

Bonner zoologische Beiträge	Band 53 (2004)	Heft 1/2	Seiten 121–133	Bonn, Juni 2005
-----------------------------	----------------	----------	----------------	-----------------

Of Phenotypes and Genotypes: Two Sides of one Coin in Taxonomy?¹

Bernhard MISOF, Cornelya F. C. KLÜTSCH, Oliver NIEHUIS & Alexandra PATT
Zoologisches Forschungsmuseum Alexander Koenig, Bonn, Germany

Abstract. We discuss the recent proposals of DNA taxonomy and review some of the basic principles necessary to assess the importance of promoted ideas. We try to show that distinctness and independence of species taxa do not necessarily correlate, at least in newly formed species, which limits the applicability of any applied pattern analysis method. We do not intend to dismiss DNA based taxonomy, but would like to emphasize that molecular characters pose completely new problems to taxonomy often not appreciated enough in the discipline. DNA taxonomy is currently promoted because of its potential for automation. We show that species identification and circumscription can not be entirely automated since the result of a species description in taxonomy is equivalent to the formulation of a locally valid hypothesis. Consequently, every delimitation of a species taxon is open for empirical falsification only on a local scale. It is impossible to extend locally derived criteria, a species delimitation based on criteria derived from single sister species, to a global scale without sacrificing relevant information. We show from a pragmatic perspective that specimen assignment based on molecular characters can be automated under restricted circumstances employing character-based approaches.

1. INTRODUCTION

Several recently published articles advocate the incorporation of molecular techniques into taxonomic protocols (for example HEBERT et al. 2001, 2004; TAUTZ et al. 2002, 2003; STOECKLE 2003, to name just a few). For many practicing taxonomists the incorporation of molecular data seems highly desirable. Without any doubt, the incorporation of as many different character sets as possible into a species taxon description will improve the fit between species taxa and real evolutionary units. What is new in recent articles advocating DNA taxonomy is the emphasis on molecular techniques over morphological approaches, which, as the authors claim, would be the only way to meet the challenges in taxonomy of the 21st century. This is indeed slightly contrary to the perception of taxonomic practice of many taxonomists and aroused obvious opposition.

“Let's face it, the morphological approach has had 250 years to advance the task, and we're only 10% of the way towards the goalposts” said HEBERT in a report on *BioMedNet News*. His solution, to which many taxonomists are strongly opposed, is to distinguish between species on the basis of similarities and differences in their DNA. If taxonomists fail to embrace molecular technology, HEBERT is clear about the consequences: “There is no more likely death of a discipline than the failure to innovate.” (HEBERT, *BioMedNet News*, DNA: the barcode of life? 8 January 2003)

It is certainly true that taxonomy still faces a major task. It has been estimated that only one tenth of assumed existing species has been recognized and described so far. Considering the long tradition in taxonomy, which has been conducted over centuries, this is not really encouraging. Taxonomists estimate that hot spots of biodiversity, tropical rain forests, and also deep sea biotas are still only fragmentarily known. The biodiversity of tropical rain forests is threatened by massive destruction and consequently eradication of largely unknown biodiversity. In many densely populated areas, like Europe or Africa, massive anthropogenic destruction has left deserted areas with presumably significant loss of species and biodiversity. It is comprehensible that taxonomy has to search for ways of speeding up the process of identifying and describing species in the face of these challenges. Here surfaces the second conundrum. Taxonomists generally feel underfunded and it appears that the profession is a vanishing one among biologists. It is hard to find students interested in taxonomic activities, and for several invertebrate groups specialists are already absent. Consequently, training in taxonomy is not being provided, further promoting the loss of knowledge and momentum. Traditional taxonomy thrives from the life-long experience of specialists with trained intuition. The delimitation of species taxa by dedicated specialists and their intuition is often inaccessible for the outsider, making the process of species description quite mysterious. There is no single character with which species taxa can be recognized across phyla. The taxonomic challenge is the discovery of the relevant individual criteria given a population of individuals. Consequently, taxonomists are sometimes seen outside of the realm of

¹ In commemoration of Clas Michael Naumann zu Königsbrück (26.06.1939 – 15.02.2004)

natural sciences presenting themselves as book keepers of mental projections. This seemingly subjective and highly artificial approach to recognizing taxa often intimidates the novice student.

Another immanent problem is the publication process in taxonomy. Currently, species descriptions are most often not placed in peer reviewed journals and do not enjoy a high standing among many biologists. It would be ill advised for students hoping for a respectable career to enter the field of taxonomy. As can easily be envisaged, a massive brain drain away from taxonomy is the result.

Among others, molecular taxonomy has been proposed to cushion this development.

But what could be the advantages of using molecular techniques in taxonomy?

The advocates of DNA taxonomy, as it is also called, claim that the genotype provides a universal code with which species could be automatically identified and classified. This indeed seems to draw taxonomic investigation out of the realm of mystery and subjectivity. Automatic detection of species taxa based on genotypic differences would draw taxonomy out of its classroom corner and would transfer taxonomy to a technical enterprise solely executable by machines (TAUTZ et al. 2003; GASTON & O'NEILL 2004).

It is indeed true that the composition of DNA presents a universal resource of variation by the combination of only four character states. Including the phenomenon of indels (a composite word derived from insertion and deletion), we have five clearly distinct characters states with which we can discriminate between DNA sequences. Consequently, genotypes offer the potential of automatic detection of differences between DNA sequences. Different to the plethora of morphological variation, this seems indeed a big advantage as it might alleviate the problem of educating taxonomic specialists. Computers are inexhaustible co-workers doing the job while we are sleeping.

Proponents of DNA taxonomy claim that the simplicity of molecular characters help to solve the problems with which current taxonomy is confronted. Specialists are not needed anymore and identification of new sequence groups can be computerized, greatly speeding up the process of describing the nine-tenths of the unknown biodiversity. It has been already put forward that the description of new species must include a mandatory analysis of selected target sequences thus giving a barcode (HEBERT et al. 2001) of every newly described taxon. To be fair, we have to note that most proponents of DNA taxonomy do not call for a complete disregard of phenotypes but instead argue for a priority of the genotype in species descriptions. Once a species taxon

has been recognized based on genotypes, the corresponding phenotypes should be recorded afterwards completing species descriptions (HEBERT et al. 2001; TAUTZ et al. 2003). In other words, the shift from phenotypes to genotypes could provide the answer to the current taxonomic challenges.

