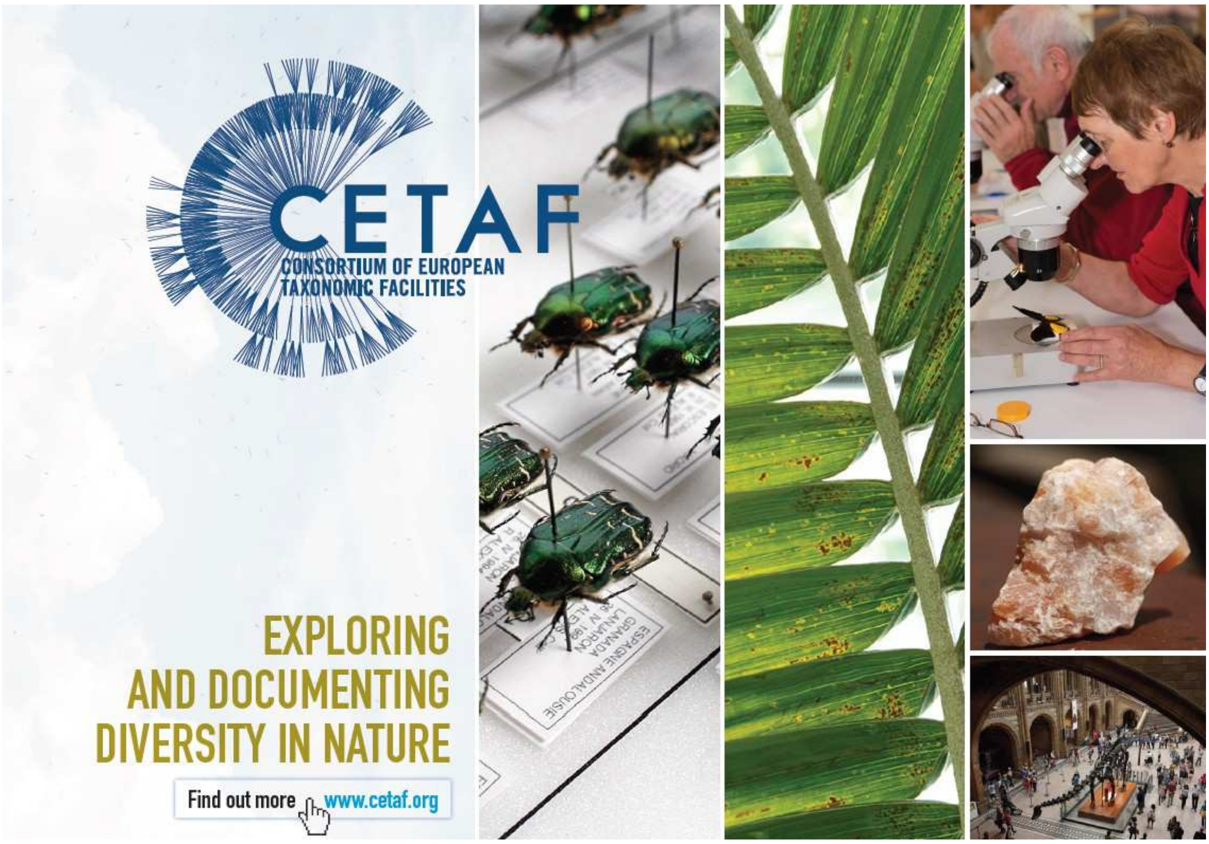

CETAF 43 – Natural History Museum, London

26th April 2018

Workshop: CETAF Joint Research Agenda for Biodiversity and Geodiversity sciences

CETAF Executive Committee Summary Report

1 November 2018

As a response to the upcoming announcement of the 9th European Framework Programme, Horizon Europe, CETAF intends to define its own *Joint Research Agenda for Biodiversity and Geodiversity sciences*. The CETAF Executive Committee identified four questions that were addressed to the CETAF Directors and Representatives attending CETAF 43 (see Annex 1) as a first step towards the development of a common vision.

- What are the relevant scientific questions that collections data can answer?
- What are the new paradigms of research that new technologies will allow us to answer?
- Where do we want CETAF collective scientific research to be in 10 years?
- What do our collections contribute to scientific development that is unique?

Four break-out groups were formed and each dealt with a single question, brainstorming and discussing the issues related to each one. The convenor and rapporteur for each group guided the discussions and recorded the discussion points. The challenges set for each group were to identify the crucial aspects of their topic that are necessary for inclusion in the *CETAF Joint Research Agenda for Biodiversity and Geodiversity sciences*, to detect commonalities among CETAF members with respect to their research objectives, and to explore a long-term vision for the community. At the end of the break-out sessions the convenor or rapporteur presented the main points and outcomes to the full meeting.

Outcomes from the group discussions

QUESTION 1: What are the relevant scientific questions that collections data can answer?

