Bar-HRM for identification of cryptic earthworm species

Gaspard Baudrin, Virginie Roy, Agnès Gigon, Lise Dupont

Short communication

Bar-HRM for identification of cryptic earthworm species

Gaspard Baudrin[1,2,3], Virginie Roy[1,2,3], Agnès Gigon[1,2,3] & Lise Dupont[1,2,3*]

[1] *Univ Paris-Est Creteil, CNRS, INRAE, IRD, IEES-Paris, F-94010 Creteil France*

[2] *Sorbonne Université, IEES-Paris, F-75005 Paris, France*

[3] *Université de Paris, IEES-Paris, F-75013 Paris, France*

[*] Corresponding author: lise.dupont@u-pec.fr

Abstract

Accurate species identification is crucial in ecological research on sentinel species such as earthworms. Here, we developed a reliable and cost-efficient method for the rapid identification of taxa within the *Allolobophora chlorotica* complex of cryptic earthworm species. We combined high resolution melting (HRM) analysis with DNA barcoding (Bar-HRM) in a three-step approach: (i) selection of a mini-barcode in the 16S mitochondrial gene, (ii) test of the method on a panel of 16 reference individuals and (iii) identification at the species level of 24 unknown specimens. Thus, we proved the efficacy of Bar-HRM to discriminate between cryptic sister (or sibling) earthworm species and we provided a standardized method that can be easily adapted to other taxa.

Key Words

Allolobophora chlorotica; complex of species; high resolution melting analysis; mitochondrial lineages; species identification

Accurate species identification is a fundamental issue of current biological and ecological researches (e.g. Elphick, 2008). Until quite recently, earthworm species identification was almost only conducted on morphological characters (Bartlett et al., 2010). Nowadays, molecular techniques are more frequently used and multiple studies revealed cryptic diversity (i.e. two or more distinct species that were classified as a single one due to morphological similarity, Bickford et al., 2007) even within long known earthworm morphospecies (e.g. King et al., 2008; James et al., 2010; Novo et al., 2010; Taheri et al., 2018). Biologically relevant differences were identified between cryptic species, highlighting that correct identification of earthworm specimens is crucial for further study of these ecologically important species (review in Marchan et al., 2018).

In the last decade, the DNA barcoding approach (Hebert et al., 2003) using the mitochondrial gene coding for cytochrome c oxidase subunit I (COI) has become the most used molecular taxonomy tool in earthworms (e.g. Decaens et al., 2016; Porco et al., 2018). In addition to allowing the detection of cryptic or overlooked diversity cases in earthworms, this approach is a way to recover specific level data for juvenile specimens and to process numerous specimens without the intervention of a taxonomist (Porco et al., 2018). While this approach holds many advantages, it can turn out to be relatively expensive when applied to large sample sizes and requires an access to a Sanger-sequencing or a high-throughput next-generation sequencing (NGS) platform.

Recently, DNA-based identification of genetic variants was made possible without any sequencing step. The high resolution melting (HRM) analysis is a highly sensitive method allowing to discriminate DNA sequences differing of only one substitution or one base pair indel (Wittwer, 2009). Following a real-time PCR, DNA fragments (from 80 to 250 bp long) are denatured with increasing temperature and the resulting changes in fluorescence caused by the release of an intercalating dye from the DNA duplex are monitored. Combining HRM

63 analysis to DNA barcoding (Bar-HRM) is thus a way to rapidly distinguish genetically
64 differentiated groups based on the thermal denaturation curves of amplified mini-barcodes
65 (Behrens-Chapuis et al., 2018; Chen et al., 2019). Thereby, Bar-HRM has appeared to be an
66 effective, simple, cheap (in long term and large scale investigation) and time-saving tool for
67 species identification (Fernandes et al., 2017; Fidler et al., 2017; Sun et al., 2017;
68 Osathanunkul et al., 2018).

69 The goal of the present study was to propose a standardized method of species
70 identification using Bar-HRM, from the evaluation of mini-barcode discriminating power to
71 the species assignation of unknown individuals, and to test its usefulness to distinguish cryptic
72 earthworm species using the *Allolobophora chlorotica* aggregate as a model. Two colour
73 morphs were described in the *Allolobophora chlorotica* aggregate, a green morph
74 representing a single taxon although composed of two divergent mitochondrial lineages (L2
75 and L3), and a pink colour morph, composed of at least two taxa (L1 and L4, Dupont et al.,
76 2016). The two colour morphs are known to have contrasting ecological preferences linked to
77 soil moisture (Satchell, 1967; Lowe and Butt, 2007) and to present postzygotic reproductive
78 isolating mechanisms (Lowe and Butt, 2008). Confirmation of species status for the L1,
79 L2/L3 and L4 lineages was obtained with microsatellite markers by Dupont et al. (2016) who
80 revealed that hybridization was rare among these taxa. The status of three other lineages (i.e.
81 L5, L6 and L7) remains however unclear (King et al., 2008; Dupont et al., 2011).

