**In-silico validation of a broad-spectrum metabarcoding assay for insect pest diagnostics.**

Alexander M. Piper1,2, Paul Cunningham1, Noel O.I. Cogan1,2, Mark J. Blacket1

1 Agriculture Victoria Research, AgriBio, Bundoora, Victoria, Australia

2 School of Applied Systems Biology, La Trobe University, Bundoora, Victoria, Australia

**Corresponding author:**

**Alexander M. Piper**

Email: [alexander.piper@agriculture.vic.gov.au](mailto:alexander.piper@agriculture.vic.gov.au)

**Running title:** Metabarcoding for insect pest detection

**Target journal:**

* Molecular Ecology Resources
* Methods in Ecology and Evolution
* Scientific Reports
* PeerJ

Figures:

1. Upset plot of species overlap between pestlist datasets – Could also include PCA of differences between datasets
2. Summary of sequences / Unique species kept through dataset filtering. Need to write out at each stage? (Supplementary) – Have some icons for the different stages (ie phmm icon, etc)
3. Plot of primer positions/entropy within COI, as well as summary of identification success for different primers – for pest taxa only
4. - Figure summarising failed classifications. Ie exact match different taxonomy, 99% match different taxonomy. Whether they were in a big group or just a couple of sequences (ie could majority rules taxonomy removal solve this) etc
5. Plot of predicted mismatch between different primers and entire insect pes tgroups with dendogram on left to show that mismatch is concentrated on some clades
6. In-silico predicted off target effects of primers – trees

# **Introduction**

For diagnostics of invasive pests and pathogens, early diagnosis and a rapid response are crucial to reduce the risk of establishment and spread within an area. Historically inspections and other surveillance activities have been based on targeted lists of regulated organisms, an approach that is fundamental to quarantine legislation and risk prioritisation around the world. However, this list-based legislation can often lag behind the speed with which new pests can emerge and spread across borders. Furthermore, with growing investment in environmental biosecurity rather than just agriculture the detection, keeping a broad range of pests out becomes more essential due to the time lag that can occur between introduction of a new species and perceptible damage to the environment. However, if all these are to be monitored for or to allow a more holistic approach to surveillance, current targeted diagnostics provide to much a workload.

Non-targeted or ‘meta’ HTS diagnostics have recently gained traction in many fields of molecular diagnostics to overcome this. By targeting a mixed population of genomic DNA or RNA these approaches allow the simultaneous identification of multiple species within large mixed communities. If the samee assay can be used across diverse taxonomic or physiological groups it has the potential to broaden the diversity of organisms within the scope of a diagnostic laboratory, as well as decrease the costs of implementation. Metagenomic and Metatranscriptomic approaches have been used to identify large reportable ranges of viral and bacterial and fungal pathogens in clinical situations (Chiu & Miller, 2019). However, for targeting many eukaryotic organisms, the size of the genomes make these processes expensive (Piper et al., 2019) and instead enrichment of a taxonomically informative marker is first used. This ‘Metabarcoding’ approach has been used for diagnostics of forensic, endangered, and pest species.

The use of non-targeted diagnostics can r and can reveal the presence of previously unrecorded introduced species that have been missed by previous targeted surveys (Batovska et al 2020). (Batovska, Mee, Lynch, Sawbridge, & Rodoni, 2019)

Despite the obvious advantages that non-targeted HTS assays offer to plant pest and pathogen diagnostics, ensuring the accuracy of these detections is paramount, in light of the complexity of NGS pipelines have introduced more ways for false positives to be introduced (Piper et al., 2019). Conventionally, plant pest diagnostics laboratories have been accredited under fixed scope standards where the range of tests covered by the laboratory’s accreditation are defined clearly and unambiguously, usually in the form of specific organism/matrix combinations (e.g. PCR test for detection of *Thrips palmi* in matrix). Before a new test can be introduced or modified (i.e. for a new organism or matrix), the assay must undergo a validation process in order to provide objective evidence to all end users that the assay is fit for purpose and performance parameters such as analytical sensitivity, analytical specificity, reproducibility and repeatability for every individual target designated in this predefined scope. However, this validation framework does not readily fit with novel assays such as non-targeted diagnostics, due to the sheer number of potential pathogens/pests, hosts and matrices that need to be tested makes this framework insufficient.