In this essay, we will try to sort potentials and pitfalls of the molecular approach in taxonomy. We are particularly interested in theoretical implications, which are surprisingly often omitted in the actual discussion. We will not discuss political implications which have been put forward (LIPSCOMB et al. 2003; MALLET & WILLMOTT 2003) and we will also refrain from discussing the developments in internet technologies which are often promoted concurrently with matters of DNA taxonomy (GODFRAY 2002; SEBERG et al. 2003). Our starting point is a taxonomy which seeks the description of evolutionary units, characterized by reproductive isolation. In this sense, taxonomists try to delimit species taxa by matching phenotypic and possibly genotypic classes with these evolutionary units. This prescription differentiates the taxonomic activity from a mere classification of biological variation for our convenience and text books. DNA taxonomists claim that the genotype is better suited for the delimitation of species taxa matching evolutionary units than any other biological properties of organisms. This proclamation is the focus of our analysis. Consequently, we will first review the principles of genetic variation and variability. Second, we will spend time to review the role of species in evolution. We will not review all of the ideas once promoted but instead give a biased view on what species and speciation probably are and is. Taxonomy as a discipline is often surprisingly ignorant of theoretical issues behind species definitions and the process of speciation. Third, we will discuss the inherently assumed parallelism of character distinctness and population independence, which is at the theoretical foundation of taxonomy in general. But we will also try to show where DNA taxonomy can indeed help to recover lost ground in describing and (re-)identifying specimens. Finally, we will develop ideas of how to incorporate molecular and morphological methods in an automatic way. A final word: we do not claim to present radically new ideas about the incorporation of molecular techniques in taxonomy, but instead try to draw the reader's attention to the limits of DNA based approaches. This is by no means intended to damn DNA taxonomy or glorify the proposals. We feel that the ongoing discussion on DNA taxonomy and taxonomy in general is becoming unbalanced with a strong impetus towards political points of view instead of scientific reasoning.

2. INTRA- VS INTERSPECIFIC VARIABILITY/VARIATION

Considering genetic consequences of speciation, one has to ask how many, what kind of, and in which rate genomic organization of loci change during the process of speciation. Observed genetic differences between species may accumulate not only during or after the speciation process when populations have become independent, but also within populations in terms of intra-specific variation. In the following, we will use the term 'variability' to designate the differences actually present between individuals within a species. The term 'variation', which is often used in a synonymous way, is used to describe differences of individuals of different species.

2.1. Intraspecific variability

As DNA-Taxonomy is often conducted on mtDNA sequences, we will concentrate on the evolutionary patterns of this DNA type, which is nevertheless quite similar to genomic DNA with the exception of its usually maternal inheritance and its haploid occurrence. Genetic differences, as mentioned before, result from four different nucleotides, conventionally symbolized A, T, G, and C. These four character states (or rather five by counting indels) are potentially assumed at each position given enough time for substitution to occur. Genetic differentiation occurs not only between subspecies or species but also within populations as populations are affected by mutation, migration, random genetic drift and natural selection (HARTL & CLARK 1997). Mutation is the source of genetic variation, which arises and increases divergence between mtDNA genotypes. The input of mutations creates and maintains a certain degree of variability within populations (AVISE 2000). Within finite populations in sexually reproducing organisms, new mutations can become fixed after some time without the effect of selection. The rate at which random genetic drift changes allele frequencies or haplotypes varies inversely with the effective population size (N_e) (WRIGHT 1931; FISHER 1930). The smaller a population, the stronger are effects of random genetic drift and the faster the populations will lose its genetic variation. The effective population size is much smaller for mtDNA than for genomic DNA because of its haploid state and its generally maternal inheritance. Now imagine a population with a large distributional range. Gene flow within this population is restricted and consequently, it will be geographically structured. This will affect genetic differentiation between demes (defined as geographical units within which mating is effectively random). The degree of differentiation depends not only on the effective population sizes of demes but also on the pattern of gene flow among them (WRIGHT 1951). Given a finite Mendelian population, which splits into

allopatric subpopulations we can expect to see genetic differences within several loci after a certain number of generations without the force of natural selection. The consequence is genetic variability between populations driven by chance. In addition, natural selection plays a role in differentiation between populations of a widely distributed species. Let us assume the beginning of a speciation process within such a species. If populations become separated, genetic variability will be inherited to subpopulations. In this situation, haplotypes within subpopulations are potentially paraphyletic or even polyphyletic (FUNK & OMLAND 2003). Until monophyly of haplotypes will be achieved time must pass by, at least more than $4N_e$ generations, as estimated in simulations (AVISE 1987). Lineage sorting refers to a related problem. A reciprocal monophyly of haplotypes will be generated after complete lineage sorting has been achieved (ENGLBRECHT et al. 2000; cf. below).

Not in every case of recent speciation events do genetic distinctness and population independence go in parallel. We intuitively assume that the formation of species will arrange the total amount of genetic variation among them into non overlapping packages of genetic variation. This assumption is indeed appealing, but not necessarily realized. DNA taxonomy has to deal with issues of population genetics, as it often relies on presumably selectively neutral gene loci. Paraphyly of selectively neutral haplotypes or the absence of genetic differentiation among those pose major problems to taxonomy and clearly restrict the applicability of molecular tools. Imagine speciation events, which lead to a clear differentiation of phenotypes due to, for example, sexual selection but without fixation of selectively neutral alleles in species taxa, because either populations had been large or there was no time yet to fix allele frequencies. The incongruence of genetic and phenotypic distinctness is not an artificial construction. Examples are given by TAUTZ et al. (2003) who mentioned the situation of the African great lakes cichlid fauna. In Lake Victoria, cichlids are strongly differentiated in phenotypic groups but lack a clear genetic differentiation in sequenced mitochondrial markers (MEYER et al. 1990, compare also TAKAHASHI et al. 2001). These considerations illustrate the limits of DNA taxonomy quite nicely; since, despite reproductive isolation, there must be enough time to build up genetic variability between populations.

2.2. Interspecific variation

The degree of genetic differences between populations, subspecies, or species can be described by using genetic distance methods (e.g. NEI 1978). Dealing with a certain degree of differentiation between populations, some researchers are going to classify species by the percentage of genetic divergence (for example MALTAGLIATI et al. 2001; JIGGINS & DAVIS 1998). This requires a constant

rate of sequence differentiation at neutral loci, which is independent of population genetic factors mentioned above. However, the influence of small population sizes on the rate of evolution (OHTA 1976) was assessed in a study by DE SALLE & TEMPLETON (1988). They compared interspecific levels of sequence divergence within a *Drosophila* taxon, comprised of several species, supposed to be generated by a founder event and another *Drosophila* taxon, again comprised of several species, which were thought to have originated from a large population. They showed that the second species group presented a much less pronounced molecular differentiation than the first one. Another example in which intraspecific sequence divergence is as great as such among distinct species is given by WAYNE et al. (1990) who analyzed genetic distances of Jackal populations. In summary, the degree of genetic differences will not always match the biological species status one to one. It is indeed tempting to say that there is still a rough correlation between genetic distance and extent of reproductive isolation. If this holds true, a convention on a minimal distance measure could be successful in identifying new species. Alpine snails of the species *Arianta arbustorum* (HAASE et al. 2003) show COI haplotypes of more than 20% differences within a population. A genetic distance of 20% as a general convention to delimit species is useless in many other taxa. This issue will be taken up later again.