The natural history collections, alternatively termed natural science collections, that are curated and enhanced by CETAF member institutions, are the source of specimen-associated primary data that document geological and biological diversity on Earth. CETAF institutions, as custodians of specimens and of the scientific expertise linked to them, are at the forefront of biodiversity and geodiversity orientated research. By linking named taxa and primary data, scientists from CETAF member institutions conduct research in **taxonomy and Earth sciences**, which constitute the very first level of an integrated biodiversity and geodiversity knowledge system, itself structured from an evolutionary perspective through phylogenies that provide the ultimate systematic framework for all subsequent questions in evolutionary biology.

Taxonomy is a science that is fundamental to the understanding of biological and geological diversity and ecosystem functioning as it seeks to describe, document, classify and interpret the natural world at different taxonomic (species, genera, families), phylogenetic (population, species and beyond) or biological levels (community, habitat, ecosystem). Species discovery, description, documentation and hierarchical or phylogenetic classification unlock data that helps researchers to understand evolutionary events as well as to reconstruct the history of life on Earth. Taxonomy represents a fundamental hypothesis-testing science in its own right, one that should not be regarded as merely a service for other disciplines, because the naming of species and their subsequent circumscriptions are hypothesis-based with these hypotheses being tested over time via successive observations and its acceptance into the *corpus* of taxonomic knowledge. It is thus essential that the community is clear on its scientific objectives: it must develop its own scientific identity and develop its own scientific questions.

Three main lines of research in taxonomy were identified: Evolution, Global changes and Conservation. They derive naturally from the unique scientific expertise developed by scientists in CETAF member institutions, and based on the immense reservoir of biological and geological data associated with some of the largest collections in the world. These three research areas build on the core taxonomic activities of CETAF institutions which respond to such primary taxonomic questions as: **Which species is it? What is its correct name? Should it be described as new to science? What are its primary anatomical, morphological, molecular, behavioural, ecological and distributional characteristics? Is it an alien or invasive species?** At a broader level we could ask the following: **How many species are there on Earth? What are their characteristics? Where and how do they live? With which species do they interact?**

A. Systematics and evolution (taxonomy as a scientific discipline in its own right)

Phylogeny provides an evolutionary structured scientific framework within which the understanding and interpretation of biological attributes can be placed and better understood across time and space. It is the responsibility of CETAF members to provide such a scientific reference, based on as much complete, accurate and up-to-date taxonomic data as possible. With exception of a few charismatic taxa, the phylogenies of most major groups are based on often non-testable or even arbitrary views of their classification, on datasets that are too small or that are biased by incomplete or taxonomically inaccurate sampling.

About 2 million species have been described to date which is an estimated 20-25% of all biodiversity estimated to live on Earth. Revisiting our classifications and phylogenies, including the use of fossils and with a better sampling of taxon diversity using data sets of different origins (anatomical-morphological and molecular characters, behavioural or ecological traits), is crucial to enable us to better refine and confirm our current knowledge.

In addition to new techniques for monitoring data in the field and the automated collection of data on collections (within the DiSSCo agenda), new and more powerful computerized analysis techniques applied to taxonomy (big data, deep learning and Artificial Intelligence-AI, biodiversity informatics,...), the search for new characters (sensorial, structural, chemical and biochemical data, ...) and the integration of new disciplinary fields (EvoDevo, biophysics, ...) into systematic approaches will be important for the establishment and revisiting of phylogenies. We could ask, **How are species classified? What are their phylogenetic relationships? How is a given taxa phylogenetically structured/organized? How do new species discoveries influence our current global knowledge? How do newly discovered fossils impact current phylogenies?** are the typical questions that could be addressed by CETAF institutions within a phylogenetic context.

These phylogenies provide the evolutionary reference framework that is needed to interpret the biological data associated with the specimens in the form of knowledge (= data integration) about taxa (taxonomic and behavioural or ecological macro-patterns). Once acquired, these phylogenies can be regularly updated and completed, becoming a pertinent tool with which we can question our general biological knowledge and test information that is obtained independently from other analyses, but from ecological approaches in particular. Questions like: **What are the behavioural and/or ecological characteristics of a given taxon? What is its role and/or importance in a given ecosystem?** can be integrated into our scientific thinking, as phylogenies provide the basis for a comparative approach, confronting knowledge and data acquired from old and new sources.

The comparative approach in the analysis of data and knowledge on geological and biological diversity is at the heart of the scientific activities carried out in CETAF institutions. Modelling, based on this knowledge, is thus strongly influenced by the accuracy and quality of the results of systematics research. For instance, new molecular calibration methods are becoming more and more sophisticated and the discovery of new fossils allows researchers to anchor evolutionary events in times, and revisit dates of the purported origin of any given taxon.

