The Australasian Virtual Herbarium: Tracking data usage and benefits for biological collections

David J. Cantrill1,2

PREMISE OF THE STUDY: Globally, natural history collections are focused on digitizing specimens and information and making these data accessible. Usage information on National Herbarium of Victoria data made available through the Atlas of Living Australia and The Australasian Virtual Herbarium (AVH) is analyzed to understand how and by whom herbarium data are being used.

METHODS: Since 2010, AVH data usage information has been gathered from users and supplied to data custodians as a spreadsheet that includes number of download events, number of records downloaded, and user reasons for downloading data in predefined categories.

RESULTS: Since 2010, in excess of 268,000 download events of 194 million records (excluding testing events) have been recorded for the National Herbarium of Victoria data set. This means, on average, every record has been downloaded 220 times in the past nine years. Data use grew continuously from 2010 to 2015 but decreased in 2016 due to fewer ecological projects.

DISCUSSION: Data have primarily been used for ecological research, but there is an emerging trend for use in education including citizen science projects. Information about data use demonstrates relevance to funding agencies and helps inform the development of collections and prioritization of resources when digitizing material.

KEY WORDS: bioinformatics; data usage; herbarium records.

The origin of herbaria can be traced back to the Italian botanist Luca Ghini in the mid-1500s, who developed and promoted the practice of pressing plants into herbaria (Arber, 1938). These early herbaria were mounted in books and use was closely tied to herbals. Over time, the nature of herbaria has changed from bound books of pressed plants used by individuals to loose sheets housed in cabinets. Both the size and number of collections have grown, and today approximately 3000 herbaria and more than 380 million specimens are housed globally in these collections (Thiers, 2017). These collections are a vast resource for the botanical research community but until recently have been neither easily accessible nor linked.

Digitization of information about the specimens and the specimens themselves is rapidly changing this situation.

The use of herbaria has also changed over time from the initial role in herbal medicine and plant identification through to today where more than 72 uses have been documented (Funk, 2004). Recent innovative uses have included: demonstrating that herbarium collections document plant responses to historical changes in carbon dioxide (e.g., Woodward, 1987), tracking isotopic signatures of environmental change (e.g., Pedicino et al., 2002), and extracting DNA for analysis (e.g., Staats et al., 2011) to name but a few. All of these examples rely on access to the physical specimens; however, digitization
opens up new possibilities including access to large data sets (Soltis, 2017). Aggregated data from many herbaria provide large spatial- and temporal-scale data sets that have been used, for example, to model plant distributions and assess the impact of climate change (e.g., González-Orozco et al., 2016). The changes in use over the past 500 years raises the question: What new uses will herbarium specimens and, in particular, digitized material be put to in the future?

Herbaria specimens document the occurrence of species at a particular place and a specific point in time. Unlike observational data, and because they are backed up by physical specimens, these data points can be updated as taxonomic concepts change. Each generation of scientists adds to these vast data repositories, improving them for the next generation. Indeed, it could be argued that, as scientific infrastructure, herbaria are one of the few that actually get better with time as knowledge attached to the specimens is gathered and improved. These attributes, together with the length of time that herbaria have been assembled, mean that herbaria represent one of the best-quality temporal and spatial data sets we have of biodiversity.

In the 1980s, the herbarium community began capturing information about specimens into in-house databases, and ultimately these were made accessible through the World Wide Web. Large-scale efforts to make information housed in natural history collections accessible have been the focus of recent investment in digitization at both national (e.g., Australasian Virtual Herbarium, iDigBio) and international (e.g., Global Biodiversity Information Facility) scales. More than 25 years ago, the Australian plant taxonomic community identified the opportunity that digitization of collection information offered to improve collection management and streamline taxonomic processes. In 2000 a pilot project at the Royal Botanic Gardens Victoria was followed by an initiative led by the Council of Heads of Australasian Herbaria (CHAH), which resulted in the creation of Australia’s Virtual Herbarium in 2001. This Australia-wide resource containing specimen information from all the major state collections of Australasian Herbaria (CHAH), which resulted in the creation of the Australasian Virtual Herbarium (AVH).

