From accidental citizen-science observations to genetic confirmation: how to spot new hidden invaders

Article in Hystrix · October 2017
DOI: 10.4404/hystrix-28.2-12421

5 authors, including:

Emiliano Mori
Università degli Studi di Siena
130 PUBLICATIONS 420 CITATIONS

Andrea Sforzi
Maremma Natural History Museum
86 PUBLICATIONS 548 CITATIONS

Andrea Galimberti
Università degli Studi di Milano-Bicocca
84 PUBLICATIONS 1,057 CITATIONS

Some of the authors of this publication are also working on these related projects:

- Eco-ethology of osprey (Pandion haliaetus) in the Mediterranean basin View project
- Online repository of Italian mammals occurrence data View project
From accidental citizen-science observations to genetic confirmation: how to spot new hidden invaders

Emiliano Mori¹,*, Alberto Baeri¹, Andrea Sforzi², Alessandro Vitale³, Andrea Galimberti²

¹Dipartimento di Scienze della Vita, Università di Siena, Via P.A. Mattioli, 4 - 53100 Siena, Italy
²ZooPlantLab, Department of Biotechnology and Biosciences, University of Milano-Bicocca, P.zza della Scienza 2, 20126-I Milano, Italy
³Museo di Storia Naturale della Maremma, Strada Corsini, 5 – 58100 Grosseto, Italy

Abstract

The crested porcupine *Hystrix cristata* is the largest rodent in Italy, where it is strictly protected according to national and international laws. The species is almost exclusively nocturnal and quite elusive: its presence may be yet recorded through the detection of quills lost on the ground. In the last 40 years, the crested porcupine showed a remarkable range expansion in Italy. In 2013, in the framework of the www.naturaesocialmapping.it recording site, a web page devoted to *H. cristata* was set up. The aim of the project was to give people the opportunity to contribute to a continuous updating of the distributional map of this rodent. Lost quills were also collected by citizen-observers, as being easily available sources of DNA. This web page allowed us to detect the first evidence of the presence of another species of porcupine free-ranging in Italy (photos and quills): the Indian porcupine *H. indica*. Species identification was carried out through DNA barcoding analysis, since it cannot be performed only through quill morphology. We increased the international genetic database with sequences of both species, so to allow a rapid identification of samples through molecular analyses. Both species were introduced to Italy by humans. Our findings also support the relevance of citizen science contribution in the process of updating species’ distribution range and the potential of spotting new invasive species.

Keywords:

*Hystrix cristata*

*Hystis indica*

molecular identification

DNA barcoding

Acknowledgements

Authors thank two anonymous reviewers for their useful suggestions on an early draft

The systematic status and field identification of many rodent taxa are still controversial and debated, especially when species belonging to the same genus share similar morphological features and geographical range (Colangelo et al., 2009; Castiglia et al., 2016; Mazzamuto et al., 2016; Wauters et al., 2017). In most cases, samples (e.g., hairs, bones and faeces) found during field surveys could not provide sufficient information to achieve unequivocal identification. In this context, advances in molecular DNA-based tools represent a basic element to improve knowledge in taxonomy (Corti et al., 2005; Colangelo et al., 2009; Abiadad et al., 2010; Padial et al., 2010; Ermakov et al., 2015), that complement monitoring activities. Many investigations in the field of bioinvasions are currently conducted under an integrative vision, where variability of molecular markers is used in combination with morphological traits, geographical origin and other sources of information (Gotzek et al., 2012; Pisano et al., 2013; Mazzamuto et al., 2016). In the last years, DNA barcoding triggered a real revolution inside species identification, with applications ranging from pure taxonomy to cases of conservation or forensics concern (Hebert et al., 2003) This identification approach has been successfully applied on a vast plethora of case studies involving mammals (Galimberti et al., 2015). Publicly accessible archives including reference DNA barcoding sequences (targeting especially at region of the mitochondrial COI) are growing, thus offering the possibility to confirm doubtful field observations, even using non-invasive sampling approaches.

