

# SCIENTIFIC REPORTS



OPEN

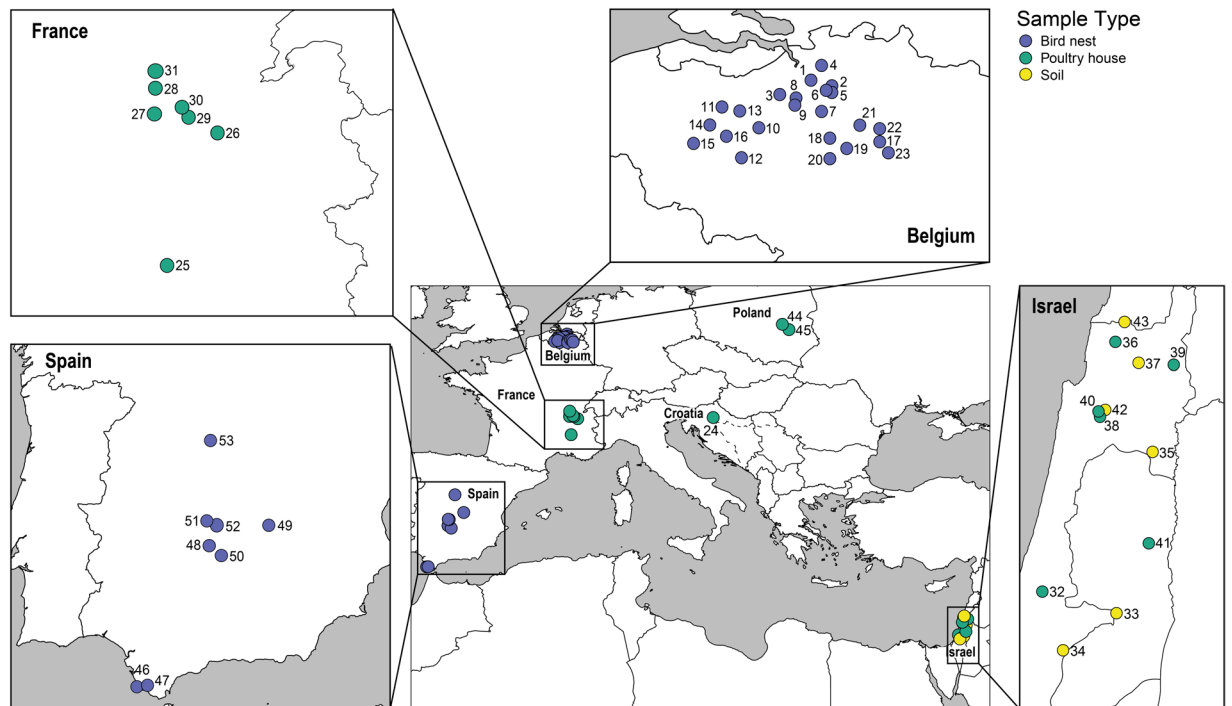
## Linking morphological and molecular taxonomy for the identification of poultry house, soil, and nest dwelling mites in the Western Palearctic

Monica R. Young<sup>1</sup>, María L. Moraza<sup>2</sup>, Eddie Ueckermann<sup>3</sup>, Dieter Heylen<sup>4,5</sup>, Lisa F. Baardsen<sup>6</sup>, Jose Francisco Lima-Barbero<sup>7,8</sup>, Shira Gal<sup>9</sup>, Efrat Gavish-Regev<sup>10</sup>, Yuval Gottlieb<sup>11</sup>, Lise Roy<sup>12</sup>, Eitan Recht<sup>13</sup>, Marine El Adouzi<sup>12</sup> & Eric Pavlevsky<sup>9</sup>

Because of its ability to expedite specimen identification and species delineation, the barcode index number (BIN) system presents a powerful tool to characterize hyperdiverse invertebrate groups such as the Acari (mites). However, the congruence between BINs and morphologically recognized species has seen limited testing in this taxon. We therefore apply this method towards the development of a barcode reference library for soil, poultry litter, and nest dwelling mites in the Western Palearctic. Through analysis of over 600 specimens, we provide DNA barcode coverage for 35 described species and 70 molecular taxonomic units (BINs). Nearly 80% of the species were accurately identified through this method, but just 60% perfectly matched (1:1) with BINs. High intraspecific divergences were found in 34% of the species examined and likely reflect cryptic diversity, highlighting the need for revision in these taxa. These findings provide a valuable resource for integrative pest management, but also highlight the importance of integrating morphological and molecular methods for fine-scale taxonomic resolution in poorly-known invertebrate lineages.

DNA barcoding<sup>1</sup> alleviates many of the challenges associated with morphological specimen identification by comparing short, standardized fragments of DNA – typically 648 bp of the cytochrome *c* oxidase I (COI) gene for animals – to a well-curated reference library. The success of this method relies on the presence of a clearly defined ‘barcode gap’, where intraspecific divergences are much more constrained than interspecific divergences. Its presence not only enables rapid specimen identification, but also facilitates species delineation through molecularly defined taxonomic units, a process automated through the barcode index number (BIN) system<sup>2</sup>. BINs correspond well with morphologically recognized species in lineages with well-curated taxonomy<sup>2–4</sup> and can improve taxonomic resolution by elucidating hidden diversity<sup>5,6</sup>. Consequently, BINs are a powerful tool for characterizing diversity in poorly-known, hyperdiverse, invertebrates<sup>7–9</sup>, but have seen limited validation in these taxa.