Consequently, the concept of flexible scope has been developed. A flexible scope of accreditation allows a laboratory to undertake certain tests, and to report the results as accredited, even though these tests or taxa are not explicitly stated in the laboratory’s scope. This involves a distinct shift in attitude and requirements from being an accredited diagnostic for a specific organism/matrix combination, for example “thrips palmi in leaves” instead to an acredited diagnostic for a broader range of taxa such as “insects”, for which every taxa would not have seen “complete” validation according to PM 7/98. Flexible scope accreditation has obvious benefits for implementation of non-targeted HTS diagnostics, particularly the flexibility to apply an already accredited test to a new set of pest/host matrices after validation. This shifts the emphasis from strict a-priori validation towards ongoing quality assurance and demonstrated expertise of the laboratory and its staff. Importantly, this does not mean freedom from validation and traditional performance criteria such as specificity, sensitivity, selectivity as well as repeatability and reproducibility will still need to be established. However, it allows for more flexibility in obtaining these parameters, particularly in light of a lack of reference materials for many organisms targeted by HTS assays.

The incorporation of In-silico methods into assay validation and QC have become increasingly important within flexible-scope accreditation, especially for non-targeted assays. This allows a dramatic scale up.

A decision framework should be developed for evaluating incidental detections that sets out steps for further characterization and risk assessment for the detected organisms in order to establish whether eradication or other management actions are appropriate or achievable. An essential first step of this decision framework is knowing whether the assay has the appropriate specificity and sensitivity to accurately regard the detection as suspect, or is it merely the detection of a closely related or problem group (see Batovska 2020). Having validation data on all the possible detected pests a-priori will dramatically speed up diagnostic turnaround and response.

For metabarcoding assays, appropriate design and selection of PCR primers play the most important role in determining the sensitivity and specificity of the assay, and therefore whether a detection should be trusted. While the 658-bp ‘folmer’ region of COI (Folmer, Black, Hoeh, Lutz, & Vrijenhoek, 1994) has been widely employed for single species diagnostics of insect pests via conventional DNA barcoding (Ashfaq, Hebert, & Naaum, 2016), modern HTS platforms impose strict limitations on molecule length that can be sequenced. Therefore, alternative (refs) or smaller stretches of the conventional barcode region or ‘mini-barcodes’ must instead be used (Brandon-Mong et al., 2015). This has resulted in the proliferation of new primer sets that have not yet been validated for effectiveness, or for their ability to diagnose pest insect species. Ensuring the accuracy of these detections is paramount as both false positive or negative detections of pest species can lead to severe environmental and economic consequences. Despite several studies now applying metabarcoding to diagnostics of insect pests (refs), the applicability of these smaller barcode regions to a broad diversity range of insect pests has not yet been systematically evaluated. If universal mini primers are to be used are to be used in a situation where there is potential for the incidental detection of unexpected and unvalidated taxa, at the minimum it should be known in advanced if that detection can be relied upon. In addition to concerns of taxonomic resolution. Because COI is a protein-coding gene, the third position of codons can be variable, leaving no strictly conserved nucleotide sites for design of universal PCR primers. PCR bias can cause deviation from expected proportions. While most biosecurity studies are aiming just for detection, quantification can provide an advantage. Furthermore, when bias is too high it can swamp out low abundance taxa.

Here we provide a framework for in-silico evaluation of non-targeted metabarcoding assays for regulated organisms, using a diverse list of globally listed invasive and pest insects as our case study. Firstly we locate the optimal diagnostic window within the target loci for placement of mini primers. Secondly we evaluate the ability of mini primers to differentiate a list of pest taxa from others in their group (sensitivity and specificity). We then look at predicted mismatch and bias for a list of pests. Finally we provide recommendations on primers to use for diagnostics of pest insect taxa, and provide a comprehensive list of pest taxa for which a metabarcoding assay using these priemrs are likely to work and not. We believe that the most rapid way to get this technology into the hands of diagnostics labs is to propose a protocol that is as universal as possible, then the user validates the assay on their target pest, we leave that to the user.

**Methods**

*Assembly of global pest list*

To assemble a global list of insect pests, all records from a range of global invasive species databases (Supplementary S1) were retrieved and filtered to retain only those taxa with genus species binomials and remove duplicadte entries. Higher taxonomic levels for all species on the list were retrieved from the NCBI and GBIF reference taxonomies using the R package taxize (S. A. Chamberlain & Szöcs, 2013), and any taxa from outside the class Insecta were removed.

