species
MSC_HE_04	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
CMC_HC_02	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
CMC_HC_04	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
CMB_HC_08	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
CMC_HC_08	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMC_HC_02	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
CMC_HC_04	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
CMB_HC_08	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
MSC_HE_04	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
MSC_HE_06	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
MSW_HE_03	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
MSC_HE_06	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
MSC_HD_06	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
MSW_HD_08	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
MSW_HD_06	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMW_HC_A	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMW_HC_B	Semi-Custom	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
BM_4	Semi-Custom	ITS2	historical	1922	Hymenoptera, Apoidea	*Apis mellifera*

BM_5	Semi-Custom	ITS2	historical	1922	Hymenoptera, Apoidea	*Apis mellifera*
BM_32	Semi-Custom	ITS2	historical	1963	Hymenoptera, Apoidea	*Apis mellifera*
BM_33	Semi-Custom	ITS2	historical	1993	Hymenoptera, Apoidea	*Apis mellifera*
BM_4	Semi-Custom	rbcL	historical	1922	Hymenoptera, Apoidea	*Apis mellifera*
MSW_HE_03	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
BM_5	Semi-Custom	rbcL	historical	1922	Hymenoptera, Apoidea	*Apis mellifera*
BM_32	Semi-Custom	rbcL	historical	1963	Hymenoptera, Apoidea	*Apis mellifera*
BM_33	Semi-Custom	rbcL	historical	1993	Hymenoptera, Apoidea	*Apis mellifera*
VB_BB_02	Semi-Custom	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
VB_BB_06	Semi-Custom	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
VB_BB_08	Semi-Custom	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_04	Semi-Custom	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_07	Semi-Custom	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_03	Semi-Custom	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
VB_BB_02	Semi-Custom	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
MSC_HD_06	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
VB_BB_06	Semi-Custom	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
VB_BB_08	Semi-Custom	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_04	Semi-Custom	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_07	Semi-Custom	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_03	Semi-Custom	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
409 51D	Semi-Custom	ITS2	historical	1947	Hymenoptera, Apoidea	*Apis mellifera*
409 50D	Semi-Custom	ITS2	historical	1947	Hymenoptera, Apoidea	*Apis mellifera*
409 51D	Semi-Custom	rbcL	historical	1947	Hymenoptera, Apoidea	*Apis mellifera*
409 50D	Semi-Custom	rbcL	historical	1947	Hymenoptera, Apoidea	*Apis mellifera*
MSW_HD_08	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
MSW_HD_06	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMW_HC_A	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMW_HC_B	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMC_HC_08	Semi-Custom	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
MSC_HE_04	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
CMC_HC_02	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
CMC_HC_04	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
CMB_HC_08	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
CMC_HC_08	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMC_HC_02	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
CMC_HC_04	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
CMB_HC_08	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Ischiodon aegyptius*
MSC_HE_04	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
MSC_HE_06	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
MSW_HE_03	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
MSC_HE_06	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
MSC_HD_06	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
MSW_HD_08	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
MSW_HD_06	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMW_HC_A	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMW_HC_B	Nextera XT	rbcL	recently collected	2022	Syrphidae	*Toxomerus floralis*
BM_4	Nextera XT	ITS2	historical	1922	Hymenoptera, Apoidea	*Apis mellifera*
BM_5	Nextera XT	ITS2	historical	1922	Hymenoptera, Apoidea	*Apis mellifera*
BM_32	Nextera XT	ITS2	historical	1963	Hymenoptera, Apoidea	*Apis mellifera*
BM_33	Nextera XT	ITS2	historical	1993	Hymenoptera, Apoidea	*Apis mellifera*
BM_4	Nextera XT	rbcL	historical	1922	Hymenoptera, Apoidea	*Apis mellifera*
MSW_HE_03	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
BM_5	Nextera XT	rbcL	historical	1922	Hymenoptera, Apoidea	*Apis mellifera*
BM_32	Nextera XT	rbcL	historical	1963	Hymenoptera, Apoidea	*Apis mellifera*
BM_33	Nextera XT	rbcL	historical	1993	Hymenoptera, Apoidea	*Apis mellifera*
VB_BB_02	Nextera XT	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
VB_BB_06	Nextera XT	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
VB_BB_08	Nextera XT	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_04	Nextera XT	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_07	Nextera XT	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_03	Nextera XT	ITS2	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*