3. SPECIES AND SPECIATION

"The primary aims of taxonomy are to name, circumscribe, describe and classify species" (SEBERG et al. 2003, p. 63). We probably all agree that the generation and maintenance of biodiversity is closely linked with the establishment of evolutionary units and our understanding of the processes will rely on our ability to recognize these evolutionary units. In sexually reproducing organisms, species are exactly those evolutionary units without which biological diversity is only incompletely understood. It is therefore paramount to reflect what processes lead to the formation of evolutionary units and what characteristics can be used to identify them. The discussion on species and speciation traditionally includes the following questions: 1) What are proper species definitions and what is the ontological status of species? 2) What are reproductive isolation and isolating mechanisms? 3) What is the predominant mode of speciation and why are there species? We add the taxonomically relevant question of how species are best classified. Subsequently, we will pick up the issues and comment on them from the perspective of a taxonomist. The history of species definitions is legacy. MAYR (1982) gave a most authoritative and readable account on species definitions in taxonomy and evolutionary biology. Definitions are either acceptable or not, but do

not represent explanations of the more or less distinctly separate arrays of organisms (DOBZHANSKY 1937). We think that the disputes over proper species definitions are not interesting and we will not follow up this route. Secondly, we will deliberately omit the discussion whether species are classes, natural kinds, or whether species constitute individuals (for a review, see again MAYR 1982). The discussion on the ontological status of species is not important here. (However, it is clear that the question of whether species constitute natural kinds or individuals is much more interesting than disputes over definitions as it confronts us with a problem in need of an explanation). With the modern synthesis in evolutionary biology, species in sexually reproducing groups of organisms are seen as groups of individuals constituting the largest possible inclusive Mendelian population (DOBZHANSKY 1935, 1937). In contrast to prevailing typological species concepts, representatives of the modern synthesis shifted the emphasis from patterns to processes. Inclusive Mendelian populations do not exist in parthenogenetic and asexual organisms and therefore, the term species in the biological sense is not applicable for these groups of organisms. Many taxonomists and natural scientists felt this limitation of the biological species concept as much too restrictive and opted for more inclusive concepts covering also asexuals and parthenogenetic forms. To our knowledge, the most farfetched example is the general lineage concept of species, which applies the term population to sexual and asexual groups of organisms (DE QUEIROZ 1999). We will restrict the term population to a group of sexually reproducing organisms; anything else blurs the essential differences and produces the illusion of unified theories of species. Of course, taxonomy is also dealing with asexually reproducing individuals, but for now, we will deliberately exclude asexuals and parthenogenetics from our discussion.

The formulation of the biological species concept shifted the emphasis from definitorial problems to real scientific inquiry. As DOBZHANSKY put it so easily the reasoning runs as following: first is the observable pattern of distinctive separate arrays of organisms and secondly there is the theory that these separate arrays seem to be maintained by established reproductive barriers between groups. Mayr added a new twist to this theory. He concluded, starting from the question of how characters can evolve despite mixing of genotypes in sexually reproducing organisms, that the formation of Mendelian groups which are reproductively isolated against other such groups should be clearly advantageous. The new twist is the explicit inclusion of the phenomenon of selection albeit in a slightly diffuse way. MAYR was not the first emphasizing isolating mechanism in defining species. Credit has to go to DOBZHANSKY who introduced the term isolating mechanism in his book *Genet-*

ics and the Origin of Species (1937) (The idea of recognition of “proper” reproductive partners is central to species and even much older. It finds its explicit predecessors in Buffon (in MAYR 1982).) DOBZHANSKY (1937) reserved an entire chapter for the discussion of isolating mechanisms. He mentions that “Darwin and Lamarck pointed out that interbreeding of groups of individuals, which are hereditarily distinct, results in dissolution and swamping of the differences by crossing.” (DOBZHANSKY 1937, p. 228). He correctly notes that the statement is somewhat misleading since single allelic differences will not diffuse in crosses but segregate again in the offspring. Only distinct character complexes of polygenic origin can break down in crosses leaving the hybrids less competitive compared to parents. For DOBZHANSKY, isolating mechanisms are mechanisms, which prevent the production of hybrid zygotes reaching the reproductive stage. Most importantly, citing Mayr again, isolating mechanisms are seen as “biological properties of individuals which prevent the interbreeding (fusion) of populations” (MAYR 1970). Therefore, sole geographical isolation of populations does not constitute a necessary and sufficient criterion for recognizing reproductively isolated species. The basic principle behind this idea is the insight that Mendelian populations are the units of evolution. (It is interesting to see that the origin of panmictic populations is not important in this respect.) Scientists like DARWIN, LAMARCK, DOBZHANSKY, MAYR, and many more have been convinced that differential adaptation to heterogeneous environments and ecological needs would have been impossible without the partitioning of panmictic populations. For that reason, we can call species independent modules in the evolutionary process.

An analogous picture can be drawn from the field of developmental biology. There it became clear after ground breaking work starting in the early 20th century (BERTALANFFY 1932; WADDINGTON 1957) that evolvability of complex characters can only be maintained by reducing pleiotropic interactions. For example, imagine genes, which effect the formation of epithelia. Specific alleles might promote the formation of a thick keratinous outer layer, thus reducing the risk of dehydration in dry habitats. This could clearly be an advantage for the bearer of this allele. However, thick keratinous layers on bronchial epithelia are certainly counterproductive as they restrict the intake of oxygen. The pleiotropic interactions prevent the fixation of an allele in this case. The situation leads to a reduced adaptability under strong pleiotropic interactions. The solution is the break up of pleiotropic interactions and the recovery of adaptability. The break up of pleiotropic interactions results in the formation of genetic and developmental units, which we also can call modules. The evolution of modules is the key invention (compare for example

RIEDL 1978; BONNER 1988; RAFF 1996; WAGNER 1996; WAGNER & ALTENBERG 1996). Similarly, species can be seen as genetically independent modules protecting their harmonious gene pool and thus, maintaining distinctness. The biological species concept refers to exactly these properties of inclusive Mendelian populations. Biological species are characterized by distinctness and independence. Whether distinctness or independence comes first is asking for the priority of the chicken or egg. MAYR himself favored for a long period allopatric speciation processes as the predominating mode of speciation, which implies that independence is acquired first and distinctness is a secondary phenomenon. MAYR’s idea predicts a linear order of phenomena from geographic isolation to reproductive isolation. However, theoretical considerations suggest that parapatric speciation and its derivatives (peripatric up to sympatric speciation) are at least conceivable, which makes the decision less obvious (TURELLI et al. 2001).

Polytypic species are species in which individual subpopulations can be characterized by clear phenotypic differences, often given subspecific rank. It is assumed that a reduction of gene flow between geographically distant populations maintains local differences. However, most of the genotypic composition within the species is still homogeneously distributed throughout the entire range. Such a situation is also found in species with clinal variation along an ecological gradient. This even distribution of phenotypic variation can be stabilized by the effects of natural selection, resulting in local fixation of alleles under stabilizing selection (local adaptation). However, selectively neutral alleles will travel freely across large ranges (DIECKMAN & DOEBELI 1999; DOEBELI & DIECKMAN 2003). In such a case, partial independence due to restricted vagility of individuals is driving the evolution of distinctness despite the absence of reproductive isolation. This example illustrates that reproductive isolation is neither a sufficient nor a necessary condition for adaptive change in populations. Local adaptive changes can be accomplished by the action of natural selection alone, even under para-(sympatric conditions (DOEBELI & DIECKMAN 2003). *Speciation* is the maintenance of phenotypic differences by means of reproductive isolation. Reproductive isolating mechanisms are biological properties of individuals preventing the fusion of gene pools. A compatible definition would be “biological properties, which lead to the interruption of the gene flow between populations”.