*Retrieval and curation of public reference data*

All COI records were retrieved from BOLD and GenBank for the unique pest families present on this list using the bold (S. Chamberlain, 2017) and Rentrez (Winter, David, 2019) packages respectively. In addition, all mitochondrial genomes were downloaded. Following download all sequences with duplicated accession numbers were removed and sequences went through a series of filtering stages. Firstly, in order to remove non-homologous sequences and extract the COI locus from mitogenomes all sequences were aligned to a profile hidden markov model (PHMM) (Eddy, 1998) of the COI locus using the viterbi algorithm (HOLMES & DURBIN, 2009), and only sequences aligning with a score of over 500 were retained. This PHMM was generated using a manually curated version of the Midori-unique (Machida, Leray, Ho, & Knowlton, 2017) dataset and the aphid R package (S. Wilkinson, 2019). Following alignment, all bases outside the 711bp amplicon of the conventional ‘folmer’ barcode region were shaved off. While this PHMM alignment successfully removed non-homologous sequence data, it did not sufficiently control for closely related pseudogenes, and therefore the sequences were further checked for stop codons. Next, records containing terms indicating insufficient identification (i.e. sp., nr., aff. – see Appendix S1 for full list of terms) were removed. In order to identify taxonomically mislabelled sequences, all sequences were clustered at 99% and species level taxonomies compared within clusters, with any sequences that differed from at least 60% of other sequences in the cluster removed from the dataset. Next, taxonomic synonyms were resolved using the Global Biodiversity Information Facility database and the taxize R package. Finally, to reduce reference bias due to incomplete sampling of the taxonomic tree, all large groups of taxa were pruned to 5 representatives of each species, with sequences discarded from smallest to largest until only 5 sequences remained.

*Identification of optimal diagnostic windows*

To identify optimal diagnostic windows for placement of mini-barcodes within the 658bp folmer region, a sliding window analysis was conducted on the curated reference sequences for each family using the SPIDER r package (Brown et al., 2012) and the number of diagnostic nucleotides and proportion of species that were monophyletic were summarised for each 220bp window. A density plot of the 10 highest ranked window positions within each family was then used to identify globally optimal regions for diagnostics of all pest insect families on the list. Published primer sets overlapping these regions were retrieved, and a number of additional sets were designed using Primer3 implemented in Geneious Prime v2019.2 (Geneious, New Zealand).

*In-silico validation of primer sets*

To evaluate the resolution of published primer and novel primer sets across pest taxa, sequence alignments for each pest family were trimmed to the amplified region of each primer set using the insect R package (S. P. Wilkinson, Davy, Bunce, & Stat, 2018), and pairwise genetic distances were generated using the raw model in SPIDER. In order to summarise the taxonomic groups likely to be misidentified or ambiguously identified when using each mini-barcode primer set, summary statistics of identification success were generated using the nearNeighbour, BestCloseMatch, and reciprocal monophyly identification methods in SPIDER, and those failing identification were further explored using neighbour joining trees. In addition to the resolution of primer sets, the level of mismatch between the primers and template must be considered when selecting a primer set as this is thought to be the primary driver of bias in quantification and detection success (Piñol, Mir, Gomez-Polo, & Agustí, 2015; Piñol, Senar, & Symondson, 2019). To evaluate the primers for mismatch, a mismatch score for each primer and insect family was derived using the in-silico PCR function in PrimerMiner (Elbrecht & Leese, 2017a). While highly degenerate primers reduce primer-template mismatch and therefore perform well for the above measure, high degeneracy can also cause undesired off target amplification (Mioduchowska, Jan, Gołdyn, Kur, & Sell, 2018). To evaluate possible off-target effects of the primers, the primer tree R package (Cannon et al., 2016) was used to pick 1000 sequences at random from the trimmed datasets for each primer set and conduct a primer blast against the ncbi nt database, and neighbour joining trees were conducted to summarise the taxonomic scope of each primer.

# **Results**

Database assembly

In total, n unique taxa were retrieved from databases (supplementary fig 1) and following filtering n unique species across n families and n orders were retained. All sequences for these families were retrieved from bold and genbank, for a total of n and n sequences respectively. Following all stages of filtering, a total of n sequences were retained.

*Identifying optimal diagnostic windows*

220bp sliding window analysis of diagnostic nucleotides proportion of monophyletic trees identified 2 regions located 255-475bp and 340-560bp into the folmer DNA barcode region for which the highest density of top 10 ranked windows for all families on the pest list occurred. A range of published primers overlapping these regions were retrieved from the literature, and supplemented with novel primers (Table 1).

*Evaluation of primer sets*

All primer pairs showed comparable resolution across the, with the exception of the BF1-AgpestR2 primer (Figure 2). For all those that failed, generally the result of synonyms and indicate the need for curation of synonyms in reference sequence databases. When taxonomic synonyms were corrected for, the large majority of remaining failed to differentiate taxa came from the family Tephritidae, for which the full DNA barcode region also struggles with (Jiang, Jin, Liang, Zhang, & Li, 2014; Krosch, Schutze, Strutt, Clarke, & Cameron, 2017).