Understanding usage of AVH initially relied on web statistics (e.g., page views), which demonstrated the volume of use but gave little information on how the data were being used. Before 2010, use cases were also collected from AVH users who wanted access to large data sets. This was achieved by setting a download limit (10,000 records) that then required users to email and request access to the data. Through this method, individual use cases, such as predicting the invasiveness of Australian Acacia outside its native range (Hui et al., 2011), were gathered. However, these individual use cases gave no insight into the majority of AVH use. The creation of the ALA and the redevelopment of the AVH interface enabled a simple monitoring tool to be built that reduced the reliance on web statistics. Before a user can download any data query, they are required to answer some questions about how those data will be used, and this information is gathered automatically and provided back to data custodians. In this study, I present results on the use of Royal Botanic Gardens Victoria data (National Herbarium of Victoria [MEL] data set) extracted by users through the AVH interface. The uses are classified into several categories defined by the ALA.

### METHODS

Records from the Royal Botanic Gardens Victoria are continuously contributed to the AVH and so the pool of data available for queries has grown from 791,106 records in 2009 to 881,328 records by June 2017. These small incremental increases are not considered to impact the data use requests. Queries against the AVH data set are provided by the ALA through the Collectory pages (Atlas of Living Australia, 2017a) on each individual institution’s home page under the “download usage stats” button (Atlas of Living Australia, 2017b). These data are collected as part of the conditions of download, whereby users who wish to download data sets are required to provide a reason for the download from a menu of options (Table 1). Once the rationale for the download is provided, the data are emailed to the user, and the query recorded and added to the spreadsheet under the relevant data use category.

Data use statistics are provided to data custodians broken down on a monthly basis and include: number of download events (i.e.,

| Category                        | No. of download events |
|---------------------------------|------------------------|
|                                 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 |
| Biosecurity                     | 79   | 129  | 58   | 63   | 215  | 164  |
| Citizen science                 |      |      |      |      | 388  | 311  |
| Collection management           | 33   | 62   | 83   | 83   | 112  | 66   |
| Conservation management         | 502  | 1057 | 769  | 160  |      |      |
| Ecological research             | 1521 | 8738 | 20,609 | 87,043 | 8279 | 54,237 |
| Education                       | 226  | 513  | 806  | 704  | 995  | 742  |
| Environmental assessment        | 271  | 500  | 355  | 403  | 950  | 520  |
| Other                           | 199  | 197  | 297  | 208  | 446  | 887  |
| Other scientific                | 13   | 132  | 87   | 105  | 205  | 23   |
| Restoration/remediation         |      |      |      |      | 79   | 46   |
| Scientific research             | 1141 | 11,902 | 19,544 | 23,683 | 5948 | 3877 |
| Systematics/Taxonomy            | 101  | 485  | 165  | 293  | 505  | 296  |
| Testing                         | 146  | 209  | 6226 | 6453 | 913  | 629  |
| Unclassified                    | 6    | 348  | 495  | 112  | 620  | 1101 | 403  |
| Total                           | 6    | 348  | 4727 | 24,030 | 48,999 | 119,818 | 20,136 | 62,201 |
| Total: testing                  | 6    | 348  | 4581 | 23,821 | 42,773 | 113,365 | 19,223 | 61,572 |
queries) and number of records downloaded; each of these is broken down by data category. Data use categories have changed over time. From October 2010 to February 2012, download events and records downloaded were recorded in one general use category (“unclassified”). In March 2012, seven categories were added (“collection management,” “scientific research,” “education,” “environmental assessment,” “biosecurity,” “conservation management,” “other”). In April 2012, three further categories were added (“ecological research,” “systematics/taxonomy,” “testing”), and in May 2012 the category “other scientific research” was added. These categories remained stable until May 2016 when two more categories were added (i.e., “citizen science,” “restoration/remediation”). In total, 12 categories are currently available but 14 categories of data use are being reported to data providers. This appears to be due to old instances of the ALA still being available and to categories that are no longer in use (e.g., “unclassified”) in the current version still being recorded by the data logging system.

