

Haplotypes, median networks, and diagnostic characters as tools to elucidate the intraspecific genetic and taxonomic structure of bumblebees, with a description of *Bombus cryptarum pallidocinctus* new subspecies (Hymenoptera: Apidae: *Bombus*)

With 5 figures and 2 tables

ANDREAS BERTSCH¹, GERHARD K. H. PRZEMECK² and MARTIN HRABÉ DE ANGELIS²

¹ Department of Biology, Philipps-University Marburg, Karl-von-Frisch Straße 8, 35032 Marburg, Germany – bertsch@staff.uni-marburg.de

² Institute of Experimental Genetics, Helmholtz-Zentrum München, Ingoldstädter Landstrasse 1, 85764 Neuherberg – przemeck@helmholtz-muenchen.de, hrabe@helmholtz-muenchen.de

Published on 2014-12-15

Summary

An alignment of 168 sequences of mitochondrial DNA of *Bombus cryptarum* reveals 29 haplotypes. A median network shows five distinct clusters, haplogroups (HG) A-E, which can be defined by diagnostic characters. Four of these HGs can be assigned to well-known taxa: HG-A = *B. cryptarum cryptarum*; HG-B = *B. cryptarum florilegus*; HG-D = *B. cryptarum albocinctus*; and HG-E = *B. cryptarum moderatus*. Haplogroup C represents a new taxon *B. cryptarum pallidocinctus* in the rank of a subspecies. This new subspecies is described. The geographic distribution of the Eurasiatic taxa is described and the implications of geographic distribution and median network for the phylogeny are discussed.

Zusammenfassung

Ein Alignment von 168 Sequenzen mitochondrialer DNA von *Bombus cryptarum* liefert 29 Haplotypen. Ein Median-Network zeigt fünf klar abgrenzbare Cluster, die Haplogruppen (HG) A-E, die durch diagnostische Positionen definiert werden. Vier dieser HG lassen sich bekannten Taxa zuordnen HG-A = *B. cryptarum cryptarum*, HG-B = *B. cryptarum florilegus*, HG-D = *B. cryptarum albocinctus*, und HG-E = *B. cryptarum moderatus*. Die Haplogruppe C entspricht einem neuen Taxon *B. cryptarum pallidocinctus* im Rang einer Unterart. Diese neue Unterart wird beschrieben. Die Verbreitung der Eurasiatischen Taxa wird untersucht und die Folgerungen aus Median-Network und geographischer Verbreitung für die mögliche Phylogenie werden diskutiert.

Key words

Hymenoptera, Apidae, *Bombus*, mitochondrial DNA, new subspecies, geographic distribution.

Introduction

Since the publications of PEDERSEN (1996, 2002), KAWAKITA et al. (2004) and CAMERON et al. (2007) we have a quite-reliable picture of the phylogenetic relationships of most bumblebee species. Recent comprehensive surveys for the subgenus *Bombus* sensu stricto are available (BERTSCH 2010; WILLIAMS et al. 2012). However, the main point of view of these investigations always was the definition and identification of 'species'. Since the problematic identification of critical species by difficult morphological characters could be replaced by the use of mtDNA sequences of cytochrome oxidase I (COI), a safe identification of specimens is possible. But although 'species' are important reference points and significant episodes in the course of evolution, with intensive collecting and the study of a large number of individuals from a wide spread of geographic proveniences, the centre of interest is gradually passing from the systematist's 'species' to the 'natural population' from which the species is abstracted. Two recent examples with a detailed analysis of the haplotype structure of the populations are DUENNES et al. (2012) for the Central-American *Bombus* (*Pyrobombus*) *ephippiatus* and LECOCQ et al. (2013) for the European *Bombus* (*Melanobombus*) *lapidarius*. In this field probably lies the most promising potential for the DNA-based research on bumblebees: the possibility to study the intraspecific genetic and taxonomic variability of populations.

In his seminal publication 'Les bourdons du genre *Bombus* LATREILLE sensu stricto en Europe Occidentale et Centrale', RASMONT (1984) refers to an unnamed subspecies of *B. (Bombus) cryptarum* FABRICIUS 1775 with a question mark, obviously unsure about this taxon, based on 2 worker specimens from Finland. In PAMILO et al. (1997), we find the remark that males of *B. cryptarum* from Finland did not show the morphological characters developed by RASMONT (1984) for specimens from Belgium. A strange COI sequence from an unidentified specimen from Norway was published by PEDERSEN (2002, Genbank AY181116), which, after re-sequencing, proved to be a sequence of *B. cryptarum* (see discussion 'detecting new taxa?' BERTSCH 2009, p. 302) and was identical with a sequence from St. Petersburg (Genbank EF362729) published in MURRAY et al. (2008). CAROLAN et al. (2012) investigated polymorphic sequences of *B. cryptarum* and detected 2 differing haplotypes. Based on three sequences from Southern Finland a *cryptarum-haplotype 2* was described (Genbank JN872566, JN872567, JN872568), identical with the sequences published by PEDERSEN 2002 and BERTSCH 2009 from Norway and MURRAY et al. 2008 from Russia. Meanwhile 10 more sequences of this *cryptarum-haplotype 2* are available at Genbank (Blast search 20.ii 2014), with a geographical spread from Norway to Western Siberia (see appendix Table 2). A detailed analysis of the intraspecific structure of populations may help to understand these findings. *B. cryptarum* with a circum-

polar distribution has proven to be especially genetically variable (BERTSCH et al. 2010); COI sequences of specimens from Eurasia and North America will be taken as a case study to test the potential of the genetic information available and to investigate what a median network of haplotypes and their geographical distribution may tell us about the intraspecific genetic and taxonomic structure and recent history of *B. cryptarum*.