<sup>1</sup>Centre for Biodiversity Genomics, University of Guelph, Guelph, Canada. <sup>2</sup>Departamento de Biología Ambiental, Facultad de Ciencias, Universidad de Navarra, Pamplona, Spain. <sup>3</sup>Unit for Environmental Sciences and Management, Potchefstroom Campus, North-West University, Private Bag X6001, Potchefstroom, 2520, South Africa. <sup>4</sup>Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, Diepenbeek, Belgium. <sup>5</sup>Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, United States of America. <sup>6</sup>Evolutionary Ecology Group, University of Antwerp, Wilrijk, Belgium. <sup>7</sup>Health and Biotechnology Group (SABIO), Institute for Game and Wildlife Research (CSIC-UCLM), Ciudad Real, Spain. <sup>8</sup>Sabiotec, Ciudad Real, Spain. <sup>9</sup>Newe-Ya’ar Research Center, Agricultural Research Organization, Ramat Yishay, Israel. <sup>10</sup>The National Natural History Collections, The Hebrew University of Jerusalem, Jerusalem, Israel. <sup>11</sup>Koret School of Veterinary Medicine, The Hebrew University of Jerusalem, Rehovot, Israel. <sup>12</sup>UMR 5175 CEFE, CNRS – Université de Montpellier – Université Paul-Valéry Montpellier – EPHE, Route de Mende, 34199, Montpellier, Cedex 5, France. <sup>13</sup>Plant Protection and Inspection Services, Bet Dagan, Israel. Correspondence and requests for materials should be addressed to M.R.Y. (email: [myoung02@uoguelph.ca](mailto:myoung02@uoguelph.ca))



**Figure 1.** Map of the 53 sampling sites in seven countries across the Western Palearctic. The location markers correspond with site numbers specified in Table 1; sample type (bird nest, poultry house, soil) is indicated by the colour of the marker.

The mites (Acari) may exceed one million species, but remain poorly known because of their small size and cryptic morphology<sup>10</sup>. While BIN-based surveys have expedited surveys of this hyperdiverse group<sup>7,11,12</sup>, the rapidly growing collection of mite barcodes generally lack lower-level taxonomy. For example, just 18% of the >12,400 mite BINs (from nearly 120,000 DNA barcode sequences) on the Barcode of Life Data System (BOLD, v4.boldsystems.org) are linked with a species name (accessed August 2018). Nonetheless, successful species delineation through DNA barcodes has been documented in several mite lineages, including the Ixodida<sup>13</sup>, Mesostigmata<sup>14</sup>, Sarcoptiformes<sup>15</sup>, and Trombidiformes<sup>16</sup>. DNA barcodes have also helped resolve issues like lumping due to cryptic morphology<sup>17</sup>, and splitting due to heteromorphy<sup>18</sup>. However, concordance between species and BINs has only been tested in a single mite lineage: medically important ticks from Canada<sup>19</sup>.

While many species of mites have detrimentally impacted human health and agriculture<sup>20,21</sup>, others are recognized for their benefits as biological control agents<sup>22</sup>. The poultry red mite (PRM; *Dermanyssus gallinae* (De Geer, 1778), for example, is a widespread pest with significant economic costs<sup>23</sup>. Since the PRM is now resistant to most acaricides, the need for novel biocontrol methods is greater than ever<sup>24,25</sup>. From this perspective, natural mite communities in soil and bird nests may provide novel predators for conservation biological control of the PRM, but have seen limited investigation<sup>26–28</sup>. In the present study we begin the development of a DNA barcode reference library for the identification of poultry litter, soil, and nest dwelling mites in the Western Palearctic. Specifically, we test the correspondence between BINs and traditionally recognized species, and analyze intraspecific divergences at COI to identify potentially cryptic taxa.

## Methods

**Specimen Collection and Preparation.** Samples of poultry litter and soil from the vicinity of poultry houses, as well as wild bird nests, were collected between 2015 and 2016 from 53 locations in Croatia, Belgium, France, Israel, Poland, and Spain (Fig. 1, Table 1). Mites were extracted from approximately 0.5 kg of substrate into 99% ethanol (EtOH) using modified Berlese-Tullgren funnels for five days. From each unique collection event (denoted by exact site and collection date), all mites, regardless of life stage or sex, were sorted to morphotype and identified to order using a standard stereomicroscope setup and keys in Krantz and Walter<sup>29</sup>. Up to five specimens per morphotype were selected for molecular analysis. Each specimen was imaged using a Leica DVM6 microscope and arrayed into a 96-well microplate (Eppendorf) containing 30  $\mu$ L of 99% EtOH, with one blank well serving as a negative control. The museum identification code (Sample ID), collection details, order level taxonomy, and specimen images were uploaded to BOLD, available in the dataset DS-SMRPM through at <https://doi.org/10.5883/DS-SMRPM>.

**Molecular Analysis.** The specimens were sequenced for the barcode region of COI using standard invertebrate DNA extraction<sup>30,31</sup>, amplification<sup>32</sup> and sequencing protocols<sup>33</sup> at the Canadian Centre for DNA barcoding (CCDB; <http://ccdb.ca/>). However, DNA extraction was modified following Porco *et al.*<sup>34</sup> to facilitate the recovery of voucher specimens. A cocktail (1:1 ratio) of LepF1/LepRI<sup>1</sup> and LCO1490/HCO2198<sup>35</sup> primers were chosen to