The addition and deletion of categories create some limitations for the study here, especially for comparison across years, where the atomization of categories will have undoubtedly impacted the use category a user would select when providing information. For example, increased atomization of data use should reduce the number of queries in the “other” category, as more choices are available to users. Other impacts are also likely. For example, the addition of the “citizen science” category in 2012 may have impacted the “education” category, the “scientific research” category, or the “other scientific” category. These impacts cannot be easily accounted for in any analysis through time. In order to minimize the impact of changes in data use categories, the following approach was taken. Data categories were aggregated into scientific research and non-research uses. Non-research use includes: “biosecurity management,” “conservation management,” “collection management,” “environmental assessment,” “citizen science,” “restoration/remediation,” “unclassified,” “other.” Research use includes: “scientific research,” “ecological research,” “systematics/taxonomy,” “other scientific research.” Testing events and records downloaded (Tables 1, 2) represent internal data use and so are discussed separately.

Although the data use is captured on a monthly basis, for this study it has been consolidated into an annualized total of queries (download events; Table 1) and records downloaded (Table 2). Data were aggregated in non-research use and scientific research use and are reported as annualized frequencies (Tables 1, 2). Monthly frequency data for ecological research are also presented as an aggregated monthly mean across the 2012 to 2016 period to illustrate the annual pattern of use. For 2017, data were extracted at the end of June and numbers have been doubled to give an estimated annualized total.

RESULTS

Testing

The robustness of the search interface is tested through the external download function through a series of standard searches. These protocols ensure that new data uploads have happened correctly and that data supplied by the data custodians are mapped to the correct fields in the AVH cache. Data downloaded for testing are given in Tables 1 and 2 but are excluded from further analysis as these data represent internal use by the data custodians. In 2014 and 2015, major system upgrades to the Royal Botanic Gardens Victoria collection management system and the AVH system occurred. As part of these upgrades, additional testing events were run to check that both the data migration to the new collection management system and the redeveloped AVH portal were returning correct search results. Consequently, large peaks in testing events are observed in 2014 and 2015.

General data use

Overall queries increased from 2011 to 2013, declined in 2014 and 2015, and have increased again in 2016 and 2017 to more than 4500 per year (Fig. 1A). This contrasts with the number of records downloaded, which peaked in 2014 (Fig. 1B).

### Table 2. Number of records downloaded between October 2010 and July 2017 separated into download categories for data provided by the Royal Botanic Gardens to The Australasian Virtual Herbarium. Note that not all cells are filled because data use categories were added and deleted during the period considered.

| Category                   | 2010  | 2011  | 2012  | 2013  | 2014  | 2015  | 2016  | 2017  |
|----------------------------|-------|-------|-------|-------|-------|-------|-------|-------|
| Biosecurity                | 8955  | 12484 | 2053  | 4230  | 5420  | 6724  | 7619  | 8363  |
| Citizen science            | 2218  | 3473  | 4929  | 5740  | 6310  | 7052  | 8035  | 8777  |
| Collection management      | 2562  | 2430  | 3100  | 3300  | 3300  | 3300  | 3300  | 3300  |
| Conservation management    | 1047  | 1154  | 1262  | 1320  | 1320  | 1320  | 1320  | 1320  |
| Ecological research        | 1178  | 1200  | 1258  | 1300  | 1300  | 1300  | 1300  | 1300  |
| Education                  | 67000 | 75000 | 78000 | 80000 | 80000 | 80000 | 80000 | 80000 |
| Environmental assessment   | 75690 | 79580 | 82580 | 85480 | 85480 | 85480 | 85480 | 85480 |
| Other                      | 1047  | 1154  | 1262  | 1320  | 1320  | 1320  | 1320  | 1320  |
| Other scientific           | 1178  | 1200  | 1258  | 1300  | 1300  | 1300  | 1300  | 1300  |
| Restoration/remediation    | 67000 | 75000 | 78000 | 80000 | 80000 | 80000 | 80000 | 80000 |
| Scientific research        | 2118  | 2401  | 2673  | 2901  | 2901  | 2901  | 2901  | 2901  |
| Systematics/taxonomy       | 27900 | 34730 | 42300 | 49290 | 54200 | 67240 | 76190 | 83630 |
| Testing                    | 4562  | 4612  | 4662  | 4712  | 4712  | 4712  | 4712  | 4712  |
| Unclassified               | 8720  | 48720 | 58288 | 68219 | 78250 | 88280 | 98210 | 108240 |
| Total                      | 8720  | 48720 | 58288 | 68219 | 78250 | 88280 | 98210 | 108240 |
| Total: testing             | 8720  | 48720 | 58288 | 68219 | 78250 | 88280 | 98210 | 108240 |
Non-research use