Material & Methods

Polymerase chain reaction (PCR) and DNA sequencing of mitochondrial COI

Total DNA was extracted from legs using the QIAamp® DNA Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's specifications for tissue (for details see BERTSCH et al. 2010). For sequence analysis five overlapping fragments of mitochondrial COI were amplified (for primer pairs see appendix Table 1). The PCR products were purified using an AMPure® PCR Purification Kit (Agencourt, Beverly, MA, USA). Sequencing reactions were performed using ABI® BigDye Terminator version 3.1 chemistry (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's instructions; all fragments were sequenced from both strands and then analysed on an ABI 3100 sequencer (Applied Biosystems). The ABI trace files were edited with 4PEAKS (from MekenTosj) and aligned manually using CLUSTALX. To produce an equal sequence length for all individuals, sequences were trimmed to 1533 bp (encoding 511 amino acids). As COI sequences for bumblebees from different authors are of different length and starting points a comparison of positions is often difficult. To enable a precise referencing of distinct positions and an unambiguous comparison of sequences from different sources we aligned the sequences with the complete COI gene of *B. ignitus* (position 1643-3202 in GenBank DQ870926) (CHA et al. 2007); the first position of the start codon ATA is POSITION 1 and the last position of the stop codon TAA is position 1560. All other positions are referenced accordingly.

The Bold Project with a sequence length of 658 nt is a minimalistic approach, originally created as a *novel system for rapid and accurate species identification* (HEBERT & GREGORY 2005). In this respect, the project has been very successful, and a large amount of data is now available for analysis (about 4800 sequences for the genus *Bombus*, 1370 thereof public). Since the interpretation of results is more and more expanding from the primary purpose of species identification to questions of relationships and phylogeny, the question arises whether this minimalistic approach is adequate for such far-reaching problems. As already discussed in respect to degraded sequences from type material (BERTSCH 2009, p. 303) sequences of

greater length produce more information and give more security into conclusions derived by such an analysis.

A total of 168 sequences (Bold length 658 nt) from Eurasian *B. cryptarum* (available at GenBank, at Bold-Public-Data-Portal and unpublished material from our laboratory) has been used for an analysis of genetic population structure; they will be referenced as Group-1 sequences. By using six pairs of primers, we produced overlapping partial sequences, which can be combined to nearly full-length (1533 of 1560 nt) sequences of the COI gene. In all, 64 such full-length sequences are currently available; they will be referenced as Group-2 sequences.

Analysis of sequence divergence, haplotype frequency and calculation of median network

The absolute numbers of substitutions were counted based on a pairwise comparison of COI sequences. The analysis of the genetic distance of sequences was performed using Mega 4.0 (TAMURA et al. 2007; KUMAR et al. 2008). The aligned data matrix of 168 sequences was collapsed to 29 haplotypes (HT) using DnaSPv5, a program distributed by LIBRADO & ROZAS (2009). Median-joining haplotype networks (BANDELT et al. 1999), both with and without maximum-parsimony-post-processing (MARDULYN 2012) were calculated, with $\epsilon = 0$ and all variable sites weighed equally using NETWORK, a program distributed by Fluxus-Technology. Median-joining networks have been recommended over maximum parsimony approaches in intraspecific studies as they capture a greater degree of ambiguity, thus enabling more realistic interpretations. Diagnostic characters were detected by visual inspection and checked by MACCLADE version. 4.08 (MADDISON & MADDISON 2002).

Results

DNA polymorphism, haplotypes, and median network

The aligned data matrix of 168 GROUP-1 sequences included 35 variable sites. One of these variable positions was parsimony uninformative (noise) and 34 variable positions parsimony informative (signal). The nucleotide frequencies were 34.00 (A), 42.05 (T), 11.77 (C) and 12.18 (G), which proves the known strong A + T bias typical for sequences of Hymenoptera. The data set comprised 29 haplotypes, a haplotype diversity (Hd) of 0.946 and a nucleotide diversity (π) of 0.01147. A median network (Fig. 1) is well suited to visualise the genetic structure for genetic variability of these 168 *B. cryptarum* sequences. It represents so to speak, a view from top at the last branches of the phylogenetic tree visualising the lumping of clades for the present-time stadium of evolution.

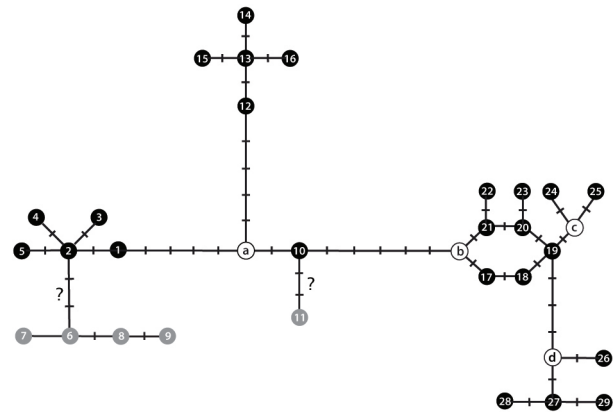


Fig. 1: Median network for 168 GROUP-1 length sequences of *B. cryptarum* (black circles with numbers 1-29 = haplotypes), slashes across network branches = separating mutations, white circles a-d = ancient nodes. ? = deficiency of data.

Taxonomic units defined by genetic divergence

The haplotypes of a median network may be grouped by *coherence* or by *divergence*. Structuring populations into HAPLOGROUPS (groups of haplotypes) relies on genetic coherence: all haplotypes connected by steps of one mutation cluster to a haplogroup. Recently it has become fashionable to use MOTUS (Molecular Operational Taxonomic Units), which rely on genetic divergence: all haplotypes separated by a certain cut-off value for genetic divergence belong to a MOTU. With our 168 sequences in search of intraspecific taxa of *B. cryptarum* things are quite obvious, groups of haplotypes are separated from the next cluster by at p-distance of about 0,006 (cut-off 0,6 %), both approaches result in five distinct groups of haplotypes (HG A-E) (Fig. 2). Haplotypes 6-9 (Turkey, Iran, and Kyrgyzstan) and haplotype 11 (Etorofu/Iturup Kuril Islands) do not fit into this procedure; they are separated by branches with two mutations. As only short GROUP-1 sequences are available, most probably longer GROUP-2 sequences will have increasing numbers of separating mutations, and both clades may merit treatment as separate haplogroups.