In the example given above, a population differentiates along an ecological gradient. This gradual differentiation is maintained by the effects of natural selection. Reproductive isolation is not necessary to maintain the gradual difference. But as soon as, for example, female preferences for certain male characteristics become genetically coupled, the gradual differentiation of pheno-

types can be translated into an abrupt change in phenotypic properties within now genetically separated populations. This may eventually lead to reproductive isolation (DOEBELI & DIECKMANN 2003). In such a situation, individual selection favors the acquisition of reproductive isolation. Starting from distinct phenotypes, the acquisition of genetic independence becomes a secondary phenomenon. The example above illustrates quite well that reproductive isolation is not necessarily a phenomenon of single gene loci but is best seen as a feature at the organismic level (compare MAYR 2001; WU 2001; WU & TING 2004). In some cases, only a small number of loci might suffice to lead to complete reproductive isolation, in some cases not. Albeit the notion of harmonious gene pools implies a genome wide differentiation, this is not even intended in MAYR's writings (MAYR 1996). We could formulate that genes responsible for reproductive isolation are taking the lead in the differentiation process of the genotype and it will take some time until unique alleles will be fixed for every locus as soon as reproductive isolation has been acquired. Within sexually reproducing populations individuals will have genes (alleles) in common defining the biological properties of reproductive isolation. We call these "species producing genes". It might indeed be possible to give a comprehensive genetic delimitation of a species taxon if based on these species producing genes (MAYR 1996).

The situation is somewhat different in allopatric populations. Change of allele frequencies are expected to be randomly distributed within the genome assuming comparable selection regimes in allopatric populations. From a theoretical point of view, this process, well known under the term random genetic drift inevitably reshapes the composition of genotypes and gene pools, finally generating incompatible gene pools between allopatric populations; incompatible in the sense that after secondary contact of populations, out-crossing individuals will suffer reduced fitness. The acquisition of species status by random drift is assumed to be slow and probably of polygenic origin. Again, it might still be possible to give a comprehensive genetic species delimitation based on genes causing isolating properties. The phenomenon of genetic incompatibility deserves closer attention since it has been considered as the most important factor of the speciation processes. Random genetic incompatibilities have been called Dobzhansky – Muller incompatibilities (ORR & TURELLI 2002). ORR (1995, 2001) showed that if the number of genetic differences between individuals of separated populations increases linearly with time, the number of incompatibilities should increase as the square of their divergence times. This has been called the "snow ball" effect (ORR 1995). The snow ball effect predicts that we will probably often overestimate the number of genetic incompatibilities between populations responsible for postzy-

gotic reproductive isolation. Therefore, we would expect to see many cases of clear genetic incompatibility in experimental crosses of sister species. This is clearly not the case (compare experiments in CHARLESWORTH 1995). It appears that even an occurrence of multiple bottle necks is unable to produce enough genetic incompatibilities between populations to establish effective postzygotic isolation. This experimental evidence casts considerable doubt on the importance of random genetic drift in speciation and it also sheds some light on the interpretation of genetic variation recorded for selectively neutral marker loci. The experimental results tell us to expect substantial differentiation without postzygotic isolation.

We have reviewed the cases in which "species defining genes" take the lead in the genetic differentiation process of newly formed populations. This does not have to be the case. Depending on speciation rates, selectively neutral loci might even take the lead in the differentiation process without causing reproductive isolation. This phenomenon is quite common and leads to geographically structured gene pools with range limited mitochondrial haplotypes, for example, in the Eurasian ground dwelling fish *Cottus gobio* (ENGLBRECHT et al. 2000): several old lineages of *Cottus gobio* recolonized central Europe and readily mixed where they met. In *Cottus gobio*, genetic distinctness and even monophyly of haplotype groups restricted to geographical areas do not indicate reproductive isolation based on biological properties of the individuals. The above mentioned experimental results and theoretical considerations should caution us in accepting selectively neutral marker loci as ideal tools of DNA taxonomy. However, all of the molecular taxonomic work so far is based on selectively neutral marker loci (compare HEBERT et al. 2001).

4. CHARACTERS: CLASSES AND RELATIONS

Populations of polytypic species and allopatric populations in general, capable of fusing after secondary contact force one to accept that the biological properties of individuals, which prevent fusion of populations, are the hallmark of biological species. Therefore, it would be ill advised to base species delimitations on the recognition of any apomorphic character set shared by a number of individuals. This practice has been introduced by people defending derivatives of the phylogenetic species concept (in MAYDEN 1997; WHEELER & MEIER 2000). It is based on the argument that every reproductively isolated population will evolve autapomorphic character states. This is certainly true in the long run, but completely ignores the initial conditions within reproductively isolated populations. There is clearly a different quality to sets of characters whether they constitute "species producing" characters or not. The phylogenetic species concept does not take into account this differ-

ence in character qualities. If different quality of information in characters is ignored, any observed difference gains weight. Representatives of typological species concepts refer to the apparent success in applying typological criteria in delimitating species taxa (compare for example HEBERT et al. 2001). We have already shown that there are many conceivable situations in which DNA based taxonomies will lead to erroneous conclusions.

Reproductive isolation is a quality of relations between populations and ultimately of individuals. This means that character attributes of reproductive isolation are from a different world compared to characters displayed by let's say pinned individuals in an insect collection. Imagine a pinned insect in front of you displaying a certain suite of characters: relatively small and elongated, red with black dots on its back. In the drawer, in which the pinned insect was spotted, we discover several other individuals looking alike. In some other drawers, we manage to locate additional specimens with similar phenotypes. A decent hypothesis would be that all those specimens belong to one species characterized as relatively small, elongated, and red with black dots on its back. We could now perform a thought experiment and search for all other possible combinations of the characters size, shape, and dorsal coloration. For example, we could look for big, stout, black and red spotted insects among the pinned myriads of specimens. This is certainly not the only combination we can think of. Maybe we will even be successful and discover a suite of specimens that look exactly like the predicted organism. If we have been successful, we would have filled the class "big, stout, black and red spotted". Does this imply that classes are real entities and not just projections of our mental condition? All organisms in a class truly display the class defining characters and for this trivial reason there is a first reality to the phenomenon. Nevertheless, there is more to classes of characters. Phenotypic characters are signs of the interaction of phenotypes and the environment, evolutionarily and ecologically. Consequently, a class of characters will share a set of environmental interactions that is definitely real. The causal relationship of form and function in this case enables the prediction of form from function and vice versa. It is clear that the causal relationship of form and function is not a linear mapping in biological systems, but we think that in principle, this statement will hold. In contrast, reproductive isolation as a quality of individual (population) interaction emerges as a character of relation separating two sets of individuals. It is in this sense an emergent quality of organisms with historical-factual properties tied to its bearers. Biological properties of individuals preventing fusion of populations are classes of characters, which can be expected to be mutually exclusive in reproductively isolated populations.