82 Here, we focused on the relatively well-defined L1, L2/L3 and L4 taxa. The
83 development of the bar-HRM method was led in three main steps. In a first time, we tested
84 whether the set of EwD/EwE primers developed by Bienert et al. (2012) for earthworms and
85 allowing to amplify a 111 bp mini-barcode region in the mitochondrial 16S rRNA gene was
86 adequate to discriminate these taxa. In order to estimate the distribution of the theoretical
87 melting temperatures (i.e. temperature at which 50% of DNA double-strands are dissociated)

of this mini-barcode among the L1, L2/L3 and L4 taxa, we used 41 published *A. chlorotica* agg. 16S sequences (Genbank Accession Numbers AM774359-93, KJ912500-505, JN869755-756). To compute the expected melting temperatures (Tm_{THEO}) from DNA sequences of the EwD/EwE region, we implemented the model described in Khandelwal & Bhyravabhotla (2010). Results showed significantly differentiated means between taxa (L1 = 78.4 ± 0.30, L2/L3 = 79.7 ± 0.47 and L4 = 80.6 ± 0.38) but slightly overlapping Tm_{THEO} ranges (Fig. 1). Because HRM takes into account the shape of the whole curve, in addition to the expected melting temperature, we assumed that it should still be possible to correctly identify specimens even from different DNA fragments displaying close melting temperatures. Interestingly, despite the divergence between L2 and L3 haplotypes in a phylogenetic tree built using a longer fragment of the 16S gene (Supplementary material, Fig S1), the Tm_{THEO} of these lineages, belonging to a single species, were not distinguishable using the EwD/EwE mini-barcode.

In a second time, we composed a panel of 16 reference individuals belonging to each of the three taxa (4 L1, 8 L2/L3 and 4 L4). These reference individuals were selected based on their already published COI sequences (Genbank Accession Number -AN- in supplementary material Table S1) in order to capture both intra and inter-lineage diversity and a 445 bp fragment of the 16S gene was subsequently sequenced using EwA/EwF primers (Bienert et al., 2012; AN in Table S1). The HRM analysis of these reference panel individuals was carried out in triplicates using MeltDoctor HRM Master Mix (Applied Biosystems) according to the manufacturer protocol in 10 µL reaction volume and using EwD/EwE primers. The fluorescence data was normalized by the use of the exponential background removal method described in Palais and Wittwer (2009) in order to obtain the curves describing the decrease of double-stranded DNA fractions in response to the temperature. We observed a clear

112 separation of the melting curves of the three taxa (Fig. 2a). The experimental melting
113 temperatures (Tm_{OBS}) ranges, i.e. between the lower and the higher Tm_{OBS}, did not overlap.
114 Finally, species identification was tested on 24 unknown specimens sampled in
115 Parisian parks and morphologically identified as members of the *A. chlorotica* aggregate.
116 Because the samples were stored in ethanol before DNA extraction (using the DNeasy ®
117 Blood & Tissue Kit, Qiagen, according to manufacturer protocol), and have therefore lost
118 their pigmentation, their colour morph could not be determined. The HRM analyses of the
119 unknown samples and the reference panel samples were conducted at the same time. The
120 distances between each unknown sample melting curve and every reference panel curves were
121 computed as the mean integral of the absolute difference between the two curves considered
122 (Palais and Wittwer, 2009). The unknown specimens could thus be assigned to one of the taxa
123 based on the lowest distance between their melting curve and the one of the reference samples
124 (Fig. 2b). Then, the identification of the 24 specimens was controlled by the sequencing of
125 their 16S gene using EwA/EwF primers (Table S1). Among them, 11 different haplotypes
126 were identified (Table S1) using the Dna SP V 5.10.01 software (Librado and Rozas, 2009).
127 HRM results matched with the 16S sequencing results for 10 of these haplotypes (i.e. 22
128 individuals). An erroneous identification comes from a haplotype (present in 2 individuals,
129 Table S1) assigned to the L4 taxa using the bar-HRM method while it clustered with L2
130 haplotypes in a phylogenetic tree (Supplementary material, Fig S1). This 16S haplotype
131 sequence was indeed expected to display a melting temperature closer to the L2/L3 group
132 than to the L4 group (Fig. 3). So, there is a discrepancy, in this particular case, between the
133 expected and observed melting temperatures and this haplotype appears to be a significant
134 outlier (Bonferroni test). Although this 16S haplotype clustered with the L2 lineage, it is
135 nevertheless noteworthy that the phylogenetic tree revealed a sequence quite divergent
136 (Supplementary material, Fig S1), a result that could partly explain the discrepancy. In order