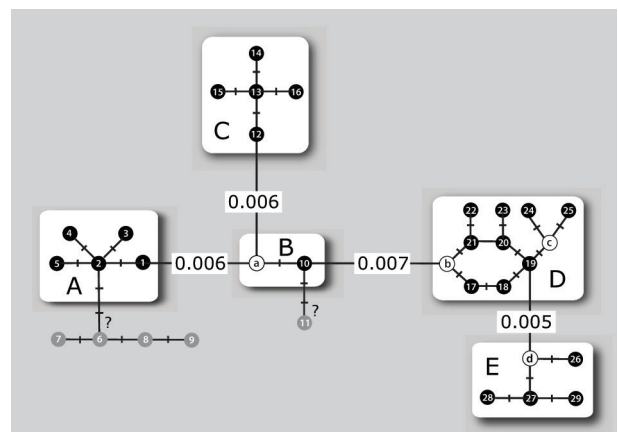


Fig. 2: Haplogroups A-E (slashes across network branches = separating mutations), Molecular Operational Taxonomic Units A-E with p-distances, respectively. ? = deficiency of data.

Taxonomic units defined by diagnostic characters

A definition of taxa by genetic coherence or divergence may be 'operational', but as classical taxonomy defines taxa by characters a straightforward linkage of genetic divergence and taxonomic characters is often problematic. As there are no gaps in the alignment of the COI sequences, single nucleotide sites can be used as positional homologs (HILLIS 1994; BROWER & SCHAWARROCH 1996), therefore, single nucleotide positions can be treated like characters and a direct comparison of genetics and taxonomies gets feasible (Fig. 3).

B. cryptarum is characterised by eight diagnostic characters (Fig. 3 green boxed), which separate this species from the closely related and morphologically similar species *B. magnus*, *B. patagiatus* and *B. lucorum* (comparison-group '*cryptarum-magnus-patagiatus-lucorum*'). Haplogroups A–E can be separated by private diagnostic characters (Fig. 3 red) or a combination of a few diagnostic positions. Since there is no ambiguity or overlap, Fig. 3 can be used as a 'key' to identify unknown specimens.

										<i>1</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>1</i>												
	<i>0</i>	<i>1</i>	<i>2</i>	<i>2</i>	<i>3</i>	<i>3</i>	<i>4</i>	<i>4</i>	<i>5</i>	<i>7</i>	<i>7</i>	<i>7</i>	<i>8</i>	<i>8</i>	<i>9</i>	<i>0</i>	<i>0</i>	<i>0</i>	<i>1</i>	<i>1</i>	<i>1</i>	<i>3</i>	<i>3</i>	<i>3</i>	<i>3</i>	<i>4</i>	<i>4</i>	<i>5</i>	<i>5</i>		
	<i>8</i>	<i>2</i>	<i>4</i>	<i>5</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>0</i>	<i>0</i>	<i>8</i>	<i>0</i>	<i>5</i>	<i>5</i>	<i>3</i>	<i>5</i>	<i>0</i>	<i>0</i>	<i>7</i>	<i>7</i>	<i>1</i>	<i>3</i>	<i>7</i>	<i>1</i>	<i>4</i>	<i>5</i>	<i>8</i>	<i>4</i>	<i>5</i>	<i>0</i>	<i>0</i>	
	<i>4</i>	<i>9</i>	<i>3</i>	<i>3</i>	<i>8</i>	<i>0</i>	<i>8</i>	<i>5</i>	<i>8</i>	<i>8</i>	<i>8</i>	<i>0</i>	<i>1</i>	<i>7</i>	<i>5</i>	<i>6</i>	<i>8</i>	<i>1</i>	<i>7</i>	<i>6</i>	<i>1</i>	<i>6</i>	<i>4</i>	<i>7</i>	<i>6</i>	<i>9</i>	<i>0</i>	<i>0</i>	<i>5</i>	<i>0</i>	<i>7</i>
A	T	A	T	T	C	T	T	T	C	C	A	G	T	T	T	T	T	T	C	C	T	C	T	T	G	C	T	C	G	G	
B	C	A	T	T	C	T	T	T	T	T	T	A	C	A	G	T	T	T	T	T	T	C	C	T	T	A	T	C	G	G	
C	C	A	C	T	C	T	T	T	T	T	T	A	C	A	A	T	C	C	T	T	T	C	T	T	T	A	T	C	G	G	
D	C	G	T	C	T	T	T	T	T	T	T	G	C	A	G	T	T	T	T	T	T	C	T	T	T	A	T	C	G	G	
E	C	G	T	C	T	T	T	T	T	T	G	C	A	G	C	T	T	T	T	T	T	C	T	T	T	A	T	C	G	G	
Mag	T	A	T	A	T	A	T	T	T	T	T	T	A	T	A	T	T	T	T	T	T	T	T	T	A	A	T	T	A	A	
Pat	T	A	T	C	T	A	T	T	T	T	T	T	A	T	A	T	T	T	T	T	T	T	T	T	A	A	T	T	A	A	
Luc	T	A	T	C	T	A	T	T	T	T	T	T	A	T	A	T	T	T	T	T	T	T	T	T	T	A	A	T	T	A	

Fig. 3: Variable and Diagnostic characters with position numbers (italics) for *B. cryptarum* (green boxed), and the five haplogroups A–E with private diagnostic characters (red). Taxa used for comparison include *B. magnus* (Mag), *B. patagiatus* (Pat), and *B. lucorum* (Luc).

