Novel application of species richness estimators to predict the host range of parasites

David M. Watson [a,*], Kirsty V. Milner [b], Andrea Leigh [b]

[a] School of Environmental Science, Charles Sturt University, Albury 2640, Australia
[b] School of Life Science, University of Technology Sydney, 15 Broadway, Ultimo 2007, Australia

ABSTRACT

Host range is a critical life history trait of parasites, influencing prevalence, virulence and ultimately determining their distributional extent. Current approaches to measure host range are sensitive to sampling effort, the number of known hosts increasing with more records. Here, we develop a novel application of results-based stopping rules to determine how many hosts should be sampled to yield stable estimates of the number of primary hosts within regions, then use species richness estimation to predict host ranges of parasites across their distributional ranges. We selected three mistletoe species (hemiparasitic plants in the Loranthaceae) to evaluate our approach: a strict host specialist (*Amyema lucasii*, dependent on a single host species), an intermediate species (*Amyema quandang*, dependent on hosts in one genus) and a generalist (*Lysiana exocarpi*, dependent on many genera across multiple families), comparing results from geographically-stratified surveys against known host lists derived from herbarium specimens. The results-based stopping rule (stop sampling bioregion once observed host richness exceeds 80% of the host richness predicted using the Abundance-based Coverage Estimator (ACE)) worked well for most bioregions studied, being satisfied after three to six sampling plots (each representing 25 host trees) but was unreliable in those bioregions with high host richness or high proportions of rare hosts. Although generating stable predictions of host range with minimal variation among six estimators trialled, distribution-wide estimates fell well short of the number of hosts known from herbarium records. This mismatch, coupled with the discovery of nine previously unrecorded mistletoe-host combinations, further demonstrates the limited ecological relevance of simple host-parasite lists. By collecting estimates of host range of constrained completeness, our approach maximises sampling efficiency while generating comparable estimates of the number of primary hosts, with broad applicability to many host-parasite systems.

© 2016 Published by Elsevier Ltd on behalf of Australian Society for Parasitology.

1. Introduction

When sampling any group of organisms, it has long been noted that sampling effort confounds estimates of species richness (Magurran, 2004; Chao and Jost, 2012). A comparable issue arises in parasitology, where sampling effort influences estimates of both host range (the number of host species infected by a particular parasite) or parasite species richness (number of parasite species that infect a particular host; Poulin, 1992; Walther et al., 1995; Guégan and Kennedy, 1996; Walther and Morand, 1998). Estimates of host range increase with the number of individual hosts examined per host species, and with the number of surveys per parasite species (Walther et al., 1995). Estimated host range is also confounded by distributional extent: as more of the geographic distribution of the parasite is sampled, the greater the proportion of host species that are encountered (e.g., metazoan parasites in freshwater fish; Poulin, 1992).

To minimise the confounding effects of sampling effort and generate reliable and comparable estimates of host range, sampling effort needs to be standardised (Poulin, 1992; Walther et al., 1995; Grenfell and Burns, 2009; Krasnov et al., 2011). Standardisation methods vary in their approach and suitability to a given context. Extrapolation methods compare completeness of sampled sets (Walther et al., 1995; Watson, 2003; Chao and Jost, 2012). Post hoc methods (rarefaction) can be applied to data already collected (Grenfell and Burns, 2009; Kavanagh and Burns, 2012) to remove the confounding factor of sampling effort. This approach to standardisation necessarily results in removal of data to ensure equivalent completeness across sample targets (Watson, 2003) and

* Corresponding author.
  E-mail address: dwatson@csu.edu.au (D.M. Watson).

therefore the loss of potentially biologically meaningful information. Loss of data is more problematic for small datasets (including rare species that are infrequently encountered). Also, this approach may not be appropriate where areas rather than individuals have been sampled or where individuals are clumped (Smith and van Belle, 1984), as is often the case with parasite data. Even where sampling effort bias may be corrected using post hoc extrapolation, the collection bias (especially for herbarium and other museum specimens) remains (Downey, 1998; Norton and De Lange, 1999; Grenfell and Burns, 2009; Kavanagh and Burns, 2012).

Watson (2003, 2010) developed the standardised search to estimate species richness of birds, combining whole-of-patch sampling with results-based stopping rules to generate richness estimates of constrained and comparable completeness. Results-based stopping rules shift emphasis away from the effort expended and focus on accuracy and comparability of results (Peterson and Slade, 1998; Hopps, 2012). Some a priori knowledge of the system under study is required as sample methods, sampling effort and stopping rule are chosen before sampling commences. A robust rule defining the precision of sampling completeness required to estimate richness is determined, then applied during sampling and when the data satisfy the rule, sampling is complete (Watson, 2003). Hence, the effort applied to sample each study site is immaterial—by calculating richness estimates iteratively, effort is scaled to sample completeness, sampling all sites to the same pre-determined degree of precision. In addition to ensuring comparability of estimates, this approach maximises sampling efficiency.