to increase the specific identification accuracy and therefore correctly assign the identified outliers, we propose to use a multi-locus bar-HRM method. For instance, a mini-barcode designed in the NADH dehydrogenase subunit 1 (ND1) mitochondrial gene allowed to discriminate between L1/L2 and L3/L4 groups of lineages (Primers designed for this study: ND1F374 – TRGCTGGATGAAGHTCAAA and ND1R519 – GCAAGYCARGCATGRAAA; 163 bp amplicon, T_a = 52°C, Fig. S2 and S3).

To conclude, we demonstrated the effectiveness of HRM for distinguishing cryptic species of earthworms. We developed a simple and rapid bar-HRM method allowing to successfully identify most of the tested specimens belonging to the *A. chlorotica* agg. We believe that this Bar-HRM method could be easily adapted to other earthworm species following the three-step procedure described here. It could be used for accurate species identification, which is an essential step in many fields of ecology. For instance, rearing of earthworms may be necessary for laboratory experiments in ecotoxicology or soil restoration (Lowe and Butt, 2005). To ensure a successful rearing, it is necessary to mix specimens from the same species (i.e. in order to avoid reproductive isolation). The bar-HRM method may be particularly useful to rapidly confirm morphological species distinction. This method is indeed much faster than sequencing because PCR and HRM steps are performed in the same reaction tube without any other post-PCR manipulations. For species confirmation, melting profiles of samples can be compared to those of reference species in a database. The prerequisite is thus to have a good knowledge of the taxonomical status of the target taxa, as it is now the case for several complexes of earthworm species (e.g. *A. chlorotica*, Dupont et al., 2011; *Pontoscolex corethrurus*, Taheri et al. 2018).

Acknowledgements

This work was partly funded by the French National Research Agency (ANR) as a part of the project Ecoville No. ANR-14-CE22-0021. The authors are grateful to H. Audusseau for her help in sampling earthworms in Parisian parks and to the PRAMMICS platform of the OSU Efluve for making qPCR thermocycler available.

References

Bartlett, M.D., Briones, M.J.I., Neilson, R., Schmidt, O., Spurgeon, D.,Creamer, R.E., 2010. A critical review of current methods in earthworm ecology: From individuals to populations. European Journal of Soil Biology 46, 67-73.

Behrens-Chapuis, S., Malewski, T., Suchecka, E., Geiger, M.F., Herder, F.,Bogdanowicz, W., 2018. Discriminating European cyprinid specimens by barcode high-resolution melting analysis (Bar-HRM)-A cost efficient and faster way for specimen assignment? Fisheries Research 204, 61-73.

Bickford, D., Lohman, D.J., Sodhi, N.S., Ng, P.K.L., Meier, R., Winker, K., Ingram, K.K.,Das, I., 2007. Cryptic species as a window on diversity and conservation. Trends in Ecology & Evolution 22, 148-155.

Bienert, F., De Danieli, S., Miquel, C., Coissac, E., Poillot, C., Brun, J.J.,Taberlet, P., 2012. Tracking earthworm communities from soil DNA. Molecular Ecology 21, 2017-2030.

Chen, Y.Q., Wang, B., Wang, P., Huang, B.S., Liu, H.G., Xiong, C., Wu, M.L., Wang, N.L., Liu, D., Hu, Z.G.,Chen, S.L., 2019. Rapid identification of common medicinal snakes and their adulterants using the Bar-HRM analysis method. Mitochondrial DNA Part A 30, 367-374.

Decaens, T., Porco, D., James, S.W., Brown, G.G., Chassany, V., Dubs, F., Dupont, L., Lapied, E., Rougerie, R., Rossi, J.-P.,Roy, V., 2016. DNA barcoding reveals diversity patterns of earthworm communities in remote tropical forests of French Guiana. Soil Biology & Biochemistry 92, 171-183.

