DNA Barcoding and Its Relevance to Pests, Plants and Biological Control

R.H. Hanner, J. Lima and R. Floyd
Biodiversity Institute of Ontario & Department of Integrative Biology
University of Guelph
Guelph, Ontario, N1G 2W1
Canada

Abstract

Sustainability of modern food production requires the ability to confidently monitor, identify and distinguish a wide variety of insect and plant species. Virtually all industries are susceptible to direct damage from insect pests and weeds, and also from associated diseases transmitted by them. Additionally, in order to limit reliance on pesticide usage, the importation and use of biocontrol agents will increase as an option to reduce economic losses while limiting any negative impacts on the environment. Many insect or plant species encountered by farmers, gardeners, government officials or researchers cannot be easily identified in a timely manner, and in cases where a non-indigenous species is discovered, delays in deploying appropriate management strategies can have serious economic and environmental consequences. A novel methodology known as DNA barcoding has the potential to mitigate the challenges posed by identification of insect and plant species. By comparing mitochondrial DNA sequences derived from an unknown specimen against a reference database of DNA profiles of known species, genetic identifications are possible. This approach has several benefits over the existing system of morphological identification: it is applicable to all sexes and life stages, as well as to fragmentary remains, and is typically more accurate, rapid and cost-effective than rearing immature stages to adulthood for identification. The effectiveness of DNA barcoding as an identification tool will continue to improve as the reference database is further expanded. Here we discuss DNA barcoding as a method which aims to facilitate species discrimination using standardised and largely automated techniques, and which promises both to accelerate insect and plant diagnostics and to broaden the accessibility of data.

INTRODUCTION

The movement of animal and plant species as an unintended consequence of human activity can only be expected to increase in the future, due to growing international trade, climate change, and increasing human mobility. In some cases, species become established outside of their native ranges, with adverse environmental or economic consequences (Pimentel et al., 2000, 2005; Colautti et al., 2006). Such non-indigenous species (NIS) are frequently termed "invasive", and famous North American examples include insect pest species such as bark beetles (Coleoptera: Curculionidae) and the gypsy moth *Lymantria dispar* (Lepidoptera: Lymantriidae). More recently, an invasive population of the potato psyllid (*Bactericera cockerelli*; Hemiptera: Psyllidae) was responsible for extensive damages on tomatoes in the south-western United States (Liu and Trumble, 2007). Unfortunately, predicting which arthropod species could potentially become invasive outside of their natural range is not a straightforward process (Worner and Gevrey, 2006).

Tomato processing, like virtually all industries based on food production and agriculture, suffers significant economic impact from pathogens and pests, including insects. Tomatoes may be directly damaged by the feeding activities of a variety of insects such as tomato fruitworms and cabbage loopers (Lepidoptera: Noctuidae), vegetable leafminers (Diptera: Agromyzidae), flea bettles (Coleoptera: Chrysomelidae),

and stink bugs (Hemiptera: Pentatomidae) (Letourneau and Goldstein, 2001) or by disease transmission such as Tomato Spotted Wilt Tospovirus vectored by thrips (Thysanoptera: Thripidae) (Kahn et al., 2005). An additional economic cost of tomato production is associated with the necessary identification of any insects or mites that are found on tomatoes crossing the border since they must be identified before the restrictions on the movement of the goods is cleared. The excessive amount of time necessary for morphological identification of unknown insects, often requiring days to weeks, can be disastrous for such shipments of perishable goods. Sustainability of modern agriculture will need to reduce pesticide-intensive activities while incorporating the ability to confidently monitor and identify a wide variety of insects, including pests, biocontrol agents, and surrounding native species.