VB_BB_02	Nextera XT	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
MSC_HD_06	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Betasyrphus adligatus*
VB_BB_06	Nextera XT	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
VB_BB_08	Nextera XT	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_04	Nextera XT	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_07	Nextera XT	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
KBHB_BE_03	Nextera XT	rbcL	recently collected	2022	Hymenoptera, Apoidea	*Apis mellifera*
409 51D	Nextera XT	rbcL	historical	1947	Hymenoptera, Apoidea	*Apis mellifera*
409 50D	Nextera XT	rbcL	historical	1947	Hymenoptera, Apoidea	*Apis mellifera*
409 51D	Nextera XT	ITS2	historical	1947	Hymenoptera, Apoidea	*Apis mellifera*
409 50D	Nextera XT	ITS2	historical	1947	Hymenoptera, Apoidea	*Apis mellifera*
MSW_HD_08	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
MSW_HD_06	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMW_HC_A	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMW_HC_B	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*
CMC_HC_08	Nextera XT	ITS2	recently collected	2022	Syrphidae	*Toxomerus floralis*

S5: Detailed description of the custom pipeline for metabarcoding pollen

There are two Amplicon PCR, the first one will be done with ITS2 primers on the 24 samples of pollen DNA and a second one will be done with rbcL primers on the same 24 samples of pollen DNA.

After the first PCR Clean-up, 48 amplicon PCR samples will be processed.

1st Amplicon PCR:

Consumables

Item

Genomic DNA

Amplicon PCR Reverse Primer (20 μM)

Amplicon PCR Forward Primer (20 μM)

Taq PLATINUM

MgCl2 PLATINUM

dNTP

MQ water

96 - well 0.2 ml PCR plate 1 plate

Procedure

1 Set up the following reaction of DNA, Taq PLATINUM mix, and primers:

Volume	
5µl (or double with low quantities)	Genomic DNA (2ng/µl – 0,1ng/µl)
2,50µl	PCR Buffer(10x)
0,75µl	MgCl PLATINUM 50mM
2,50µl	dNTP 2mM
0,50µl	Primer1= LF1 20µM
0,50µl	Primer2=LR1 20µM
0,15µl	Taq PLATINUM 5 Units/µl
13,10µl	MQ H2O
Total	25 µl

2 Perform PCR in a thermal cycler using the following program:

• 94° C for 3 minutes

• 35 cycles of:

— 94° C for 30 seconds

— 55° C for 30 seconds

— 72° C for 30 seconds

• 72° C for 5 minutes

• Hold at 4° C

- Check the success of amplification with electrophoresis gel. Redo every samples than didn't work

1st PCR Clean up:

Consumables

Item Quantity Storage

MilliQ water

HighPrep™ PCR beads

Freshly Prepared 80% Ethanol (EtOH)

96 - well 0.2 ml PCR plate 1 plate

Preparation

• Bring the HighPrep™ PCR beads to room temperature.

• Prepare the plate according to plan with 20 µl and a balance

Procedure

1 Centrifuge the Amplicon PCR plate at 1,000 × g at 20° C for 1 minute to collect condensation, carefully remove seal. (Not necessary if next day)

2 Vortex the HighPrep™ PCR beads for 30 seconds to make sure that the beads are evenly dispersed. Add an appropriate volume of beads to a trough depending on the number of samples processing.

3 Using a multichannel pipette, add 20 µl of HighPrep™ PCR beads (1:1 proportion) to each well of the Amplicon PCR plate. Change tips between columns.