Dupont, L., Lazrek, F., Porco, D., King, R.A., Rougerie, R., Symondson, W.O.C., Livet, A., Richard, B., Decaens, T., Butt, K.R.,Mathieu, J., 2011. New insight into the genetic structure of the *Allolobophora chlorotica* aggregate in Europe using microsatellite and mitochondrial data. Pedobiologia 54, 217-224.

Dupont, L., Porco, D., Symondson, W.O.C.,Roy, V., 2016. Hybridization relics complicate barcode-based identification of species in earthworms. Molecular Ecology Resources 16, 883-894.

Elphick, C.S., 2008. How you count counts: the importance of methods research in applied ecology. Journal of Applied Ecology 45, 1313-1320.

Fernandes, T.J.R., Costa, J., Oliveira, M.,Mafra, I., 2017. DNA barcoding coupled to HRM analysis as a new and simple tool for the authentication of Gadidae fish species. Food Chemistry 230, 49-57.

Fidler, G., Kocsube, S., Leiter, E., Biro, S.,Paholcsek, M., 2017. DNA Barcoding coupled with High Resolution Melting analysis enables rapid and accurate distinction of *Aspergillus* species. Medical Mycology 55, 642-659.

Hebert, P.D.N., Cywinska, A., Ball, S.L.,Dewaard, J.R., 2003. Biological identifications through DNA barcodes. Proceedings of the Royal Society of London Series B: Biological Sciences 270, 313-321.

James, S.W., Porco, D., Decaens, T., Richard, B., Rougerie, R.,Erseus, C., 2010. DNA Barcoding reveals cryptic diversity in *Lumbricus terrestris* L. 1758 (Clitellata): resurrection of *L. herculeus* (Savigny, 1826). Plos One 5, e15629.

Khandelwal, G.,Bhyravabhotla, J., 2010. A phenomenological model for predicting melting temperatures of DNA sequences. Plos One 5, e12433.

King, R.A., Tibble, A.L.,Symondson, O.C., 2008. Opening a can of worms: unprecedented sympatric cryptic diversity within British lumbricid earthworms. Molecular Ecology 17, 4694-4698.

Librado, P.,Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25, 1451-1452.

Lowe, C.N.,Butt, K.R., 2005. Culture techniques for soil dwelling earthworms: A review. Pedobiologia 49, 401-413.

Lowe, C.N.,Butt, K.R., 2007. Life-cycle traits of the dimorphic earthworm species *Allolobophora chlorotica* (Savigny, 1826) under controlled laboratory conditions. Biology and Fertility of Soils 43, 495-499.

Lowe, C.N.,Butt, K.R., 2008. *Allolobophora chlorotica* (Savigny, 1826): Evidence for classification as two separate species. Pedobiologia 52, 81-84.
Marchan, D.F., Cosin, D.J.D.,Novo, M., 2018. Why are we blind to cryptic species? Lessons from the eyeless. European Journal of Soil Biology 86, 49-51.
Novo, M., Almodovar, A., Fernandez, R., Trigo, D.,Cosin, D.J.D., 2010. Cryptic speciation of hormogastrid earthworms revealed by mitochondrial and nuclear data. Molecular Phylogenetics and Evolution 56, 507-512.
Osathanunkul, M., Osathanunkul, K., Wongwanakul, S., Osathanunkul, R.,Madesis, P., 2018. Multiuse of Bar-HRM for *Ophiocordyceps sinensis* identification and authentication. Scientific Reports 8
Palais, R.,Wittwer, C.T., 2009. Mathematical algorithms for high-resolution DNA melting analysis. In: Johnson, M.L., Brand, L. (eds) Methods in Enzymology: Computer Methods, Vol 454, Pt A, pp. 323-343
Porco, D., Chang, C.-H., Dupont, L., James, S., Richard, B.,Decaëns, T., 2018. A reference library of DNA barcodes for the earthworms from Upper Normandy: Biodiversity assessment, new records, potential cases of cryptic diversity and ongoing speciation. Applied Soil Ecology 124, 362-371.
Satchell, J.E., 1967. Colour dimorphism in *Allolobophora chlorotica* Sav. (Lumbricidae). Journal of Animal Ecology 36, 623-630.
Sun, W., Yan, S., Li, J.J., Xiong, C., Shi, Y.H., Wu, L., Xiang, L., Deng, B., Ma, W.,Chen, S.L., 2017. Study of commercially available *Lobelia chinensis* products using Bar-HRM technology. Frontiers in Plant Science 8
Taheri, S., James, S., Roy, V., Decaens, T., Williams, B.W., Anderson, F., Rougerie, R., Chang, C.H., Brown, G., Cunha, L., Stanton, D.W.G., Da Silva, E., Chen, J.H., Lemmon, A.R., Lemmon, E.M., Bartz, M., Baretta, D., Barois, I., Lapied, E., Coulis, M.,Dupont, L., 2018. Complex taxonomy of the 'brush tail' peregrine earthworm *Pontoscolex corethrurus*. Molecular Phylogenetics and Evolution 124, 60-70.
Wittwer, C.T., 2009. High-Resolution DNA Melting analysis: advancements and limitations. Human Mutation 30, 857-859.