**Discussion**

* Mini barcodes can successfully differentiate the majority of insect pests
* Those that could not be differentiated by mini-barcodes were in classic problem groups and problems with taxonomy
* Limitations, we haven’t tested new long read sequenceing, and sequencing of the full mitochondria might be useful
* While this would necessitate running the analysis again for new primers, we have now defined a process for evaluation
* Similarly, we could take this process and move it to new taxa and gene regions. Biosecurity diagnostics is becoming important for Fungi, bacteria etc
* Of course, this process should not fully replace laboratory validation on targets, however it allows you to be more selective
* Therefore we recommend that this process be followed up by validation to be conducted on a phylogenetically and physiologically diverse group of organisms. As well as high priority targets.
* Luckily, many of these published primers have been previously tested on insects and other arthropods, which provides further confidence.
* However, matrix effects must also be validated etc.

**Acknowledgments**

**Availability of data and materials:**

All scripts required to reproduce the analyses are available at <https://github.com/alexpiper/Drosophila_metabarcoding>. The scripts to curate and format reference databases have been contained within an R package available at <https://github.com/alexpiper/taxreturn>.

# References

Ashfaq, M., Hebert, P. D. N. P. D. N., & Naaum, A. (2016). DNA barcodes for bio-surveillance: Regulated and economically important arthropod plant pests. *Genome*, *59*(11), 933–945. doi:10.1139/gen-2016-0024

Batovska, J., Mee, P. T., Lynch, S. E., Sawbridge, T. I., & Rodoni, B. C. (2019). Sensitivity and specificity of metatranscriptomics as an arbovirus surveillance tool. *Scientific Reports*, *9*(1), 1–13. doi:10.1038/s41598-019-55741-3

Brandon-Mong, G.-J. J., Gan, H.-M. M., Sing, K.-W. W., Lee, P.-S. S., Lim, P.-E. E., & Wilson, J.-J. J. (2015). DNA metabarcoding of insects and allies: An evaluation of primers and pipelines. *Bulletin of Entomological Research*, *105*(6), 717–727. doi:10.1017/S0007485315000681

Brown, S. D. J., Collins, R. A., Boyer, S., Lefort, M.-C., Malumbres-Olarte, J., Vink, C. J., & Cruickshank, R. H. (2012). SPIDER: an R package for the analysis of species identity and evolution, with particular reference to DNA barcoding. *Molecular Ecology Resources*, *12*, 562–565.

Cannon, M. V., Hester, J., Shalkhauser, A., Chan, E. R., Logue, K., Small, S. T., & Serre, D. (2016). In silico assessment of primers for eDNA studies using PrimerTree and application to characterize the biodiversity surrounding the Cuyahoga River. *Scientific Reports*, *6*(December 2015), 1–11. doi:10.1038/srep22908

Chamberlain, S. (2017). bold: Interface to Bold Systems API. Retrieved from https://cran.r-project.org/package=bold

Chamberlain, S. A., & Szöcs, E. (2013). taxize: taxonomic search and retrieval in R. *F1000Research*, *2*, 191. doi:10.12688/f1000research.2-191.v2

Chiu, C. Y., & Miller, S. A. (2019). Clinical metagenomics. *Nature Reviews Genetics*, *20*, 341–355. doi:10.1038/s41576-019-0113-7

Eddy, S. R. (1998). Profile hidden Markov models. *Bioinformatics*. doi:10.1093/bioinformatics/14.9.755

Elbrecht, V., & Leese, F. (2017a). PrimerMiner: an r package for development and in silico validation of DNA metabarcoding primers. *Methods in Ecology and Evolution*, *8*(5), 622–626. doi:10.1111/2041-210X.12687

Elbrecht, V., & Leese, F. (2017b). Validation and Development of COI Metabarcoding Primers for Freshwater Macroinvertebrate Bioassessment. *Frontiers in Environmental Science*, *5*, 11. doi:10.3389/fenvs.2017.00011

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. doi:10.1371/journal.pone.0013102

HOLMES, I., & DURBIN, R. (2009). Dynamic Programming Alignment Accuracy. *Journal of Computational Biology*. doi:10.1089/cmb.1998.5.493

Jiang, F., Jin, Q., Liang, L., Zhang, A. B., & Li, Z. H. (2014). Existence of species complex largely reduced barcoding success for invasive species of Tephritidae: A case study in Bactrocera spp. *Molecular Ecology Resources*, *14*(6), 1114–1128. doi:10.1111/1755-0998.12259

Jordaens, K., Sonet, G., & Richet, R. (2013). Identification of forensically important Sarcophaga species ( Diptera : Sarcophagidae ) using the mitochondrial COI gene, 491–504. doi:10.1007/s00414-012-0767-6

Krosch, M. N., Schutze, M. K., Strutt, F., Clarke, A. R., & Cameron, S. L. (2017). A transcriptome-based analytical workflow for identifying loci for species diagnosis: A case study with Bactrocera fruit flies (Diptera: Tephritidae). *Austral Entomology*, *58*, 395– 408. doi:10.1111/aen.12321

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, 1–14.