Bombus subgenus Bombus

Bombus cryptarum
 253T 328C 330T 751C 1314C 1359T 1500C 1507G

A = cryptarum
 W-Europe, Denmark, Alps, Turkey, Iran,? Kyrgyzstan
 408C 588C 708C 837G 1116C 1131C 1356C 1380G 1440C

B = florilegus
 Japan (Hokkaido), Russia (Kuril Islands)
 1347C

C = new ssp.
 Scandinavia, Finland, European Russia, W-Siberia
 243C 348C 1008C 1071C 1077C 1455C

D = albocinctus
 Russia (Far East), China, Mongolia, Altai
 261C 723T

E = moderatus
 USA (Alaska), Northern Canada
 406C 906C 1176C

Bombus patagiatus patagiatus
 174G 336T 387A 490A 520C 561C 582C 639C 672G 975C 981C 1062T 1318A 1362A 1423G 1453T 1478G 1485C

Bombus magnus
 115C 147C 189G 253A 255T 312A 453T 546C 822A 964C 1515C

Bombus lucorum
 100C 174T 219C 343T 363T 381C 396A 414C 441C 540C 645C 666T 711G 763G 1104G 1140T 1167T

Fig. 4: Nested clades with diagnostic characters for species *B. cryptarum* and private diagnostic characters for subspecies = Haplogroups A–E (bold). Taxa used for comparison include *B. patagiatus*, *B. magnus*, and *B. lucorum*.

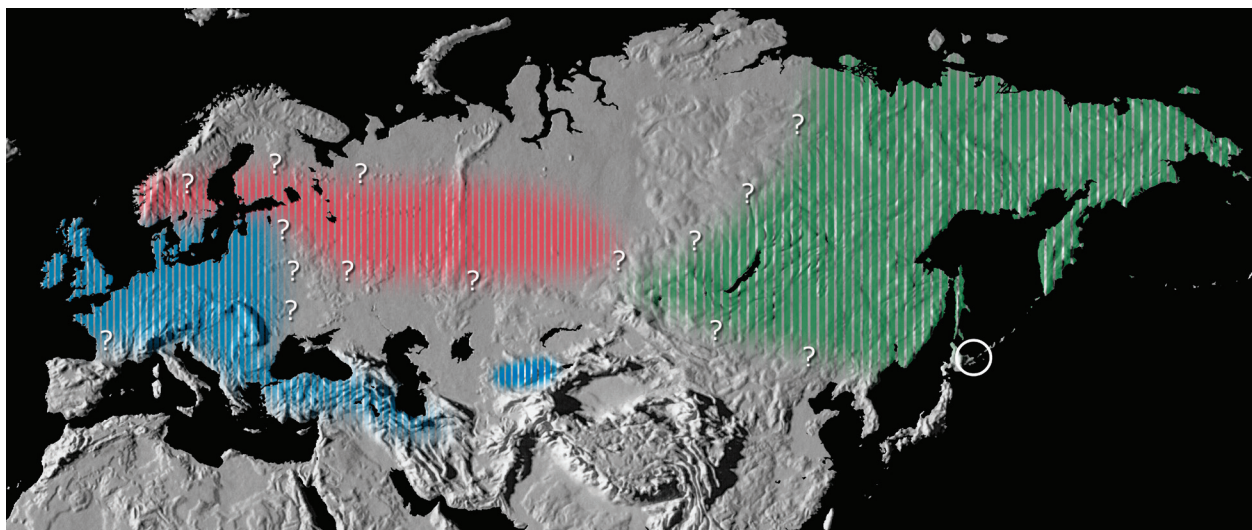


Fig. 5: Distribution of the 3 Eurasiatic subspecies of *B. cryptarum*. *B. cryptarum cryptarum* (blue), *B. cryptarum* new ssp. (orange), and *B. cryptarum albocinctus* (green). ? = Range borders and overlap uncertain. Circle = 'Refugio' of *B. cryptarum florilegus*.

A graphical display of nested clades (TEMPLETON et al. 1995; TEMPLETON 1998) representing the taxonomic hierarchy proves especially useful to visualise the complex order, which represents the genetic intraspecific structure and their relationships with taxonomy (Fig. 4).

Intraspecific taxonomy: geographic subspecies

All sequences from North America belong to Haplogroup E, which characterises *B. cryptarum moderatus* CRESSON 1863. All sequences from Hokkaido (specimens with a black ‘tail’) belong to Haplogroup B, which characterises *B. cryptarum florilegus* PANFILOV 1956. For the rest of Eurasia (Fig. 5), we have three different taxa: Haplogroup A characterises *B. cryptarum cryptarum* FABRICIUS 1775 from Western Europe (originally described as a dark form of *B. lucorum*), Haplogroup D characterises *B. cryptarum albocinctus* SMITH 1854 from the Russian Far East, China, and Mongolia (originally described as a light form of *B. lucorum* from Kamchatka), and Haplogroup C is a thus-far unrecognised taxon with the rank of a subspecies distributed from Scandinavia, Finland and European Russia to Western Siberia. Instead of two taxa distributed in Eurasia (see distribution map Fig. 1 BERTSCH et al. 2010), that is *B. cryptarum cryptarum* at the Atlantic Western Part (with undetermined eastern-distribution border) and *B. cryptarum albocinctus* at the Pacific Eastern Part (with undetermined western-distribution border) we have now to deal with three taxa (Fig. 5) with unknown borders and areas of overlap. In the northern plains of the Eurasian landmass, the new subspecies has a distribution separating the western *B. cryptarum cryptarum* from the eastern *B. cryptarum albocinctus*, whereas in the southern mountains of Kyrgyzstan and the Russian Altai *B. cryptarum cryptarum* and *B. cryptarum albocinctus* may come into contact.

