Article

Building a Plant DNA Barcode Reference Library for a Diverse Tropical Flora: An Example from Queensland, Australia

Craig M. Costion 1,*, Andrew J. Lowe 2,†, Maurizio Rossetto 3, Robert M. Kooyman 3,4, Martin F. Breed 2, Andrew Ford 5 and Darren M. Crayn 6,†

1 Botany Department, National Museum of Natural History, MRC 166, Smithsonian Institution, PO Box 37012, Washington, DC 20013-7012, USA
2 School of Biological Sciences and the Environment Institute, University of Adelaide, Adelaide 5005, Australia; andrew.lowe@adelaide.edu.au (A.J.L.); martin.breed@adelaide.edu.au (M.F.B.)
3 National Herbarium of NSW, Royal Botanic Garden Sydney, Mrs Macquaries Road, Sydney 2000, Australia; maurizio.rossetto@rbgsyd.nsw.gov.au
4 Department of Biological Sciences, Macquarie University, Sydney 2109, Australia; robert.kooyman@mq.edu.au
5 CSIRO Land and Water, Tropical Forest Research Centre, Atherton 4883, Australia; andrew.ford@csiro.au
6 Australian Tropical Herbarium, James Cook University, Smithfield 4878, Australia; darren.crayn@jcu.edu.au
* Correspondence: ccostion@gmail.com; Tel.: +1-202-633-0920 (ext. 123); Fax: +1-202-786-2563
† These authors contributed equally to this work.

Academic Editor: Michael Wink
Received: 19 October 2015; Accepted: 24 February 2016; Published: 29 February 2016

Abstract: A foundation for a DNA barcode reference library for the tropical plants of Australia is presented here. A total of 1572 DNA barcode sequences are compiled from 848 tropical Queensland species. The dataset represents 35% of the total flora of Queensland’s Wet Tropics Bioregion, 57% of its tree species and 28% of the shrub species. For approximately half of the sampled species, we investigated the occurrence of infraspecific molecular variation in DNA barcode loci rbcL, matK, and the trnH-psbA intergenic spacer region across previously recognized biogeographic barriers. We found preliminary support for the notion that DNA barcode reference libraries can be used as a tool for inferring biogeographic patterns at regional scales. It is expected that this dataset will find applications in taxonomic, ecological, and applied conservation research.

Keywords: biogeography; wet tropics; rbcL; matK; trnH-psbA intergenic spacer

1. Introduction

Tropical rain forests present unique challenges to identifying plant species. They are characteristically diverse, with an area as small as two hectares potentially containing 300+ vascular woody species of trees, shrubs, and lianas [1]. Many of these species can be extremely rare and/or poorly known [2] and often their most diagnostic characters for identification (e.g., leaves, fruit and flowers) occur high in the canopy out of sight and reach. In many cases fruits and/or flowers are required for accurate identification, which can hinder progress on identification of tropical species in remote localities for years or even decades. This combination of factors renders identification of species in tropical forests to be characteristically slow and dependent on trained experts in taxonomy.

Taxonomy, the scientific discipline of identifying and assigning names to species, may be one of the world’s oldest professions, yet today it is undoubtedly an uncommon trade. Although there has been much recent discussion about the discipline of taxonomy being in a state of decline [3], a recent
paper [4] suggested the opposite is true. Between 1864 and 2010, the number of authors describing new species, articles describing new species and total new species described in the zoological record has increased. This may be attributed to the development of new research methods and approaches or the “publication landscape” changing as science publications have become more interdisciplinary in nature. This trend, plus the rise of the international DNA barcoding initiative, are shifting the ways new species are discovered and expanding the traditional horizons of the discipline of taxonomy. The International Barcode of Life Project (iBOL) was established in 2009 and uses standardized portions of a species genome to identify species, with the aim to aid traditional taxonomy and species identification. Now species can be identified from plant leaf fragments and cambium [5], roots [6], herbal medicine preparations bought in stores [7], meat bought at the market [8], fecal remains of animals [9], and even from digested plant tissues inside the guts of insects [10]. These new and innovative methods hold much promise for addressing some of the challenges of identifying plant species in the tropics and are expected to accelerate the discovery of new species.