254 **Figure legends**

255

256 **Figure 1**: Expected melting temperatures (Tm_{THEO}) of the 16S EwD/EwE fragment according
257 to the three *Allolobophora chlorotica* agg. taxa, computed from published sequences.

258

259 **Figure 2**: Melting curves profiles. a) panel of 16 reference individuals (L1 yellow, L2/L3
260 blue, L4 orange) and b) panel of 16 reference individuals (dotted line) and 24 assigned
261 unknown specimens (solid line) with non-matching control individuals in purple. HRM
262 analysis of each individual was carried out in triplicates.

263

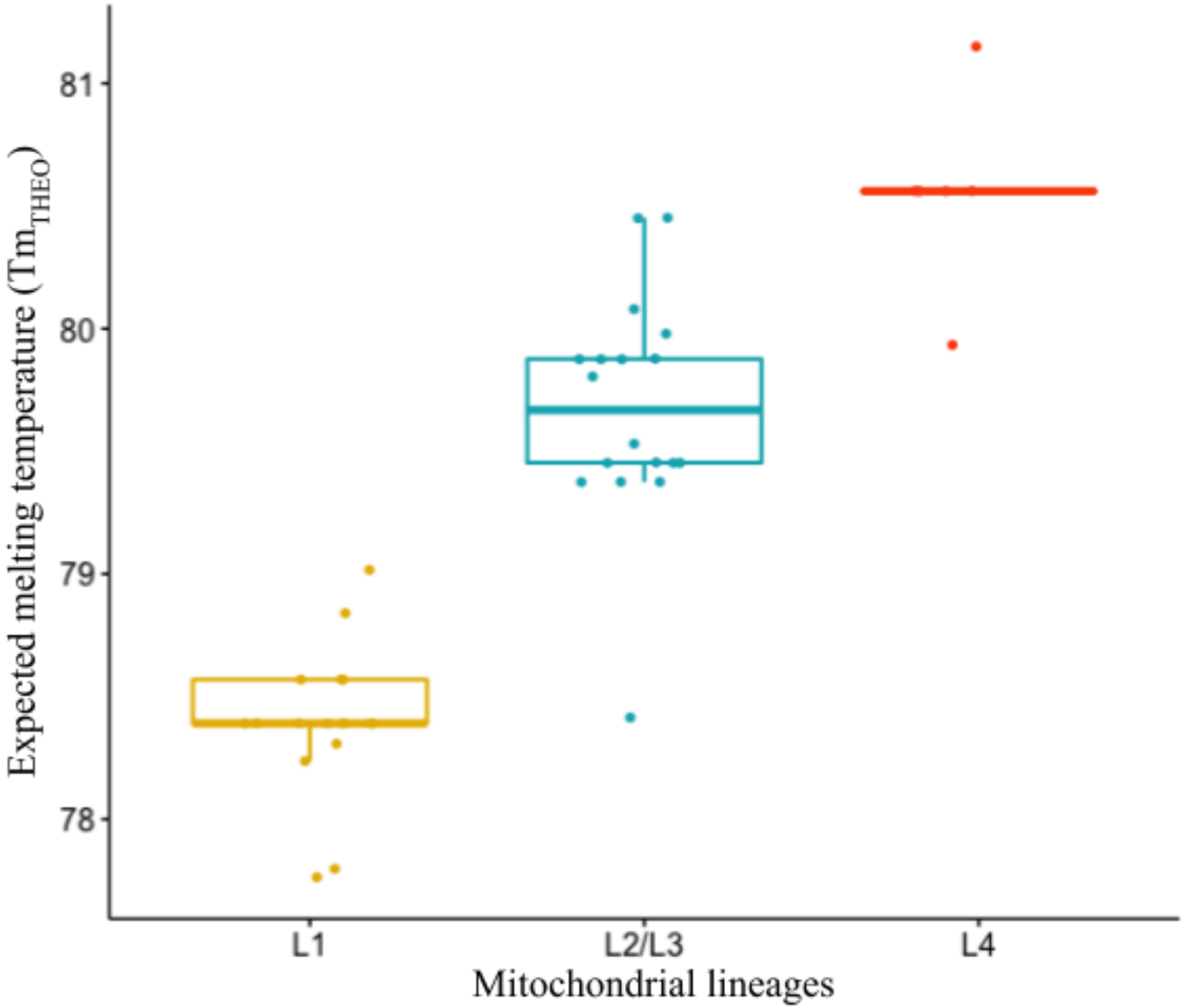
264 **Figure 3**: Relationship between expected melting temperatures (computed) and experimental
265 melting temperatures (measured by HRM) for all the individuals (reference panel individuals
266 and assigned individuals) with the assignment to the three species indicated by colours (L1
267 yellow, L2/L3 blue, L4 orange). The 2 non-matching control individuals sharing the same
268 haplotype are indicated in purple.