Porcupines are nocturnal, elusive rodents and most sightings are related to observations of (i) quills lost on the ground, (ii) dens and (iii) diggings on the ground to get food (Mori et al., 2013). The subgenus *Hystrix* includes three genetically related species of porcupines, distributed throughout Africa, Southern Europe and Asia with areas of almost complete allopatry (Mohr, 1965; Walid, 2011; Trucchi et al., 2016). The crested porcupine *H. cristata* L., 1758 shows a discontinuous geographical range. It is distributed in a portion of Southern Europe (peninsular Italy and Sicily, introduced in Sardinia), Northern Africa (from Morocco to Libya) and sub-Saharan Africa (from Senegal to Tanzania). In Central and Southern Tanzania, the range of *H. cristata* partly overlaps with the range of the Cape porcupine *H. africaeaustralis* Peters, 1852 (Skinner and Smithers, 1990; Barthelmess, 2006). The Indian porcupine *H. indica* Kerr, 1792 ranges from Turkey and Israel to Western India and Nepal (Mohr, 1965). These species are morphologically very similar, and very difficult to set apart when syntopic. Main diagnostic features are summarized in Tab. 1 (Corbet and Jones, 1956; Barthelmess, 2006; Aulagnier et al., 2010).

The range expansion of the crested porcupine in Italy has been recorded through a citizen-science approach, still ongoing. In 2013, in the framework of the www.naturaesocialmapping.it recording site, a web page devoted to *H. cristata* was set up. The aim of the project was to give people the opportunity to contribute to a continuous updating of the distributional map of this rodent. This approach has been proven to be very successful, as the species is unmistakable, with just a 0.03% of identification error. In September 2016, an individual showing the outer morphology of an Indian porcupine, i.e. a black crest, was accidentally recorded in the surroundings of a den in Abbadia di Fiastra (Province of Macerata, Central Italy), representing the first formal observation of this species in the wild outside its natural range. Quills were collected and genetically analyzed through DNA barcoding for the identification of the species.

We therefore produced the first reference DNA barcoding records for *H. cristata*. DNA barcoding sequences of *Hystric cristata* were not included in the international GenBank (https://www.ncbi.nlm.nih.gov) and the Barcode of Life Data System (http://www.boldsystems.org/).
whereas only four COI sequences of *H. indica* were available (databases last accessed on March 2017). Thus, a robust reference DNA barcoding dataset was created for both *Hystrix* species to achieve a reliable identification of the samples found in Central Italy (hereafter referred to as “Macerata”). A total of eight samples belonging to *H. cristata, H. indica* and the quills from the individual found at Abbadia di Fiastra, were collected (Tab. 2). Reference sequences belonged to individuals of *Hystrix cristata* collected in Italy (captures or road-killed), and to individuals of *Hystrix indica* from Turkey and Israel (road-killed). To maximize the chance of observing intraspecific geographic variation, conspecifics were sampled from distant sites (Table 2). Samples were voucherized and then stored in 99% ethanol at -20°C.

Total DNA was extracted from 25 mg of tissue or from the basal portion (4 mm) of three quills by using the DNeasy Blood & Tissue Kit (Qiagen, Milan, Italy) following manufacturer’s instructions with minor modifications in the case of quills. In particular, lysis was performed overnight and final DNA was eluted in 40 µl of deionized sterile water warmed up to 65°C. Purified DNA concentration of each sample was estimated fluorometrically with a NanoDrop™ 1000 Spectrophotometer (Thermo Scientific, USA). DNA barcoding characterization of samples was conducted by comparing nucleotide sequence differences at the standard barcode region for metazoans (i.e., 648 bp at the 5’ end of the mitochondrial *COI*; Herz et al. 2003), following the same procedures described in Mazzamuto et al. 2016. Sequence data were submitted to the European Bioinformatics Institute of the European Molecular Biology Laboratory (EMBL-EBI). In addition to the samples analyzed in this study, we considered the four DNA barcoding sequences of *H. indica* publicly available in GenBank and an outgroup *COI* sequence of *Atherurus africanus* (a.n. KJ192744). Sequences were aligned using the ClustalW option implemented in Bioedit (Hall, 1999) with default options. The taxonomic status of the Macerata quills samples was also tested by comparing the COI sequence with reference DNA barcoding stored in the Barcode of Life Database using the Identification Engine tool (IDS) (http://www.boldsystems.org/index.php/IDS_OpenIdEngine; Species Level Barcode Records database), which returns unique species assignments based on 99% sequence similarity of the barcode sequence. Considering the comprehensive dataset, average K2P genetic sequence divergences (and relative standard errors, SE) between and within distinct lineages were calculated and a NJ reconstruction was performed using MEGA 6 with the same settings described in Galimberti et al., 2012.