This includes uses in the “biosecurity,” outreach (“education” and “citizen science”), “conservation management,” and “collection management” categories. The number of download events increased over the seven-year period, although with a slight decline in 2014 and 2015 (Fig. 1A). This contrasts with the number of records downloaded, which peaked in 2014 (Fig. 1B).

Scientific research use

Use for research rose to a peak in 2015 when a large number of small queries for ecological research occurred. A dramatic fall in queries occurred in 2016, but this decrease has been reversed in 2017 with projected number of queries to exceed 2015 (Fig. 2A). Similar to the queries, the number of records downloaded peaked in 2015, declined in 2016 (but not as markedly as the queries), and is projected to recover in 2017 but not to 2015 levels (Fig. 2B).

DISCUSSION

Prior to October 2010, our understanding of AVH data use was restricted to requests for large data sets and web statistics on page views. The AVH had a limit on the number of records that could be downloaded (10,000). If users needed data sets larger than the data download limit, they were required to contact an administrator to request permission. By designing smaller queries, users could get around this feature and so extract large data sets. Requests for large data sets highlighted the value placed on the data that were being made available but was largely ad hoc, resulting in a limited understanding of how the data were being used more broadly (e.g., Saintilan, 2009; Waters et al., 2010; Hui et al., 2011). The use cases for large data sets obtained through this process allowed data custodians to demonstrate the impact of making the herbarium specimen data free and easily accessible. They were important for demonstrating value when seeking further financial support to build on the original AVH. However, they only captured a small subset of the total data usage. The advantage of this manual approach to understanding how herbarium data were being used was that the impact of the data use could be demonstrated. However, it came with a time and administrative cost.

The creation of the ALA allowed further development of AVH, including expanding it to cover Australasian herbaria as well as implementation of the data use system that is reported here. The data use system is automated, resulting in a lower administrative burden but at the cost of information richness, particularly the impact of the data use. The use data were broken down into two main groups, scientific research and non-research uses.

Non-research uses

In 2014, a large peak in records downloaded was seen, which has been interpreted as users taking the whole data set and loading it into local systems to reduce the need to repeat queries in the AVH, or for other analysis that required data modifications. Recent publications confirm this approach by the research community (e.g., Dodd et al., 2015, 2016). Large data downloads may mask underlying trends in data use that will become more evident with time. Collection management was a small component of data use, being
under 1% of the download events between 2012 and 2017. However, the number of records downloaded has increased as a percentage of downloads from 0.45% in 2012 to 9.8% in 2017. This suggests that queries of the whole AVH data set offer some benefit for collection management and may relate to an increase in data-cleaning activity by the community.

Uses in the categories “education” and, more recently, “citizen science” can be viewed as outreach activity. Download events in these categories have grown from 226 in 2012 to 1383 in 2017 (Fig. 3A). This contrasts with the number of records downloaded, which has fluctuated from year to year with peaks in 2014 and 2016 (Fig. 3B). The overall trend is for a decrease in the size of downloads, suggesting that users are becoming more discriminating in the data extracted.

The categories “conservation management,” “environmental assessment,” and “restoration/remediation” can be viewed as conservation activity. Download events in these categories rose to a peak in 2013 and declined in 2014 and 2015 before rising again (Fig. 4A). This contrasts with the records downloaded, which comprised some very large downloads in 2013 and 2014 (Fig. 4B). The size of these downloads suggests that some users downloaded the whole data set, likely to populate in-house database systems (e.g., AVH or subsets of AVH are part of the Australian National Heritage Assessment Tool [ANHAT] Database; Laity et al., 2015; Burley et al., 2016a, 2016b). The decline in use from 2013 coincides with the downturn in mining activity and development of new mines, therefore reducing the regulatory requirements for conservation assessments. This interpretation suggests the AVH data are being used to inform planning for new developments.