Biological control is a method commonly employed in regulating the population of harmful or undesirable species by the use of their natural enemies such as predators and parasitoids (biocontrol agents). Biological control is increasingly seen as desirable since it minimizes the use of chemicals, such as pesticides and herbicides, which may have adverse environmental effects; and also because the regulatory effect can persist for a long time with little to no cost or subsequent human management input required (Losey and Vaughan, 2006). The importance of biocontrol will increase as NIS continue to be discovered and their negative effect on economic growth becomes noticeable (Pemberton, 2003). However, biocontrol agents can themselves have unintended negative impacts, for example if an introduced parasitoid attacks another biocontrol agent that had been introduced previously to control an insect or plant species Once an introduced biocontrol agent is released into an environment, it cannot be recalled if it disperses into new habitats or causes non-target effects by attacking novel species (Kuhlmann et al., 2006). Therefore, like surveillance of NIS, biocontrol programs serve as an example of a practice whose success depends on a thorough and accurate knowledge of species taxonomy, biology and ecology (Gillespie et al., 2006; Heraty et al., 2007; Gariepy et al., 2008). Yet deployment of biocontrol through farmers, gardeners or government officials, is mediated by those with little specialist biological expertise, making it essential that accurate information is available and accessible to the general public. Furthermore, many biocontrol agents are sold on a commercial basis, yet there is little standardisation ensuring, for example, that what is sold by different companies under the same name is in fact the same species (Bely and Weisblat, 2006).

All effort to monitor and control harmful insect species depends upon our ability to quickly and accurately identify problematic species and beneficial biocontrol agents among the broad range of biodiversity that may be encountered. A rapid and widely accessible identification system is essential to implement management strategies such as quarantine or eradication with minimal delay after an NIS is detected (Carnegie et al., 2006). Here we discuss DNA barcoding as a method which aims to facilitate species discrimination by using standardised and largely automated techniques which promises to both accelerate insect diagnostics and broaden the accessibility of data. Furthermore, though DNA barcoding is more commonly employed with studies involving animal species (Floyd et al., 2002; Hebert et al., 2003), we also suggest the importance of using a similar molecular diagnostic technique in identifying invasive plant species as early as possible.

DNA BARCODING

The concept of molecular barcoding - proposed by Floyd et al. (2002), and developed into a global biological identification system (Hebert et al., 2003) - is that it is possible to rapidly and accurately identify a species by amplifying and sequencing a short, standardised region of its genome, specifically the mitochondrial gene, cytochrome *c* oxidase subunit 1 (COI) (Hebert et al., 2003). Because this method depends on DNA and not morphology, it is applicable to any life stage, from egg to adult. It also aims to employ standardised protocols that may be applied to a wide range

of organisms with a minimum of technical expertise and without requiring extensive knowledge of traditional morphological taxonomy. Used as a component of regular monitoring, DNA barcoding has the potential to detect invasions early, providing authorities with the time to develop containment and eradication strategies before populations of a NIS become unmanageable (Armstrong and Ball, 2005; Chornesky et al., 2005; Carnegie et al., 2006; Haack, 2006; Scheffer et al., 2006). For commercially-available biocontrol agents, barcoding could provide a standard ID system to ensure that the correct species are sold, and that agents sold under the same species name are indeed the same species (Bely and Weisblat, 2006).

The typical workflow of constructing a barcode reference database is shown in Figure 1. From each specimen to be barcoded, a tissue sample (e.g. insect leg) is taken, and used to generate a DNA sequence. The original specimen is photographed and kept as a morphological voucher. The image and all collateral information (identification, collection data, etc) are stored along with the DNA barcode sequence in a custom database. In this way, a reference database of COI barcode sequences matched to known species names and images is built up; any novel COI sequence from an unknown specimen may be searched against the database and, if a match is found, species identification may be rapidly made.

The reference database of known COI sequences is a vital component of barcoding. The Barcode of Life Data Systems, or BOLD (Ratnasingham and Hebert, 2007) is a web-based platform which acts as both a universal repository for DNA barcode data and an online workbench with tools for data analysis. The DNA barcoding enterprise also requires rigorous standardisation of both methodology and data handling if it is to be successful in its goals. In addition to the COI sequence, fully validated barcode records also require complete taxonomy of the specimen from which they derive, photographs, GPS-coordinate data specifying the location from which it was collected, and a trackable ID linking the record to a voucher specimen which must be deposited in a suitable repository (museum). The frozen DNA extract is also typically archived for potential future analysis.