Site No.	Country	State/Province	Exact Site	Lat	Lon	Sample Type
1	Belgium	Antwerp	Antwerpen	51.1911	4.4267	Bird nest
2	Belgium	Antwerp	Boechout (Boshhoek)	51.1241	4.5228	Bird nest
3	Belgium	Antwerp	Bornem	51.1106	4.2284	Bird nest
4	Belgium	Antwerp	Brasschaat	51.2718	4.4852	Bird nest
5	Belgium	Antwerp	Hove (Boshhoek)	51.1367	4.5096	Bird nest
6	Belgium	Antwerp	Lint	51.1266	4.4933	Bird nest
7	Belgium	Antwerp	Mechelen	51.0229	4.4848	Bird nest
8	Belgium	Antwerp	Niel	51.1010	4.3409	Bird nest
9	Belgium	Antwerp	Puurs	51.0847	4.3244	Bird nest
10	Belgium	East Flanders	Aalst	50.9397	4.0578	Bird nest
11	Belgium	East Flanders	Destelbergen	51.0539	3.8203	Bird nest
12	Belgium	East Flanders	Gerardsbergen	50.7741	3.9422	Bird nest
13	Belgium	East Flanders	Kalken	51.0358	3.9221	Bird nest
14	Belgium	East Flanders	Merelbeke	50.9446	3.7177	Bird nest
15	Belgium	East Flanders	Oudenaarde	50.8456	3.6093	Bird nest
16	Belgium	East Flanders	Zottegem	50.8951	3.8262	Bird nest
17	Belgium	Flemish Brabant	Boutersem	50.8561	4.8739	Bird nest
18	Belgium	Flemish Brabant	Kortenberg	50.8739	4.5413	Bird nest
19	Belgium	Flemish Brabant	Oud-Heverlee	50.8198	4.6675	Bird nest
20	Belgium	Flemish Brabant	Overijse	50.7701	4.5389	Bird nest
21	Belgium	Flemish Brabant	Rotselaar	50.9451	4.7487	Bird nest
22	Belgium	Flemish Brabant	Tielt-Winge	50.9243	4.8641	Bird nest
23	Belgium	Flemish Brabant	Tienen	50.8045	4.9321	Bird nest
24	Croatia		Zagreb	45.8248	15.969	Poultry litter
25	France	Auvergne Rhones Alpes	Issirac	44.7233	5.0411	Poultry litter
26	France	Auvergne Rhones Alpes	Lhuis	45.7482	5.5416	Poultry litter
27	France	Auvergne Rhones Alpes	Mionnay	45.8948	4.9199	Poultry litter
28	France	Auvergne Rhones Alpes	Relevant	46.0897	4.9450	Poultry litter
29	France	Auvergne Rhones Alpes	Relevant	45.8791	5.2552	Poultry litter
30	France	Auvergne Rhones Alpes	Rignieux	45.9491	5.1788	Poultry litter
31	France	Auvergne Rhones Alpes	Saint Etienne du Bois	46.2330	4.9319	Poultry litter
32	Israel	Central Coastal Plain	Moshav Satria	31.8915	34.8403	Poultry litter
33	Israel	Jerusalem	Jerusalem	31.7947	35.2410	Soil
34	Israel	Jerusalem	Nehusha	31.6284	34.9523	Soil
35	Israel	Northern	Bét Alfa	32.5176	35.4364	Soil
36	Israel	Northern	'En Ya'aqov	33.0093	35.2352	Poultry litter
37	Israel	Northern	Kammon	32.9154	35.3608	Soil
38	Israel	Northern	Kefar Yehoshua'	32.6747	35.1519	Poultry litter
39	Israel	Northern	Korazim	32.9070	35.5506	Poultry litter
40	Israel	Northern	New é Ya'ar	32.7056	35.1801	Soil
41	Israel	Northern	Ramat Zevi	32.1079	35.4158	Poultry litter
42	Israel	Northern	Sede Ya'aqov	32.6989	35.1439	Poultry litter
43	Israel	Northern	Zar'it	33.0985	35.2847	Soil
44	Poland	Masovian	Deba	51.4387	22.1781	Poultry litter
45	Poland	Masovian	Zygmunt	51.7810	21.6713	Poultry litter
46	Spain	Andalusia	Castilnovo	36.2530	-6.0803	Bird nest
47	Spain	Andalusia	La Barca de Vejer	36.2605	-5.9613	Bird nest
48	Spain	Castilla-La Mancha	Abenojar	38.8958	-4.4366	Bird nest
49	Spain	Castilla-La Mancha	Alcazar de San Juan	39.3899	-3.2109	Bird nest
50	Spain	Castilla-La Mancha	Almodovar del Campo	38.7312	-4.1880	Bird nest
51	Spain	Castilla-La Mancha	Cabaneros National Park	39.2852	-4.3392	Bird nest
52	Spain	Castilla-La Mancha	El Rostro	39.2971	-4.4165	Bird nest
53	Spain	Comunidad de Madrid	Rascafria	40.8717	-3.8982	Bird nest

**Table 1.** Summary of the 53 collection locations including the type of sample collected at each locality.

amplify and sequence a 652 bp fragment of DNA from the barcode region of COI because of their prior success in a broad array of mite taxa<sup>11</sup>. The DNA extracts were archived in  $-80^{\circ}\text{C}$  freezers at the Centre for Biodiversity Genomics (CBG; biodiversitygenomics.net), and the specimen vouchers were stored in 95% EtOH and returned to the Newe-Ya'ar Research Center and the Centre d'Ecologie Fonctionnelle & Evolutive for morphological preparations.

Taxon	Specimens	PCR Products	Sequences
Mesostigmata	456	373 (81.8%)	254 (55.7%)
Opilioacarida	4	3 (75.0%)	0 (0.0%)
Sarcoptiformes: Astigmatina	106	48 (45.3%)	0 (0.0%)
Sarcoptiformes: Oribatida	10	10 (100%)	4 (4.0%)
Trombidiformes	76	65 (85.5%)	40 (52.6%)
Total	652	499 (76.5%)	298 (45.7%)

**Table 2.** Summary of the number of specimens analysed, with the number of PCR products and barcode compliant sequences generated for each order. Success rates are provided in brackets.