**Scientific research**

Despite the importance of herbarium specimens as vouchers for systematic and taxonomic research, data use by this research group forms a small percentage of the total use, ranging from 0.39% to 2.63% of download events and 0.32% to 3.30% of the records downloaded in any year (Fig. 2A, B). This contrasts with ecologists, who are the major users of AVH data. Downloads for ecological research peaked in 2015 at 87,043 queries and nearly 19 million records (average 218 records per query) (Tables 1, 2). This contrasts with use in other years where there were fewer queries of a larger size. For example, in 2016 there were 8279 query events extracting 7.3 million records (average 882 records per query). The peak use in 2015 suggests a large project with multiple small queries, perhaps for species niche modeling. From the published literature, we can identify studies that have taken large downloads and then modified (or removed) records for analysis, such as the study of plant invasions and naturalizations in Australia (Dodd et al., 2015, 2016). The combination of literature searches and the download statistics gives us some insights into herbarium data use.

Other patterns of use are apparent when the data for ecological research are examined on a monthly basis (Fig. 5). Here, data queries peak in January, coinciding with the major break in academic teaching (December to March) and grant preparation for submission in March. There is also a marked lack of use in August and September, probably coinciding with the main period of fieldwork in spring. This usage pattern could be used to inform decisions about when to upgrade the system in order to minimize disruption to users.

**Data use and recommendations**

The manual approach (2001 to 2006) to data use reveals the AVH data were being used for a diversity of purposes, including the generation of species lists for large areas to aid assessment in planning...
for new national parks and also for World Heritage applications. Other data uses included species niche modeling for Australian taxa that had become weedy overseas (e.g., Acacia species; Hui et al., 2011), for which the climate envelope occupied by the taxon in Australia was used to predict the potential range in the country where they had become weedy. This information was used to prioritize taxa for eradication and target areas for control to reduce the projected impact of the target taxa. Although the manual method provided a rich source of the impact of AVH data, it was limited to a very small set of data uses.

In contrast, the automated tool provides a useful overview of all data uses by predefined category. This approach has limitations because it assumes the data custodians understand all of the use cases and can generate meaningful predefined categories, which may not be the case. On average, around one third of the queries (8.4% to 51.7%; average 34.82%) and records downloaded (9.2% to 71.00%; average 37.7%) are allocated to general categories (“other,” “unclassified,” “scientific research”). This limits our understanding of the research fields utilizing the resource and suggests that the use categories set by the data custodians need refinement. Alternatively, a different methodology is needed to better capture information on the purposes that users are downloading data and the user audience. Options may include further refinement of and addition of user categories. Alternatively, another alternative may be using a free text field when the general categories are selected that can be analyzed using machine learning or natural language processing to classify the groups.

Although the automated data use system provides important insights into overall data use, we currently rely on additional methods, such as publication searches, to understand use cases and likely impact. Based on literature searches across a range of academic databases, we have added to the use cases collected prior to the implementation of the automated data use system. Taxon distribution maps are generated and published routinely for systematic and taxonomic (e.g., Wilkins and Whitlock, 2016 as just one of many examples) studies. The distribution information is used in a variety of ways, including phylogeographic studies (e.g., Barrett et al., 2014; Bayly et al., 2016); continental-scale diversity patterns and visualization of species richness (Stevenson et al., 2012; Nagalingum et al., 2015), endemism (Mishler et al., 2014), and using these data to examine spatial patterns of species turnover (González-Orozco et al., 2014a). The integration of distributional data with phylogenetic information has been used to generate patterns of phylogenetic diversity (González-Orozco et al., 2016; Thornhill et al., 2016) and turnover (Laffan et al., 2016) and indices developed from these studies can be used to inform conservation policy (Laity et al., 2015). The development of measures of phylogenetic diversity and endemism has been used to understand ecology of rainforests and to inform conservation actions (e.g., Kooyman et al., 2013). AVH data are also being used as part of analyses to: understand spatial biases in biodiversity data sets (Haque et al., 2017); relate distribution patterns to physical variables such as soil and climate (Prentice et al., 2017); examine the spatial distribution of traits (Moles et al., 2014; Tindall et al., 2017); combine species distributions with phylogeographic and environmental variables to identify biogeographic regions and understand evolutionary patterns of plant distribution (e.g., González-Orozco et al., 2013, 2014b), or to relate diversity to primary productivity (Burley et al., 2016a) and ecosystem function (Burley et al., 2016b). Herbarium data have been used to test invasion models (e.g., Taylor and Kumar, 2012), to identify invasion hotspots (O’Donnell et al., 2012), and have been used in tandem with genetics to identify introduction events and invasion history (Dormontt et al., 2014). These studies have also been used to examine threats to biodiversity under climate change using forward modeling of native (e.g., González-Orozco et al., 2016) and introduced (O’Donnell et al., 2012) species.