Machida, R. J., Leray, M., Ho, S. L., & Knowlton, N. (2017). Data Descriptor: Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. *Scientific Data*, *4*, 170027. doi:10.1038/sdata.2017.27

Marquina, D. (2019). New mitochondrial primers for metabarcoding of insects , designed and evaluated using in silico methods, (August 2018), 90–104. doi:10.1111/1755-0998.12942

Meusnier, I., Singer, G. A. C., Landry, J. F., Hickey, D. A., Hebert, P. D. N., & Hajibabaei, M. (2008). A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*, *9*, 4–7. doi:10.1186/1471-2164-9-214

Mioduchowska, M., Jan, M., Gołdyn, B., Kur, J., & Sell, J. (2018). Instances of erroneous DNA barcoding of metazoan invertebrates: Are universal cox1 gene primers too “universal”? *Plos One*, *13*(6), e0199609. doi:10.1371/journal.pone.0199609

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*(4), 819–830. doi: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 determines when DNA metabarcoding can be quantitative. *Molecular Ecology*, *28*(May), 407– 419. doi:10.1111/mec.14776

Piper, A. M., Batovska, J., Cogan, N. O. I., Weiss, J., Cunningham, J. P., Rodoni, B. C., & Blacket, M. J. (2019). Prospects and challenges of implementing DNA metabarcoding for high-throughput insect surveillance. *GigaScience*, *8*(8), 1–22. doi:10.1093/gigascience/giz092

Rennstam Rubbmark, O., Sint, D., Horngacher, N., & Traugott, M. (2018). A broadly-applicable COI primer pair and an efficient single tube amplicon library preparation protocol for metabarcoding. *Ecology and Evolution*, *8*, 12335– 12350. doi:10.1016/S1872-2067(12)60761-X

Vamos, E. E., Elbrecht, V., & Leese, F. (2017). Short COI markers for freshwater macroinvertebrate metabarcoding, 1–20. doi:10.3897/mbmg.1.14625

Wilkinson, S. (2019). aphid: an R package for analysis with profile hidden Markov models. *Bioinformatics*. doi:10.1093/bioinformatics/btz159

Wilkinson, S. P., Davy, S. K., Bunce, M., & Stat, M. (2018). Taxonomic identification of environmental DNA with informatic sequence classification trees. *PeerJ Preprints*, *6*, e26812v1. doi:10.7287/peerj.preprints.26812v1

Winter, David, J. (2019). rentrez: An R package for the NCBI eUtils API. *The R Journal*, *9*(2), 520. doi:10.32614/rj-2017-058

Zeale, M. R. K., Butlin, R. K., Barker, G. L. A., Lees, D. C., & Jones, G. (2011). Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*. doi:10.1111/j.1755-0998.2010.02920.x