The forward and reverse chromatograms were assembled into consensus sequences for each specimen and edited using CodonCode Aligner v. 4.2.7 and uploaded to BOLD. Each sequence meeting minimum quality criteria ( $\geq 500$  base pairs,  $< 1\%$  ambiguous nucleotides, free of contamination and stop codons) was assigned a BIN by BOLD. The sequences were further validated by inspecting their placement in a Neighbor-Joining tree (K2P distance model, BOLD alignment) and corresponding specimen images using the ‘Taxon ID Tree’ function in BOLD (Supplementary Figs 1 and 2). Taxa with unexpected placement in the tree (i.e. conflicting identifications within a cluster, conspecifics forming outgroups, etc.) were blasted against all barcode records on BOLD using the ‘Identification Engine’ tool whereupon instances of contamination (i.e. bacteria, Insecta, etc.) were flagged and filtered from the reference library.

**Specimen Identification.** Following BIN assignment, up to five vouchers per BIN were prepared for light microscopy by either mounting the specimens directly into Hoyer’s medium, or in the case of Oribatida, placing the specimen in lactic acid on a cavity slide. Since the specimens were sufficiently cleared during the tissue lysis stage of DNA extraction, the typical clearing procedures were not necessary. All remaining vouchers were prepared for SEM imaging on a Hitachi TM3000 TableTop Scanning Electron Microscope, with standard drying and coating procedures.

Each specimen was identified to the lowest possible level of taxonomy, and compared to identifications of other members of the same BIN. Some specimens were not slide mounted because of redundancy, or morphologically identified when precluded by their life stage, sex or voucher quality, and were thus assigned the lowest level of taxonomy in agreement with other members in the BIN. Specimens identified in this way were denoted by ‘BIN Taxonomy Match’ in the Identification Method field.

**Data Analysis.** Sampling completeness was assessed by constructing a BIN accumulation curve and by estimating total BIN richness using the incidence coverage estimator (ICE) in EstimateS<sup>36</sup>. Maximum intraspecific and minimum interspecific p-distances were calculated for all morphologically identified specimens using the ‘Barcode Gap Analysis’ tool on BOLD. Species correspondence with BINs were characterized by one of four categories: matches (perfect correspondence between one species and one BIN), splits (one species is represented by more than one BIN), merges (two or more species are assigned to a single BIN), and mixtures (a combination of splits and merges) as described in Ratnasingham and Hebert<sup>2</sup>.

## Results

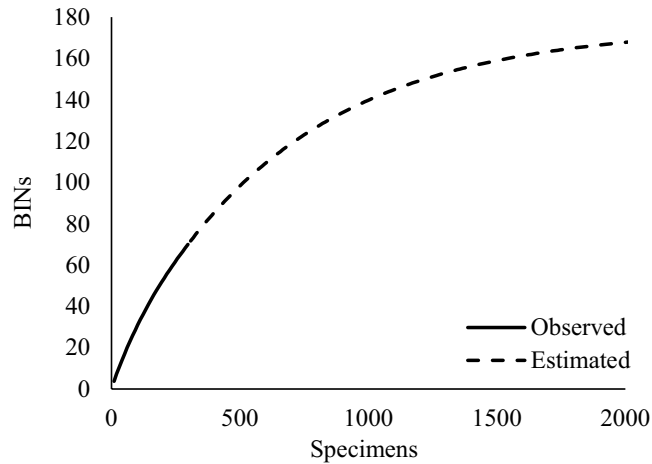
**Sequence Recovery.** Barcode compliant sequences were recovered from 298 of the 652 specimens analysed, with an overall PCR success rate of 76.5% and sequencing success rate of 45.7%. Success varied greatly among the major lineages. PCR success, for example, ranged from a high of 85% in the Trombidiformes, to a low of 45% in the Astigmatina (Sarcoptiformes). Sequencing success, on the other hand, ranged from a high of 56% in the Mesostigmata to a low of 0% in the Astigmatina (Sarcoptiformes) and Opilioacarida (Table 2). Non-target amplification was detected in 28 sequences, including cross-mite contamination, insects, and occasionally bacteria. These sequences were flagged on BOLD, removed from the BOLD identification engine, and excluded from subsequent analyses.

**DNA Barcode Reference Library and Sample Completeness.** Minimum quality requirements for BIN assignment were met by 298 sequences representing 70 BINs in total ( $\bar{x} = 4.2$  specimens/BIN). Of these 70 BINs, 48 (68.6%) were morphologically identified to the species level, while genus was the lowest identification for six BINs (8.6%), family for 15 BINs (21.4%), and one BIN was identified only to the order level (1.4%). In total, 35 species, 27 genera, 24 families, and three orders were identified in our barcode reference library (Table 3). The slope of the BIN accumulation curve remains steep, indicating incomplete sampling of the fauna (Fig. 2), and the estimate of total BIN richness was more than double the current observations (ICE = 172 BINs).