AVH data are also being integrated into other products. AVH data are readily imported into the ALA for further analysis using a range of ALA tools. Other digital products, such as VicFlora, generate species distribution maps on the fly from AVH data as part of the species profile pages (VicFlora, 2018a). Checklists are also generated by querying the AVH data set with a spatial query and compiling all the taxa supported by herbarium records in that region (VicFlora, 2018b).

The study reported on here has examined the use of the Royal Botanic Gardens Victoria data housed in the AVH. An interesting question for the future is: How representative are the data use and use cases identified for the National Herbarium of Victoria data set? It should be considered that each herbarium is unique, with its own collecting history both in time and space. It is likely that most user queries are searching across the whole AVH data set and so are blind to the data providers (unless they filter on particular herbaria), which means that the pattern of queries and, therefore, data use is likely to be the same as that reported on here. However, individual herbaria are different sizes with different taxonomic and geographic coverage. For example, the Queensland Herbarium (BRI) only supplies records of Queensland taxa and so will not appear in any searches looking at non-Queensland distributions. The Western Australian Herbarium (PERTH) collection is largely focused on Western Australian taxa and so will feature less prominently in studies of east coast taxa. Furthermore, legislation for environmental assessments has different requirements in different states; therefore, herbaria from different jurisdictions are likely to have slightly different query profiles in the conservation management area. The stringency of Western Australian environmental requirements means that this data set may well be used more extensively by consultants in Western Australia than other data sets. These are areas that remain to be examined and are beyond the scope of this study.

CONCLUSIONS

Digitization and making data accessible from herbarium collections has opened up new opportunities that demonstrate the relevance of

**FIGURE 5.** Monthly use of Royal Botanic Gardens Victoria data for ecological research based on download events.
these resources for the wider scientific community. Herbaria are no longer specimen repositories used by a limited number of taxonomists; they are now rich data sources that have many applications. While specimens remain critical for fixing names and creating a stable taxonomic framework that underpins all of biology, the use of natural history data for other scientific endeavours is just beginning. It is critical for custodians of natural history collections to understand how data and digital objects are being used and how often. The data use function developed for the data custodians within the ALA gives important insights into data use and data use patterns. Nevertheless, this system could be improved given that approximately one third of all queries are not currently classified into specific use cases. Future implementation of these types of systems needs to address this gap to allow a more comprehensive understanding of data use to be gained. Although we broadly understand the data use, we do not have a good handle on the impact of these projects due to the general nature of the data collected. In order to demonstrate relevance, we need tools to understand both the totality of the use and, importantly, the impact of the use case on end users.

DATA ACCESSIBILITY

Data are accessible via the Atlas of Living Australia web portal to the Collectory pages found at (http://collections.ala.org.au/). The data set used in this study is available on the Royal Botanic Gardens Victoria page (http://collections.ala.org.au/public/show/in21) under the National herbarium of Victoria data set (http://collections.ala.org.au/public/show/co55) and can be freely downloaded by clicking on the “download usage stats” button.

LITERATURE CITED

Atlas of Living Australia. 2017a. Atlas of Living Australia – Collectory [online]. Website http://collections.ala.org.au/ [accessed 30 July 2017].

Atlas of Living Australia. 2017b. Atlas of Living Australia – National herbarium of Victoria [online] Website http://collections.ala.org.au/public/show/co55 [accessed 30 July 2017].

Arber, A. 1938. Herbals their origin and evolution, 2nd ed. Cambridge University Press, Cambridge, United Kingdom.

