Institution: Imperial College London

Unit of Assessment: 5 Biological Sciences

Title of case study: 6 - DNA Barcoding of Life: Development of DNA-based Species Identification Technologies

1. Summary of the impact (indicative maximum 100 words)

Taxonomy is of key relevance to the environment, agriculture, food production, and human health. However, describing all living organisms is such a daunting task that it calls for new approaches. A DNA-based system for species identification, called 'DNA Barcoding', is one such solution. Imperial researchers identified DNA barcodes for plants in 2008, which have since had impacts on the environment, health and welfare and in commerce. The plant DNA barcodes have been endorsed by the Consortium for the Barcoding of Life and have led to multiple applications ranging from facilitating biodiversity inventories, helping authentication of material (herbal medicine) for trade control in Malaysia, South Africa, India and Nigeria, and combating invasive species and smuggling in Africa.

2. Underpinning research (indicative maximum 500 words)

- A DNA-based system for species discovery, identification and delimitation could complement, or even substitute, the existing centuries-old taxonomic system. It would 'democratise' species identification even for non-specialists, while being applicable to all life stages and partially preserved specimens. This approach - either called DNA Barcoding or DNA Taxonomy - requires a reference sequence database at species level to ultimately catalogue and provide identifications for the estimated 10 million species on Earth. The implementation of such system requires the discovery of suitable genetic markers for species identification (i.e. 'DNA Barcodes'), biologically meaningful methods for defining species entities, and a proof-of-principle of the methods in complex ecosystems. Researchers in the Division of Ecology and Evolution at Imperial have been instrumental in designing and trialling a DNA-based taxonomic approach.

- Research was carried out between 2005 and 2012.

- **Key researchers:**
  Professor Vincent Savolainen, Professor of Organismic Biology, 01/10/07 - present
  Professor Timothy Barraclough, Professor of Evolutionary Biology, 01/11/96 – present
  Professor Alfried Vogler, Professor of Molecular Systematics, 01/01/95 – present

- **Key research insights**: Professor Alfried Vogler was instrumental in developing the basic principle for animals and, together with Professor Vincent Savolainen (then based at the Royal Botanical Gardens, Kew), was involved in developing an international consortium of major natural history museums [3]. Professor Timothy Barraclough's key contribution was the design of a method for DNA-based species delimitation [4], which was trialled in insect tropical communities. Finally, Professor Vincent Savolainen and Professor Barraclough's research led to development of a universal DNA barcode for plants [1, 2], a particularly challenging problem - e.g., as reported in *Science*: 'Wanted: a DNA barcode for Plants' (318:190, 2007). Specifically, Professor Savolainen combined results from research in Costa Rica and Africa and, together with Professor Barraclough, demonstrated the feasibility of DNA barcoding plants using the *matK* locus. They found that this genetic marker exhibited the necessary 'barcoding gap'; that is, it is well conserved within a plant species but varies extensively between them. This DNA barcode was trialled in (sub)tropical plant communities, leading to the discovery of a new species of orchids and demonstrating its use for identifying taxa covered by the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) regulations [1]. The corresponding paper was endorsed by the International Consortium for the Barcoding of Life (e.g. [2]). The papers above have been very quickly highly cited and the methodologies have been applied across most groups of animals and plants. Ongoing research is validating methods for DNA taxonomy from environmental DNA samples in animals [5]. We published additional technical protocols to facilitate the use of *matK* by the broadest possible stakeholder audience [6].

3. References to the research (* References that best indicate quality of underpinning research)

[1] *Lahaye R., van der Bank M., Bogarin D., Warner J., Pupulin F., Gigot G., Maurin O., Duthoit S., Barraclough T. G., Savolainen V., '*DNA barcoding the floras of Biodiversity hotspots',* PNAS 105: 2923-2928 (2008). DOI, **209 citations (as at 12/11/12)**

[2] CBOL Plant Working Group (incl. Savolainen V.)., '*A DNA barcode for land plants',* PNAS 106:12794-12797 (2009). DOI, **256 citations (as at 12/11/12)**

[3] *Savolainen, V., Cowan, R.S., Vogler, A.P., Roderick, G., Lane, R., '*Towards writing the encyclopaedia of life: an introduction to DNA barcoding'*. Phil. Trans. Roy. Soc. B 360, 1805-1811 (2005). DOI, **182 citations (as at 12/11/12)**

[4] *Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. '*Sequence-based species delimitation for the DNA taxonomy of undescribed insects',* Syst. Biol. 55, 595–609 (2006). DOI, **237 citations (as at 12/11/12)**

[5] Tang, C.Q., Leasi, F., Obertegger, U., Kieneke, A., Barraclough, T.G., Fontaneto, D., '*The widely used small subunit 18S rDNA molecule greatly underestimates true diversity in biodiversity surveys of the meiofauna*', PNAS 109: 16208-16212 (2012). DOI, **1 citation (as at 25/4/13)**