DNA extracted from the ethanol preserved tissues and quills was of high quality (ratios of absorbance, A_{260/280} and A_{260/230} > 1.80 and >1.90, respectively) and provided good yields (∼30 ng/µl). Amplification with the selected primer pairs was successful and resulting DNA concentration of purified amplicons was ∼50 ng/µl. High quality sequences showing a strong chromatogram signal along the entire read were obtained for all the individuals sampled for the present study. Due to different lengths of GenBank sequences, we trimmed the alignment to the same final size of 645 bp. The COI DNA barcoding dataset contained sequences with no insertion/deletions (indels), stop codons or biased by NUMT (sensu Bensasson et al. 2001) interference. The alignment was characterized by 64 variable positions, of which 62 were parsimony-informative. Despite different sampling localities, all samples of *H. cristata* shared the same haplotype (HAP1), whereas, four haplotypes were found for *H. indica*, one of which shared by three Genbank entries (HAP5) and the other three exclusive to the remaining accessions (i.e., HAP6 from GenBank and HAP3 and HAP4 from the two samples analyzed in this study, including the sample of Macerata). The Macerata *Hystrix* sample showed an exclusive COI haplotype (HAP4). Sequences accessions and haplotypes are provided in Tab. 2. When COI sequence obtained from the Macerata sample was used as query in GenBank-NCBI and BOLD-IDS tool, both systems returned a reliable match (i.e., similarity % > 99%) with reference DNA barcoding sequences of *H. indica*. The analysis of the K2P distance matrix (Table S1) revealed a moderate intraspecific divergence among lineages at the *COI* locus for *H. indica* (mean K2P distance ± standard error: 1.25% ± 0.31%, range: 0 - 2.06%) and a large interspecific divergence between *H. cristata* and *H. indica* (mean K2P distance ± standard error: 9.38% ± 1.26%, range: 9.06 - 9.80%). Despite the investigation of the other five *Hystrix* species should be recommended to assess the actual genetic distances, the existence of a clear barcoding gap (Wiemers and Fiedlers, 2007) may confirm the efficacy of DNA barcoding in unambiguously identifying *Hystrix* species.

The Macerata *Hystrix* sample had HAP2 and HAP6 haplotypes belonging to *H. indica* as its nearest neighbors (K2P distance ± standard error: 0.16% ± 0.15%), whereas a much higher genetic divergence occurred between the Macerata sample and *H. cristata* haplotypes (K2P distance ± standard error: 8.88% ± 2.23%).

The NJ reconstruction (Fig. 1) better summarized the pattern of molecular variability and unequivocally supported the assignment of the Macerata quill sample to the species *H. indica*. In our work, we confirmed the efficacy of a standard molecular method, which allows for a rapid identification of porcupine species where they occur syntopically. The high molecular divergence at the barcode locus *COI* between the two *Hystrix* species (i.e. > 8% K2P) makes DNA barcoding approach reliable to assess the occurrence of the invasive taxon also starting from samples retrieved from non-invasive approaches. Such a molecular divergence is comparable to values obtained in previous studies targeting at Cyr b (K2P distance 10.38%) and D-loop (K2P 8.22%—9.21%) mitochondrial regions (Trucchi et al., 2008, 2009). However, such values are biased by the analysis of a single *H. indica* sample (compared to the 7 used in our DNA barcoding approach) and further investigations are needed to better support genetic distances at these markers. In addition, the high number of nucleotide differences at the barcode locus between *H. cristata* and *H. indica* permits the development of fast (and sequencing free) detection

---

Table 1 – Main diagnostic features in the subgenus *Hystrix*. MAR, midline along the rump; NB, nasal bones; ONL, occipito-nasal length; P, premaxilla; PW, premaxilla width; NBW, nasal bone width (from Corbet and Jones, 1965; Barthelmes, 2006; Audigier et al., 2010).