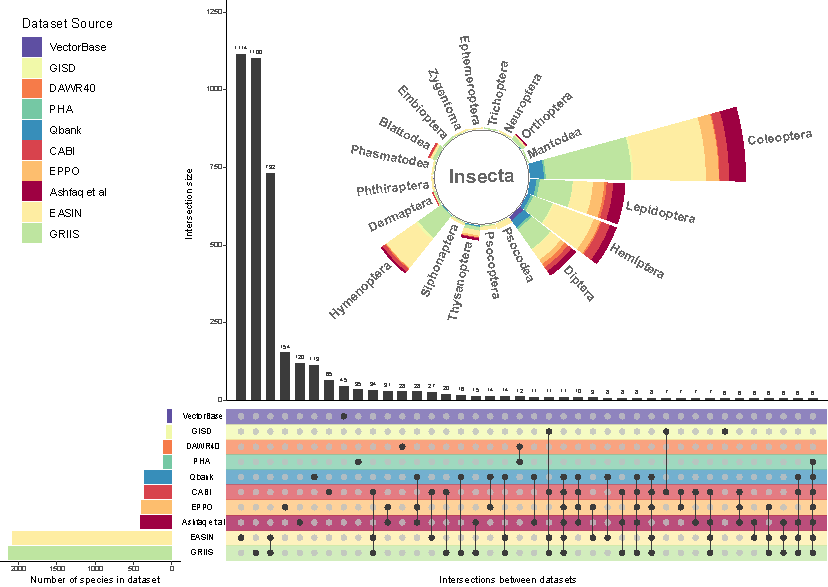
# Tables

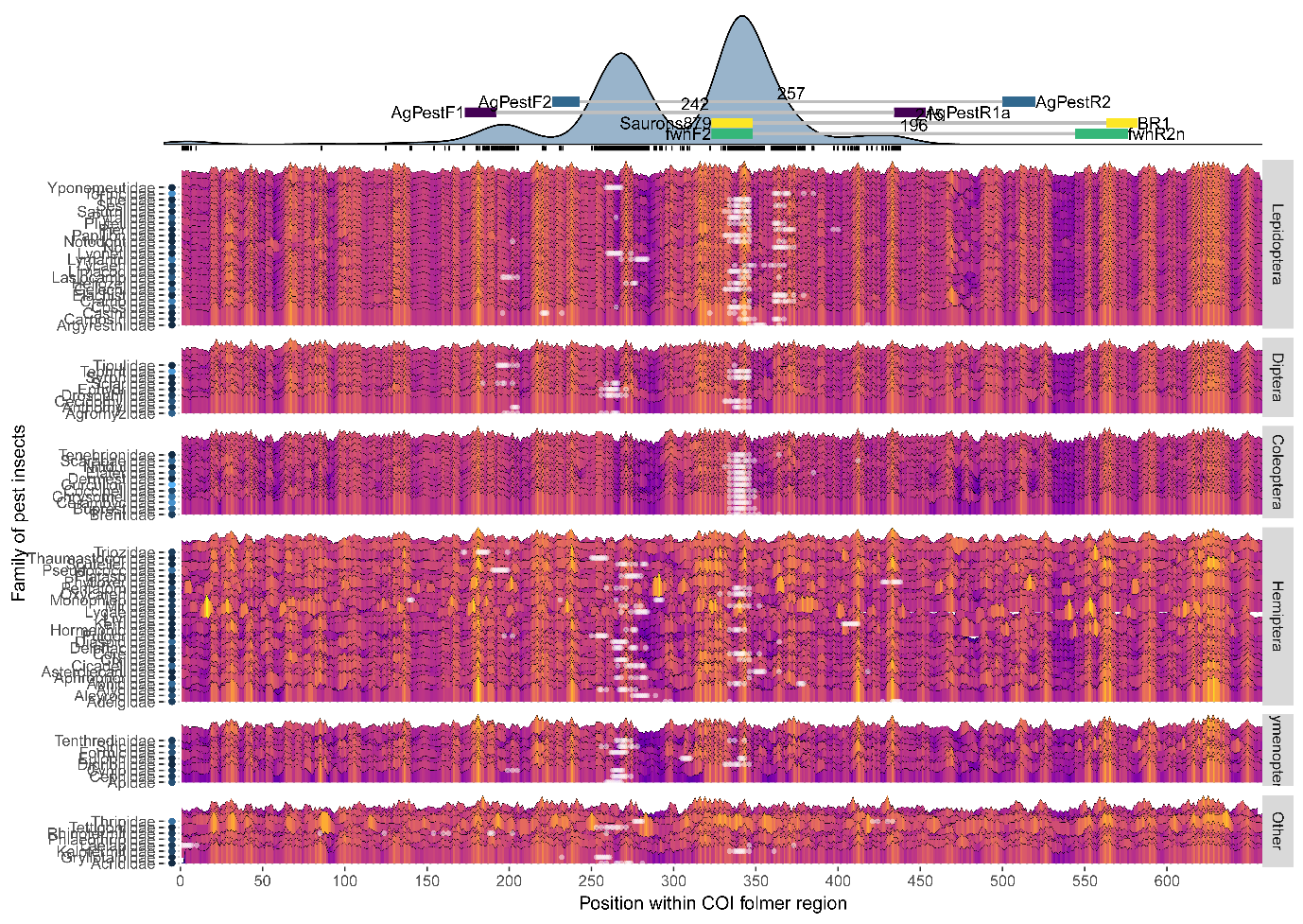
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Name | F.seq | F.Start | R.seq | R.Start | amplicon | study |
| BF1-BR1 | ACWGGWTGRACWGTNTAYCC | 326 | ARYATDGTRATDGCHCCDGC | 563 | 218 | (Elbrecht & Leese, 2017b) |
| fwhF2-fwhR2n | GGDACWGGWTGAACWGTWTAYCCHCC | 323 | GTRATWGCHCCDGCTARWACWGG | 544 | 196 | (Vamos, Elbrecht, & Leese, 2017) |
| fwhF1-fwhR1 | YTCHACWAAYCAYAARGAYATYGG | -23 | ARTCARTTWCCRAAHCCHCC | 182 | 182 | (Vamos et al., 2017) |
| SternoCOIF1 - SternoCOIR1 | ATTGGWGGWTTYGGAAAYTG | 179 | ATRAARTTRATWGCTCCTA | 432 | 234 | Batovska et al (in prep) |
| mtCOIintF-BR1 | GGWACWGGWTGAACWGTWTAYCCYCC | 323 | ARYATDGTRATDGCHCCDGC | 563 | 215 | (Leray et al., 2013) |
| SauronS878-BR1 | GGDRCWGGWTGAACWGTWTAYCCNCC | 323 | ARYATDGTRATDGCHCCDGC | 563 | 215 | (Rennstam Rubbmark, Sint, Horngacher, & Traugott, 2018) |
| Uni-MinibarF1-Uni-MinibarR1 | TCCACTAATCACAARGATATTGGTAC | -22 | GAAAATCATAATGAAGGCATGAGC | 134 | 131 | (Meusnier et al., 2008) |
| UniMinibarF1−d-UniMinibarF1−d | TCYACTAATCATAAAGATATTGGYAC | -22 | AAAATTATAATAAARGCRTGRGC | 134 | 131 | (Jordaens, Sonet, & Richet, 2013) |
| ZBJ−ArtF1c-ZBJ−ArtR2c | AGATATTGGAACWTTATATTTTATTTTTGG | -8 | WACTAATCAATTWCCAAATCCTCC | 182 | 161 | (Zeale, Butlin, Barker, Lees, & Jones, 2011) |
| AgPestF1-AgPestR1a | ATYATWATTGGDGGDTTYGG | 173 | GTRATRAARTTDAYWGMHCC | 434 | 242 | This Study |
| AgPestF1-AgPestR1b | ATYATWATTGGDGGDTTYGG | 173 | ARAATWGADGADAYWCCWGC | 413 | 221 | This Study |
| AgPestF2-AgPestR1a | HGAYATRGCHTTYCCHCG | 226 | GTRATRAARTTDAYWGMHCC | 434 | 191 | This Study |
| AgPestF2-AgPestR1b | HGAYATRGCHTTYCCHCG | 226 | ARAATWGADGADAYWCCWGC | 413 | 170 | This Study |