The primary mission of iBOL is to provide a platform for the construction of a global DNA barcode reference library of species, the Barcode of Life Data Systems (BOLD) [11], and to promote increased geographic and taxonomic coverage of the species currently represented. BOLD is an online resource developed by the Canadian Centre for DNA Barcoding (CCDB), which stores the DNA barcode records, all supporting trace file sequence data as well as the data on the voucher collections for each DNA sample. BOLD also enables global community access to the data with online tools for visualization, species validation, and analysis. It is now a central data repository and informatics hub for DNA barcoding projects worldwide.

This project aims to construct a DNA barcode library for the tropical flora of Australia in collaboration with iBOL and the CCDB. As a starting point we present analyses of 1572 DNA barcodes from 848 species from Queensland. In addition to providing a broad coverage of the species that occur in the region, the project sampled multiple individuals per species for approximately half of the sampled species (473) to investigate the occurrence of infraspecific molecular variation in DNA barcode loci.

The Australian state of Queensland is the world’s sixth largest sub-national political region, spanning more than 1,850,000 km². Over half of Queensland lies in the southern tropics. Great diversity of landscapes and climates has allowed a large number of plant species to evolve. The 2014 Census of the Queensland Flora [12] lists 14,174 native plant species, making Queensland the most species rich state of Australia, plus it contains three out of the 12 major Australian centres of plant endemism [13]. This project focuses on plants occurring in the tropical northern part of the state, specifically within the Wet Tropics Bioregion and the Iron Range-McIlwraith Range region of Cape York Peninsula. However, it also includes some species with ranges that extend outside this area into Queensland’s western monsoonal, arid regions and the southern subtropical zone (Figure 1).

Figure 1. Distribution of BATF DNA barcode records across the state of Queensland, Australia.
The global importance of the Wet Tropics Bioregion for biodiversity is well recognized and most of the region is included within the Queensland Wet Tropics World Heritage Area [14]. The region is considered one of the best-preserved living museums, containing assemblages of species representing multiple different eras of the Earth’s evolutionary history including lineages of relict and recently radiated origins. Thus progress towards a complete DNA barcode library for this bioregion may be considered a priority and a valuable asset to Australia and to the world.

The current DNA barcode library consists of three plastid loci: \textit{rbcLa}, \textit{matK}, and the \textit{trnH-psbA} intergenic spacer region and is hosted on the BOLD online database. The research was initiated through a project to generate DNA barcodes for 500 Australian tropical tree species. Additional species were subsequently added through contributions of postgraduates and research collaborations with the Australian Tropical Herbarium. This project lays the foundation for a more complete plant DNA barcode library for Australian tropical flora to be completed over time. It is expected that the compilation of this genetic resource will accelerate both academic research in the region and applied uses of DNA barcode data in fields such as quarantine, forensics, ecological restoration, climate change impacts, and citizen science.

2. Methods

DNA samples were obtained from a combination of fresh dried leaves stored on silica gel and herbarium specimens. For most species, fresh leaf material was obtained from field research plots, biodiversity survey expeditions, and local arboreta. Multiple samples (two to six) of 473 species were collected. For each collection, one leaf was stored on silica gel. For species that could not be located in the field, herbarium specimens from the Australian Tropical Herbarium were destructively sampled. Specimens no more than 20 years of age were selected, with preference given to specimens collected within the last ten years. Leaf tissue fragments were loaded into 96 well plates using sterilized forceps, and sent to the Canadian Center of DNA Barcoding (CCDB) for DNA extraction following their protocols [15]. All DNA samples are vouchered by herbarium specimens held at the Australian Tropical Herbarium (CNS). Voucher specimen data were submitted to BOLD via their standard submission template. We followed the standard data submission protocol from BOLD for specimen data entry [16]. A total of 773 specimens in the dataset include photographs of the voucher specimens. Specimen data were linked to each sequence via the BOLD web platform under the project folder titled Barcoding Australia’s Tropical Flora (BATF). PCR amplifications for the two official DNA barcode loci \textit{rbcLa} and \textit{matK}, plus a third, \textit{trnH-psbA}, which is popular for its high PCR amplification success were conducted jointly between the CCDB and the Australian Tropical Herbarium following the CCDB amplification protocols for plant and fungi [17] and sequencing was conducted at the CCDB following their protocols [18]. DNA samples of all specimens are held at the Australian Tropical Herbarium with aliquots of some species also held at CCDB and the Smithsonian Institution’s National Museum of Natural History. Aliquots are available to the scientific community by request.