Barrett, R. A., M. J. Bayly, M. F. Duretto, P. I. Forster, P. Y. Ladiges, and D. J. Cantrill. 2014. A chloroplast phylogeny of Zieria (Rutaceae) in Australia and New Caledonia shows widespread incongruence with species-level taxonomy. Australian Systematic Botany 27: 427–449.

Bayly, M. J., G. D. Holmes, P. I. Forster, J. Munzinger, D. J. Cantrill, and P. Y. Ladiges. 2016. Phylogeography, classification and biogeography of Halfordia (Rutaceae) in Australia and New Caledonia. Plant Systematics and Evolution 302: 1457–1470.

Burley, H. M., K. Mokany, S. Ferrier, S. W. Laffan, K. J. Williams, and T. D. Harwood. 2016a. Primary productivity is weakly related to floristic alpha and beta diversity across Australia. Global Ecology and Biogeography 25: 1294–1307.

Burley, H. M., K. Mokany, S. Ferrier, S. W. Laffan, K. J. Williams, and T. D. Harwood. 2016b. Macroecological scale effects of biodiversity on ecosystem functions under environmental change. Ecology and Evolution 6: 2579–2593.

Dodd, A. J., M. A. Burgman, M. A. McCarthy, and N. Ainsworth. 2015. The changing patterns of plant naturalization in Australia. Diversity and Distributions 21: 1038–1050.

Dodd, A. J., M. A. McCarthy, N. Ainsworth, and M. A. Burgman. 2016. Identifying hotspots of alien plant naturalisation in Australia: Approaches and predictions. Biological Invasions 18: 631–645.

Dormontt, E. E., M. G. Gardner, M. F. Breed, J. G. Rodger, P. J. Prentis, and A. J. Lowe. 2014. Genetic bottlenecks in time and space: Reconstructing the invasions from contemporary and historical collections. PLoS One 9: e106674.

Funk, V. A. 2004. 100 uses for an Herbarium (well at least 72). Website http://botany.si.edu/documents/2003_Funk_100Uses.pdf [accessed 31 July 2017].

González-Orozco, C. E., S. Laffan, N. J. Knerr, and J. T. Miller. 2013. A biogeographical regionalization of Australian Acacia species. Journal of Biogeography 40: 2156–2166.

González-Orozco, C. E., M. C. Ebach, S. Laffan, A. H. Thornhill, N. Knerr, A. N. Schmidt-Lebuhn, C. C. Cargill, et al. 2014a. Quantifying phyogeographical regions of Australia using geospatial turnover in species composition. PLoS One 9: e92558.

González-Orozco, C. E., A. H. Thornhill, N. J. Knerr, S. Laffan, and J. T. Miller. 2014b. Biogeographical regions and phyogeography of the eucalypts. Diversity and Distributions 20: 45–68.

González-Orozco, C. E., L. J. Pollock, A. H. Thornhill, B. D. Mishler, N. J. Knerr, S. W. Laffan, J. T. Miller, et al. 2016. Phylogenetic approaches reveal biodiversity threats under climate change. Nature Climate Change 6: 1110–1114.

Haque, M. M., D. A. Nipperess, R. V. Gallagher, and L. J. Beaumont. 2017. How well documented is Australia’s flora? Understanding spatial bias in vouchered plant specimens. Austral Ecology 42: 690–699.

Hui, C., D. M. Richardson, M. P. Robertson, J. R. U. Wilson, and C. J. Yates. 2011. Macroecology meets invasion ecology: Linking the native distributions of Australian acacias to invasiveness. Diversity and Distributions 17: 872–883.

Kooyman, R. M., M. Rossetto, H. Sauquet, and S. W. Laffan. 2013. Landscape patterns in rainforest phylogenetic signal: Isolated islands of refugia or structured continental distributions. PLoS One 8: e80685.

Laffan, S. W., D. F. Rosauer, G. Di Virgilio, J. T. Miller, C. E. González-Orozco, N. Knerr, A. H. Thornhill, and B. D. Mishler. 2016. Range-weighted metrics of species and phylogenetic turnover can better resolve biogeographic transition zones. Methods in Ecology and Evolution 7: 580–588.