DNA barcoding does not strive to be the sole method for identifying animal life, and once standardized COI reference sequences are collected, submitted, and approved, supplementary nuclear gene(s) for specific taxa can be incorporated to strengthen a global identification system. It is clear that DNA barcoding complements traditional morphology-based and nuclear gene-based identification methods for achieving maximum accuracy and efficiency of arthropod species identification.

DISCUSSION

Diagnosis of pest species is clearly an important step as strategies for the control of NIS require, at minimum, the ability to detect and identify an invader in an accurate and timely manner. Yet too often, this is a slow and inefficient process, and officials such as port inspectors typically do not have the expertise to recognise the vast range of harmful species which might potentially be intercepted. Furthermore, many arthropod species have life-cycle stages, such as eggs and larvae, which cannot be identified by conventional means (Pogue and Schaefer, 2007). DNA barcoding revealed that a single species of tomato psyllid (*Bactericera cockerelli*, Hemiptera: Psyllidae) may contain two distinct genetic and geographic populations (Liu et al., 2006). One of the populations is considered native and conforms to its normal range while the other population is considered to be invasive and is expanding its geographic range. Due to analysis of the DNA barcodes, coupled with other molecular and ecological data, studies are continuing to understand the similarities and differences between these distinct biotypes of the tomato psyllid (Liu and Trumble, 2007).

Key advantages of the DNA barcoding protocol are its adoption of rigorous data standards (Hanner, 2005), and traceability at all stages of the process. Barcoding could also improve traceability in other aspects of commercial industry; for sample, tomatoes may be grown in one country, shipped to another for canning, and sold in numerous

other countries. Suppose some shipment becomes contaminated, such as an insect or fungal infection being found in a can. It becomes important to determine at what stage of production that contaminant was introduced. DNA barcoding could rapidly inform food inspectors whether, for example, the insect belonged to a species native to the country in which the tomatoes were grown (and therefore was most likely introduced before processing), or the country in which canning took place; such forensic determinations could have both financial and legal implications.

Though the majority of studies on DNA barcoding thus far have been with animals, researchers are also exploring the idea of similarly using common molecular markers for identification of plants (Chase et al., 2005; Ivanova et al., 2008). As with animals intercepted at ports of entry, many plant species have life-cycle stages, such as seeds, which cannot be identified by conventional means (Pogue and Schaefer, 2007). Unfortunately, it has proven less straightforward to establish a universal marker for plants, because the COI gene in plants has a far lower rate of nucleotide substitutions and hence shows insufficient levels of variation to discriminate species (Newmaster et al., 2006). As a consequence, researchers have instead suggested a tiered approach involving a number of different genes to arrive at a conclusion of plant species identity. For example, the plastid gene *rbcL* has been recommended as a "core" marker, being present and easily amplified from all land plants, which should show sufficient variability to allow identification to family level, and then alternative markers would be sequenced to permit identification to species level as required (Newmaster et al., 2006; Chase et al., 2007; Kress and Erickson, 2007).

THE FUTURE

The effectiveness of barcoding as an identification tool for unknowns will only improve as the reference database is further expanded. It is therefore important to prioritise the data injection of the most economically important species, such as pests, NIS and biocontrol agents. Given the promise that DNA barcoding holds for early detection of NIS and to provide monitoring for biological control programs, we advocate the development of policy at the national and international levels to create the global database, and to enhance its universal access (e.g. an international Barcode of Life project; see: www.dnabarcoding.org). The international mobility of goods, vessels, and people necessitates that any DNA barcoding and management program designed for a given country be incorporated into a common global database to facilitate species identification, knowing their region of origin, and risk assessment. We view this information as critical for minimizing a variety of non-tariff barriers to trade.

While the barcoding protocol currently requires a well-equipped molecular lab to generate DNA sequences, technological advances continue to be made in speed, cost -effectiveness and miniaturisation. A number of groups are actively working on portable DNA sequencing devices, and thus a point-of-contact or handheld barcoding device is envisioned for the future (Janzen, 2004; Savolainen et al., 2005). Such a device would revolutionise the ability to make immediate biological identifications in the field.