4 Gently pipette entire volume up and down 10 times if using a 96 - well PCR plate

5 Incubate at room temperature without shaking for 5 minutes.

6 Place the plate on a magnetic stand for 2 minutes or until the supernatant has cleared.

7 With the Amplicon PCR plate on the magnetic stand, use a multichannel pipette to remove and discard the supernatant. Change tips between samples.

8 With the Amplicon PCR plate on the magnetic stand, wash the beads with freshly prepared 80% ethanol as follows:

a Using a multichannel pipette, add 200 µl of freshly prepared 80% ethanol to each sample well.

b Incubate the plate on the magnetic stand for 30 seconds.

c Carefully remove and discard the supernatant.

10 With the Amplicon PCR plate on the magnetic stand, perform a second ethanol wash as follows:

a Using a multichannel pipette, add 200 µl of freshly prepared 80% ethanol to each sample well.

b Incubate the plate on the magnetic stand for 30 seconds.

c Carefully remove and discard the supernatant.

d Use a P20 multichannel pipette with fine pipette tips to remove excess ethanol.

11 With the Amplicon PCR plate still on the magnetic stand, allow the beads to air - dry for 10 minutes or until there is not ethanol left.

12 Remove the Amplicon PCR plate from the magnetic stand. Using a multichannel pipette, add 52.5 µl of MilliQ water to each well of the Amplicon PCR plate.

13 Gently pipette mix up and down 10 times, changing tips after each column (or seal plate and shake at 1800 rpm for 2 minutes). Make sure that beads are fully resuspended.

14 Incubate at room temperature for 2 minutes.

15 Place the plate on the magnetic stand for 2 minutes or until the supernatant has cleared.

16 Using a multichannel pipette, carefully transfer 50 µl of the supernatant from the Amplicon PCR plate to a new 96 - well PCR plate. Change tips between samples to avoid cross - contamination.

17 Using the last 2,5µl in plate, check the result of the cleaning with a gel.

SAFE STOPPING POINT

If you do not immediately proceed to *Index PCR*, seal plate

2nd Index PCR:

This step attaches dual indices and Illumina sequencing adapters using the Nextera XT Index Kit.

Consumables

Item Quantity Storage

Nextera XT Index 1 Primers (N7XX) from the Nextera XT Index kit (FC - 131 - 1001 or FC - 131 - 1002)

Nextera XT Index 2 Primers (S5XX) from the Nextera XT Index kit (FC - 131 - 1001 or FC - 131 - 1002)

Taq PLATINUM

MgCl2 PLATINUM

dNTP

MQ water

96 - well 0.2 ml PCR plate 1 plate

Microseal 'A' film

Procedure

1 Using a multichannel pipette, transfer 5 µl from each 48 well to a new 96 - well plate. The remaining 45 µl is not used in the protocol and can be stored for other uses.

2 Arrange the Index 1 and 2 primers in a rack using the following arrangements as needed:

a Arrange Index 2 primer tubes (white caps, clear solution) vertically, aligned with rows A through D. (Half a plate as the other half is going to be Nextera XT kit).

b Arrange Index 1 primer tubes (orange caps, yellow solution) horizontally, aligned with columns 1 through 12.

3 Place the 96 - well PCR plate with the 5 μl of resuspended PCR product DNA.

4 Set up the following reaction of DNA, Taq PLATINUM mix, and primers:

Volume	
5 µl	Amplicon DNA
2,50 µl	PCR Buffer(10x)
1,5 µl	MgCl PLATINUM 50mM
5,0 µl	dNTP 2mM
5 µl	Nextera XT Index Primer 1 (N7xx)
5 µl	Nextera XT Index Primer 2 (S5xx)
0,3 µl	Taq PLATINUM 5 Units/µl
34 µl	MQ H2O
Total	50 µl

5 Gently pipette up and down 10 times to mix.

6 Cover the plate with Microseal 'A'.

7 Centrifuge the plate at 1,000 × g at 20° C for 1 minute.

8 Perform PCR on a thermal cycler using the following program:

- 95° C for 3 minutes
- 8 cycles of:
  - 95° C for 30 seconds
  - 55° C for 30 seconds
  - 72° C for 60 seconds
- 72° C for 10 minutes
- Hold at 4° C