| AgPestF2-AgPestR2 | HGAYATRGCHTTYCCHCG | 226 | RACWGMTCAVAYAAATARDGG | 500 | 257 | This Study |
| ArF5-AgPestR1a | GCICCIGAYATRKCITTYCCICG | 221 | GTRATRAARTTDAYWGMHCC | 434 | 191 | This Study |
| ArF5-AgPestR1b | GCICCIGAYATRKCITTYCCICG | 221 | ARAATWGADGADAYWCCWGC | 413 | 170 | This Study |
| ArF5-AgPestR2 | GCICCIGAYATRKCITTYCCICG | 221 | RACWGMTCAVAYAAATARDGG | 500 | 257 | This Study |
| BF2-AgPestR1a | GCHCCHGAYATRGCHTTYCC | 221 | GTRATRAARTTDAYWGMHCC | 434 | 194 | This Study |
| BF2-AgPestR1b | GCHCCHGAYATRGCHTTYCC | 221 | ARAATWGADGADAYWCCWGC | 413 | 173 | This Study |
| BF2-AgPestR2 | GCHCCHGAYATRGCHTTYCC | 221 | RACWGMTCAVAYAAATARDGG | 500 | 260 | This Study |
| SternoCOIF1-AgPestR1a | ATTGGWGGWTTYGGAAAYTG | 179 | GTRATRAARTTDAYWGMHCC | 434 | 236 | This Study |
| SternoCOIF1-AgPestR1b | ATTGGWGGWTTYGGAAAYTG | 179 | ARAATWGADGADAYWCCWGC | 413 | 215 | This Study |
| AgPestF1-SternoCOIR1 | ATYATWATTGGDGGDTTYGG | 173 | ATRAARTTRATWGCTCCTA | 432 | 240 | This Study |
| AgPestF2-SternoCOIR2 | HGAYATRGCHTTYCCHCG | 226 | ATRAARTTRATWGCTCCTA | 432 | 189 | This Study |
| BF1-AgPestR2 | ACWGGWTGRACWGTNTAYCC | 326 | RACWGMTCAVAYAAATARDGG | 500 | 155 | This Study |
| HexCOIF4- AgPestR2 | HCCHGAYATRGCHTTYCC | 223 | RACWGMTCAVAYAAATARDGG | 500 | 260 | (Marquina, 2019) |
| HexCOIF4- AgPestR1a | HCCHGAYATRGCHTTYCC | 223 | GTRATRAARTTDAYWGMHCC | 434 | 194 | (Marquina, 2019) |
| BF1-HexCOIR4 | ACWGGWTGRACWGTNTAYCC | 326 | TATDGTRATDGCHCCNGC | 563 | 218 | (Marquina, 2019) |
| fwhF2-HexCOIR4 | GGDACWGGWTGAACWGTWTAYCCHCC | 323 | TATDGTRATDGCHCCNGC | 563 | 215 | (Marquina, 2019) |
| SauronS878-HexCOIR4 | GGDRCWGGWTGAACWGTWTAYCCNCC | 323 | TATDGTRATDGCHCCNGC | 563 | 215 | Marquina, 2019) |

