

Technical and Financial Progress Report

Data Content Grants for the Digitalization of Data Species and Specimen Thematic Network (SSTN)



Prepared by:

Dr. Hugo Navarrete, project director
&
Priscilla Muriel, Ph.D., principal investigator
QCA Herbarium, Pontificia Universidad Católica del Ecuador

01/01/2010

1. Executive Summary / Resumen ejecutivo

This project aims to digitize, georeference and disseminate data from cloud forest collections hosted at QCA Herbarium, in order to strengthen and disseminate the knowledge about the species found in these forest, thus collaborating with future conservation and management plans.

Since the amended agreement was sent and signed on October 30th 2009, both the activities and the original budget were reprogrammed. The database at QCA was checked to include the DarwinCore fields specified on the signed agreement, excluding nine fields that do not apply to herbarium dried specimens. Two new altitude ranges were defined to select cloud forest specimens for digitization into the database: 1500–4000 m for Andean collections, and 400–1300 m for specimens from the Coastal chains. Geographic coordinates were preliminary searched; however geographic data need to go through major editing, in order to be properly georeferenced.

4944 specimens (50% of the data agreed to be delivered) from Ecuadorian cloud forests were digitized. From these, 4807 records include geographic coordinates. The 4944 collections accessed belong to 1980 plant taxa: 1681 species, subspecies and varieties are represented, grouped into 162 different plant families and 615 genera.

Melastomataceae (414 records, 129 species) and Orchidaceae (215 records, 119 species) are the best-represented and most abundant families. In contrast, more than 61% of the families are represented by less than 50 species each, and 1651 species and infraspecific taxa have ten or less recorded specimens.

The major challenge encountered to the development of the project is that the amended agreement between GS/OAS and PUCE was signed on October 2009, even though the first agreement was sent on April 2009. This caused that all planned activities were re-scheduled to be achieved before April 2010. However, as counterpart funds were available from May 2009, we present on this report 4944 digitized specimens (ca. 50% of the agreed 10 000).

The georeferencing process needs to continue for these and the remaining collections in order to complete the aims of the project.

El presente proyecto tiene como objetivos el digitalizar, georeferenciar y divulgar los datos provenientes de las colecciones albergadas en QCA provenientes de los bosques de neblina del Ecuador, para fortalecer y publicar el conocimiento sobre las especies botánicas presentes en estos bosques contribuyendo con su conservación y planes de manejo.

A partir del envío y firma del acuerdo enmendado en octubre del 2009, tanto las actividades como el presupuesto han sido modificados para cumplir estas metas. La base de datos del Herbario QCA fue revisada para incluir los campos DarwinCore especificados en el acuerdo firmado, con excepción de nueve campos que no aplican para especímenes de herbario. Dos nuevos rangos altitudinales fueron establecidos para seleccionar de entre la colección los especímenes de bosque nublado a digitalizarse: 1500–4000 m para colecciones

andinas 400–1300 m para especímenes provenientes de las cordilleras costeras. Se buscó de manera preliminar las coordenadas geográficas, sin embargo estos datos deben ser editados para una georeferenciación apropiada.

Se digitalizaron 4944 especímenes de bosques nublados (50% de los datos a entregar acordados). De éstos, 4807 registros ya incluyen coordenadas geográficas.

Las 4944 colecciones accesibilizadas corresponden a 1980 taxones: están representadas 1681 especies, subespecies y variedades, agrupadas dentro de 162 familias y 615 géneros de plantas.

Melastomataceae (414 registros, 129 especies) y Orchidaceae (215 registros, 119 especies) son las familias más abundantes y mejor representadas. En contraste, más del 61% de las familias está representado por menos de 50 especies cada una y 1651 especies y taxones infraespecíficos tienen diez o menos registros.

El desafío más importante encontrado para el desarrollo del proyecto es que el acuerdo enmendado entre GS/OAS y PUCE fue firmado en octubre de 2009, aun cuando el acuerdo inicial fue enviado en abril. Esto provocó que todas las actividades planificadas deban ser reprogramadas para completarse antes de abril de 2010. Sin embargo, los fondos de contrapartida estuvieron disponibles desde mayo de 2009, lo que nos permite presentar en el presente reporte 4944 especímenes digitalizados (ca. 50% de los 10 000 acordados).

La georeferenciación está en proceso y debe continuar para éstos y para los nuevos especímenes a digitalizar y así completar los objetivos del proyecto.

2. Programmed Products' Results and Reach of the Project

The main goal of this proposal is to generate data from the cloud forest collections hosted at the QCA Herbarium, in order to strengthen and disseminate the knowledge about the species found in these forest, thus collaborating with future conservation and management plans. For this purpose, the information from the cloud forest collections deposited at QCA will be digitized, their locality data will be georeferenced, and finally the information related to cloud forest collections will be disseminated through a web portal.

Since May 1st 2009¹, we have digitized into the QCA herbarium database 4944 specimens from Ecuadorian cloud forests. These correspond to 50% of the data to be delivered, as stated on the amended agreement received on October 30th, 2009. From the 4944 collections, 97,2% (4807 records) include geographic coordinates, and from these, 4787 have altitude data recorded, the remaining 19 records do not report altitude on their labels. Only 138 records do not record coordinates and need to be completely georeferenced.