DNA barcoding offers the potential not only for enhanced recognition and thereby control of harmful species in the immediate term, it also promises to help us to understand the global movement of species in a level of detail never before attainable. DNA barcoding is not a means to limit the introduction of NIS nor to increase the safe use of newly introduced biocontrol agents; however, it is a tool that may be simultaneously employed by researchers, environmentalists, and the general public for increased vigilance of human and ecosystem health. The Canadian Barcode of Life Network is dedicated to advancing barcode protocols for animals, plants, protists and fungi. The Network also supports the development and parameterization of the BOLD database, with an early emphasis on building barcode coverage for agricultural pests and biological control agents. At the time of this writing (June, 2008) some 40,000 species are represented by more than 400,000 individual specimens barcoded. At this

rate, the future of automated biological identifications appears bright indeed.

ACKNOWLEDGEMENTS

Literature Cited

Armstrong, K.F. and Ball, S.L. 2005. DNA barcodes for biosecurity: invasive species identification. Philosophical Transactions of the Royal Society B: Biological Sciences 360:1813-1823.

Bely, A.E. and Weisblat, D.A. 2006. Lessons from leeches: a call for DNA barcoding in the lab. Evolution & Development 8:491-501.

Carnegie, A.J., Matsuki, M., Haugen, D.A., Hurley, B.P., Ahumada, R., Klasmer, P., Sun, J. and Iede, E.T. 2006. Predicting the potential distribution of Sirex noctilio (Hymenoptera: Siricidae), a significant exotic pest of Pinus plantations. Annals of Forest Science 63:119-128.

Chase, M.W., Cowan, R.S., Hollingsworth, P.M., van den Berg, C., Madriñán, S., Petersen, G., Seberg, O., Jørgsensen, T., Cameron, K.M., Carine, M., Pedersen, N., Hedderson, T.A.J., Conrad, F., Salazar, G.A., Richardson, J.E., Hollingsworth, M.L., Barraclough, T.G., Kelly, L. and Wilkinson, M. 2007. A proposal for a standardised protocol to barcode all land plants. Taxon 56:295-299.

Chase, M.W., Salamin, N., Wilkinson, M., Dunwell, J.M., Kesanakurthi, R.P., Haidar, N. and Savolainen, V. 2005. Land plants and DNA barcodes: short-term and long-term goals. Philosophical Transactions of the Royal Society B: Biological Sciences 360:1889-1895.

Chornesky, E.A., Bartuska, A.M., Aplet, G.H., Britton, K.O., Cummings-Carlson, J., Davis, F.W., Eskow, J., Gordon, D.R., Gottschalk, K.W., Haack, R.A., Hansen, A.J., Mack, R.N., Rahel, F.J., Shannon, M.A., Wainger, L.A. and Wigley, T.B. 2005. Science priorities for reducing the threat of invasive species to sustainable forestry. Bioscience 55:335-348.

Colautti, R., Bailey, S., van Overdijk, C., Amunsen, K. and MacIsaac, H.J. 2006. Characterized and projected costs of nonindigenous species in Canada. Biological Invasions 8:45-59.

Floyd, R., Abebe, E., Papert, A. and Blaxter, M. 2002. Molecular barcodes for soil nematode identification. Molecular Ecology 11:839-850.

Gariepy, T., Kuhlmann, U., Gillott, C. and Erlandson, M. 2008. A large-scale comparison of conventional and molecular methods for the evaluation of host-parasitoid associations in non-target risk-assessment studies. Journal of Applied Ecology 45:708-715.

Gillespie, D.R., Mason, P.G., Dosdall, L.M., Bouchard, P. and Gibson, G.A.P. 2006. Importance of long-term research in classical biological control: an analytical review of a release against the cabbage seedpod weevil in North America. Journal of Applied Entomology 130:401-409.

Haack, R.A. 2006. Exotic bark- and wood-boring Coleoptera in the United States:

recent establishments and interceptions. Canadian Journal of Forest Research 36: 269-288.

Hanner, R. 2005. Proposed Standards for BARCODE Records in INSDC (BRIs). Consortium for the Barcode of Life. Accesed June 2008 at the following url: barcoding.si.edu/PDF/DWG_data_standards-Final.pdf

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

Heraty, J.M., Woolley, J.B., Hopper, K.R., Hawks, D.L., Kim, J. and Buffington, M. 2007. Molecular phylogenetics and reproductive incompatibility in a complex of cryptic species of aphid parasitoids. Molecular Phylogenetics and Evolution 45:480-493.