2nd PCR Clean up:

Consumables

Item	Quantity	Storage
10 mM Tris pH 8.5		
HighPrep™ PCR beads		
Freshly Prepared 80% Ethanol (EtOH)		
96 - well 0.2 ml PCR plate	1 plate	

Procedure

1 Centrifuge the Index PCR plate at 280 × g at 20° C for 1 minute to collect condensation.

2 Vortex the HighPrep™ PCR beads for 30 seconds to make sure that the beads are evenly dispersed. Add an appropriate volume of beads to a trough.

3 Using a multichannel pipette, add 56 μl of HighPrep™ PCR beads to each well of the Index PCR plate. (1:1 beads proportion for 20 µl of PCR products)

4 Gently pipette mix up and down 10 times if using a 96 - well PCR plate

5 Incubate at room temperature without shaking for 5 minutes.

6 Place the plate on a magnetic stand for 2 minutes or until the supernatant has cleared.

7 With the Index PCR plate on the magnetic stand, use a multichannel pipette to remove and discard the supernatant. Change tips between samples.

8 With the Index PCR plate on the magnetic stand, wash the beads with freshly prepared 80% ethanol as follows:

a Using a multichannel pipette, add 200 µl of freshly prepared 80% ethanol to each sample well.

b Incubate the plate on the magnetic stand for 30 seconds.

c Carefully remove and discard the supernatant.

9 With the Index PCR plate on the magnetic stand, perform a second ethanol wash as follows:

a Using a multichannel pipette, add 200 µl of freshly prepared 80% ethanol to each sample well.

b Incubate the plate on the magnetic stand for 30 seconds.

c Carefully remove and discard the supernatant.

d Use a P20 multichannel pipette with fine pipette tips to remove excess ethanol.

10 With the Index PCR plate still on the magnetic stand, allow the beads to air - dry for 10 minutes.

11 Remove the Index PCR plate from the magnetic stand. Using a multichannel pipette, add 27.5 µl of 10 mM Tris pH 8.5 to each well of the Index PCR plate.

12 If using a 96 - well PCR plate, gently pipette mix up and down 10 times until beads are fully resuspended, changing tips after each column.

13 Incubate at room temperature for 2 minutes.

14 Place the plate on the magnetic stand for 2 minutes or until the supernatant has cleared.

15 Using a multichannel pipette, carefully transfer 25 µl of the supernatant from the Index PCR plate to a new 96 - well PCR plate. Change tips between samples to avoid crosscontamination.

SAFE STOPPING POINT

If you do not immediately proceed to *Index PCR*, seal plate

Validate Library

Run 1 µl of a 1:50 dilution of the final library on a Bioanalyzer DNA 1000 chip to verify the size.

Library Quantification, Normalization, and Pooling:

Illumina recommends quantifying your libraries using a fluorometric quantification method that uses dsDNA binding dyes. Calculate DNA concentration in nM, based on the size of DNA amplicons as determined by an Agilent Technologies 2100 Bioanalyzer trace:

$$\frac{(\text{concentration in ng/µl})}{(660\ \text{g/mol} \times \text{average library size})} \times 10^6 = \text{concentration in nM}$$

Dilute concentrated final library using Resuspension Buffer (RSB) or 10 mM Tris pH 8.5 to 4 nM. Aliquot 5 μl of diluted DNA from each library and mix aliquots for pooling libraries with unique indices. Final volume of 480µl Depending on coverage needs, up to 96 libraries can be pooled for one MiSeq run.

S6: Detailed description of the Nextera XT DNA Library Prep Kit for metabarcoding pollen

The Nextera Kit can be used with PCR Amplicons. The PCR amplicon must be > 300 bp. Shorter amplicons can be lost during the library cleanup step.

Amplicon PCR on Genomic DNA

There is two PCR, the first one will be done with ITS2 primers on the 24 samples of pollen DNA and a second one will be done with rbcL primers on the same 24 samples of pollen DNA.