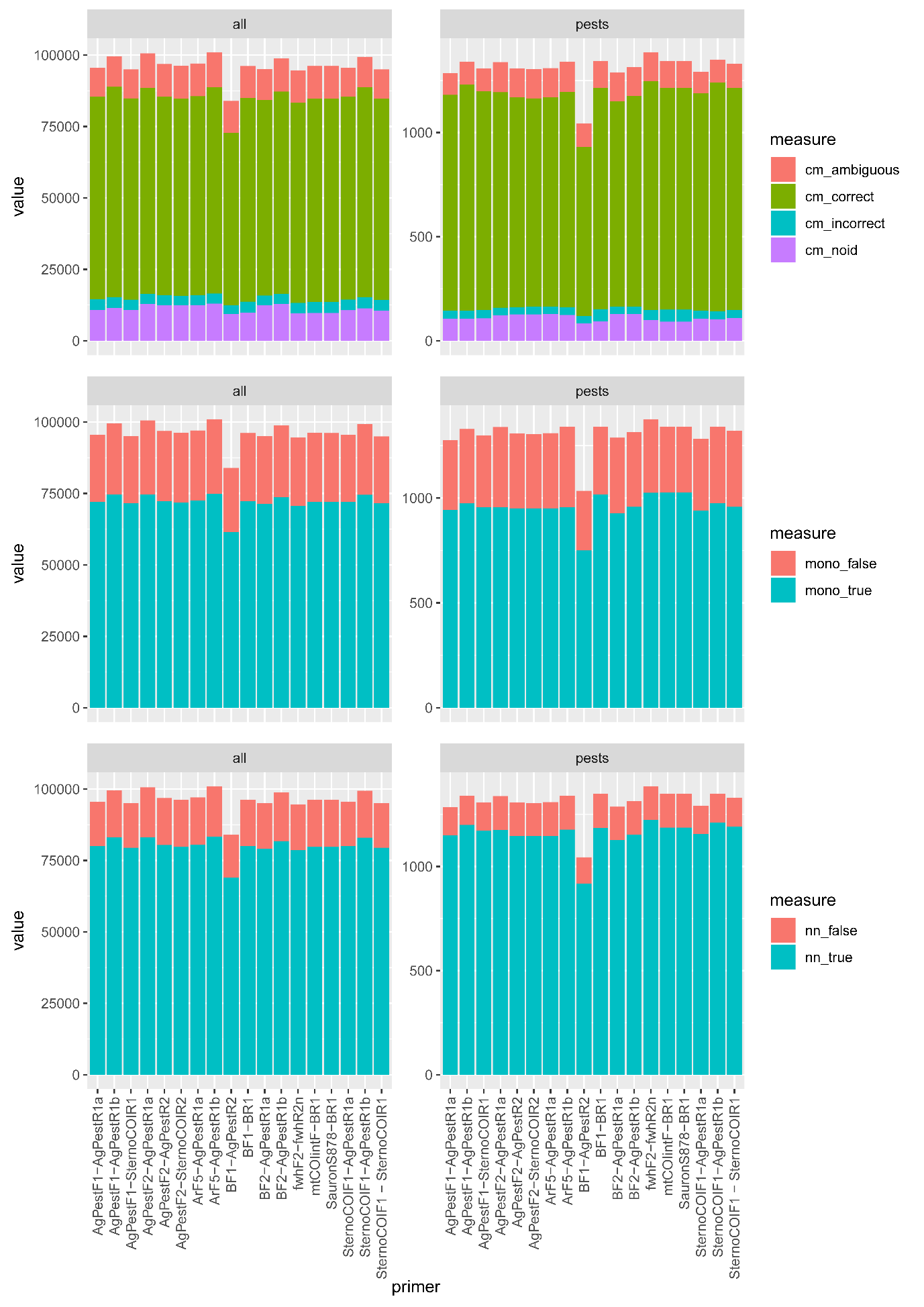
Table 1: Published and novel primers evaluated in this study

Figures:

Figure 1 – Pestlist







*Figure 2 - Evaluation of novel and existing primers for resolution across pest insect families*

Figure3 - Evaluation of individual Primers for mismatch