Description of B. cryptarum pallidocinctus new subspecies

Type material: Holotype: 1 ♀, Labels: printed ‘Finland, Nordsavo, Kuopio, Puijo Hill, 14.vi.2001 A Bertsch’. Genbank KJ668128. Additional material. 1 ♀, Labels: printed ‘Russia, St. Petersburg, Smolensk cemetery 13.v.1993 A. Bertsch, Genbank KJ668126’; 1 ♀, Labels: printed ‘Russia, St. Petersburg, Pavlovo 14.vi.1993 A. Bertsch, Genbank KJ668127’; 1 ♀, Labels: printed ‘Russia, Moscow, Botanical Garden Academy of Sciences, 17.vi.1993 A. Bertsch, Genbank KJ668129’. Deposited at the SDEI (Senckenberg Deutsches Entomologisches Institut).

Description: Diagnostic characters of mtDNA, position numbers referenced to COI gene, NUMBER 1 = first Nucleotide of the Start codon ATA at position 1643 of the complete mitochondrial genome of *B. ignitus* (Genbank DQ870926, CHA et al. 2007): ²⁴³C ³⁴⁸C ¹⁰⁰⁸C ¹⁰⁷¹C ¹⁰⁷⁷C ¹⁴⁵⁵C.

Medium-size species, queen length 14–18 mm. The mandible distally broadly rounded with two pronounced anterior teeth and a weak posterior tooth, which is separated from the rounded anterior margin by a broad, shallow-rounded excision or ‘incisura’; labrum with the lateral tubercles low with many coarse punctures on their proximal side, the median labral furrow broad ± rounded; clypeus convex, shining with large punctures widely scattered; ocello-ocular area along inner eye margin with scattered large punctures and with many small punctures, which are only very narrowly reduced in density opposite the ocelli, the area of dense small punctures reaching the eye margin just behind the eye dorsally; mid basitarsus with the distal posterior corner acute but broadly rounded; hind tibia with the outer surface smooth and brightly shining; hind basitarsus with the proximal posterior process produced by more than its proximal breadth, the posterior margin nearly straight; metasomal tergum 2 clearly chagrined, punctures ± distinct; metasomal sternum 6 with a weak central keel. Color pattern: the head with the short hair black, black thorax with a well-defined broad pale yellow collare, going down on the episternum, with a distinct S-band of dark hair at the posterolateral border of the pronotallobus; hairs of the leg bases and hind femur black, hairs of the corbicular fringes black; wings light brown. Hairs of metasomal tergum 1 black, tergum 2 pale yellow, with scattered black hairs intermixed near the posterior margin, especially laterally; tergum 3 predominantly black with lateral white fringes, with white hairs posteriorly; terga 4–6 white with few black hairs intermixed; sternal posterior fringes whitish.

Name: Derived from the characteristic pale yellow-greyish coloration of Collare and Tergum 2, from Latin *pallido* = pale and *cinctus* = striped, with cingulum.

Discussion

Number vs. length of sequences

We need more information through either a greater number of sequences and/or sequences of greater length. In the case of the haplotypes 6 – 9 from Turkey, Iran, and Kyrgyzstan only more sequences of greater length may help to clarify the situation. At a first sight the *B. cryptarum cryptarum* from the Austrian Alps and the specimens from Turkey look rather isolated from each other, but there are ‘connecting specimens’ from the mountains of the Balkans in museum collections. We need sequences from the Balkan mountains to see whether there is a connection by one-step mutation branches or not. Furthermore, it might be possible that by longer GROUP-2 sequences the divergence from now two mutations with GROUP-1 length may increase to 4-5 mutations, which would imply an additional taxon also in the rank of a subspecies, *B. cryptarum iranicus* KRÜGER 1954. And, as for all *B. cryptarum* specimens we need

much more detailed information about morphological attributes, which might help to lump or separate taxa.

How to use diagnostic characters

Their number recognizes diagnostic characters: they are useful tools, which help to identify unknown specimens, and, contrary to most morphological characters they are unambiguous, having no gradual change or overlap (see discussion ‘facing the facts: morphology vs. molecules’, BERTSCH 2009, p. 304). The number (i.e., the position in a sequence) defines a diagnostic character. As long as authors number their diagnostic characters individually their usability is heavily restricted. MURRAY et al. (2008) numbered their diagnostic characters in comparison with a reference sequence (Genbank AY181162), CAROLAN et al. (2012) and WILLIAMS et al. (2012) numbered their diagnostic characters from the beginning of their partial sequences. As this partial sequences have different starting positions (in reference to the COI gene), their diagnostic characters get different numbers. None of these publications, even though some have identical co-authors, are comparable in numbering of diagnostic characters. As described in ‘methods and materials’ we recommend to number sequences and diagnostic characters in reference to the complete COI gene, NUMBER 1 is defined as the first nucleotide of the start codon ATA of COI.

Diagnostic characters must be treated like other characters: if a typical individual of *B. cryptarum* can be separated from *B. lucorum* by the labral furrow (narrow and wedge-shaped v for *B. cryptarum* and broad and rounded u for *B. lucorum*) this distinction is only valid for the comparison group ‘*cryptarum-lucorum*’. If, however, an undetected species with intermediate form of the labral furrow is involved one may see instead of two operational distinct attributes a continuum, and most probably doubt the existence of two separate taxa. Diagnostic characters of sequences show no intergrading a nucleotide is unambiguous. The diagnostic character ³²⁸C separates *B. cryptarum* from *B. magnus*, *B. patagiatus*, and *B. lucorum*, this is only valid for the comparison group ‘*cryptarum-magnus-patagiatus-lucorum*’. Things may be different if other taxa of bumblebees are possibly included.