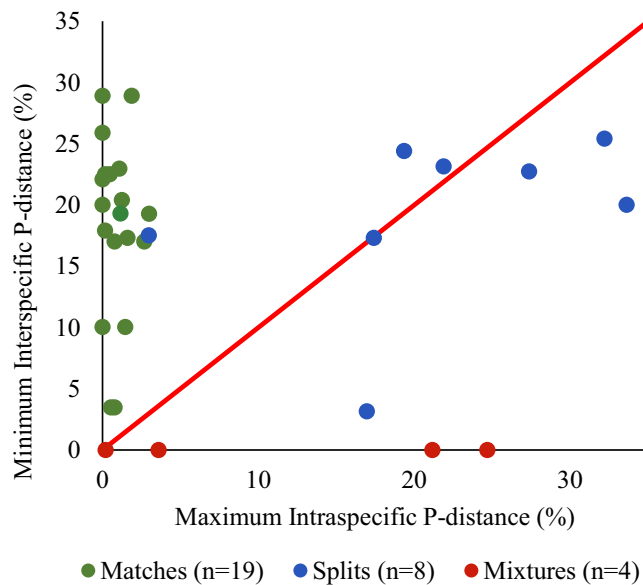
**Barcode Gap and BIN Analysis.** Of the 35 morphologically identified species with BINs, 19 (61%) perfectly corresponded with BIN assignments, while eight (26%) resulted in BIN splits, and two cases of BIN mixtures affecting four species (13%) were detected (Fig. 3, Table 3). The barcode gap analysis revealed nine species in which maximum interspecific p-distance exceeded minimum intraspecific p-distance (Fig. 3), all of which were involved in BIN splits or mixtures. Maximum intraspecific p-distances averaged 7.7%, and dropped to 0.9% when BIN splits and mixtures were excluded from analyses.

Order	Family	Genus/Species	Specimens	Specimens with BINs	BINs	BIN Category	Maximum Intraspecific P-distance (%)	Minimum Interspecific P-distance (%)
Mesostigmata		Mesostigmata spp.	167	2	1			
	Ameroseiidae	<i>Ameroseius eumorphus</i> Bregetova, 1077	5	5	1	Match	1.87	28.92
		<i>Ameroseius macrochela</i> (Westerboer, 1963)	1	1	1	Singleton		
		<i>Ameroseius</i> sp.	2	2	1			
		Ascidae	Ascidae sp.	1	1	1		
	Blattisociidae	<i>Lasioseius floridensis</i> Berlese, 1916	9	8	2	Split	19.36	24.42
	Dermanyssidae	<i>Dermanyssus carpathicus</i> Zeman, 1979	5	5	1	Match	1.15	19.29
		<i>Dermanyssus gallinae</i> (De Geer, 1778)	7	7	1	Match	2.99	19.29
	Digamasellidae	Digamasellidae sp.	4	4	1			
		<i>Dendrolaelaps longisculus</i> (Leitner, 1949)	8	8	1	Mixture	0.19	0
		<i>Dendrolaelaps presepum</i> * (Berlese, 1918)	19	15	3	Mixture	24.70	0
	Laelapidae	Laelapidae spp.	3	1	1			
		<i>Androlaelaps casalis</i> * (Berlese, 1887)	19	12	2	Split	33.63	20.02
		<i>Androlaelaps</i> sp.	1	1	1			
		<i>Gaeolaelaps aculeifer</i> (Canestrini 1883)	5	5	1	Match	0	22.11
		<i>Stratiolaelaps scimitus</i> (Berlese, 1892)	4	4	1	Match	0	20.02
	Macrochelidae	<i>Macrocheles matrius</i> * (Hull, 1925)	5	5	1	Match	0.15	22.52
		<i>Macrocheles merdarius</i> * (Berlese, 1889)	16	15	2	Split	21.89	23.16
		<i>Macrocheles muscaedomesticae</i> * (Scopoli, 1772)	21	21	1	Match	0.46	22.52
		<i>Macrocheles penicilliger</i> (Berlese, 1904)	5	5	2	Split	32.21	25.42
		<i>Macrocheles scutatiformis</i> Petrova, 1967	1	1	1	Singleton		
	Macronyssidae	<i>Ornithonyssus sylviarum</i> (Canestrini & Fanzago, 1877)	5	5	1	Match	0	25.89
	Melicharidae	<i>Proctolaelaps</i> sp.	4	4	1			
		<i>Proctolaelaps nr. parascolyti</i> *Costa, 1963	7	7	4	Mixture	21.17	0
		<i>Proctolaelaps pygmaeus</i> (Müller, 1859)	5	6	2	Mixture	3.60	0
		<i>Proctolaelaps scolyti</i> Evans, 1958	8	8	2	Split	16.95	3.15
	Parasitidae	<i>Cologamasus</i> sp.	1	1	1			
		<i>Gamasodes spiniger</i> * (Oudemans, 1936)	13	13	1	Match	2.67	17.02
		<i>Parasitus fimetorum</i> * Hyatt, 1980	16	15	2	Split	17.40	17.32
		<i>Parasitus hyalinus</i> (Willmann, 1949)	12	12	1	Match	1.24	20.41
		<i>Poecilochirus carabi</i> G. Canestrini & R. Canestrini, 1882	7	7	1	Match	1.60	17.32
		<i>Vulgarogamasus burchanensis</i> (Oudemans, 1903)	13	13	1	Match	0.77	17.02
Polyaspididae	<i>Uroseius</i> sp.	2	2	1				
Rhodacaridae	<i>Protogamasellopsis corticalis</i> Evans & Purvis, 1987	11	11	3	Split	27.38	22.74	
	<i>Rhodacarellus silesiacus</i> Willmann, 1935	3	3	2	Split	2.97	17.52	
Trematuridae	Trematuridae sp.	1	1	1				
	<i>Nenteria floralis</i> Karg, 1986	5	—	—				
	<i>Trichouropoda orbicularis</i> (C.L. Koch, 1839)	4	4	1	Match	0	10.03	
	<i>Trichouropoda ovalis</i> (C.L. Koch, 1839)	4	4	1	Match	0.16	17.90	
Urodinychidae	<i>Uroobovella fimicola</i> * (Berlese, 1903)	8	6	1	Match	1.08	22.96	
	<i>Uroobovella marginata</i> * (C. L. Koch, 1839)	13	1	1	Singleton			
Uropodidae	<i>Uropoda orbicularis</i> (Müller, 1776)	4	4	1	Match	1.46	10.03	
Opilioacarida	Opilioacarida sp.	4	—	—				
Sarcoptiformes	Astigmatina spp.	65	—	—				
	Oribatida spp.	47	—	—				
Oppiidae	Oppiidae sp.	4	4	1				
Trombidiformes	Trombidiformes spp.	35	—	—				
	Anystidae	Anystidae sp.	1	1	1			
	Bdellidae	Bdellidae spp.	8	8	4			
	Cheyletidae	<i>Cheletomorpha lepidopterorum</i> (Shaw, 1794)	2	2	1	Match	0	28.92
		<i>Cheyletus bidentatus</i> Fain and Nadchatram, 1980	14	14	1	Match	0.54	3.47
		<i>Cheyletus malaccensis</i> * (Oudemans, 1903)	5	5	1	Match	0.77	3.47
	Cunaxidae	Cunaxidae spp.	4	4	3			
	Erythracaridae	Erythracaridae sp.	1	1	1			
	Eupodidae	Eupodidae sp.	1	—	—			
	Scutacaridae	Scutacaridae sp.	1	1	1			
Tetranychidae	<i>Tetranychus urticae</i> (C.L. Koch, 1833)	1	1	1	Singleton			
Tydeidae	Tydeidae sp.	3	3	1				