1 Set up the following reaction of DNA, Taq PLATINUM mix, and primers:

Volume	
2.5 µl	Genomic DNA (2ng/µl – 0,1ng/µl)
2,50µl	PCR Buffer(10x)
0,75µl	MgCl PLATINUM 50mM
2,50µl	dNTP 2mM
0,50µl	Primer1= LF1 20µM
0,50µl	Primer2=LR1 20µM
0,15µl	Taq PLATINUM 5 Units/µl
15,75µl	MQ H2O
Total	25 µl

2 Perform PCR in a thermal cycler using the following program:

- 95° C for 3 minutes
- 25 cycles of:
  - 95° C for 30 seconds
  - 55° C for 30 seconds
  - 72° C for 30 seconds
- 72° C for 5 minutes
- Hold at 4° C

After this first PCR, 48 samples of PCR amplicons will be processed.

Tagment Amplicon DNA

This step uses the Nextera transposome to tagment gDNA, which is a process that fragments DNA and then tags the DNA with adapter sequences in a single step.

Consumables

- ATM (Amplicon Tagment Mix)
- TD (Tagment DNA Buffer)
- NT (Neutralize Tagment Buffer)
- gDNA (0.2 ng/µl per sample)/ Amplicon DNA
- Hard-Shell 96-well PCR plate, skirted
- Microseal 'B' adhesive seals

Preparation

1. Prepare the following consumables:

Item Storage Instructions
gDNA -25°C to -15°C Thaw on ice. Invert the thawed tubes 3–5 times, and then centrifuge briefly.
ATM -25°C to -15°C Thaw on ice. Invert the thawed tubes 3–5 times, and then centrifuge briefly.
TD -25°C to -15°C Thaw on ice. Invert the thawed tubes 3–5 times, and then centrifuge briefly.
NT 15°C to 30°C Check for precipitates. If present, vortex until all particulates are resuspended.

2. Save the following tagmentation program on the thermal cycler:

- Choose the preheat lid option
- 55°C for 5 minutes
- Hold at 10°C

Procedure

1. Add the following volumes in the order listed to each well of a new Hard-Shell skirted PCR plate. Pipette to mix.
   - TD (10 µl)
   - Normalized gDNA (5 µl)
2. Add 5 µl ATM to each well. Pipette to mix.
3. Centrifuge at 280 × g at 20°C for 1 minute.
4. Place on the preprogrammed thermal cycler and run the tagmentation program. When the sample reaches 10°C, immediately proceed to step 5 because the transposome is still active.
5. Add 5 µl NT to each well. Pipette to mix.
6. Centrifuge at 280 × g at 20°C for 1 minute.
7. Incubate at room temperature for 5 minutes.

The PCR plate contains 25 µl tagmented and neutralized gDNA, all of which is used in the next step.

Amplify Libraries

This step amplifies the tagmented DNA using a limited-cycle PCR program. PCR adds the Index 1 (i7), Index 2 (i5), and full adapter sequences to the tagmented DNA from the previous step. The index adapters and Nextera PCR Master Mix are added directly to the 25 µl of tagmented gDNA from the previous step.

Consumables

- NPM (Nextera PCR Master Mix)
- Index 1 adapters (N7XX)
- Index 2 adapters (S5XX)
- TruSeq™ Index Plate Fixture
- Microseal 'A' film

Preparation

1 Prepare the following consumables:

Item Storage Instructions
Index adapters
(i5 and i7)
-25°C to -15°C Only prepare adapters being used. Thaw at room temperature for 20 minutes.
Invert each tube to mix. Centrifuge briefly.
NPM -25°C to -15°C Thaw on ice for 20 minutes.