The 4944 collections accessed belong to 1980 plant taxa, with 89 specimens determined to family level. 4855 specimens are determined to genus, species, subspecies and variety levels as described in Table 1:

Table 1. Number of records digitized, according to their taxonomic rank

<i>Rank</i>	<i>No. of records</i>	<i>No. of taxa</i>
Family	89	299 (162 ²) (615 ³)
Genus	1163	
Species	3649	1651
Subspecies	13	9
Variety	30	21
TOTALS	4944	1980

The best-represented families are: Melastomataceae (414 records), Rubiaceae (237), Orchidaceae (215), Ericaceae (211), Araceae (210), Piperaceae (194), Campanulaceae (172), Asteraceae (149), Polypodiaceae (134), Elaphoglossaceae (123) and Poaceae (117), with 44 % of the records (2176). Orchidaceae and Melastomataceae are also the most diverse families with 129 and 119 species each (representing 7,7% and 7,1% of the species, respectively). However, more than 61% of the families (1029) are represented by less than 50 species each.

Only 30 species have more than 10 records each, from these the most abundant are: *Miconia theaezans* (57 records), *Brugmansia sanguinea* (34), *Viola arguta*

¹ Although the amended agreement between GS/OAS and PUCE was sent and signed on October 30th, 2009, the first agreement was received on April 1st 2009. Hence, PUCE arranged for the use of the counterpart funds from May 1st 2009.

² This is the total number of families recorded for the 4944 specimens

³ This is the total number of genera recorded for the 4944 specimens

(23), *Saurauia pseudostrigillosa* (22), *Guarea kunthiana* (21), and *Valeriana bracteata* (21). The remaining 1651 species and infraspecific taxa have ten or less recorded specimens.

The province with the highest number of specimens recorded is Pichincha, with 53,6 % of the records (2650). Also, Pichincha is the most diverse province, with 800 species recorded on its cloud forests.

3. Employed methodology and activities done to achieve the programmed products.

The first agreement between GS/OAS and PUCE was received on April 1st 2009, but the amended agreement was sent and signed on October 30th, 2009. Thus, the activities initially planned were reprogrammed in order to fit the new timeline (six months instead of twelve, see Table 2).

1. As stated on the timeline detailed on the table, the first activity was to check the database and include all missing fields in order to include the DarwinCore fields specified on article 1.1 of the signed agreement⁴. The following fields were excluded after analysis, as they do not apply to herbarium dried specimens and to this project in particular: *WaterBody*, *IslandGroup*, *Island*, *MinimumDepthInMeters*, *MaximumDepthInMeters*, *CollectingMethod*, *ValidDistributionFlag*, *Sex*, and *LifeStage*.
2. Based on literature data⁵ and the collections already recorder into the database, a new altitude range was selected (1500–400 m for the Andes Cordillera, and 400–1300 m for the Coastal chains), in order to extract cloud forest specimens from the herbarium cabinets. Additionally, habitat and locality data reported on labels helped discriminating collections not related to Ecuadorian cloud forests (for instance, collections made on Guayllabamba valley, at 2500 m altitude were not included, as the habitat cited on label states “natural and disturbed dry scrub”).
3. After the specimens were extracted from the herbarium, they were inputted into the database, following the usual guidelines and protocols standardized at QCA. Scientific names were added to the database as needed, checking for correct spelling, authorships and abbreviations, according to international

⁴ Darwin Core GlobalUniquelIdentifier, DateLastModified, BasisOfRecord, InstitutionCode, CollectionCode, CatalogNumber, InformationWithheld, Remarks, ScientificName, HigherTaxon, Kingdom, Phylum, Class, Order, Family, Genus, SpecificEpithet, InfraspecificRank, InfraSpecificEpithet, AuthorYearOfScientificName, NomenclaturalCode, IdentificationQualifer, HigherGeography, Continent, Country, StateProvince, County, Locality, MinimumElevationInMeters, MaximumElevationInMeters, EarliestDateCollected, LatestDateCollected, DayOfYear, Collector, Attributes, ImageURL, RelatedInformation, CatalogNumberNumeric, IdentifiedBy, DateIdentified, CollectorNumber, FieldNumber, FieldNotes, VerbatimCollectingDate, VerbatimElevation, VerbatimDepth, Preparations, TypeStatus, OtherCatalogNumbers, RelatedCatalogedItems, Disposition, IndividualCount, DecimalLatitude, DecimalLongitude, GeodeticDatum, CoordinateUncertaintyInMeters, PointRadiusSpatialFit, VerbatimCoordinates, VerbatimLatitude, VerbatimLongitude, VerbatimCoordinateSystem, GeoreferenceProtocol, GeoreferenceSources, GeoreferenceVerificationStatus, GeoreferenceRemarks, FootprintWKT, FootprintSpatialFit

⁵ Muriel, P. 2007. La diversidad de ecosistemas en el Ecuador. Pp. 28–38. En: de la Torre, L., H. Navarrete, P. Muriel, M. J. Macía & H. Balslev (eds.). 2007. Plantas Útiles del Ecuador. Herbario QCA de la Escuela de Ciencias Biológicas de la Pontificia Universidad Católica del Ecuador & Herbario AAU del Departamento de Ciencias Biológicas de la Universidad de Aarhus. Quito & Aarhus.

standards. Finally, the specimens were processed (new determinations and folder labels printed and fixed) and filed back in the herbarium.

4. The geographic coordinates for the specimens were preliminary searched on maps, gazetteers, and the database at QCA; however the geographic data need to go through a deep, major editing, using more up-to-date references (to be acquired using IABIN funding) in order to be properly georeferenced.