Traditional taxonomy tries to identify those classes of characters responsible for the quality of reproductive isolation. By doing this, taxonomists often just intuitively sort and weight characters drawn from the general pattern (classes) displayed by individuals to approximate the phenomenon of reproductive isolation. Our sensory apparatus and experience as taxonomists makes us successful in identifying and circumscribing biological species based on phenotypic characters. This is common sense among taxonomists and accepted outside the taxonomic community as well. It would be fascinating to collect complete genetic information on characters leading to reproductive isolation in circumscribing biological species taxa (see also MAYR 1996), but unfortunately, this would probably be a hopeless activity. The quality of reproductive isolation is not restricted to a general class of phenotypic characters and we can expect all kinds of genotypic combinations in this respect. It is obvious that an automated identification and circumscription of species taxa based on the genetics of reproductive isolation is not feasible in the near future. In contrast, DNA taxonomy tries to infer species taxa from analyzing more or less randomly drawn genetic information. With randomly drawn genetic information we refer to the selection of genetic marker systems, which has until now never been guided by the needs of taxonomy. The inclusion of genotypic characters was traditionally based on economical and technical considerations. The limits of the genetic markers currently in use are easily highlighted. A monophyly of sequenced haplotypes will coincide with species taxa only if there was enough time for the population to complete fixation before the onset of a new speciation event. It is therefore an empirical question whether the application of DNA taxonomy becomes restricted to just "old" or equivalently fast divergence processes. FUNK & OMLAND (2003) estimated the frequency of species-level paraphyly or polyphyly at 23% based on 2319 published reports. This frequent occurrence of paraphyly and polyphyly should warn the DNA-taxonomist.

5. SPECIES TAXON AND SPECIMEN IDENTIFICATION

5.1. Introduction

Species taxa are most properly identified by their species-producing class of characters, called biological properties of the individuals preventing the fusion of populations. This is of course often impossible to achieve. The art in taxonomy is the selection of character sets which come close to that goal. At the phenotypic level, we operate by comparing character sets, weighting the information at hand, and develop a hypothesis about the relevance of different character sets. In an ideal case, we would filter a certain set of characters re-

sponsible for reproductive isolation. The hypothesis that the selected set of characters is responsible for reproductive isolation between a pair of species is at least theoretically open for empirical falsification. This is a most important strength of the biological species concept. It separates the biological species concept from all other concepts, which emphasize the population lineage properties of species. The introduction of a historical component in species definitions, like specimens belong to a certain species because they are directly linked with other species by descent, creates an hypothesis which is not directly open for empirical falsification. It is clear that organisms are linked by descent. There is nothing wrong with this. But we are convinced that it is not necessary to jumble explanations of historical-narrative with hypothetical-deductive qualities in delimiting species taxa. It is like mixing of taxonomy and phylogenetic systematics without noticing the essential different qualities of the disciplines. Admittedly, the empirical test is seldom conducted and therefore, most species descriptions have to be regarded as points of views instead of critically evaluated hypothesis. But how can the taxonomic enterprise still be a success? It all rests on the shoulders of the specialists familiar with their group of organisms. These specialists develop a feeling for the relevant characters delimiting species taxa in their groups. The success of the traditional taxonomic method is therefore dependent on life-long experience and intuitive solutions. These are certainly not ideal preconditions for the automated inventarisation of biodiversity.

5.2. Specimen identification by means of molecular tools

In general, DNA taxonomy does not deal with species-producing sets of genes. It could be hoped for that "species-producing genes" are not randomly distributed among at least closely related taxa enabling a semi-automated detection of those genes. But detailed genetic analyses of speciation processes revealed that even among closely related sister taxa of *Drosophila*, species-specific genes greatly vary and can not be compared (see review in WU & TING 2004). The manifold differences in courtship display alone seem to support this view. The contradicting argument might be raised that several studies document a driving force of sexual selection in speciation processes (for example OWENS et al. 1999; PANHUIS et al. 2001; ARNQVIST et al. 2000; MISOF 2002). If sexual selection is fueling speciation, it will mostly be restricted to pre-mating isolation and variation within courtship displays. The extent of phenotypic species-differences should be directly translated into a set of genotypic differences comparable between species taxa. To us it appears that we just do not know enough about speciation genes and comparative analyses of speciation processes. We will have to wait for answers from combined empirical and comparative analy-

ses of speciation phenomena until we can develop subtle hypotheses of the "genic view of speciation" (WU 2001).

In any reproductively isolated group, genetic differences will accumulate. Eventually, genetic differences will become fixed for selectively neutral loci as well. Subsequently, the genetic variation for selected loci will gain characteristic autapomorphic characters. But those populations will not by necessity evolve isolating mechanisms. Again, in delimitating and circumscribing species taxa we need to identify the biological properties of individuals causing reproductive isolation. It is not at all sufficient to show even in a sympatric situation that two subgroups of individuals maintain phenotypic differences independently of sex. Just think of genetically fixed polymorphisms within species, text book examples are some orchid flowers or *Anopheles* mosquitos. Selectively neutral marker loci will not be among the candidates. We refer the reader to the illuminating experimental data on effects of random drift on postzygotic reproductive isolation (CHARLESWORTH 1995). Admittedly, autapomorphic characters of, for example, mitochondrial haplotypes within a clear biological species can be used *a posteriori* to filter specimen identity. This filtering of specimens is equivalent to species assignment. Assignment is restricted in its applicability to situations, in which haplotypes within sister taxa have reached monophyly and diagnostic (autapomorphic) character states (for an empirical analysis compare FUNK & OMLAND 2003). It is important to stress the reciprocal monophyly of sister taxa. It is not sufficient to show that the biological species under consideration shows a monophyletic set of haplotypes. For example, consider the scenario of a peripatric speciation event. We start from a population with a large distributional range with a certain degree of genetic variability. In a peripatric situation, geographically marginal subpopulations might evolve reproductive isolation. Their haplotypes of selectively neutral marker loci will likely be monophyletic. However, the stem population from which the peripatric population became isolated was driven in the status of paraphyly. Nevertheless, individuals from the stem population will remain reproductively compatible. It will take time until the stem population again reaches monophyly for the haplotypes under consideration. The paraphyletic situation in the stem population poses clear problems to specimen identification and assignment. Some haplotypes within the stem population will be more closely related to haplotypes of the reproductively isolated population. The distribution of genetic distances will not correspond with species taxa and consequently, can not be used to infer species taxa. It is particularly surprising that genetic distance measures are advocated in many recent publications for the delimitation of species taxa in phenotypically indistinguishable forms. The character based as-

signment and identification is the only way to go. If, and only if, a set of diagnostic characters has been correlated with reproductively isolated populations, these diagnostic characters can be used in specimen assignment and identification.

A theoretical example will hopefully illustrate the argument above. Assume we are naïve DNA taxonomists and received an unknown invertebrate larva collected in marine sediments. The assignment of the larva to a known taxon will be of great importance because it can bridge gaps in the analysis of biodiversity. We might use genetic information to assign the larva to a certain taxon for which genetic markers have already been characterized. From this larva, we sequence fragments of three different gene loci and the observed sequence information indicates that the sequences are only 5% different from sequences of specimens within a known taxon. As naïve taxonomists, we would very quickly say that the larva is the yet undescribed larva of a new taxon. Another detailed analysis of phenotypes and biological properties of individuals within the species taxon suggests that the species has a very heterogeneous genetic distance distribution within its range. We find that genetic distances are indeed observed between 0 - 5 % within a Mendelian group, but a geographically marginal subpopulation appears reproductively isolated judged from additional evidences. Although, genetic distances within this subpopulation are small and the average genetic distance of this subpopulation to members of the stem population fall within the recorded range of the presumably Mendelian stem population. In this situation, genetic distances will be of no help in specimen identification and assignment. However, if we concentrate our search on the identification of autapomorphic characters, which describe the recently derived reproductively isolated subpopulation, we could circumvent the problem of overlapping genetic distances and paraphyly of the stem population. The likelihood of specimen assignment will depend on the extent of previous sampling within the stem and subpopulations; the better the sampling, the better our chance of correct re-identification (compare POZHITKOV & TAUTZ 2002). POZHITKOV & TAUTZ (2002) developed an algorithm to select ideally suited probes for specimen identification in large through-put-assays. This strategy seems very promising, in particular for small organisms in great abundance. The assignment of specimens will of course obey the rules of probability. A very nice example has been developed by the *www.DNA-surveillance* project. It constitutes a web-based molecular identification of whales, dolphins and porpoises (BAKER et al. 2003). Obviously, this DNA based specimen identification tool has helped to review species descriptions and type assignments in this group of organisms. The taxonomy of cetaceans is plagued by the problem of some extremely

rare species for which morphological descriptions are often very incomplete. In this case, DNA based taxonomy has undoubtedly helped clarifying issues.