2 Save the following program on the thermal cycler:

- Choose the preheat lid option.
- 72°C for 3 minutes
- 95°C for 30 seconds
- 12 cycles of:
- 95°C for 10 seconds
- 55°C for 30 seconds
- 72°C for 30 seconds
- 72°C for 5 minutes
- Hold at 10°C

Procedure

1 [24 libraries] Arrange the index adapters in the TruSeq Index Plate Fixture as follows.

u Arrange Index 1 (i7) adapters in columns 1–6 of the TruSeq Index Plate Fixture.

u Arrange Index 2 (i5) adapter in rows A–D of the TruSeq Index Plate Fixture.

3 Using a multichannel pipette, add 5 µl of each Index 1 (i7) adapter down each column. Replace the cap on each i7 adapter tube with a new orange cap.

4 Using a multichannel pipette, add 5 µl of each Index 2 (i5) adapter across each row. Replace the cap on each i5 adapter tube with a new white cap.

5 Add 15 µl NPM to each well containing index adapters. Pipette to mix.

6 Centrifuge at 280 × g at 20°C for 1 minute.

7 Place on the preprogrammed thermal cycler and run the PCR program. The volume is 50 µl.

SAFE STOPPING POINT

If you are stopping, seal the plate and store at 2°C to 8°C for up to 2 days. Alternatively, leave on the thermal cycler overnight.

Clean Up Libraries

This step uses HighPrep™ PCR beads to purify the library DNA and remove short library fragments.

Consumables

u RSB (Resuspension Buffer)

u HighPrep™ PCR beads

u Freshly prepared 80% ethanol (EtOH)

u 96-well midi plate

u Hard-Shell 96-well PCR plate, skirted

About Reagents

u The HighPrep™ PCR beads are a user-supplied consumable.

u Vortex HighPrep™ PCR beads before each use.

u Vortex HighPrep™ PCR beads frequently to make sure that beads are evenly distributed.

u Always prepare fresh 80% ethanol for wash steps. Ethanol can absorb water from the air, impacting your results.

Preparation

1 Prepare the following consumables:

Item Storage Instructions

RSB -25°C to -15°C Thaw at room temperature.

RSB can be stored at 2°C to 8°C after the initial thaw.

HighPrep™ PCR Beads 2°C to 8°C Let stand on the benchtop for 30 minutes to bring to room temperature.

2 Prepare fresh 80% ethanol from absolute ethanol.

Procedure

1 Centrifuge at 280 × g at 20°C for 1 minute.

2 Transfer 50 µl PCR product from each well of the PCR plate to corresponding wells of a new midi plate.

NOTE: The ratio of PCR product to volume of beads is 3:2. For example, 50 µl PCR product to 30 µl AMPure. If you pull less than 50 µl of PCR product, adjust your ratio of AMPure beads accordingly.

3 Add 30 µl HighPrep™ PCR beads to each well. Smaller amplicons in Nextera XT library preps typically yield smaller insert size ranges. To maximize recovery of smaller fragments from the bead cleanup step, use the following conditions.

Input Size (bp) HighPrep™ PCR Recommendation HighPrep™ PCR Volume (µl)

300–500 1.8x HighPrep™ PCR 90

> 500 0.6x HighPrep™ PCR

(0.5x HighPrep™ PCR for ≥ 2 x 250

cycles)*

30

(25 µl for ≥ 2 x 250 cycles)*

gDNA or other genomic input 0.6x HighPrep™ PCR 30

*Applicable only to the MiSeq™ or HiSeq™ 2500 using HiSeq Rapid v2 reagents.

4 Shake at 1800 rpm for 2 minutes.
5 Incubate at room temperature for 5 minutes.
6 Place on a magnetic stand and wait until the liquid is clear (~2 minutes).
7 Remove and discard all supernatant from each well.
8 Wash two times as follows.
    a Add 200 µl fresh 80% EtOH to each well.
    b Incubate on the magnetic stand for 30 seconds.
    c Remove and discard all supernatant from each well.
9 Using a 20 µl pipette, remove residual 80% EtOH from each well.
10 Air-dry on the magnetic stand for 15 minutes.
11 Remove from the magnetic stand.
12 Add 52.5 µl RSB to each well.
13 Shake at 1800 rpm for 2 minutes.
14 Incubate at room temperature for 2 minutes.
15 Place on a magnetic stand and wait until the liquid is clear (~2 minutes).
16 Transfer 50 µl supernatant from the midi plate to a new Hard-Shell PCR plate.
SAFE STOPPING POINT
If you are stopping, seal the plate and store at -25°C to -15°C for up to seven days.