Laity, T., S. W. Laffan, C. E. González-Orozco, D. P. Faith, D. F. Rosauer, M. Byrne, J. T. Miller, et al. 2015. Phylodiversity to inform conservation policy: An Australian example. Science of the Total Environment 534: 131–143.

Mishler, B. D., N. Knerr, C. E. González-Orozco, A. H. Thornhill, S. W. Laffan, and J. T. Miller. 2014. Phylogenetic measures of biodiversity and neo-paleo-endemism in Australian Acacia. Nature Communications 5: 4473.

Moles, A. T., S. E. Perkins, S. W. Laffan, H. Flores-Moreno, M. Awasthy, M. L. Tindall, L. Sack, et al. 2014. Which is a better predictor of plant traits: Temperature or precipitation? Journal of Vegetation Science 25: 1167–1180.

Nagalingum, N. S., N. Knerr, S. W. Laffan, C. E. González-Orozco, A. H. Thornhill, J. T. Miller, and B. D. Mishler. 2015. Continental scale patterns and predictors of fern richness and phylogenetic diversity. Frontiers in Genetics 6: 132.

O’Donnell, J., R. V. Gallagher, P. D. Wilson, P. O. Downey, L. Hughes, and M. R. Leishman. 2012. Invasion hotspots for non-native plants in Australia under current and future climates. Global Change Biology 18: 617–629.

Pedicino, L. C., S. W. Leavitt, J. L. Betancourt, and P. K. Van de Water. 2002. Historical variations in δ13Cleaf of herbarium specimens. Journal of Vegetation Science 13: 549–557.

Prentice, E., N. Knerr, A. N. Schmidt-Lebuhn, C. E. González-Orozco, E. N. Bui, S. Laffan, and J. T. Miller. 2017. Do soil and climate properties drive biogeography of Australasian Proteaceae. Plant Soil 417: 317–329.

Saintilan, N. 2009. Biogeography of Australian saltmarsh plants. Austral Ecology 34: 929–937.

Soltis, P. S. 2017. Digitization of herbaria enables novel research. American Journal of Botany 104: 1281–1284.

Staats, M., A. Cuenca, J. F. Richardson, R. Vreliink-van Ginkel, G. Petersen, O. Seder, and F. T. Bakker. 2011. DNA damage in plant herbarium tissue. PLoS One 6: e28488.

Stevenson, L. A., C. E. González-Orozco, N. Knerr, D. C. Cargill, and J. T. Miller. 2012. Species richness and endemism of Australian bryophytes. Journal of Bryology 34: 101–107.

Taylor, S., and I. Kumar. 2012. Sensitivity analysis of CLIMEX parameters in modelling potential distribution of Lantana camara L. PLoS One 7(7): e40969.
Thiers, B. 2017. Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden’s Virtual Herbarium. Website http://sweetgum.nybg.org/science/ih/ [accessed 27 November 2017].

Thornhill, A. H., B. D. Mishler, N. J. Knerr, C. E. González-Orozco, C. M. Costion, D. M. Crayn, S. W. Laffan, and J. T. Miller. 2016. Continental-scale spatial phylogenetics of Australian angiosperms provides insights into ecology, evolution and conservation. Journal of Biogeography 43: 2085–2098.

Tindall, M. L., F. J. Thomson, S. W. Laffan, and A. T. Moles. 2017. Is there a latitudinal gradient in the proportion of species with spinescence? Journal of Plant Ecology 10: 294–300.

VicFlora. 2018a. VicFlora. Website https://vicflora.rbg.vic.gov.au/flora/taxon/83b57182-7112-49e7-923e-c667f89f89bf [accessed 29 January 2018].

Waters, J. M., T. Wernberg, S. D. Connell, M. S. Thomsen, G. C. Zuccarello, G. T. Kraft, J. C. Sanderson, et al. 2010. Australia’s marine biogeography revisited: Back to the future? Austral Ecology 35: 988–992.

Wilkins, C. F., and B. A. Whitlock. 2016. Seringia revised to include Keraudrenia (Lasiopetalae; Malvaceae s.l.). Australian Systematic Botany 28: 265–325.

Woodward, F. I. 1987. Stomatal numbers are sensitive to increases in CO₂ from preindustrial levels. Nature 327: 617–618.