**Table 3.** Breakdown of the 652 specimens analysed including the number of sequences with BIN assignments and summary of BINs for each taxon. Species are characterized into BIN categories with estimates of intra- and interspecific distances. The species previously associated with the poultry red mite are denoted by asterisks (\*).



**Figure 2.** The observed (solid line) and estimated (dashed line) accumulation of BINs with increasing sample size for the 298 specimens with BIN assignments.



**Figure 3.** Comparison of maximum intraspecific and minimum interspecific divergences (p-distances) of the 35 morphologically identified species. Data points are coloured based on species correspondence with BINs, and the diagonal red line indicates the 1:1 ratio of divergences. The barcode gap is present in species that fall above the line, and absent in those below.

## Discussion

Through the integration of morphological and molecular taxonomic methods, we provide DNA barcode coverage for 35 described species and 70 mite BINs from soil, bird nest, and poultry house-associated assemblages in the Western Palearctic. The integrity of most vouchers was sufficiently maintained for morphological identification, and SEM imaging of diagnostic characters (see the following BIN page for example: [BOLD:ADA3054](#)). While only 13 of these species have been previously associated with the poultry red mite<sup>27,37</sup>, additional species are undoubtedly present in our dataset but remain undetected because of low sequencing success combined with several BINs lacking identifications. Our failure to generate any sequences for Astigmatina (Sarcoptiformes) may be explained by low primer affinity, considering amplification rates were also lowest in this group. Primer affinity, however, does not justify the low successes in other lineages with higher amplification rates. Comparable methods, for example, have yielded much higher successes (77%) among soil and leaf litter mites (including Astigmatina) from subarctic Canada<sup>11</sup>, demonstrating the broad applicability of these primers among a diverse array of taxa. Since 40% of the amplification products generated uninterpretable chromatograms, poor quality DNA template may be responsible for low sequencing successes among taxa.

The concordance between BINs and mite species was much lower than in some well-studied invertebrates (e.g. perfect concordance in 92% of beetles<sup>4</sup> and ticks<sup>19</sup>). However, similar concordance levels have been reported



for many taxa including geometrid moths<sup>38</sup> (67%), true bugs<sup>39</sup> (70%), and spiders<sup>5</sup> (54%). Low concordance is mainly driven by species with large intraspecific divergences (>3% p-distance) resulting in the assignment of two or more BINs. While this does not preclude accurate barcode-based identification, it highlights potentially cryptic species because most BIN splits formed widely separated clades (e.g. >15% p-distance) lacking intermediate haplotypes. In fact, 16S and 18S rRNA gene topologies for *Androlaelaps casalis* (Berlese, 1887) and *Proctolaelaps scolyti* Evans, 1958 were congruent with BIN splits, further supporting our cryptic species hypothesis in these taxa<sup>27</sup>. *Rhodacarellus silesiacus* Willmann, 1936, on the other hand, also formed two distinct but narrowly separated clades (<3% divergence), with divergences similar to those in species with concordant BINs (e.g. *Dermanyssus gallinae* and *Gamasodes spiniger* (Oudemans, 1936)), such that additional sampling may reveal intermediate haplotypes causing the BINs to collapse into one<sup>2</sup>.

More problematic for the barcode based identification of mites are the two cases of shared barcodes confounded by BIN splits (BIN mixtures) affecting four species: *Dendrolaelaps longiusculus* (Leitner, 1949)/*D. presepum* (Berlese, 1918), and *Proctolaelaps parascolyti* Costa, 1963/*P. pygmaeus* (Müller, 1859). Since multiple species are assigned to the same BIN, mixtures impede accurate identifications, but may also represent taxonomic errors<sup>2</sup>. Misidentification is unlikely, since procedures were in place to evaluate and correct such errors. However, both cases of BIN mixtures involve closely allied congeners which may be subjected to hybridization or incomplete lineage sorting<sup>40</sup>. Given the large intraspecific divergences observed, though, a more probable explanation is the presence of cryptic diversity compounded by inadequate species descriptions. Future work should scrutinize the morphology of genetic clusters from both mixtures and splits for more effective characters to discriminate these potentially cryptic species.