5.3. Specimen identification by means of phenotypic tools

There are other approaches for the rapid (re-)identification of species beside DNA based taxonomy. One possibility is automated species identification through digital image analysis (WEEKS et al. 1997; WEEKS et al. 1999a, 1999b; WATSON et al. 2004; ROTH et al. 1999; STEINHAGE et al. 1997; STEINHAGE 2000). In this automated approach, images of insect wings (e. g. of moths, bees, and wasps) are used to distinguish between species. Images are analyzed concerning venation, color or scale pattern. The system DAISY (Digital Automated Identification System; O'NEILL & GAULD 2001; WATSON et al. 2004), for example, was able to identify 83 % of unknown moth species from wing images. Taken into account that digital images were taken in the field, where light conditions are not always standardized and recorded specimens are moving, this is an impressively high success rate. A similar approach is run by the ABIS (Automated Bee Identification System, ROTH et al. 1999; STEINHAGE et al. 1997; STEINHAGE 2000) project. Here, images of forewings of bees, are taken and their venation is used to identify species on different hierarchical levels like genus, species group within a genus and species. Both projects have in common that the programs have to be trained with sets of images to have some kind of "experience" in identifying species. Thus, the systems are only able to discriminate between already digital recorded species. The systems are able to re-identify specimens of known taxa, but not able to recognize new species. In consequence, this can lead to misidentifications, as the system always tries to match an unknown image to a species. GASTON & O'NEILL (2004) pointed out that an ideal automated species identification system should be able to reject species identification if similarities to known species in the database are too small. Hence, this could help to identify unknown phenotypes within a set of known species in the future. Please notice, this is not similar to saying that the automated systems can be used for recognizing new species. We suggest that the systems can be used to recognize unknown phenotypic variation, which can then subsequently be evaluated by taxonomists. Further improvements of the systems could include better images, more morphological characters (e. g. images of additional parts of the body like limb or head structures), and a database with information about distribution, food plants, flight time and other biological data (WATSON et al. 2004). The project DORSA (German Orthoptera Collections, RIEDE et al. 2004) is one attempt to incorporate distribution maps, links to collections, images of type specimens, and last but not least sound recordings

in a single database. A rapid assessment tool was developed for automated song recognition and a sound library will be established soon. In this case, it is possible to re-identify orthopteran species due to the characteristic song pattern of single species. The integration of the mentioned distributional and biological data allows a verification of the identified species. But again, this effort is primarily suitable for re-identifications and not for recognition of new species since the only way to confirm a specimen as an unknown species is to exclude all other known species, which can be difficult if not all known species are included in the database.

In summary, although auspicious efforts have been made to automate species identification through digital image analysis, it is still a long way to go. Two major aspects have to be emphasized in this context. First, all of the presented computed systems are some kind of semi-automated since the program has to be trained; database establishment and updates as well as verification of recorded species will always remain in human hands. Second, these projects are tools for re-identification and not for recognition of new species. But computer-based routine re-identification systems could be at least a great facility to speed up the re-identification of species in biodiversity research. Of course, if species can only be differentiated by genital morphology, which needs first a preparation step, the automated re-identification system will be again of limited use.

6. CONCLUSIONS

The proponents of DNA taxonomy try to convince that without relying on molecular characters we will not be able to register a significant share of biodiversity in reasonable time. Automated DNA based taxonomy is the key word here. We have tried to argue that taxonomy can not become a robot activity and speed up taxonomy for several reasons. The most important one is that describing species taxa is a scientific activity resulting in testable hypotheses despite the prevailing absence of empirical falsifications. The hypotheses are based on universal theories (species are reproductively isolated groups of individuals) and identified local realizations (species-specific characters causing reproductive isolation). Biological properties of individuals preventing fusion of populations are as numerous as there are biological species and are emergent qualities of species taxa. These qualities can not be predicted and consequently elude an automatic detection. But still, we expect a significant impetus from molecular studies, which analyze genetic variability within and among populations. New data on genetic variability can potentially direct our interests in taxonomic studies if genetic and phenotypic data do not correlate. For example, investigations are published, which report on the application of molecular techniques revealing hitherto masked differ-

entiation among seemingly identical phenotypes (PROUDLOVE & WOOD 2003). Obviously, DNA based methods help to discover patterns of differentiation that are possibly correlated with reproductive isolation. For some groups of organisms, in which morphology will not be of great help, the application of molecular tools can certainly be advantageous and become mandatory. However, similarly to phenotype-based species descriptions, assessments of species status based on molecular characters remain points of view as well and are even more limited in their relevance compared to phenotypic characters. Molecular differentiation based on selectively neutral marker loci is definitively not a test of reproductive isolation between populations. Given a set of "well" circumscribed species taxa, DNA taxonomy will speed up descriptive work in biodiversity assessment in the sense that the number of genetically diagnosable species for a given collection event or area can be automatically recorded. This work will profit from automation since hypotheses testing is not involved. However, recent developments in automatic specimen identification based on phenotypes are encouraging as well. These developments should not become underfunded in competition with molecular tools, because the ultimate application of phenotypic identification seems more versatile than any DNA based method. We emphasize that the current practice of using genetic distances to delimit species is ill advised and should be replaced by character-based methods. We see that DNA taxonomy has its clear limitations in specimen (re-) identifications and species taxa circumscriptions. We should stress that we do not even have a grip on the problems facing the description of the 90% undescribed taxa and we speculate that many of them are not described yet because they are close sister taxa of known species. If this is true we are gearing towards the area in which DNA taxonomy is most limited. We should be aware of this problem.

ADDENDUM

It is clear to us that we largely ignored the distinction between sympatric/parapatric and allopatric situations. Some readers might argue that our argumentations are more or less acceptable in allopatric situations, in which assignment of species status is arbitrary anyway, but are not applicable under sympatric/parapatric conditions.

It is indeed tempting to say that in a sympatric situation at least two unlinked selectively neutral markers, which segregate independently of sex, indicate separate species status of their bearers. We would like to argue that even in a sympatric case just two selectively neutral markers rarely tell us who is who.