Check Libraries

1 Run 1 µl of undiluted library on an Agilent Technology 2100 Bioanalyzer using a High Sensitivity DNA chip.

Normalize Libraries

This process normalizes the quantity of each library to ensure more equal library representation in the pooled library.

Consumables

- LNA1 (Library Normalization Additives 1)
- LNB1 (Library Normalization Beads 1)
- LNW1 (Library Normalization Wash 1)
- LNS1 (Library Normalization Storage Buffer 1)
- 0.1 N NaOH (fewer than 7 days old) (3 ml per 96 samples)
- 96-well midi plate
- Hard-Shell 96-well PCR plate, skirted
- 15 ml conical tube
- Microseal 'B' adhesive seals

About Reagents

- Vortex LNA1 vigorously to make sure that all precipitates have dissolved. Inspect in front of a light.
- Vortex LNB1 vigorously, with intermittent inversion (at least 1 minute). Repeat until all beads are resuspended and no beads are present at the bottom of the tube when it is inverted.
- Always use a wide-bore pipette tip for LNA1.
- Mix only the required amounts of LNA1 and LNB1 for the current experiment. Store the remaining LNA1 and LNB1 separately at the recommended temperatures.
- Aspirate and dispense beads slowly due to the viscosity of the solution.

WARNING
This set of reagents contains potentially hazardous chemicals. Personal injury can occur through inhalation, ingestion, skin contact, and eye contact. Wear protective equipment, including eye protection, gloves, and laboratory coat appropriate for risk of exposure. Handle used reagents as chemical waste and discard in accordance with applicable regional, national, and local laws and regulations. For additional environmental, health, and safety information, see the SDS at support.illumina.com/sds.html.

Preparation

1 Prepare the following consumables:

Item Storage Instructions
LNA1 -25°C to -15°C Prepare under a fume hood.
Bring to room temperature. Use a 20°C to 25°C water bath as needed.
LNB1 2°C to 8°C Bring to room temperature. Use a 20°C to 25°C water bath as needed.
LNW1 2°C to 8°C Bring to room temperature. Use a 20°C to 25°C water bath as needed.
LNS1 Room temperature Bring to room temperature.

Procedure

1 Transfer 20 µl supernatant from the Hard-Shell PCR plate to a new midi plate.
2 Add 44 µl LNA1 per sample to a new 15 ml conical tube. Calculate about 5% extra sample to account for sample loss due to pipetting. For example: for 96 samples, add 4.4 ml LNA1 to the tube (100 samples × 44 µl = 4.4 ml).
3 Thoroughly resuspend LNB1. Pipette to mix.
4 Transfer 8 µl LNB1 per sample (including the 5% extra) to the 15 ml conical tube containing LNA1. Invert to mix. For example: for 96 samples, transfer 800 µl LNB1 to the tube of LNA1 (100 samples × 8 µl = 800 µl).
5 Pour the bead mixture into a trough.
6 Add 45 µl combined LNA1 and LNB1 to each well containing libraries.
7 Shake at 1800 rpm for 30 minutes.
8 Place on a magnetic stand and wait until the liquid is clear (~2 minutes).
9 Remove and discard all supernatant from each well.
10 Wash two times as follows.
    a Add 45 µl LNW1 to each well.
    b Shake at 1800 rpm for 5 minutes.
    c Place on a magnetic stand and wait until the liquid is clear (~2 minutes).
    d Remove and discard all supernatant from each well.
11 Add 30 µl 0.1 N NaOH to each well.
12 Shake at 1800 rpm for 5 minutes.
13 During the 5 minute elution, label a new 96-well PCR plate SGP for storage plate.
14 Add 30 µl LNS1 to each well of the SGP plate. Set aside.
15 After the 5 minute elution, make sure that all samples in the midi plate are resuspended. If they are not, resuspend as follows.
    a Pipette to mix or lightly tap the plate on the bench.
    b Shake at 1800 rpm for 5 minutes.
16 Place on a magnetic stand and wait until the liquid is clear (~2 minutes).
17 Transfer the supernatant from the midi plate to the SGP plate.
18 Centrifuge at 1000 × g for 1 minute.
NOTE
After denaturation, the libraries are single-stranded DNA, which resolves poorly on an agarose gel or Bioanalyzer chip. For quality control, use the double-stranded DNA saved from step 16 of the cleanup procedure.
SAFE STOPPING POINT
If you are stopping, seal the plate and store at -25°C to -15°C for up to seven days.