This study represents the first step towards development of a DNA barcode reference library for the identification of poultry litter, soil, and nest dwelling mites from the Western Palearctic, which may in turn reveal natural enemies key to the control of PRM. Although sequencing success rates should be improved, we demonstrate that nearly 80% of the species analysed can be accurately identified through DNA barcodes. Our BIN analysis, however, indicates a high proportion of cryptic diversity and some potential taxonomic confusion. This method consequently presents a powerful tool not only for the identification of unknown specimens, but as the foundation for integrative taxonomy and diversity estimation in hyperdiverse invertebrates such as mites.

## Data Availability

All specimen and sequence data is available in the BOLD dataset DS-SMRPM through the following, <https://doi.org/10.5883/DS-SMRPM>. Valid sequences were also deposited in GenBank under the following accessions: MH983560-MH983861.

## References

1. Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. Biological identifications through DNA barcodes. *Proc. R. Soc. Lond. B.* **270**, 313–321 (2003).
2. Ratnasingham, S. & Hebert, P. D. N. A DNA-based registry for all animal species: The barcode index number (BIN) system. *PLoS ONE* **8**, e66213 (2013).
3. Huemer, P., Mutanen, M., Sefc, K. M. & Hebert, P. D. N. Testing DNA barcode performance in 1000 species of European Lepidoptera: Large geographic distances have small genetic impacts. *PLoS ONE* **9**, e115774 (2014).
4. Hendrich, L. *et al.* A comprehensive DNA barcode database for Central European beetles with a focus on Germany: adding more than 3500 identified species to BOLD. *Mol. Ecol. Resour.* **15**, 795–818 (2015).
5. Blagoev, G. A. *et al.* Untangling taxonomy: a DNA barcode reference library for Canadian spiders. *Mol. Ecol. Resour.* **16**, 325–341 (2016).
6. Ren, J.-M. *et al.* Barcode index numbers expedite quarantine inspections and aid the interception of nonindigenous mealybugs (Pseudococcidae). *Biol. Invasions* **20**(2), 449–460 (2017).
7. Telfer, A. *et al.* Biodiversity inventories in high gear: DNA barcoding facilitates a rapid biotic survey of a temperate nature reserve. *Biodivers. Data J.* **3**, e6313 (2015).
8. Hebert, P. D. N., Hollingsworth, P. M. & Hajibabaei, M. From writing to reading the encyclopedia of life. *Philos. Trans. Royal Soc. B.* **371**, 20150321 (2016).
9. D'Souza, M. L. & Hebert, P. D. N. Stable baselines of temporal turnover underlie high beta diversity in tropical arthropod communities. *Mol. Ecol.* **27**, 2447–2460 (2018).
10. Walter, D. E. & Proctor, H. C. Mites and biological diversity. in *Mites: Ecology, evolution & behaviour* 447–459 (Springer, 2013).
11. Young, M. R., Behan-Pelletier, V. M. & Hebert, P. D. N. Revealing the hyperdiverse mite fauna of subarctic Canada through DNA barcoding. *PLoS ONE* **7**, e48755 (2012).
12. Beaulieu, F. *et al.* Acari of Canada. *ZooKeys* **819**, 77–168 (2019).
13. Zhang, R. L. & Zhang, B. Prospects of using DNA barcoding for species identification and evaluation of the accuracy of sequence databases for ticks (Acari: Ixodida). *Ticks Tick Borne Dis.* **5**, 352–358 (2014).
14. Li, J.-B. *et al.* COI barcoding as a molecular assay for the identification of phytoseiid mites. *Sys. Appl. Acarol.* **17** (2012).
15. Dabert, J., Ehrnsberger, R. & Dabert, M. *Glaucalges tytonis* sp. n. (Analgidae, Xolalgidae) from the barn owl *Tyto alba* (Strigiformes, Tytonidae): compiling morphology with DNA barcode data for taxon descriptions in mites (Acari). *Zootaxa* **1719**(11) (2008).
16. Pesic, V. & Smit, H. *Neumania kyrgyzica* sp. nov. a new water mite from Kyrgyzstan based on morphological and molecular data (Acari, Hydrachnidia: Unionicolidae). *Sys. Appl. Acarol.* **22**, 885 (2017).
17. Stalstedt, J., Bergsten, J. & Ronquist, F. "Forms" of water mites (Acari: Hydrachnidia): intraspecific variation or valid species? *Ecol. Evol.* **3**, 3415–3435 (2013).
18. Glowka, E., Dragun-Damian, A., Broda, L., Dabert, J. & Dabert, M. DNA barcodes reveal female dimorphism in syringophilid mites (Actinotrichida: Prostigmata: Cheyletoidea): *Stibarokris phoeniconaias* and *Ciconichenophilus phoeniconaias* are conspecific. *Folia Parasit.* **61**, 272–276 (2014).
19. Ondrejicka, D. A., Morey, K. C. & Hanner, R. H. DNA barcodes identify medically important tick species in Canada. *Genome* **60**, 74–84 (2017).
20. Lipsker, D., Jaulhac, B. (eds.) *Lyme borreliosis: biological and clinical aspects*. Vol. 37. (Karger, 2009).
21. Dhooria, M. S. *Fundamentals of applied acarology*. (Springer, 2016).
22. Mcmurtry, J. A., Moraes, G. J. D. & Sourassou, N. F. Revision of the lifestyles of phytoseiid mites (Acari: Phytoseiidae) and implications for biological control strategies. *Sys. Appl. Acarol.* **18**, 297 (2013).

