



## Exploring Millennium Seed Bank collections from North America- a case study

The Royal Botanic Gardens Kew (Kew) in UK maintains seed collections from over 4800 US plant populations at the Millennium Seed Bank (MSB). We are developing tools to analyse the quality of these collections so that Kew and our partners can identify gaps in coverage and respond to problems that may require training or research support.

### Our questions:

- 1: Which US native seed collections meet fieldwork data standards?
2. Which seed collections meet current viability targets?
3. Which families are typically not supported by comprehensive field passport data?

### The resources that we used:

- (1) The MSB Data Warehouse on line
- (2) Seed accessions data from Kew's Seed Bank Database.
- (3) A standards assessment tool for evaluating collections data quality, including field (passport) data, developed using BRAHMS software.
- (4) Tableau Public<sup>®</sup> software to create data visualisations. These can be embedded in a website to enable user interactivity, or downloaded as standalone graphics (Task 3).

### What is the MSB Data Warehouse?

It is an online tool to explore and analyse regularly updated collection data shared by participating seed banks worldwide. It helps users with the planning of seed collection programmes and in improving their seed germination testing.

Access is restricted to MSBP partners and native seed banks that provide data on their collections. The locality data of sensitive species can be withheld as specified in a Data Provider Agreement with Kew.

The MSBP Data Warehouse is powered by an underlying [BRAHMS](#) database which is compiled and managed on our secure servers at Kew. The amount of data available on the MSBP Data Warehouse continues to grow as more partners share their data within the MSBP community

Visit: <http://brahmsonline.kew.org/msbp>

## Using the geographical filter to explore collections in the MSB Data Warehouse

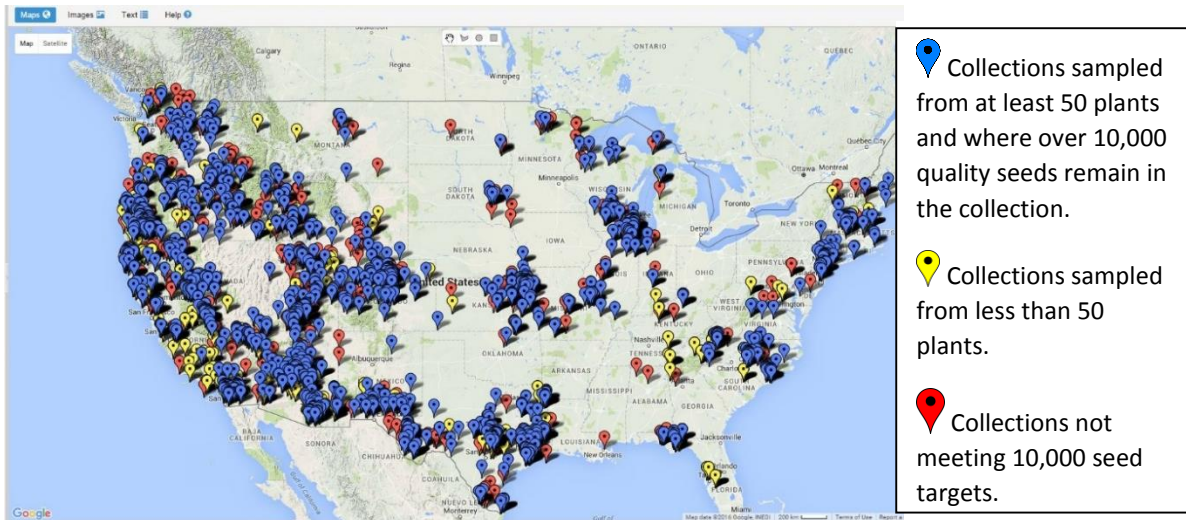
The screenshot shows the MSB Data Warehouse interface. On the left, a 'Geography' filter is set to 'USA (47)'. The main table displays the following data:

Accession	Family	Genus	Taxon	Country	Best Last Test
601887	Compositae	Artemisia	Artemisia michauxiana Besser var. hirsutus Philipson	USA	100
468233	Cyperaceae	Carex	Carex haydenii Dewey	USA	0
514794	Apiaceae	Oxypolis	Oxypolis filiformis (Walter) Britton var. filiforme (A.Rich....	USA	0
414849	Bromeliaceae	Hechtia	Hechtia texensis S.Watson	USA	100
71266	Compositae	Ambrosia	Ambrosia psilostachya DC.	USA	0
221348	Juncaceae	Juncus	Juncus bufonius L.	USA	89
537708	Simaroubaceae	Castela	Castela tortuosa	USA	0
232885	Vitaceae	Vitis	Vitis californica	USA	25