Accuracy of DNA barcodes was assessed through the BOLD taxon ID tree function, neighbor joining analyses and BLAST searches via GenBank. Species that displayed molecular variation among samples were further investigated. These included: (A) species that showed discordant results between the DNA sequence derived phylogeny and morphologically recognized species; and (B) species that showed variation within species but no discordance with other closely related species.

For species that fell into category (A), voucher specimens were carefully checked, and the raw sequence data (trace and contig files) re-examined. This process discovered a small number of vouchers that were incorrectly identified, and resolved some instances of incorrect base calls. For species that fell into category (B), the raw sequence data (trace and contig files) were re-examined. Species showing confirmed infraspecific variation were then investigated further to identify any geographic and/or ecological patterns.
3. Results and Discussion

3.1. Taxonomic Coverage

The Barcoding Australia’s Tropical Flora project (BATF) consisted of mostly angiosperms with 857 angiosperm species spanning 43 orders and 113 families (see Table 1). An additional five gymnosperm species and six pteridophyte species were also included. A subset of 473 of these species contained multiple records (two to six) from separate geographic localities in Queensland. The nomenclature used for each species name follows the Australian Plant Name Index [19].

<table>
<thead>
<tr>
<th>Order</th>
<th>Families</th>
<th>Genera</th>
<th>Species</th>
<th>Samples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alismatales</td>
<td>1</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Apiales</td>
<td>4</td>
<td>10</td>
<td>13</td>
<td>26</td>
</tr>
<tr>
<td>Aquifoliales</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Arecales</td>
<td>1</td>
<td>9</td>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>Asparagales</td>
<td>3</td>
<td>35</td>
<td>36</td>
<td>41</td>
</tr>
<tr>
<td>Asterales</td>
<td>3</td>
<td>6</td>
<td>6</td>
<td>9</td>
</tr>
<tr>
<td>Austrobaileyales</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Boraginales</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>Brassicales</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Canellales</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Caryophyllales</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Celastrales</td>
<td>1</td>
<td>8</td>
<td>9</td>
<td>22</td>
</tr>
<tr>
<td>Commelinoides</td>
<td>2</td>
<td>4</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Cucurbitales</td>
<td>2</td>
<td>5</td>
<td>6</td>
<td>10</td>
</tr>
<tr>
<td>Dillenioides</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Dioscoreales</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ericales</td>
<td>6</td>
<td>18</td>
<td>53</td>
<td>112</td>
</tr>
<tr>
<td>Escallonioides</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>Fabales</td>
<td>2</td>
<td>28</td>
<td>38</td>
<td>66</td>
</tr>
<tr>
<td>Gentianales</td>
<td>4</td>
<td>44</td>
<td>69</td>
<td>134</td>
</tr>
<tr>
<td>Huertiales</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Lamiales</td>
<td>6</td>
<td>23</td>
<td>25</td>
<td>45</td>
</tr>
<tr>
<td>Laurales</td>
<td>4</td>
<td>13</td>
<td>54</td>
<td>131</td>
</tr>
<tr>
<td>Liliales</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td>Magnolioides</td>
<td>3</td>
<td>14</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>Malpighiales</td>
<td>15</td>
<td>44</td>
<td>68</td>
<td>129</td>
</tr>
<tr>
<td>Malvae</td>
<td>3</td>
<td>15</td>
<td>24</td>
<td>61</td>
</tr>
<tr>
<td>Myrtales</td>
<td>4</td>
<td>27</td>
<td>67</td>
<td>141</td>
</tr>
<tr>
<td>Oxalidales</td>
<td>3</td>
<td>11</td>
<td>20</td>
<td>38</td>
</tr>
<tr>
<td>Pandanales</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Paracryphioides</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Piperales</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>Poales</td>
<td>2</td>
<td>26</td>
<td>28</td>
<td>29</td>
</tr>
<tr>
<td>Proteales</td>
<td>1</td>
<td>18</td>
<td>20</td>
<td>40</td>
</tr>
<tr>
<td>Ranunculales</td>
<td>1</td>
<td>8</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>Rosales</td>
<td>6</td>
<td>21</td>
<td>36</td>
<td>65</td>
</tr>
<tr>
<td>Santalales</td>
<td>4</td>
<td>9</td>
<td>9</td>
<td>14</td>
</tr>
<tr>
<td>Sapindales</td>
<td>6</td>
<td>51</td>
<td>101</td>
<td>237</td>
</tr>
<tr>
<td>Saxifragales</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Solanales</td>
<td>2</td>
<td>8</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td>Vitales</td>
<td>1</td>
<td>5</td>
<td>6</td>
<td>8</td>
</tr>
<tr>
<td>Zingiberoides</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