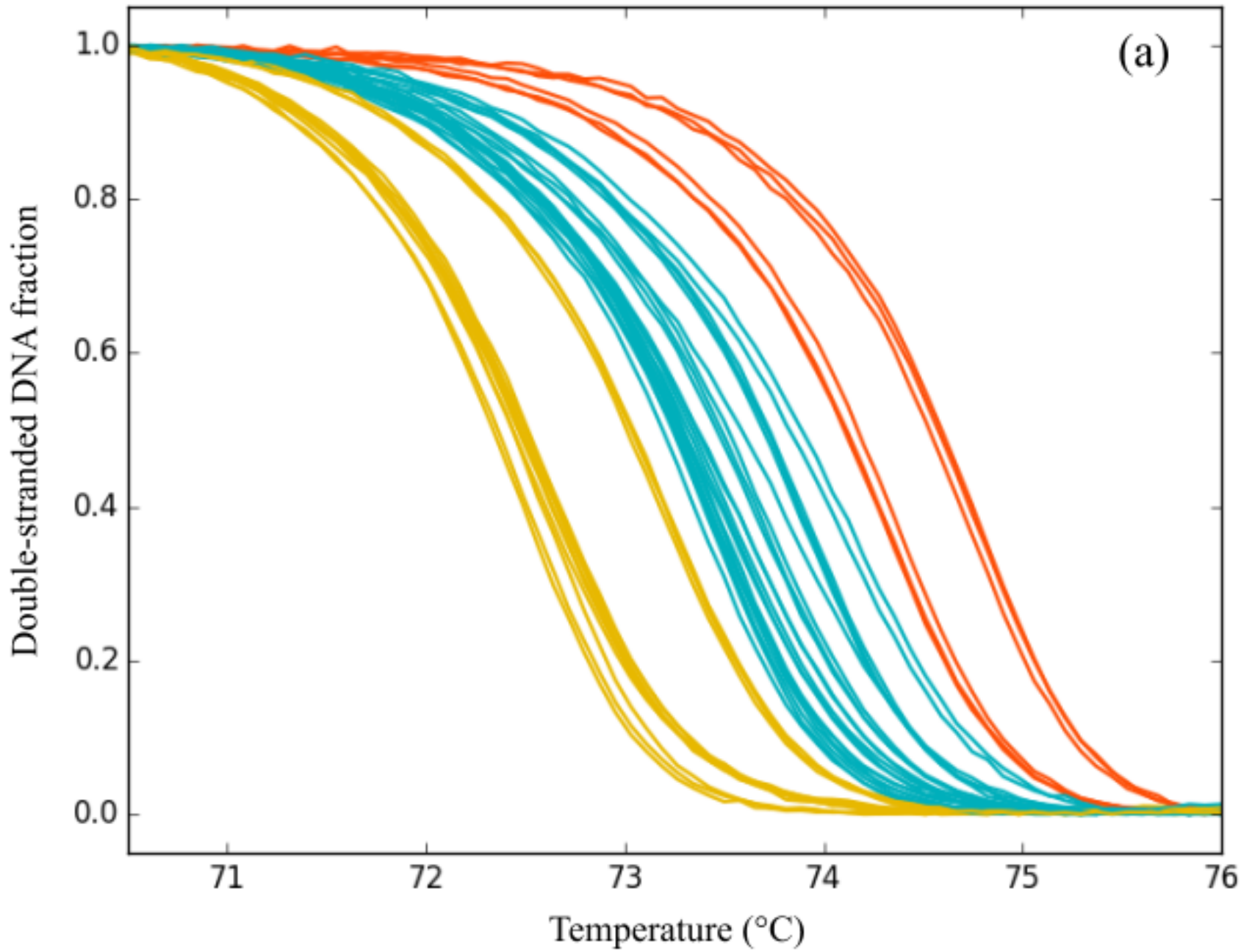
269

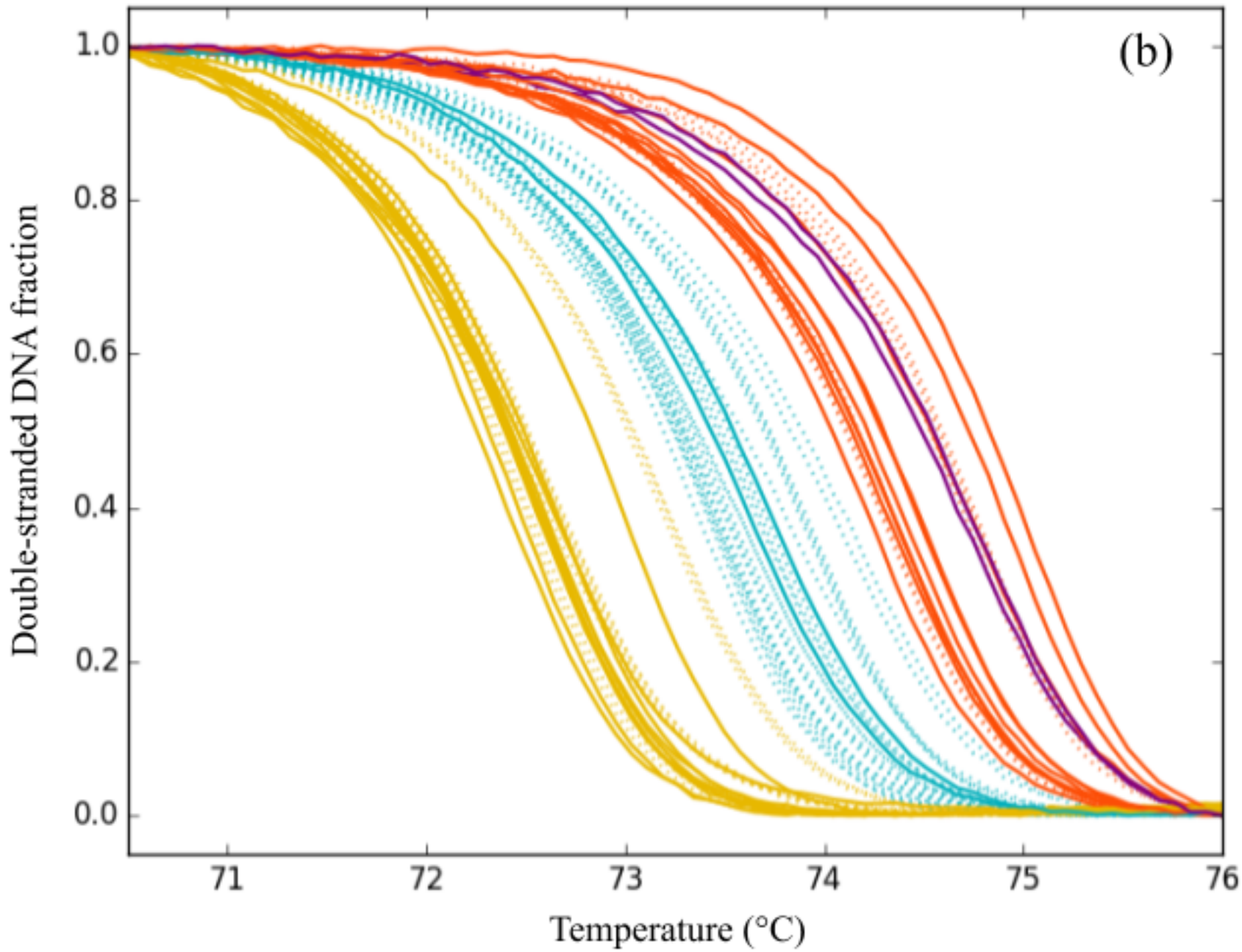
270

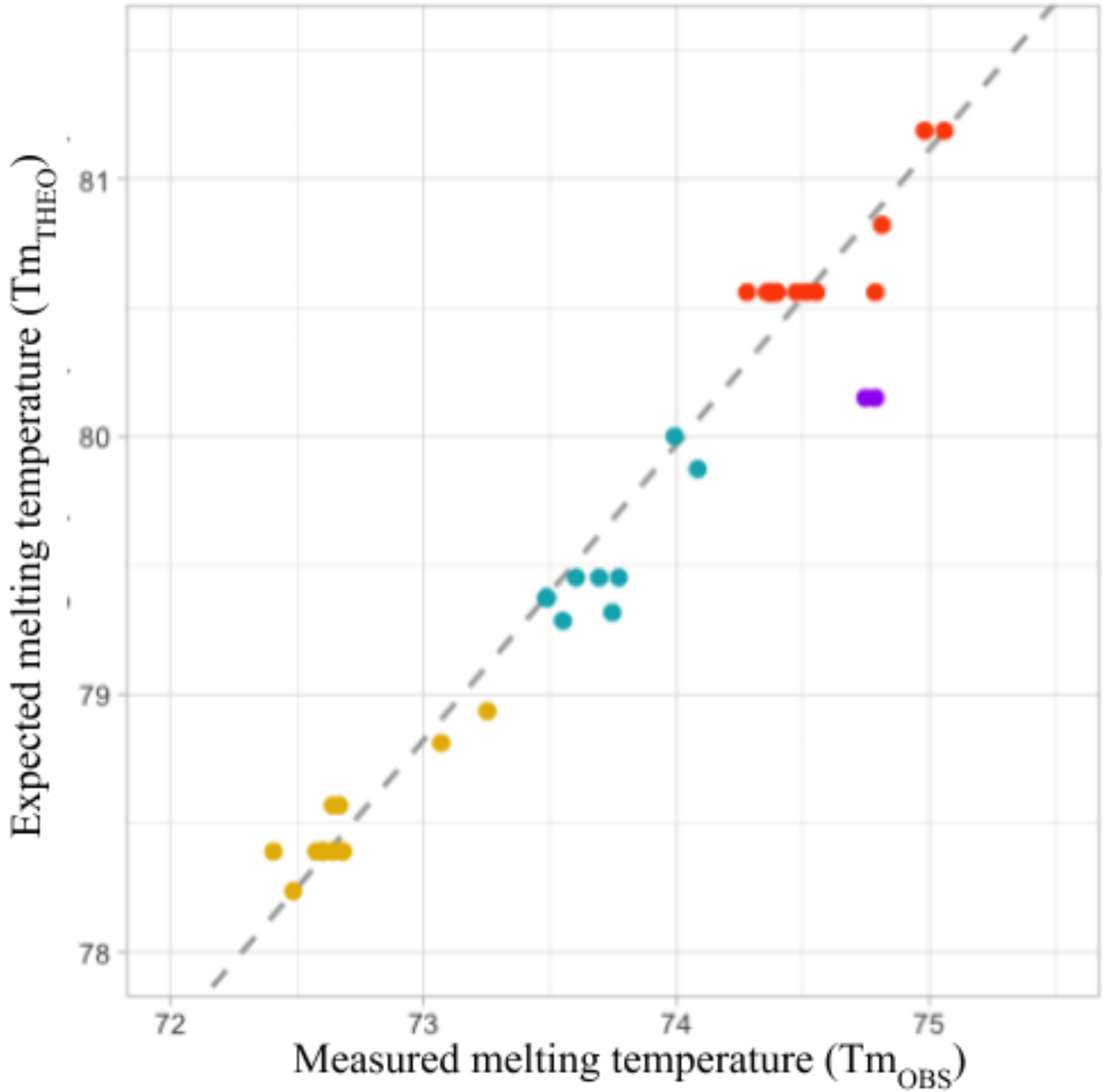
271

272


Expected melting temperature (Tm_{THEO})
81
80
79
78
L1
L2/L3
L4
Mitochondrial lineages


(a)
Double-stranded DNA fraction
1.0
0.8
0.6
0.4
0.2
0.0
71
72
73
74
75
76
Temperature (°C)


(b)
Double-stranded DNA fraction
1.0
0.8
0.6
0.4
0.2
0.0
71
72
73
74
75
76
Temperature (°C)


Expected melting temperature (Tm_{THEO})
81
80
79
78
72
73
74
75
Measured melting temperature (Tm_{OBS})