Theoretical work shows that in a sympatric speciation process the differentiation of selectively neutral markers

will lag behind. But finally, even selectively neutral markers will reach fixation. In this case they can indeed indicate species status. However, we would assume that phenotypic differentiation will be quite apparent as well and DNA taxonomy will tell us what we already guessed. On the contrary, consider a secondary contact of populations. If populations display differentiation, for example ecological or sexual differentiation, reinforcement will further promote the evolution of reproductive isolation. But since reproductive isolation is not absolutely perfect, selectively neutral marker loci can still penetrate species borders. In consequence, well differentiated selectively neutral markers will become associated with the wrong "species" without dissolving species borders. We would expect polyphyly or paraphyly of selectively neutral marker loci in relation to species borders and phenotypic differentiation (equivalent to species differentiation genes) will be the better choice in taxonomy. This is exactly what FUNK & OMLAND (2003) found in roughly 23% of published reports. Without phenotypic differentiation in sympatry, DNA taxonomy based on selectively neutral markers can only help if several selectively neutral marker loci display congruent pictures of reciprocal monophyly. We are not aware of a published report applying these considerations rigorously.

Acknowledgements. We would like to thank all members of the molecular unit of the ZFMK for extremely fruitful discussions. We dedicate this paper to the memory of Clas Michael Naumann who died on the 15th February 2004. His much to early death has taken away a person who has come to the aid of us whenever it was necessary; we have to grow up now without him.

REFERENCES

- ARNQVIST, G., EDVARDSSON, M., FRIBERG, U. & NILSSON, T. (2000): Sexual conflict promotes speciation in insects. *Proceedings of the national Academy of Sciences of the USA* **97**(19): 10460-10464.
- AVISE, J. C. (1987): Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics* **18**: 489-522.
- AVISE, J. C. (2000): *Phylogeography – The History and Formation of Species*. Harvard University Press, Cambridge, Massachusetts.
- BAKER, C. S., DALEBOUT, M. L., LAVERY, S. & ROSS, H. A. (2003): www.DNA-surveillance: applied molecular taxonomy for species conservation and discovery. *Trends in Ecology and Evolution* **18**(6): 271-272.
- BONNER, J. T. (1988): *The Evolution of Complexity*. Princeton University Press, Princeton, NJ.
- CHARLESWORTH, B. (1995): Down the bottleneck? *Current Biology* **5**: 995-996.
- DE QUEIROZ, K. (1999): The general lineage concept of species and the defining properties of the species category. Pp. 49-89 in WILSON, R. A. (ed.) *Species*. MIT Press, Cambridge, Massachusetts.
- DE SALLE, R. & TEMPLETON, A. R. (1988): Founder effects accelerate the rate of mtDNA evolution of Hawaiian *Drosophila*. *Evolution* **42**: 1076-1084.
- DOBZHANSKY, T. (1935): A critique of the species concept in biology. *Philosophy of Science* **2**: 344-355.
- DOBZHANSKY, T. (1937): *Genetics and the Origin of Species*. Columbia University Press, New York.
- DIECKMANN, U. & DOEBELI, M. (1999): On the origin of species by sympatric speciation. *Nature* **400**(6742): 354-357.
- DOEBELI, M. & DIECKMANN, U. (2003): Speciation along environmental gradients. *Nature* **421**(6920): 259-264.
- ENGLBRECHT, C. C., FREYHOF, J., NOLTE, A., RASSMANN, K., SCHLIEWEN, U. & TAUTZ, D. (2000): Phylogeography of the bullhead *Cottus gobio* (Pisces: Teleostei: Cottidae) suggests a pre-pleistocene origin of the major central European populations. *Molecular Ecology* **9**(6): 709-722.
- EYUALEM, A. & BLAXTER, M. (2003): Comparison of biological, molecular and morphological methods of species identification in a set of cultured *Panagrolaimus* isolates. *Journal of Nematology* **35**: 119-128.
- FISHER, R. A. (1930): *The Genetical Theory of Natural Selection*. Clarendon Press, London.
- FUNK, D. J. & OMLAND, K. E. (2003): Species-Level Paraphyly and Polyphyly: Frequency, Causes, and Consequences, with Insights from Animal Mitochondrial DNA. *Annual Review of Ecology, Evolution, and Systematics* **34**: 397-423.
- GASTON, K. J. & O'NEILL, M. A. (2004): Automated species identification-why not? *Philosophical Transactions of the Royal Society London, Biology* **359** (1444): 655-667.
- GODFRAY, H. C. J. (2002): Challenges for taxonomy. *Nature* **417**: 17-19.
- HARTL, D. L. & CLARK, A. G. (1997): *Principles of population genetics*, 3rd ed. Sinauer Associates, Sunderland, MA.
- HAASE, M., MISOF, B., WIRTH, T., BAMINGER, H., & BAUR, B. (2003): Genetic Evidence for Survival of a Land Snail under Pleistocene Permafrost. *Journal of Evolutionary Biology* **16**: 415-428.
- HEBERT, P. D. N., CYWINSKA, A., BALL, S. L. & DE-WAARD, J. R. (2001): Biological identifications through DNA barcodes. *Proceedings of the Royal Society London B* **270**, 313-322.
- HEBERT, P. D. N., PENTON, E. H., BURNS, J., JANZEN, D. J. & HALLWACHS, W. (2004): Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly, *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences of the USA* **101** (41): 14812-14817.
- JIGGINS, C. D. & DAVIES, N. (1998): Genetic evidence for a sibling species of *Heliconius charithonia* (Lepidoptera; Nymphalidae). *Biological Journal of the Linnean Society* **64**: 57-67.
- LIPSCOMB, D. L., PLATNICK, N. & WHEELER, Q. (2003): The intellectual content of taxonomy: a comment on DNA taxonomy. *Trends in Ecology and Evolution* **18**: 65-66.
- MALLET, J. & WILLMOTT, K. (2003): Taxonomy: renaissance or Tower of Babel. *Trends in Ecology and Evolution* **18**: 57-59.
- MALTAGLIATI, F., CAMILLI, L., LARDICCI, C. & CASTELLI, A. (2001): Evidence for morphological and genetic di-