Pool Libraries

Pooling libraries combines equal volumes of normalized libraries in a single tube. After pooling, dilute and heat-denature the library pool before loading libraries for the sequencing run.

Consumables

- Adhesive PCR foil seal
- Eppendorf LoBind microcentrifuge tubes
- PCR eight-tube strip

Preparation

1 To prepare for the sequencing run, begin thawing reagents according to the instructions for your instrument.
2 If the SGP plate was stored frozen at -25°C to -15°C, thaw at room temperature. Pipette to mix.

Procedure

1 Centrifuge at 1000 × g at 20°C for 1 minute.
2 Label a new Eppendorf tube PAL.
3 Transfer 5 µl of each library from the SGP plate to the PAL tube. Invert to mix.
4 Dilute pooled libraries to the loading concentration for your sequencing system. For instructions, see the denature and dilute libraries guide for your system.
5 Store unused pooled libraries in the PAL tube and SGP plate at -25°C to -15°C for up to 7 days.

S7: Overview of costs for "in-house" library prep based either on a commercial kit (Nextera XT) or on two custom-made pipelines. Results are also reported for outsourcing to Macrogen (96 samples) and MGI Sequencing Technology (227 samples).

JEMU Nextera XT DNA Library Preparation Kit		
n. samples		96
Nextera XT DNA Library Preparation Kit (+VAT)	€	4,431
TG Nextera® XT Index Kit v2 Set B (96 Indices) (+VAT)	€	1,479
library QC Genomics Core KUL	€	50
Sequencing - MiSeq - 300 PE (1 FC)	€	1,900
cost / plate	€	7,956
cost / sample (personnel costs @60k €/yr, gross, 1 working week)	€	1,154
cost / sample (consumables)	€	83
cost / sample (consumables + personnel)	€	95

JEMU semi-custom library prep		
n. samples		96
TG Nextera® XT Index Kit v2 Set B (96 Indices) (+VAT)	€	1,479
custom primers with hoverhangs adapters	€	110
Agencourt AMPure XP (full cost)	€	516
library QC Genomics Core KUL	€	50
Sequencing - MiSeq - 300 PE (1 FC)	€	1,900
	€	4,151
	€	1,154
cost / sample (consumables)	€	43
cost / sample (consumables + personnel)	€	55

JEMU fully-custom library prep		
n. samples		96
custom primers with hoverhangs adapters	€	110
Agencourt AMPure XP (full cost)	€	516
Custom indices with Illumina adapters (full cost)	€	462
library QC Genomics Core KUL	€	50
Sequencing - MiSeq - 300 PE (1 FC)	€	1,900
	€	3,134
	€	1,154
cost / sample (consumables)	€	33
cost / sample (consumables + personnel)	€	45

Outsourced Macrogen		
n. samples		96
metabarcoding library prep	€	2,112
Sequencing - MiSeq - 300 PE (1 FC)	€	1,900
	€	4,108
cost / sample (consumables + personnel)	€	43

Outsourced MGI Sequencing Technology (best offer RMCA)		
n. samples		227
metabarcoding library prep + Sequencing - MiSeq - 300 PE (1 FC)	€	5,000
cost / sample (consumables + personnel)	€	22