23. Sparagano, O. *et al.* Prevalence and key figures for the poultry red mite *Dermanyssus gallinae* infections in poultry farm systems. *Exp. Appl. Acarol.* **48**, 3–10 (2009).
24. Marangi, M. *et al.* Evaluation of the poultry red mite, *Dermanyssus gallinae* (Acari: Dermanyssidae) susceptibility to some acaricides in field populations from Italy. *Exp. Appl. Acarol.* **48**, 11–18 (2009).
25. Sparagano, O. A. E., George, D. R., Harrington, D. W. J. & Giangaspero, A. Significance and control of the poultry red mite, *Dermanyssus gallinae*. *Annu. Rev. Entomol.* **59**, 447–466 (2014).
26. Lesna, I., Sabelis, M. W., van Niekerk, T. G. C. M. & Komdeur, J. Laboratory tests for controlling poultry red mites (*Dermanyssus gallinae*) with predatory mites in small 'laying hen' cages. *Exp. Appl. Acarol.* **58**, 371–383 (2012).
27. Roy, L. *et al.* Arthropod communities of laying hen houses: an integrative pilot study toward conservation biocontrol of the Poultry Red Mite *Dermanyssus gallinae*. *Biol. Control* **114**, 176–194 (2017).
28. Tamara, B. H., Granich, J., Körbes, J. H., da Silva, G. L. & Ferla, N. J. Mite fauna (Acari) associated with the poultry industry in different laying hen management systems in Southern Brazil: a species key. *Acarologia* **58**, 140–158 (2018).
29. Krantz, G. W. & Walter, D. E. (eds) *A Manual of Acarology*. 3rd Edition. (Texas Tech University Press, 2009).
30. Ivanova, N. V., Dewaard, J. R. & Hebert, P. D. N. An inexpensive, automation-friendly protocol for recovering high-quality DNA: Technical Note. *Mol. Ecol. Notes* **6**, 998–1002 (2006).
31. Ivanova, N. V., deWaard, J. R. & Hebert, P. D. N. CCDB protocols, glass fiber plate DNA extraction. [ccdb.ca/site/wp-content/uploads/2016/09/CCDB\\_DNA\\_Extraction.pdf](https://ccdb.ca/site/wp-content/uploads/2016/09/CCDB_DNA_Extraction.pdf) (2007).
32. Ivanova, N. V. & Grainger, C. M. CCDB protocols, COI amplification. [ccdb.ca/site/wp-content/uploads/2016/09/CCDB\\_Amplification.pdf](https://ccdb.ca/site/wp-content/uploads/2016/09/CCDB_Amplification.pdf) (2007).
33. Ivanova, N. V. & Grainger, C. M. CCDB protocols, sequencing. [ccdb.ca/site/wp-content/uploads/2016/09/CCDB\\_Sequencing.pdf](https://ccdb.ca/site/wp-content/uploads/2016/09/CCDB_Sequencing.pdf) (2007).
34. Porco, D., Rougerie, R., Deharveng, L. & Hebert, P. Coupling non-destructive DNA extraction and voucher retrieval for small soft-bodied Arthropods in a high-throughput context: the example of Collembola. *Mol. Ecol. Resour.* **10**, 942–945 (2010).
35. Folmer, O., Hoesh, W. R., Black, M. B. & Vrijenhoek, R. C. Conserved primers for PCR amplification of mitochondrial DNA from different invertebrate phyla. *Mol. Mar. Biol. Biotechnol.* **3**, 294–299 (1994).
36. Colwell, R. K. EstimateS: Statistical estimation of species richness and shared species from samples. [viceroy.eeb.uconn.edu/estimates/](http://viceroy.eeb.uconn.edu/estimates/) (2013).
37. Brady, J. The Mites of Poultry Litter: Observations on the Bionomics of Common Species, with a Species List for England and Wales. *J. Appl. Ecol.* **7**, 331 (1970).
38. Zahiri, R. *et al.* Probing planetary biodiversity with DNA barcodes: The Noctuoidea of North America. *PLoS One* **12**, e0178548 (2017).
39. Gwiazdowski, R. A., Footitt, R. G., Maw, H. E. L. & Hebert, P. D. The Hemiptera (Insecta) of Canada: constructing a reference library of DNA barcodes. *PLoS One* **10**, e0125635 (2015).
40. Funk, D. J. & Omland, K. E. Species-level paraphyly and polyphyly. *Annu. Rev. Ecol. Evol. Syst.* **34**, 397–423 (2003).

## Acknowledgements

We would like to express our gratitude to the Canadian Centre for DNA Barcoding (CCDB) for their technical assistance and support for molecular analysis, as well as the Barcode of Life Data System (BOLD) team at the Centre for Biodiversity Genomics for their support on data management and storage. We would also like to thank Jandir Cruz Santos for critically analysing nomenclature and providing helpful comments on the manuscript. The study was carried out within the framework of the European Cooperation in Science and Technology (COST) Action (FA1404-COREMI) 'Improving current understanding and research for sustainable control of the poultry red mite *Dermanyssus gallinae*'.

## Author Contributions

M.R.Y., L.R., Y.G. and E.P. devised the study; M.A., L.B., D.H., J.F.L.B., S.G. and E.G.R. collected the specimens and prepared the mites for molecular analysis and museum archiving; M.L.M. and E.U. identified the specimens; E.P. managed and coordinated the activities of the study; M.R.Y. wrote the manuscript and analysed the data; all authors contributed to discussion of ideas and manuscript revisions.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-41958-9>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019