**Total**: 113 families, 496 genera, 857 species, 1572 samples. Taxonomy follows APG III.
3.2. Variation within Species across Geographic Breaks

There have been many DNA barcode studies to date that have discussed the accuracy and PCR amplification success of different loci, but few studies have assessed the variation within species across their geographical ranges. We were interested to explore the utility of a large DNA barcode dataset of several hundred species sampled across a large geographical area, for addressing landscape scale questions in biogeography, conservation, and ecology.

We sampled 473 species from two to six geographic locations across each species’ range in Queensland, Australia. Of these, 100 (21%) species were found to have infraspecific variation across the species’ range. The variation ranged from one to 18 base pair differences across all three barcode loci per species. The trnH-psbA intergenic spacer region was the most variable locus with on average 3.60 base pair differences within each species, followed by matK with 1.25, then rbcL with 0.31. The mean infraspecific variation per species across all three loci was 5.12.

Although some species seemed to show stochastic variation, with samples variable within a relatively close region without any obvious biogeographic barriers, many showed variation across clear biogeographic barriers. Of particular interest were three recognized biogeographic barriers: the Normanby Gap, the Black Mountain Corridor, and the Burdekin Gap (Figure 2). Figure 2A shows the current distribution of rain forest and vine thickets in Queensland (green). Both the Normanby Gap and Burdekin Gap are large dry zones separating large rain forest blocks. The Black Mountain Corridor is currently rain forest but is recognized to be regenerated rain forest since the last glacial maximum, which reconnected areas to the north and south of it that maintained continuous rain forest refugia during the glacial period [20].

![Figure 2](image-url)  
**Figure 2.** Extent of rain forest and vine thicket in Queensland shown in green (A); distribution of samples of variable species spanning the Burdekin Gap in yellow (B); the Black Mountain Corridor (BMC) in blue (C); and the Normanby Gap in red (D).

While our sampling was insufficient to discern whether observed genetic variation was continuous (clinal) across species’ ranges or changes abruptly at putative barriers, the following results identify species that would be profitable targets of follow up research into the location and significance of geographical breaks on plant population genetic structure.
3.3. Burdekin Gap

The Burdekin Gap separates the Wet Tropics Bioregion to the north from a corridor of rain forest fragments stretching from the southern-most tropical rain forest patches in Australia into the subtropical zone through New South Wales. This biogeographic break is named after the Burdekin River and is recognized to have dated geographic breaks from the Pliocene for open forest frogs, rainforest lizards and open forest lizards, the Pleistocene for rainforest birds, and an older Miocene divergence has been estimated for freshwater fish [21,22]. Differentiation across this biogeographic barrier has not been previously reported in plant species. Our dataset included 19 species sampled across the Burdekin Gap, of which five showed variation between north and south of the Burdekin Gap (See Figure 2B) with variation ranging from 1 to 4 base pairs and an average of 1.83 per species (Table 2).

Table 2. List of species showing infraspecific variation across the Burdekin Gap. Infraspecific variation (column 3) tallied by total number of base pair (bp) differences within each species across all three barcode loci.

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>Infra-Variation (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Euphorbiaceae</td>
<td>Croton acronychioides F. Muell.</td>
<td>2</td>
</tr>
<tr>
<td>Malvaceae</td>
<td>Brachychiton australis (Schott &amp; Endl.) A. Terracc.</td>
<td>1</td>
</tr>
<tr>
<td>Monimiaceae</td>
<td>Wilkiea macrophylla (A.Cunn.) A. DC.</td>
<td>1</td>
</tr>
<tr>
<td>Podocarpaceae</td>
<td>Podocarpus graje de Laub.</td>
<td>1</td>
</tr>
<tr>
<td>Rubiaceae</td>
<td>Exerista vacciniifolia (F. Muell.) S. T. Reynolds &amp; R. J. F. Hend.</td>
<td>4</td>
</tr>
</tbody>
</table>