Ivanova, N.V., Fazekas, A.J. and Hebert, P.D.N. 2008. Semi-automated, Membrane-Based Protocol for DNA Isolation from Plants. Plant Mol. Biol. Rep. (in press)

Janzen, D. 2004. Now is the time. Philosophical Transactions of the Royal Society B: Biological Sciences 359:731-732.

Kahn, N.D., Walgenbach, J.F. and Kennedy, G.G. 2005. Summer weeds as hosts for *Frankliniella occidentalis* and *Frankliniella fusca* (Thysanoptera: Thripidae) and as reservoirs for Tomato Spotted Wilt Tospovirus in North Carolina. Journal of Economic Entomology 98:1810-1815.

Kress, W.J. and Erickson, D.L. 2007. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding trnH-psbA spacer region. PLoS ONE 2, e508.

Kuhlmann, U., Mason, P.G., Hinz, H.L., Blossey, B., De Clerck-Floate, R.A., Dosdall, L.M., McCaffrey, J.P., Schwarzlaender, M., Olfert, O., Brodeur, J., Gassmann, A., McClay, A.S. and Wiedenmann, R.N. 2006. Avoiding conflicts between insect and weed biological control: selection of non-target species to assess host specificity of cabbage seedpod weevil parasitoids. Journal of Applied Entomology 130:129-141.

Letourneau, D.K. and Goldstein, B. 2001. Pest damage and arthropod community structure in organic vs. conventional tomato production in California. Journal of Applied Ecology 38:557-570.

Liu, D. and Trumble, J.T. 2007. Comparative fitness of invasive and native populations of the potato psyllid (*Bactericera cockerelli*). Entomologia Experimentalis et Applicata 123:35-42.

Liu, D., Trumble, J.T. and Stouthamer, R. 2006. Genetic differentiation between eastern populations and recent introductions of potato psyllid (*Bactericera cockerelli*) into western North America. Entomologia Experimentalis et Applicata 118:177-183.

Losey, J.E. and Vaughan, M. 2006. The economic value of ecological services provided by insects. Bioscience 56:311-323.

Newmaster, S.G., Fazekas, A.J. and Ragupathy, S. 2006. DNA barcoding in land plants: evaluation of *rbcL* in a multigene tiered approach. Canadian Journal of Botany-Revue Canadienne De Botanique 84:335-341.

Pemberton, R.W. 2003. Potential for biological control of the lobate lac scale, *Paratachardina lobata lobata* (Hemiptera: Kerriidae). Florida Entomologist 86: 353-360.

Pimentel, D., Lach, L., Zuniga, R. and Morrison, D. 2000. Environmental and economic costs of nonindigenous species in the United States. Bioscience 50:53-65.

Pimentel, D., Zuniga, R. and Morrison, D. 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecological Economics 52:273-288.

Pogue, M.G. and Schaefer, P.W. 2007. A review of selected species of *Lymantria* Hübner [1819] including three new species (Lepidoptera: Noctuidae: Lymantriinae) United States Department of Agriculture, Forest Health Technology Enterprise Team.

Ratnasingham, S. and Hebert, P.D.N. 2007. The barcode of life data systems

(www.barcodinglife.org). Molecular Ecology Notes 7:355-364.

Savolainen, V., Cowan, R.S., Vogler, A.P., Roderick, G.K. and Lane, R. 2005. Towards writing the encyclopedia of life: an introduction to DNA barcoding. Philosophical Transactions of the Royal Society B: Biological Sciences 360:1805-1811.

Scheffer, S.J., Lewis, M.L. and Joshi, R.C. 2006. DNA barcoding applied to invasive leafminers (Diptera: Agromyzidae) in the Philippines. Annals of the Entomological Society of America 99:204-210.