<table>
<thead>
<tr>
<th>Species</th>
<th>Colour of the crest</th>
<th>Colour of MAR</th>
<th>NB (% of ONL)</th>
<th>P (% of PW/NBW)</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Hystrix cristata</em></td>
<td>predominantly white</td>
<td>black or mottled</td>
<td>long (58-68)</td>
<td>intermediate (&lt; 36)</td>
</tr>
<tr>
<td><em>Hystrix indica</em></td>
<td>predominantly brown</td>
<td>white</td>
<td>short (45-50)</td>
<td>wide (&gt; 44)</td>
</tr>
<tr>
<td><em>Hystrix acaustra</em></td>
<td>predominantly white</td>
<td>white</td>
<td>intermediate (51-58)</td>
<td>narrow (&lt; 23)</td>
</tr>
</tbody>
</table>

---

Figure 1 – Neighbour joining tree based on COI sequences of *Hystrix cristata* and *H. indica*. For each samples, voucher or GenBank accession number is provided (further details can be retrieved from Tab. 2). Bootstrap support (1000 replicates) values > 90% are indicated above the nodes. The filled circle shows the Macerata quill sample. A colour version of this figure can be viewed in the online version of this article.
approaches based, for example, on High Resolution Melting through Real-Time PCR (Ramón-Laca et al., 2014). The conservation status of the crested porcupine represents a European paradox. Despite most evidences suggest that it might have been introduced to Italy in Early mediaeval times (Bertolino et al., 2017; Trucchi et al., 2016), a few dissident voices suggesting a native origin are still present (see Bertolino et al. (2017) for a review). This species is listed in Berne Convention (all II) and Habitat Directive (all IV), as well as protected by the Italian law since 1978 (Italian National Law n. 968/1977). Starting from the 1970s, crested porcupines have undergone a remarkable range expansion in Italy (Mori et al., 2013). In the province of Macerata, where the H. indica sample was found, the population of porcupines started to become consistent and stable after 1976 (for the current distribution see Supplemental Figure S2). Our findings showed for the first time the presence of free-ranging Indian porcupine in this area of Central Italy. This species is quite common in Zoological Gardens and other Wildlife Centres (at least in 4 out of 23 structures surveyed in Italy). Hybridization between Asiatic and African porcupine species has been recorded in captivity (Mohr, 1965), so we cannot exclude that hybridization with locally escaped individuals might occur. In the Sinai peninsula, the opposite situation has occurred, with the presence of potentially introduced H. cristata overlapping with native, rare H. indica (Saleh and Basuony, 1998; Trucchi et al., 2016). The skull morphology of three individuals collected in this area (Saleh and Basuony, 1998), showed intermediate features between H. cristata and H. indica, therefore suggesting a possible hybridization which needs to be confirmed. A further step of monitoring would require the development of DNA-based tools, such as microsatellite or SNPs markers to properly distinguish possible hybrids from their parental species (Randi et al., 2001; Randi and Lucchini, 2002).

The management of the crested porcupine in Italy is controversial. As the species is widely poached for its meat (Lovari et al., 2016), and it is considered to be responsible for crop and riverbank damages (Bertolino et al., 2017; Laurenzi et al., 2016). This is somehow in contrast with the level of legal protection deserved to the species. Management strategies would necessarily change if a different porcupine species, surely introduced, would similarly affect human economy. We have no current evidence on the number of free-ranging H. indica individuals in the Macerata area. Immediate actions for the monitoring of occurrence distribution and breeding evidences of the species is necessary to control its potential expansion. Further attention by government ad international agencies should be deserved to importation and detection of potentially dangerous animal species in zoological gardens and wildlife centres. We recommend to continue the monitoring of the range expansion of the crested porcupine in Italy and the collection of quills (especially where doubtful individuals are observed) for DNA analyses.

### References


How to spot new hidden invaders

### Table 2 – Reference samples of H. cristata and H. indica with provenance, coordinates, voucher codes, haplotype and GenBank accession numbers.