3.4. Black Mountain Corridor

The Black Mountain Corridor separates the northern and southern parts of the Wet Tropics Bioregion. The emergence and disappearance of this barrier could be as early as the mid Miocene based on several faunal genetic divergences across it including skinks [23], vertebrates and insects [22], snails [24], and some plants [25]. Rossetto et al. [25] found that within the family Elaeocarpaceae, the Black Mountain Corridor acted as a geographic barrier for three out of 11 species, and that the barrier was less important for species with smaller fruits than for large fruited species. Our dataset included 110 plant species with samples north and south of the Black Mountain Corridor. A total of 36 (33%) of these species showed variation across the corridor (See Figure 2C) with infraspecific-variation ranging from 1 to 15 base pairs and an average of 3.5 base pairs per species (Table 3).

Table 3. List of species showing infraspecific variation across the Black Mountain Corridor. Infraspecific variation (column 3) tallied by total number of base pair (bp) differences within each species across all three barcode loci.

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>Infra-Variation (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annonaceae</td>
<td>Goniothalamus australis Jessup</td>
<td>1</td>
</tr>
<tr>
<td>Apocynaceae</td>
<td>Tabernaemontana pandacaqui Lam.</td>
<td>4</td>
</tr>
<tr>
<td>Celastraceae</td>
<td>Hippocratea barbata F. Muell.</td>
<td>2</td>
</tr>
<tr>
<td>Eupomatiaceae</td>
<td>Eupomatia laurina R. Br.</td>
<td>1</td>
</tr>
<tr>
<td>Fabaceae</td>
<td>Archidendron ramiflorum (F. Muell.) Kosterm.</td>
<td>1</td>
</tr>
<tr>
<td>Fabaceae</td>
<td>Acacia celsa Tindale</td>
<td>2</td>
</tr>
<tr>
<td>Fabaceae</td>
<td>Castanospermum australe A. Cunn. ex Mudie</td>
<td>8</td>
</tr>
<tr>
<td>Fabaceae</td>
<td>Entada phaseoloides (L.) Merr.</td>
<td>13</td>
</tr>
<tr>
<td>Icacinaceae</td>
<td>Apodytes brachypterysis F. Muell.</td>
<td>2</td>
</tr>
<tr>
<td>Lauraceae</td>
<td>Cryptocarya livida B. Hyland</td>
<td>1</td>
</tr>
<tr>
<td>Lauraceae</td>
<td>Endiandra impressicosta C. K. Allen</td>
<td>1</td>
</tr>
<tr>
<td>Lauraceae</td>
<td>Endiandra leptodendron B. Hyland</td>
<td>1</td>
</tr>
<tr>
<td>Lauraceae</td>
<td>Cryptocarya bellendenkerana B. Hyland</td>
<td>2</td>
</tr>
<tr>
<td>Lauraceae</td>
<td>Cryptocarya densiflora Blume</td>
<td>5</td>
</tr>
<tr>
<td>Lauraceae</td>
<td>Endiandra insignis (F. M. Bailey) F. M. Bailey</td>
<td>5</td>
</tr>
<tr>
<td>Lauraceae</td>
<td>Endiandra longipedicellata C. T. White &amp; W. D. Francis</td>
<td>5</td>
</tr>
</tbody>
</table>
Table 3. Cont.