- vergence in *Perinereis cultrifera* (Polychaeta, Nereididae) from two habitat types at Elba Island. Journal of the Marine Biological Association of the United Kingdom **81**: 411-414.
- MAYR, E. (1970), Populations, Species, and Evolution. Harvard University Press, Cambridge, MA.
- MAYR, E. (1982): The growth of biological thought. Diversity, Evolution and Inheritance. The Belknap Press of Harvard University Press, Cambridge, MA etc.
- MAYR, E. (1996): What is a species and what is not? Philosophy of Science **63**: 262-277.
- MAYR, E. (2001): Wu's genic view of speciation. Journal of Evolutionary Biology **14**: 866-867.
- MAYDEN, R. L. (1997): A hierarchy of species concepts: the denouement in the saga of the species problem. Pp. 381-424 in CLARIDGE, M. F., DAWAH, H. A. & WILSON, M. R. (eds.) Species - the Units of Biodiversity. Chapman & Hall, London
- MEYER, A., KOCHER, T. D., BASASIBWAKI, P. & WILSON, A. C. (1990): Monophyletic origin of Lake Victoria cichlid fishes suggested by mitochondrial DNA sequences. Nature **347**: 550-553.
- MISOF, B. (2002): Diversity of Anisoptera (Odonata): Inferring speciation processes from patterns of morphological diversity. Zoology **105**: 355-365.
- NEI, M. (1978): Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics **89**: 583-590.
- O'NEILL, M. A. & GAULD, I. D. (2001): DAISY. Automated insect identification project. <http://www.chasseuer.usc.edu/pups/projects/daisy.html>
- OHATA, T. (1976): Role of slightly deleterious molecular evolution and polymorphism. Theoretical Population Biology **10**: 254-275.
- ORR, H. A. (1995): The population genetics of speciation: the evolution of hybrid incompatibilities. Genetics **139**: 1805-1813.
- ORR, H. A. (2001): The genetics of species differences. Trends in Ecology and Evolution **16**(7): 343-350.
- ORR, H. A. & TURELLI, M. (2002): The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. Evolution **55**: 1085-1094.
- OWENS, I. P. F., BENNETT, P. M. & HARVEY, P. H. (1999): Species richness among birds: body size, life history, sexual selection or ecology? Proceedings of the Royal Society London B **266**: 933-939.
- PANHUIS, T., BUTLIN, R., ZUK, M. & TREGENZA, T. (2001): Sexual selection and speciation. Trends in Ecology and Evolution **16**(7): 364-371.
- POZHITKOV, A. & TAUTZ, D. (2002): An algorithm and program for finding sequence specific oligo-nucleotide probes for species identification. BMC Bioinformatics **3**: 9.
- PROUDLOVE, G. & WOOD, P. J. (2003): The blind leading the blind: cryptic subterranean species and DNA taxonomy. Trends in Ecology and Evolution **18**(6): 272-273.
- RAFF, R. A. (1996): The Shape of Life: Genes, Development, and the Evolution of Animal Form. University of Chicago Press, Chicago.
- RIEDE, K., LAMPE, K.-H. & INGRISCH, S. (2004): Deutsche Orthopterensammlungen (DORSA) Aufbau eines virtuellen Museums - Verhandlungen Westdeutscher Entomologentag 2002, Löbbecke Museum, Düsseldorf, 63 (DORSA: German Orthoptera Collections; <http://www.dorsa.de/>)
- RIEDL, R. (1978): Order in Living Organisms. Wiley Press, New York.
- ROTH, V., POGODA, A., STEINHAGE, V. & SCHRÖDER, St. (1999): Pattern Recognition Combining Feature- and Pixel-based Classification Within a Real World Application. 21. Jahrestagung der Deutschen Gesellschaft für Mustererkennung, DAGM '99, Bonn, Sept. 15-17, 1999, Informatik aktuell: 120-129.
- RUNDLE, H. D. (1999): Experimental tests of founder-flush: a reply to Templeton. Evolution **53**: 1632-1633.
- SEBERG, O., HUMPHRIES, C. J., KNAPP, S., STEVENSON, D., PETERSEN, G., SCHARFF, N., ANDERSEN, N. M. (2003): Shortcuts in Systematics? - A commentary on DNA-based Taxonomy. Trends in Ecology and Evolution **18**: 63-65.
- STEINHAGE, V. (2000): Automated Identification of Bee Species in Biodiversity Information Systems. Computer Science for Environmental Protection 2000, UI 2000, Bonn, Oct. 4-6, 2000, Metropolis Publ. 1: 339-344.
- STEINHAGE, V., KASTENHOLZ, B., SCHRÖDER, S. & DRESCHER, W. (1997): A Hierarchical Approach to Classify Solitary Bees Based on Image Analysis. Mustererkennung 1997, 19. DAGM-Symposium, Informatik aktuell: 419-426.
- STOECKLE, M. (2003): Taxonomy, DNA, and the Bar Code of Life. Bioscience **53** (9): 2-3.
- TAKAHASHI, K., TERAI, Y., NISHIDA, M. & OKADA, N. (2001): Phylogenetic relationships and ancient incomplete lineage sorting among cichlid fishes in Lake Tanganyika as revealed by analysis of the insertion of retroposons. Molecular Biology and Evolution **18**(11): 2057-2066.
- TAUTZ, D., ARCTANDER, P., MINELLI, A., THOMAS, R. H. & VOGLER, A. P. (2002): DNA points the way ahead in taxonomy. Nature **418**: 479.
- TAUTZ, D., ARCTANDER, P., MINELLI, A., THOMAS, R. H. & VOGLER, A. P. (2003): A plea for DNA taxonomy. Trends in Ecology and Evolution **18**: 70-74.
- TURELLI, M., BARTON, N. H. & COYNE, J. A. (2001): Theory and speciation. Trends in Ecology and Evolution **16**(7): 330-343.
- VON BERTALANFFY, L. (1932): Theoretische Biologie. 2 vols. Gebrüder Bornträger, Berlin.
- WADDINGTON, C. H. (1957): The strategy of genes. Allen & Unwin, London.
- WAGNER, G. P. (1996): Homologues, natural kinds and the evolution of modularity. American Zoologist **26**: 36-43.
- WAGNER, G. P. & ALTENBERG, L. (1996): Complex adaptations and the evolution of evolvability. Evolution **50**: 967-976.
- WATSON, A. T., O'NEILL, M. A. & KITCHING, I. J. (2004): Automated insect identification of live moths (Macrolepidoptera) using Digital Automated Identification System (DAISY): Systematics and Biodiversity **1**(3): 287-300.
- WAYNE, R. K., MEYER, A., LEHMAN, N., VAN VALKENBURGH, B., KAT, P. W., FULLER, T. K., GIRMAN, D. & O'BRIEN, S. J. (1990): Large sequence divergence among mitochondrial DNA genotypes within populations of eastern African black-backed jackals. Proceed-

- ings of the National Academy of Sciences of the USA **87**: 1772-1776.
- WEEKS, P. J. D., GAULD, I. D., GASTON, K. J. & O'NEILL, M. A. (1997): Automating the identification of insects: a new solution to an old problem. *Bulletin of Entomological Research* **87**: 203-211.
- WEEKS, P. J. D., O'NEILL, M.A., GASTON, K. J. & GAULD, I. D. (1999a): Automating insect identification: exploring the limitations of a prototype system. *Journal of Applied Entomology* **123**: 1-8.
- WEEKS, P. J. D., O'NEILL, M.A., GASTON, K. J. & GAULD, I. D. (1999b): Species-identification of wasps using principle component associative memories. *Image and Vision Computing* **17**: 861-866.
- WHEELER, Q. D. & MEIER, R. (eds) (2000): *Species Concepts and Phylogenetic Theory: a Debate*. Columbia University Press, New York
- WRIGHT, S. (1931): Evolution in Mendelian populations. *Genetics* **16**: 97-159.
- WRIGHT, S. (1951): The genetical structure of populations. *Annals of Eugenics* **15**: 323-354.
- WU, C.-I. (2001): The genic view of the process of speciation. *Journal of Evolutionary Biology* **14**: 851-865.
- WU, C.-I. & TING C.-T. (2004): Genes and speciation. *Nature Reviews, Genetics* **5**: 114-122.
- Authors' address:** Bernhard MISOF (corresponding author), Cornelya F. C. KLÜTSCH, Oliver NIEHUIS & Alexandra PATT, Zoologisches Forschungsmuseum Alexander Koenig, Adenauerallee 160, 53113 Bonn, Germany, e-mail: b.misof.zfmk@uni-bonn.de