Worner, S.P. and Gevrey, M. 2006. Modelling global insect pest species assembalges to determine risk of invasion. Journal of Applied Ecology 43:858-867.
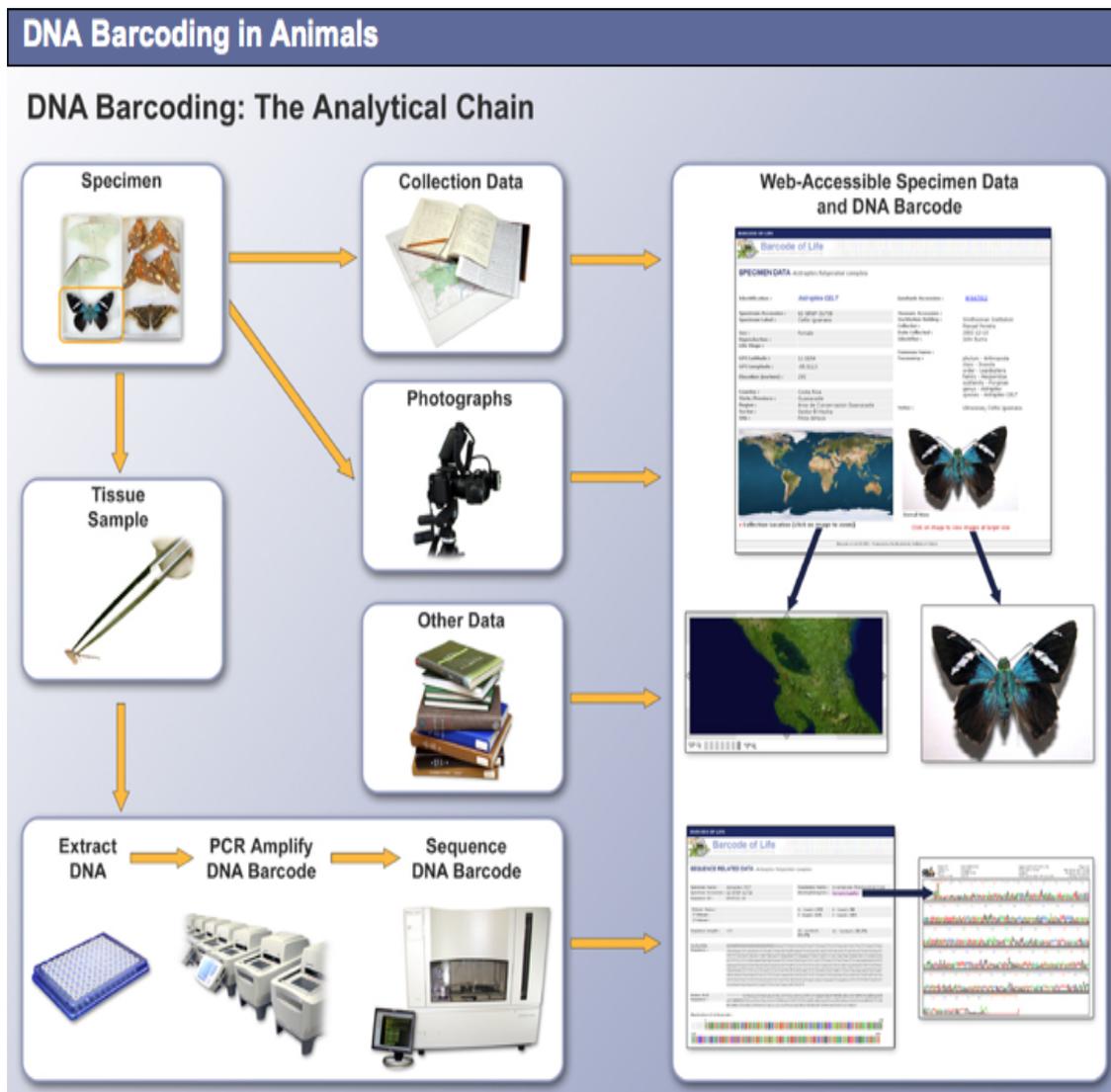
Figures

Fig. 1. The workflow of DNA barcoding. A tissue sample (e.g. insect leg) is taken from a specimen, and used to generate a DNA sequence. The original specimen is photographed and kept as a voucher. The image and all collateral information (identification, collection data, etc.) are stored along with the DNA barcode sequence in the BOLD database.