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>Infra-Variation (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lauraceae</td>
<td><em>Beilschmiedia bancroftii</em> (F. M. Bailey) C. T. White</td>
<td>15</td>
</tr>
<tr>
<td>Malvaceae</td>
<td><em>Argyroderon perlatum</em> (F. M. Bailey) Edlin ex J. H. Boas</td>
<td>1</td>
</tr>
<tr>
<td>Monimiaceae</td>
<td><em>Steganthera laxiflora</em> (Benth.) Whiffin &amp; Foreman</td>
<td>4</td>
</tr>
<tr>
<td>Moraceae</td>
<td><em>Strēlīs găber</em> (Merx.) Corner</td>
<td>1</td>
</tr>
<tr>
<td>Myristicaceae</td>
<td><em>Myristica globosa</em> Warb.</td>
<td>1</td>
</tr>
<tr>
<td>Myrtaceae</td>
<td><em>Syzygium johnsonii</em> (F. Muell.) B. Hyland</td>
<td>1</td>
</tr>
<tr>
<td>Myrtaceae</td>
<td><em>Syzygium kuranda</em> (F. M. Bailey) B. Hyland</td>
<td>1</td>
</tr>
<tr>
<td>Myrtaceae</td>
<td><em>Leptospermum wooroonooran</em> F. M. Bailey</td>
<td>2</td>
</tr>
<tr>
<td>Myrtaceae</td>
<td><em>Sphaerantia discolar</em> Peter G. Wilson &amp; B. Hyland</td>
<td>4</td>
</tr>
<tr>
<td>Myrtaceae</td>
<td><em>Rhodamnia sessiliflora</em> Benth.</td>
<td>9</td>
</tr>
<tr>
<td>Phyllanthaceae</td>
<td><em>Glochidion sessiliflorum var. pedicellatum</em> Airy Shaw</td>
<td>7</td>
</tr>
<tr>
<td>Rhamnaceae</td>
<td><em>Emmenosperma cunninghamii</em> Benth.</td>
<td>1</td>
</tr>
<tr>
<td>Rubiaceae</td>
<td><em>Atractocarpus fitzalanii</em> (F. Muell.) Puttock</td>
<td>2</td>
</tr>
<tr>
<td>Rubiaceae</td>
<td><em>Atractocarpus hirtus</em> (F. Muell.) Puttock</td>
<td>2</td>
</tr>
<tr>
<td>Rutaceae</td>
<td><em>Melicope stipitata</em> C. T. White &amp; W. D. Francis</td>
<td>2</td>
</tr>
<tr>
<td>Sapindaceae</td>
<td><em>Harpulia rhicarpa</em> C. T. White &amp; W. D. Francis</td>
<td>2</td>
</tr>
<tr>
<td>Sapindaceae</td>
<td><em>Diploglottis harpulioides</em> S. T. Reynolds</td>
<td>3</td>
</tr>
<tr>
<td>Sapindaceae</td>
<td><em>Sarcotheca cuneata</em> Radlk.</td>
<td>3</td>
</tr>
<tr>
<td>Sapindaceae</td>
<td><em>Diploglottis smithii</em> S. T. Reynolds</td>
<td>5</td>
</tr>
<tr>
<td>Sapindaceae</td>
<td><em>Rhysolutacia robertsonii</em> (F. Muell.) Radlk.</td>
<td>6</td>
</tr>
<tr>
<td>Sypmlocaceae</td>
<td><em>Sypmlocos amputiformis</em> C. T. White</td>
<td>5</td>
</tr>
</tbody>
</table>

3.5. Normanby Gap

The Normanby Gap, also recognized as the Laura Gap [22], separates the Iron Range-McIlwraith Range rain forest area of Cape York from the Wet Tropics Bioregion to the south. The area is named after the Normanby River, which is Australia’s third largest river. Although the region surrounding the river basin is seasonally flooded during the wet season, the pronounced dry season prevents the development of rain forest. The region is unsuitable for cropping and much of it is utilized for cattle grazing. Our dataset included 53 species spanning the Normanby Gap. Out of these 53, 15 species showed infraspecific variation across the Normanby Gap (See Figure 2D) with variation ranging from 1 to 9 base pairs and an average of 3.81 per species (Table 4).

Table 4. List of species showing infraspecific variation across the Normanby Gap. Infraspecific variation (column 3) tallied by total number of base pairs (bp) differences within each species across all three barcode loci.