Diagnostic characters should be applied in a two-step procedure: A naïve and schematic application of diagnostic characters is not recommended. If in a first step (with comparison group ‘*cryptarum-magnus-lucorum-patagiatus*’) the succession of diagnostic characters is ²⁵³T ³²⁸C ³³⁰T ⁷⁵¹C ¹³¹⁴C ¹³⁵⁹T ¹⁵⁰⁰C ¹⁵⁰⁷G, or part of it in case of shorter GROUP-1 sequences, this individual belongs to *B. cryptarum*. However, this is only valid if no unknown taxa outside the comparison group are involved. Only after this is settled the use of diagnostic characters to identify subspecies in a second step makes sense. Genbank sequences AY181117 and AY181119 for example are

specified as *B. cryptarum*. Therefore, it might be interesting to see what subspecies they are. To be safe, we examine in a first step their status as ‘species’, and instead of ²⁵³T ³²⁸C ³³⁰T ⁷⁵¹C, which we expect for *B. cryptarum* between positions 253 and 751 we get ²⁵³C ³²⁸T ³³⁰A ⁷⁵¹T. Both sequences do not belong to *B. cryptarum* (they are *B. lucorum*, see BERTSCH 2009 Fig. 4, but that is another story) and a further test for subspecies of *B. cryptarum* would be nonsense. Or let us test the identification of a degraded sequence identified as *B. cryptarum* by PCR-RFLP approach (Genbank KC192045, VESTERLUND et al. [2014]). A BLAST search (identification by divergence) results in 99 % *B. magnus* or *B. lucorum* (to understand this strange result you have to know that all three comparison sequences of the BLAST result AY181124, AY181123 and AY181121 are misidentified specimens of *B. cryptarum*), whereas for diagnostic characters we get ⁸³⁷G ¹¹¹⁶C ¹¹³¹C between positions 805 and 1221, which identifies the sequence as *B. cryptarum cryptarum*. Recently LOWENSTEIN et al. (2009), in an attempt to identify Tuna species in Sushi, gave a thorough report about the possibilities for identification by genetic divergence (BOLD identification or BLAST search) or by diagnostic characters, and came to the conclusion, that using characteristic attributes (= diagnostic characters) is the best solution. The discussion in LOWENSTEIN et al. (2009) of the problems involved in identification by different methods is very much recommended for reading.

Phylogenetic implications

What do we know about the evolutionary history of *B. cryptarum*? Nothing. According to theory, the phylogenetically older clades are in the centre of a median network, while the phylogenetically younger clades can be found at the peripheries. For the median network of *B. cryptarum* this would imply that the haplotypes of *B. cryptarum florilegus* must be seen as the phylogenetic oldest members, from which three different branches evolved. *B. cryptarum florilegus* then is not just a specialist island colour morph of *B. cryptarum albocinctus* (see WILLIAMS et al. 2012) but some sort of restricted Refugio distribution of a taxon with broader distribution in past times. The restricted occurrence of this subspecies at Nemuro peninsula at Hokkaido always looked more as a Refugio of an old than as a toehold for a young, expanding taxon. A genetic investigation (TAKAHASHI et al., 2008), which analysed microsatellite markers of the Nemuro peninsula population, supports this idea of a Refugio. The present distribution of the three subspecies on the Eurasian mainland makes South-western Siberia the ‘centre of diversity’, and the idea of their spread from this centre seems rather convincing. The median network treats the Turkey, Iran, and Kyrgyzstan proveniences as attached to the Central European haplotypes, but it may also be possible that they evolved the other way, from Southern Siberia via Kyrgyzstan and Iran towards Western Europe. With more data, possibly our knowledge

about these events will be more precise, and a detailed analysis of the genetic structure of populations will be central for such progress.

Colour form or subspecies?

However, it is absolutely priority that we finish the inconsiderate use of scientific names and the mixing of 'colour forms' and subspecies. *B. albocinctus* has been described as a colour form of *B. lucorum* from Kamchatka, we know meanwhile that this 'colour form' is part of *B. cryptarum*, and that specimens from the provenience 'Russian Far East' identified as *B. cryptarum albocinctus* are characterised by their specific COI sequences (Haplogroup D). Therefore the use of 'albocinctus' for all colour forms of *B. cryptarum* with bleached bands (whitish-yellow) all over mainland Eurasia is misleading (see Fig. 5 of WILLIAMS et al. [2012]), not all specimens *B. cryptarum* with bleached colour bands belong to Haplogroup D and not all specimens *B. cryptarum albocinctus* (Haplogroup D) have bleached colour bands, most specimens of *B. burjaeticus* KRÜGER (which belong to *B. cryptarum albocinctus* (Haplogroup D) are quite yellow, or even melanised. Another example could be *B. lapidarius decipiens* PEREZ 1890 described as a subspecies of *B. lapidarius* from the Pyrenees and later on expanded to the yellow-banded specimens of *B. lapidarius* from Southern Italy and Sicily. An analysis of the COI gene by LECOCQ et al. (2013) shows a difference of *B. lapidarius decipiens* from Spain to *B. lapidarius lapidarius* of only one mutation, *B. lapidarius decipiens* is a colour form, which from the genetical distance does not merit the status of a subspecies. Whereas the yellow banded specimens from Southern Italy and Sicily with a genetic distance of about 9 mutations represent a subspecies, which should get a separate name. It would be a great progress if this mixing of taxa defined by mtDNA (maternal genetic lineages) with mere colour forms would find an end.

Acknowledgments

Ralf Cartar (Calgary/Canada) supplied specimens *B. moderatus* from Alberta/Canada. Masao Ito (Sapporo) confirmed the identity of the Far Eastern GenBank sequences (AF279482/83) *B. cryptarum* with *B. albocinctus* and supplied specimens of *B. florilegus* from Hokkaido. Yuri Pesenko † (St. Petersburg/Russia) was good company during fieldwork near St. Petersburg and in the surroundings of Chita in (Russian Transbaikal) and supplied museum material. Willem Hogenes (Amsterdam/the Netherlands) supplied museum material. Bo Vest Pedersen (Copenhagen/Denmark) supplied the material of *Bombus* sp. BVP-A (AY181116) from Norway. Timofey Levchenko (Moscow/Russia) and Jan Ove Gjershaug (Trondheim/Norway) kindly let us see some unpublished material. We would like to thank them all for their valuable help.