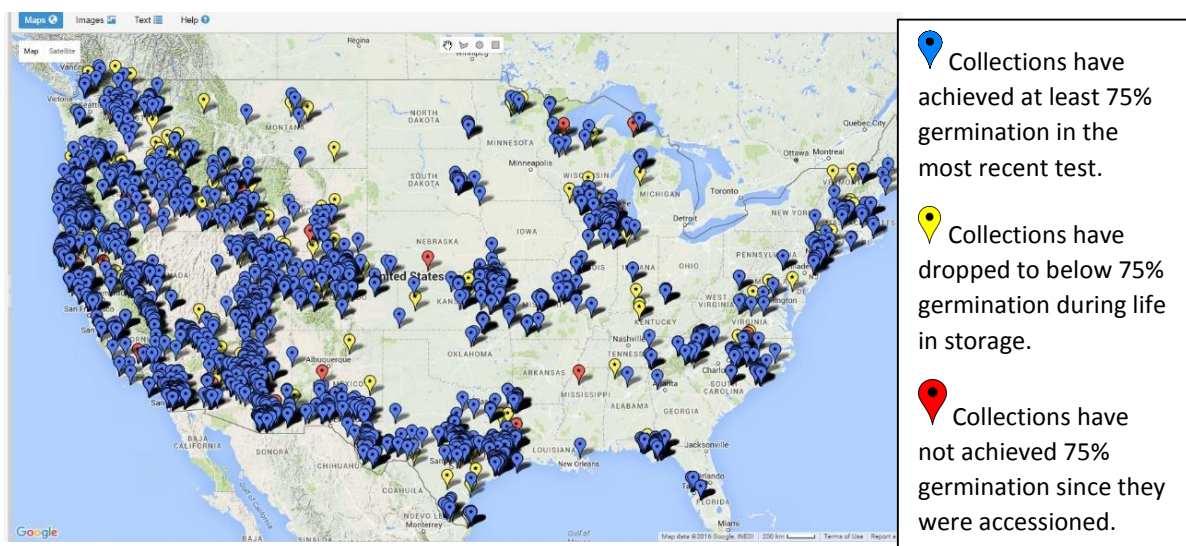
## 1. Which US native seed collections meet fieldwork standards?

Collectors have been trained to sample from large plant populations, to maximise the genetic diversity of the collection and to ensure that sufficient seed is available for research and conservation. We filtered collections that met sampling targets and exported these as kml files to be viewed either in Google Earth or within the Data Warehouse as blue pins (image 1 below). Collections not meeting targets for seed number are shown in Red and for sampling are in Yellow.