<table>
<thead>
<tr>
<th>Specimen Voucher</th>
<th>Species name</th>
<th>Sample</th>
<th>Place of origin</th>
<th>Country</th>
<th>Haplotype</th>
<th>GenBank a.n.</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIB:ZPL:0776</td>
<td>Hystrix cristata</td>
<td>Tissue</td>
<td>Prata, Grosseto, Tuscany</td>
<td>Italy</td>
<td>HAP1</td>
<td>LT476352</td>
<td>this study</td>
</tr>
<tr>
<td>MIB:ZPL:0777</td>
<td>Hystrix cristata</td>
<td>Quills</td>
<td>Monti Dauni, Foggia, Apulia</td>
<td>Italy</td>
<td>HAP1</td>
<td>LT476353</td>
<td>this study</td>
</tr>
<tr>
<td>MIB:ZPL:0778</td>
<td>Hystrix cristata</td>
<td>Quills</td>
<td>Monti Lepini, Rome, Latium</td>
<td>Italy</td>
<td>HAP1</td>
<td>LT476354</td>
<td>this study</td>
</tr>
<tr>
<td>MIB:ZPL:0779</td>
<td>Hystrix cristata</td>
<td>Quills</td>
<td>Velezzo Lomellina, Pavia, Lombardy</td>
<td>Italy</td>
<td>HAP1</td>
<td>LT476355</td>
<td>this study</td>
</tr>
<tr>
<td>MIB:ZPL:0780</td>
<td>Hystrix cristata</td>
<td>Quills</td>
<td>Saluzzo, Cuneo, Piedmont</td>
<td>Italy</td>
<td>HAP1</td>
<td>LT476356</td>
<td>this study</td>
</tr>
<tr>
<td>MIB:ZPL:0781</td>
<td>Hystrix cristata</td>
<td>Quills</td>
<td>Torrebelvicino, Vicenza, Veneto</td>
<td>Italy</td>
<td>HAP1</td>
<td>LT476357</td>
<td>this study</td>
</tr>
<tr>
<td>MIB:ZPL:0782</td>
<td>Hystrix indica</td>
<td>Quills</td>
<td>Hula Valley</td>
<td>Israel</td>
<td>HAP2</td>
<td>LT476358</td>
<td>this study</td>
</tr>
<tr>
<td>MIB:ZPL:0783</td>
<td>Hystrix indica</td>
<td>Tissue</td>
<td>Adiyaman</td>
<td>Turkey</td>
<td>HAP3</td>
<td>LT476359</td>
<td>this study</td>
</tr>
<tr>
<td>WF-RFDF-11</td>
<td>Hystrix indica</td>
<td>–</td>
<td>–</td>
<td>India</td>
<td>HAP6</td>
<td>JN171477</td>
<td>GenBank</td>
</tr>
<tr>
<td>WF-RFDF-15-1</td>
<td>Hystrix indica</td>
<td>–</td>
<td>–</td>
<td>India</td>
<td>HAP5</td>
<td>JN171482</td>
<td>GenBank</td>
</tr>
<tr>
<td>WF-RFDF-15-2</td>
<td>Hystrix indica</td>
<td>–</td>
<td>–</td>
<td>India</td>
<td>HAP5</td>
<td>JN171483</td>
<td>GenBank</td>
</tr>
<tr>
<td>WF-RFDF-15-3</td>
<td>Hystrix indica</td>
<td>–</td>
<td>–</td>
<td>India</td>
<td>HAP5</td>
<td>JN171484</td>
<td>GenBank</td>
</tr>
<tr>
<td>MIB:ZPL:0784</td>
<td>Hystrix sp.</td>
<td>Quills</td>
<td>Abbadia di Fiastra, Macerata, Marche</td>
<td>Italy</td>
<td>HAP4</td>
<td>LT476360</td>
<td>this study</td>
</tr>
</tbody>
</table>


Supplemental information
Additional Supplemental Information may be found in the online version of this article:
Table S1 Analysis of the K2P distance matrix.
Figure S2 Distribution of the crested porcupine in Central Italy.