[6] Dunning, L.T., Savolainen, V., '*Broad-scale amplification of matK for DNA barcoding plants, a technical note',* Bot. J. Linn. Soc. Vol 164:1 1-9 (2010). DOI, **16 citations (as at 25/4/13)**

Grant Support:

[G1] V. Savolainen (PI), 'Regional patterns of biodiversity and conservation in South Africa: the flora of the Kruger National Park as a case study', The Royal Society, 1/9/07-31/3/13, £240,735

[G2] A. Vogler (PI), 'Large-scale DNA sequencing in taxonomy', BBSRC, 01/07/04-30/09/07, BBS/B/04358, £329,549

[G3] T. Barraclough (PI), 'The evolutionary characterisation of bacterial diversity from DNA sequence data', BBSRC, BB/G004250/1 , 01/02/09-20/04/12, £316,984

4. Details of the impact (indicative maximum 750 words)

Nature of the impact: The DNA-based method has truly changed the approach to taxonomic questions and the use of taxonomy. Until the publication by Savolainen & Barraclough [1], the botanical community relied heavily on morphology-based species identification from well-preserved specimens, struggling to reach a consensus on a suitable DNA barcode. Paper [1] set "*the community standard*" [A] for DNA barcoding. After we showed the appropriateness of the *matK* locus as barcodes, DNA-based authentication from any part of a plant, even reduced to powder, was made possible and led to various applications as described below. We published additional technical protocols to facilitate the use of *matK* by the broadest possible stakeholder audience [6]. The Consortium for the Barcoding of Life (CBOL) endorsed *matK* as one of the barcodes for plants in its 2009 PNAS paper [2], and the technique is now used in "*hundreds of laboratories around the world*" [A].

Beneficiaries: the taxonomic community, museum scientists, botanical gardens, the International Barcode of Life Project, government bodies and agencies involved in trade control and consumer fraud.

Significance: For centuries, all users of species identification have relied on time-consuming morphology-based expertise by a few specialists of each taxonomic group. This expertise has been reduced severely over the past decades (Systematics and Taxonomy Report 2008; House of Lords Science and Technology Committee). With the advances of DNA barcoding, taxonomic identification has become easier and much more affordable to non-specialist users. As an example of the widespread utility of DNA barcoding, the International Barcode of Life Project (iBOL), a 25-nation consortium, aims to create 5 million barcode records from 500,000 species (including plants) that are endangered, of particular socio-economic importance or used in environmental assessment by 2015 [B].

Date of Impact: 2008-2013

Impact on environment:

- Biodiversity inventories have been facilitated in numerous public bodies and charities by using DNA barcoding protocols using the *matK* locus method. For example:
    - The Botanical Garden of Wales has been barcoding the flora of Wales, which has about 75% of UK flowering plants [C], intensively. The Barcode Wales project has created a reference database of DNA barcodes based on "*the 1143 native flowering plants and conifers of Wales, assembling over 5700 DNA barcodes*" [D]. This has made Wales "*the first country in the world to DNA barcode all of its flowering plants*" [D]. The Barcode Wales project has created a powerful platform for a broad range of research from biodiversity conservation to human health and "*Welsh flora DNA barcodes are freely available on the Barcode of Life Database (BOLD) for use by researchers throughout the world*" [D, E].
    - Following the completion of the Barcode Wales project, the team at National Botanic Gardens of Wales has joined forces with teams from the Royal Botanic Garden Edinburgh and the Royal Botanic Gardens Kew to DNA barcode the rest of the UK flora [E]. As a further example, The Royal Botanical Gardens at Kew confirms that it has been barcoding medical plants of Nigeria and Gabon using the *matK* marker [F].
    - In Southern Africa, Kruger National Park (South Africa) has been barcoding its woody flora (Yessoufou et al., J. Ecol., 2013, DOI). By mid-2010 "*most taxa of trees and shrubs (93 percent) [had been] completed for the rbcL and matK regions*" [G]. The barcodes of woody vegetation in the Cheringoma District in Mozambique have been collected since 2010 [G]. The African Centre for DNA Barcoding (ACDB) lists a total of 9 plant DNA barcoding projects in Southern Africa and reports that 11,451 specimens and 6,932 species have been barcoded (as of Jan 2013) [H].
- In South Africa, a DNA barcoding project has started with the aim of stopping the smuggling of endangered cycad species. The DNA database will enable customs officials to identify specimens and help prevent the illegal trade [I]. A database for 50 tree species from the Convention on International Trade in Endangered Species of Wild Flora and Fauna (CITES) that are expected to be encountered at South African ports and borders is also being built [G].
- DNA barcoding is being used to combat alien invasive species in South Africa, where over 500 naturalized plant species listed are known to have entered once pristine habitats. Alien plant species are of growing concern in South Africa due to their negative impacts on biodiversity. In 2011, a government-funded collaborative project between the ACDB, the South African National Biodiversity Institute and the Early Detection and Rapid Response Programme of the Working for Water Programme was formed as part of the iBOL project. The project collected invasive plant and animal samples in South Africa, as well as South African species known to be invasive elsewhere in the world, over an eight-month period for DNA barcoding and provided access to the barcode data through the BOLD system (see below) [J].