## 2. Which seed collections meet current viability targets?

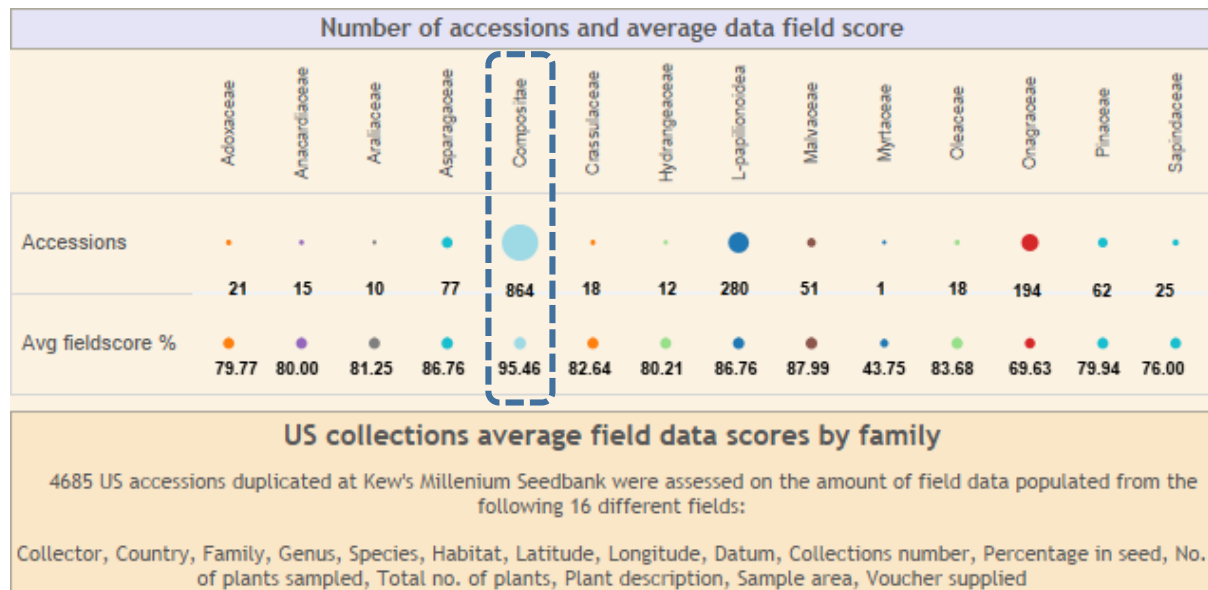
MSB Seed Collections receive an initial viability test and are monitored at intervals to ensure that samples can continue to be distributed. For this task we filtered collections for which at least 75% of the seed tested in the most recent test had successfully germinated (Blue pins). Yellow and Red pins show respectively collections which have not achieved that threshold, or which have dropped below 75% during their life in seed bank storage.





### 3. Which families are typically not supported by comprehensive field data?

Seed Collectors are trained in recording comprehensive field (passport) data that is available for subsequent users of the collections. We wished to know if some families tend to have incomplete data (i.e. from the sixteen data fields that are analysed) and thus only partially meet the MSB field data standards. From the table below we can see that the 864 Compositae accessions (the most well represented family), has an average field (passport) data score of 95.46%.



### Conclusions

We have shown that existing collection data can be easily transformed for in depth analysis of collections' quality, including highlighting potential areas for future improvement. By harnessing new technologies readily available these analyses can be given greater impact through online mapping and interactive data visualisations. The techniques used in this study can be adapted to reveal many other insights, for example by filtering on year of collection or collecting organisation.

### Acknowledgements

Thankyou to data providers including Bureau of Land Management and associated Seed of Success partners.

Department of Plant Sciences, University of Oxford. © 1985 - 2015 Botanical Research And Herbarium Management System (BRAHMS). Version 7. Available from:

<http://herbaria.plants.ox.ac.uk/bol/brahms/> (November 2015)

© 2016 Tableau Software, Incorporated and its licensors. All Rights Reserved

Study authored by Naomi Carvey [n.carvey@kew.org](mailto:n.carvey@kew.org) and Michael Way [m.way@kew.org](mailto:m.way@kew.org)

May 2016



**Index to links for further testing**

### **USA MAP 1**

**HIGH QUALITY COLLECTIONS** – High adjusted seed quantity AND high sample size

<http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/HighQualityUSA.kml>

**LOW SAMPLE SIZE** < 50 plants

<http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/LowSampleUSA.kml>

**LOW ADJ SEED COUNT** < 10,000

<http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/LowAdjSeedCountUSA.kml>

### **USA MAP 2**

**BEST LAST TEST HIGH** > 74

<http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/HighBestLastUSA.kml>

**BEST LAST TEST LOW** < 75

<http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/LowBestLastUSA.kml>

**BEST LAST TEST DECLINING** where best last < 75 AND best ever > 74

<http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/DecliningBestLastUSA.kml>

### **NATIONAL PARK BOUNDARIES**

#### **YOSEMITE**

<http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/YOSEMITE.kml>

#### **CRATER LAKE**

<http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/CRATERLAKE.kml>

#### **BRECON BEACONS**

[http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/brecon\\_beacons.kml](http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/brecon_beacons.kml)

**BRECON BEACONS Tree preservation order areas**

[http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/Brecon\\_Tree\\_Preservation\\_Orderers.kml](http://brahmsonline.kew.org/Content/Projects/msbp/resources/kml/Brecon_Tree_Preservation_Orderers.kml)