References

- BANDELT, H. J.; FORSTER, P. & RÖHL, H. 1999: Median-joining networks for inferring intraspecific phylogenies. – *Molecular Phylogenetics and Evolution* 16: 37-48.
- BERTSCH, A. 2009: Barcoding cryptic bumblebee taxa: *B. lucorum*, *B. cryptarum* and *B. magnus*, a case study (Hymenoptera: Apidae: *Bombus*). – *Beiträge zur Entomologie* 59: 287-310.
- BERTSCH, A.; HRABÉ DE ANGELIS, M. & PRZEMECK, G. K. H. 2010: Phylogenetic relationships of the bumblebees *Bombus moderatus*, *B. albocinctus*, *B. burjaeticus*, *B. florilegus* and *B. cryptarum* based on mitochondrial DNA markers: a complex of closely related taxa with circumpolar distribution. (Hymenoptera: Apidae: *Bombus*). – *Beiträge zur Entomologie* 60: 13-32.
- BROWER, A. V. Z. & SCHAWAROCH, V. 1996: Three steps of homology assessment. – *Cladistics* 12: 265-272.
- CAMERON, S. A.; HINES, H. M. & WILLIAMS, P. 2007: A comprehensive phylogeny of the bumble bees (*Bombus*). – *Biological Journal of the Linnean Society* 91: 161-188.
- CAROLAN, J. C.; MURRAY, T. E.; FITZPATRICK, U.; CROSSLLEY, J.; SCHMIDT, H.; CEDERBERG, B.; MCNALLY, L.; PAXTON, R. J.; WILLIAMS, P. H. & BROWN, M. J. F. 2012: Colour Patterns Do Not Diagnose Species: Quantitative Evaluation of a DNA Barcoded Cryptic Bumblebee Complex. – *PlosOne* 7 (1) e29251: 1-10.
- CHA, S. Y.; YOON, H. J.; LEE, H. E. M.; YOON, M. H.; HWANG, J. S.; BIN, B. R.; HAN, Y. S. & KIM, I. 2007: The complete nucleotide sequence and gene organization of the mitochondrial genome of the bumblebee, *Bombus ignitus* (Hymenoptera, Apidae). – *Gene* 392: 206-220.
- DUENNES, M. A.; LOZIER, J. D.; HINES, H. H. & CAMERON, S. A. 2012: Geographical patterns of genetic divergence in the widespread Mesoamerican bumble bee *Bombus ephippiatus* (Hymenoptera: Apidae). – *Molecular Phylogenetics and Evolution* 64: 219-231.
- FOLMER, O.; BLACK, M.; HOCH, 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: 294-299.
- HEBERT, P. D. N. & GREGORY, T. R. 2005: The Promise of DNA Barcoding for Taxonomy. – *Systematic Biology* 54: 852-859.
- HILLIS, D. M. 1994: Homology in molecular biology. – In *Homology, The Hierarchical Basis of Comparative Biology*. – Edited by B. H. HALL. Academic Press, London New York: pp. 197-227.
- KAWAKITA, A.; SOTA, T.; ITO, M.; ASCHER, J. S.; TANAKA, H.; KATO, M. & ROUBIK, D. W. 2004: Phylogeny, historical biogeography and character evolution in bumble bees (*Bombus*: Apidae) based on simultaneous analysis of three nuclear gene sequences. – *Molecular Phylogenetics and Evolution* 31: 799-804.

- KUMAR, S.; NEI, M.; DUDLEY, J. & TAMURA, K. 2008: MEGA. A biologist-centric software for evolutionary analysis of DNA and protein sequences. – *Briefings in Bioinformatics* 9: 299-306.
- LECOCQ, TH.; DELLICOUR, S.; MICHEZ, D.; LHOME, P.; VANDERPLANCK, M.; VALTEROVA, I.; RASPLUS, J.-Y. & RASMONT, P. 2013: Scent of a break-up: phylogeography and reproductive trait divergences in the red-tailed bumblebee (*Bombus lapidarius*). – *BMC Evolutionary Biology* 13: 263. doi: 10.1186/1471-2148-13-263.
- LIBRADO, P. & ROZAS, J. 2009: DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. – *Bioinformatics* 25: 1451-1452.
- LOWENSTEIN, J. H.; AMATO, G. & KOKOTRONICS, S. O. 2009: *The real maccoi*: identifying Tuna Sushi with DNA Barcodes. – contrasting characteristic attributes and genetic distance. – *PlosOne* 4 (11) e7866: 1-14.
- MADDISON, W. P. & MADDISON, D. R. 2002: *MacClade Analysis of Phylogeny and Character Evolution*. – Sunderland, Sinauer Associates: 398 pp.
- MARDULYN, P. 2012: Trees and/or networks to display intraspecific DNA sequence variation. – *Molecular Ecology* 21: 3385-3390.
- MURRAY, T. E.; FITZPATRICK, U.; BROWN, M. J. F. & PAXTON, R. J. 2008: Cryptic species diversity in a widespread bumble bee complex revealed using mitochondrial DNA RFLPs. – *Conservation Genetics* 9: 653-666.
- PAMILO, P.; Tengö, J.; RASMONT, P.; PIIRHONEN, K.; PEKKARINEN, A. & KAARNAMA, E. 1997: Pheromonal and enzyme genetic characteristics of the *Bombus lucorum* species complex in northern Europe. – *Entomologica fennica* 7: 187-194.
- PEDERSEN, B. V. 1996: A phylogenetic analysis of cuckoo bumblebees (*Psithyrus*, LEPELLETIER) and bumblebees (*Bombus*, LATREILLE) inferred from sequences of the mitochondrial gene cytochrome oxidase I. – *Molecular Phylogenetics and Evolution* 5: 289-297.
- PEDERSEN, B. V. 2002: European bumblebees (Hymenoptera: Bombini) – phylogenetic relationships inferred from DNA sequences. – *Insect Systematics & Evolution* 33: 361-386.
- RASMONT, P. 1984: Les Bourdons du genre *Bombus* LATREILLE sensu stricto en Europe Occidentale et Centrale (Hymenoptera, Apidae). – *Spixiana* 7: 135-160.
- SIMON, C.; FRATI, F.; BECKENBACH, A.; CRESPI, B.; LIU, H. & FLOCK, P. 1994: Evolution, weighting and phylogenetic utility of mitochondrial Gene sequences and a compilation of conserved polymerase chain reaction primers. – *Annals of the Entomological Society of America* 87: 651-701.
- TAKAHASHI, J.; AYABE, T.; MITSUHATA, M.; SHIMIZU, I. & ONO, M. 2008: Diploid male production in a rare and locally distributed bumblebee, *Bombus florilegus* (Hymenoptera, Apidae). – *Insectes Sociaux* 55: 43-50.
- TAMURA, K.; DUDLEY, L.; NEI, M. & KUMAR, S. 2007: Mega 4.0: Molecular Evolutionary Genetics Analysis (Mega) software version 4.0. – *Molecular Biology and Evolution* 24: 1596-1599.
- TEMPLETON, A. R. 1998: Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. – *Molecular Ecology* 7: 381-397.
- TEMPLETON, A. R.; ROUTMAN, E. & PHILIPPS, C. A. 1995: Separating population structure from population history: A cladistics analysis of the geographical distribution of mitochondrial DNA haplotypes in the tiger salamander, *Ambystoma tigrinum*. – *Genetics* 140: 767-782.
- VESTERLUND, S.-R.; SORVARI, J. & VASEMÄGI, A. 2014: Molecular identification of cryptic bumblebee species from degraded samples using PCR-RFLP approach. – *Molecular Ecology Resources* 14: 122-126.
- WILLIAMS, P. H. & 17 co-authors 2012: Unveiling cryptic species of the bumblebee subgenus *Bombus* s. str. worldwide with COI barcodes (Hymenoptera, Apidae). – *Systematics and Biodiversity* DOI:10.1080/14772000.2012.664574.