Impact on economy/commerce:

- The *matK* barcode and protocols to use it have been implemented in the BOLD systems in Canada, which provides species identification commercially. In April 2013, BOLD listed 176,905 species formally DNA barcoded (plants and animals) [K]. BOLD came into existence at the beginning of the DNA barcoding movement, funded by the Canadian government. Until the research reported here, it was only archiving DNA barcodes of animals. After references [1] and [2], it has been archiving and making available plant DNA barcodes to a large range of users.

Impact on health/welfare:

- Consumer fraud prevention in teas and herbal medicines now uses protocols based on DNA barcoding protocols (as reported at the International Meeting of DNA Barcoding of Life, Adelaide, Australia, 2011). Demonstrating the importance of accessible plant barcoding, Stoeckle et al. findings indicate unlisted ingredients are common in herbal teas (Scientific Reports, 1, 42, 2001, DOI). Government bodies, museums and botanical gardens are using these techniques, although they might ask university laboratories to carry out the work. By 2011, a library with DNA barcodes for “Malaysia's 1,200 plant species with potential medicinal value” was in development to offer “‘a quick one step detection kit’ to reduce fraud in the lucrative herbal medicine industry” [L]. At that time, DNA barcode libraries were also under construction for the medicinal plants of several other nations, including South Africa, India and Nigeria [L].
- An integrated web medicinal materials DNA database, MMDBD (Medicinal Materials DNA Barcode Database, [M]) was also created to archive plant DNA barcodes for use in authentication of herbal medicines, particularly Chinese medicines (published in *BMC Genomics* 11:402, 2010, DOI). The database can be accessed by anybody needing to identify medicinal plants with DNA barcoding and had archived 1,658 species and 31,468 sequences by May 2012 [M].

5. Sources to corroborate the impact (indicative maximum of 10 references)

[A] Letter from Chair of the Plant Working Group of the Consortium for the Barcoding of Life, Botanical Garden Edinburgh, 26/11/12 (available from Imperial on request)
[B] The International Barcode of Life project, ‘What Is iBOL?’, http://www.barcodeoflife.org/content/about/what-ibol (archived at https://www.imperial.ac.uk/ref/webarchive/sjf on 26/4/13)
[C] BBC News, ‘National Botanic Garden of Wales logs plant DNA barcode’, 26/6/12, http://www.bbc.co.uk/news/uk-wales-18590298 (archived at https://www.imperial.ac.uk/ref/webarchive/vjf on 26/4/13)
[D] Letter from Head of Conservation and Research, National Botanic Garden of Wales, 13/11/12 (available from Imperial on request)
[E] Barcode Wales, http://www.gardenofwales.org.uk/science/barcode-wales/ (archived at https://www.imperial.ac.uk/ref/webarchive/wjf on 26/4/13)
[F] Letter from Director of the Jodrell Laboratory, Royal Botanic Gardens, Kew, 28/11/12 (available from Imperial on request)
[G] Barcode Bulletin, Vol.1, No.2, June 2010, pages 16 & 14, ‘Q&A with Michelle van der Bank’, (archived here)
[H] African Centre for DNA Barcoding, http://acdb.co.za (archived at https://www.imperial.ac.uk/ref/webarchive/yjf on 26/4/13)
[I] Cape Times, ‘Scientists use DNA barcodes to fight cycad smuggling’ 12/1/10, http://acdb.co.za/uploads/File/Press/47.Cape%20Times%2012.01.2010.pdf (archived here)
[J] DNA barcoding of terrestrial and fresh water invasive species in South Africa, ACDB, http://acdb.co.za/index.php/sa-invasive-project-new/project-presentation-2.html (archived at https://www.imperial.ac.uk/ref/webarchive/1jf on 26/4/13)
[K] BOLD Systems, http://www.boldsystems.org/index.php/TaxBrowser_Home?target=Plant (archived at https://www.imperial.ac.uk/ref/webarchive/dkf on 29/4/13)
[L] News-Medical.net article, “New creative uses of DNA 'barcoding'”, 28/11/11, http://www.news-medical.net/news/20111128/New-creative-uses-of-DNA-barcoding.aspx (archived at https://www.imperial.ac.uk/ref/webarchive/5jf on 26/4/13)
[M] Medical Materials DNA Barcode Database, http://137.189.42.34/mherbsdb/ (archived at https://www.imperial.ac.uk/ref/webarchive/6jf on 26/4/13)