<table>
<thead>
<tr>
<th>Family</th>
<th>Species</th>
<th>Infra-Variation (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annonaceae</td>
<td><em>Miliusa traceyi</em> Jessup</td>
<td>2</td>
</tr>
<tr>
<td>Lauraceae</td>
<td><em>Endiandra longipedicellata</em> C. T. White &amp; W. D. Francis</td>
<td>5</td>
</tr>
<tr>
<td>Malvaceae</td>
<td><em>Argyroderon polyanum</em> L. S. Sm.</td>
<td>1</td>
</tr>
<tr>
<td>Moraceae</td>
<td><em>Ficus variegata</em> Blume</td>
<td>2</td>
</tr>
<tr>
<td>Myrtaceae</td>
<td><em>Gossia retusa</em> N. Snow &amp; Guymer</td>
<td>4</td>
</tr>
<tr>
<td>Myrtaceae</td>
<td><em>Rhodomyrtus macrocarpa</em> Benth.</td>
<td>4</td>
</tr>
<tr>
<td>Myrtaceae</td>
<td><em>Syzygium fibrosum</em> F. M. Bailey</td>
<td>1</td>
</tr>
<tr>
<td>Phyllanthaceae</td>
<td><em>Breynia cernua</em> (Poir.) Mull. Arg.</td>
<td>3</td>
</tr>
<tr>
<td>Proteaceae</td>
<td><em>Helicia australasica</em> F. Muell.</td>
<td>3</td>
</tr>
<tr>
<td>Rhamnaceae</td>
<td><em>Emmenosperma cunninghamii</em> Benth.</td>
<td>1</td>
</tr>
<tr>
<td>Rubiaceae</td>
<td><em>Ixora timorensis</em> Deene</td>
<td>3</td>
</tr>
<tr>
<td>Rutaceae</td>
<td><em>Bosistoa brassii</em> T. G. Hartley</td>
<td>9</td>
</tr>
<tr>
<td>Rutaceae</td>
<td><em>Medicosma sessilifora</em> (C. T. White) T. G. Hartley</td>
<td>9</td>
</tr>
</tbody>
</table>

4. Conclusions

4.1. DNA Barcodes as a Tool for Studying Biogeography

These results should be regarded as preliminary due to the small numbers of individuals sampled per species (two to six). Follow-up studies are required to test the congruence between the patterns of infraspecific molecular variation and putative biogeographical barriers observed in some species
through more comprehensive sampling across the entire range of each species. Our study shows that the potential value of publicly available DNA barcode libraries such as BATT for rapid assessment of landscape level and species level patterns will only be fully realised when sampling is sufficiently comprehensive to reveal underlying geographical patterns in molecular variation. In the meantime such databases have value as tools to identify taxa and areas for further study with higher resolution markers and more comprehensive sampling. Lastly it is important to highlight that the value of this work lies not only in the DNA barcode database, but the curated tissue and voucher samples (available upon request), which can enable future researchers to save a substantial amount of time and money instead of re-collecting them.

4.2. Data Resources

Data set name: BATF. Download form [26].

The plant DNA barcode reference library for tropical Queensland dataset is accessible through the public data portal BOLD [27]. All data including taxonomic information, geographic records, voucher specimen details, sequence data in the form of fasta files, and trace files (ab1 files) can be easily downloaded through the BOLD graphic user interphase. BOLD account holders can additionally access the data through the BOLD workbench platform. All voucher specimens are stored at the Australian Tropical Herbarium (CNS) and all DNA samples are jointly stored at the Australian Tropical Herbarium, the Canadian Centre for DNA Barcoding (CCDB), and The National Museum of Natural History, Smithsonian Institution (NMNH).

Acknowledgments: We thank the following for funding and support: Australian Department of the Environment through the Commonwealth Environment Research Facilities (CERF) and National Environmental Research Program (NERP); Australian Tropical Herbarium, James Cook University; and University of Adelaide. We also thank Melissa Harrison and Hugh Cross for technical support in the laboratory, and the many people who contributed botanical collections to the Australian Tropical Herbarium that made this work possible especially Gary Sankowsky, Bruce Gray, Paul Forster, and Stephanie Stuart.

Author Contributions: Craig M. Costion is a postdoctoral fellow at the Smithsonian Institution’s National Museum of Natural History. His research is focused on plant biodiversity patterns and conservation in the Australasia and Pacific region. Craig M. Costion, Darren M. Crayn, and Andrew J. Lowe conceived the study. Craig M. Costion collected the data. Craig M. Costion, Martin F. Breed, Maurizio Rossetto, and Robert M. Kooyman analyzed the data. Craig M. Costion lead the writing and Darren M. Crayn, Andrew J. Lowe, Martin F. Breed, Maurizio Rossetto, Robert M. Kooyman and Andrew Ford helped write the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

4. Tancoigne, E.; Dubois, A. Taxonomy: No Decline, but inertia. *Cladistics* 2013, 29, 567–570. [CrossRef]
8. Sin, W.M.; Tam, Y.K.; Tsui, S.; Ng, C.; Mok, C.; Ha, W. An integrated and validated DNA-based protocol developed to fight against commercial frauds: A Case of Fraudulent Substitutions for Deer Products. *DNA Barcodes* 2013, 1, 27–34. [CrossRef]