Tab. 1: Primer pairs used for PCR, ¹⁾ standard primer (FOLMER et al. 1994), used in cases when primers 0-fwd & Z-rev did not work properly, and ²⁾ (SIMON et al. 1994), all other primers designed for bumblebees.

Code			
LCO1490 ¹⁾	5'-GGTCAACAAATCATAAAGATATTGG-3'	17→	
HCO2198 ¹⁾	5'-TAAACTTCAGGGTGACCAAAAAATCA-3'		←725
0-fwd	5'-GGAATAATATATTTTATTTTTCG-3'	40→	
Z-rev	5'-CCAAAAATCAAATAAATGTTG-3'		←710
1-fwd	5'-TAGGATCACCAGATATAGC-3'	254→	
K-rev	5'-GAGCTCAAACAATAAATCC-3'		←862
4-fwd	5'-CATTTATTTTGATTTTTTGG-3'	691→	
F-rev	5'-CATAATGAAAATGTCCAACACTAC-3'		←1132
5-fwd	5'-AATGAAAGAGGTAAAAAGAAAC-3'	772→	
A-rev	5'-ATGTTGAGGGAAAAATGTTAT-3'		←1281
C1-J-2441 ²⁾	5'-CCAACAGGAATTAAAATTTTATAGATGATTAGC-3'	937→	
TL2-N-3014 ²⁾	5'-TCCAATGCACTAATCTGCCATATT-3'		←>1560

Tab. 2: List of sequences *Bombus cryptarum pallidocinctus* used in the present analysis with Genbank numbers, and collection locality information. For about 20 unpublished sequences from this new subspecies see 'map' and 'taxon-ID-tree' for Bold: ACE4135.

GenBank	Locality	Country	Region	Latitude	Longitude	Altitude	
A) Sequences from our laboratory							
KJ668125	Böverdalen	Norway	Lom	61°47'50.54"N	08°26'17.78"E	650 m	♀
KJ668126	St. Petersburg	Russia	St. Petersburg	59°56'31.70"N	30°15'05.46"E	12 m	♀
KJ668127	Pavlovo	Russia	St. Petersburg	59°48'39.00"N	30°55'27.23"E	15 m	♀
KJ668128	Kuopio	Finland	Nordsavo	62°54'31.60"N	27°39'31.04"E	215 m	♀
KJ668129	Moskow	Russia	Moscow	55°50'37.40"N	37°37'09.03"E	135 m	♀
B) Sequences from Genbank							
EF362729	St. Petersburg	Russia	St. Petersburg	59°56'31.70"N	30°15'05.46"E	12 m	♀
JQ843371	Novosibirsk	Russia	Novosibirsk Obl.	54°49'44.17"N	83°07'06.28"E	181 m	♀
JQ843374	Ozero Chany	Russia	Novosibirsk Obl.	54°38'15.54"N	78°12'59.25"E	110 m	♀
JQ843381	Ozero Chany	Russia	Novosibirsk Obl.	54°38'15.54"N	78°12'59.25"E	110 m	♀
JQ843377	Grachevo	Russia	Kurgan Obl.	55°30'39.69"E	65°14'43.77"E	80 m	♀
JQ843417	Grachevo	Russia	Kurgan Obl.	55°30'39.69"E	65°14'43.77"E	80 m	♀
JN872568	Littoinen	Finland	Varsinais-Suomi	60°26'57.94"N	22°22'53.77"E	38 m	♀
JN872567	Houtskar/Huvudland	Finland	Varsinais-Suomi	60°13'21.46"N	21°22'32.28"E	14 m	♀
JN872566	Parainen/Alön	Finland	Varsinais-Suomi	60°19'06.59"N	22°18'29.44"E	38 m	♀
JQ843392	Nipfjället	Sweden	Kopparberg	61°56'46.10"N	12°49'44.67"E	905 m	♀
JQ843398	Nipfjället	Sweden	Kopparberg	61°56'46.10"N	12°49'44.67"E	905 m	♀