Although species richness estimators have previously been applied to host-parasite data (Walther and Morand, 1998), three aspects of our work are novel. Firstly, we believe this is the first application of an iterative results-based stopping rule in real time to maximise the reliability and efficiency of estimating the number of species hosting a particular parasite. Second, unlike otherwise similar applications of results-based stopping rules for birds (Watson, 2010; Luck et al., 2013), bryophytes (Callaghan, 2012) and licence plates (Peterson and Slade, 1998) to determine whether additional samples are required to estimate species richness at that locality, the stopping rule is here applied spatially to decide whether additional localities within a bioregion need to be sampled to derive robust estimates of the number of species acting as hosts. Finally, treating bioregional inventories of known completeness as samples, richness estimates of host range across the distributional range of parasites were derived and compared with previously published inventories of known hosts to determine reliability. Although using mistletoe species as illustrative examples, the techniques developed here are broadly relevant to studies of host-parasite dynamics generally, representing a tractable approach to estimate host range free from sampling bias.

2. Materials and methods

2.1. Study species

Mistletoes are an ideal system for studying the interplay between sampling effort and host range. Most parasites are more difficult to sample, often requiring capture and careful examination of hosts, magnifying the impact of improvements to sampling efficiency and reliability of host range estimates on study design and logistics. As aerial hemiparasites, both host and parasite are sessile and the parasite is readily discernible (Reid, 1990; Overton, 1994; Mathiasen et al., 2008), allowing accurate estimates of the abundance and density of mistletoes, hosts and potential hosts. The taxonomy of Australian mistletoes is well-resolved

and field identification of species is straightforward (Watson, 2011).

From Downey's (1998) inventory of all known hosts for Australian mistletoe species drawn from herbarium specimens, three mistletoe species were selected representing narrow (dependent on a single host species), intermediate (dependent on host species in one genus) and broad (dependent on many host genera across multiple families) host ranges. Leopardwood mistletoe *Amyema lucasii* (Blakely) Danser is considered to be almost exclusively dependent on leopardwood (*Flindersia maculosa*) hosts (Barlow, 1984; Cunningham et al., 1992; Quirico, 1992; Watson, 2011) but has been recorded on a further six host species (Barlow, 1984; Downey, 1998; Watson, 2011), five of which occur in the family Rutaceae. It is found in semi-arid woodland from the Mitchell District, Queensland, Australia to the lower Darling River, New South Wales (NSW), Australia (Barlow, 1984; Fig. 1A). Grey mistletoe *Amyema quandang* (Lindl.) Tiegh. has been recorded on 53 host species, of which 40 are in the *Acacia* genus with *Acacia dealbata*, *Acacia pendula*, *Acacia harpophylla*, *Acacia homalophylla*, *Acacia aneura* and *Acacia papyrocarpa* commonly recorded as hosts (Barlow, 1984; Reid and Lange, 1988; Quirico, 1992; Keith, 2004; Barea and Herrera, 2009; Bowen et al., 2009; Watson, 2011). *Amyema quandang* occurs in semi-arid and arid woodland in all mainland states of Australia (Barlow, 1984; Fig. 1B). Harlequin mistletoe *Lysiana exocarpi* (Behr) Tiegh. has been recorded infecting 114 species in 45 genera and 21 families (Downey, 1998; Watson, 2011), with the most common hosts belonging to the genera *Acacia*, *Senna*, *Casuarina*, *Eremophila*, *Alectryon*, *Exocarpos*, *Santalum* and *Amyema* (epiparasitic on the latter three parasitic genera in the Santalales). It has the largest distribution of the study species (including eight exotic species) and grows in open forest and woodland in arid to temperate regions of all mainland states in Australia (Barlow, 1984; Fig. 1C). For these three mistletoe species, available host lists (Downey, 1998) representing a comprehensive inventory of all plant species known to host these mistletoes were used to compare predictions based on geographically-stratified surveys (Milner, K. 2014. Optimising estimates of host spectrum: Australian mistletoe as a model system. Honours thesis (unpublished), University of Technology, Sydney, Australia).

2.2. Mistletoe surveys

Field sampling was conducted in April–September 2014 across six of the 17 bioregions of NSW: Sydney Basin, Riverina, Darling Riverine Plains, Cobar Peneplain, Mulga Lands and Broken Hill Complex (Fig. 2), habitats known to support some of the greatest diversity of mistletoe species globally (Vidal-Russell and Nickrent, 2008). Due to the inherently low abundance of mistletoes in this environment (Watson, 2009) and their patchy distribution (Rawsthorne et al., 2012), roadside sampling was undertaken to maximise encounter rates. Mistletoes are more abundant on roadsides due to increased run-off and other improvements in environmental conditions for hosts (Norton and Reid, 1997; Norton and Stafford Smith, 1999; Watson, 2009). Roadside sampling also maximises the diversity of habitats sampled while controlling for land use type. Upon arriving in a given bioregion, sampling commenced as soon as woody vegetation was observed, identifying the habitat type (after Keith, 2004) and carefully scrutinising all trees and shrubs within view of the road (from a slow-moving vehicle, either two or three observers scanning the vegetation on both sides of the road). Upon detecting the target mistletoe species, walking surveys commenced, thoroughly searching roadside vegetation until 25 mistletoe hosts were recorded. This number was chosen after initial analysis based on preliminary data from two mistletoe species determined that 25 records captured variation