New questions like, **How old is a given taxon? How has it evolved in time and space? Which scenarios (dispersion, vicariance...) might best explain the evolution and current distribution of a species?** provide a new structured approached to understanding biodiversity, and providing the solid scientific basis for biodiversity knowledge.

B. Systematics and global changes (taxonomy as a scientific discipline at the service of society)

With more than half a billion specimens, CETAF institutions host a fantastic 'reference library' of datasets (biological or geological data that is associated with the specimens) that document past and present geological and biological diversity. Specimen-based data is complemented by new data from biomonitoring facilities that are tracking different parameters (pollution data, landscape use or changes...). Specimens that are collected from all around world allow for the establishment of local, regional and global patterns (distributions, niches, biological associations...), in time and space. In turn, these can be used to document and interpret local, regional and global changes. Questions like **What do collections reveal about the past states of biodiversity? Which reference 'starting point' should be used for ongoing monitoring activities in any given local area? What is the impact of global climate change on species distribution at the local, regional or global level (niche modelling)? How do species respond locally or regionally, both positively and negatively, to climate change in terms of their distribution patterns, new biological interactions, genetic traits?**, and more applied questions, such as **How can the economic impact of alien species be assessed?** or **How can the future of endangered species be shaped via conservation efforts? Should this be done?** could also be incorporated into core research activities.

From a more global perspective, questions like **What does past and present diversity tell us about the past history of the Earth in terms of biogeography, paleo-climate, paleo-environment functioning? How do these past reconstructions help us to model and predict our future? How do these results help us to take informed, environmentally-driven decisions?** will then arise.

Obviously, geographic and time scale issues occur in comparative approaches at the level of taxa (from species or genera to phyla), areas (local to global) and periods (past, current and future). There are also qualitative ones that are encountered when comparing different micro- or macrohabitats/biotopes (natural to anthropized, subterranean to high elevations, dry to humid ecosystems, from fresh water to marine habitats,...).

Tackling these issues will lead us to new research questions in biodiversity but also to questions that have a more theoretical aspect and touch on the advances that can be made based on evolving analytical techniques, like H**ow do improving analytical methods influence comparative approaches in biodiversity? How do they help with the interpretation and visualization of results?** More direct biodiversity-related questions, such as **How has a given trait has evolved over time (evolution) or in different environments (adaptation)? What is the impact of altitude, latitude (temperate versus tropical, ...) on the distribution and diversity of any given taxon? How can we compare these impacts?** may become an integral part of the scientific research agenda of CETAF institutions.

C. Systematics and Conservation issues *(taxonomy as a scientific discipline for decision-making)*

Climate change will impact all biodiversity on Earth in one way or another which in turn will raise important questions for conservation. Decision makers will have to agree on priorities and act upon them accordingly. Important questions of societal relevance will have to be documented and addressed. These research questions will be diverse and could be for local, regional or global purposes or based on short-, medium- or long-term visions. Although research should be driven pure scientific issues, it remains constrained by economic, societal or political pressures and should be tailored, in part, to provide direct answers to societal demands, but especially conservation issues: **Which species are disappearing? From where and why? Why and how do species become invasive? How should invasive or alien species be managed? What effects does climate change have on biodiversity hotspots? How can we detect new evolutionary hotspots? How should be conserve relevant areas? Should local specific diversity or regional phylogenetic diversity be favoured in conversation decisions? What do genetics-based conservation monitoring activities tell us about endangered or invasive species through time and space? What are the historical distribution and ecological niches for population restoration in any given area? Which good quality raw data are needed for the estimation of ecosystem services? What information needs to be provided to decision makers, in response to the different societal and political demands?** In the case of botanical gardens or zoos, **how are CETAF member institutions relevant in conservation efforts? Can CETAF member institutions play a role as trusted sources of material for population restoration from seed banks and/or living collections?**

Scientific issues that are relevant to CETAF members are numerous and multiple, both in fundamental biodiversity science, and more applied approaches. On a **taxonomic basis,** which itself relies on collections, they have a common and structured framework - that is **Evolution** – which constrains them in time and space.

Because we are part of a naturally changing and evolving world, and our understanding of it is continually enriched and renewed as our knowledge of biodiveristy is refined (in terms of taxonomy, phylogenies, biological scenarios), the work of CETAF scientists is being updated on a continual basis. **How much should CETAF scientists commit to other related research areas such as ecology, applied research, etc.? How should this be accommodated or planned?** Questions such as these imply the balancing of limited resources and the need to ensure that institutions are not being forced to deviate from their fundamental collections-based mission in taxonomy.

QUESTION 2: What are the new paradigms of research that new technologies will allow us to answer?

The biggest potential for establishing and implementing new paradigms and new research avenues will come from integrating the opportunities provided by **new technologies** into the exploring, describing, documenting and analysing of biodiversity. Integrating and linking data at large scale will bring us to new levels of knowledge where the whole is more than the sum of its parts. Reflecting biological and geological diversity in a naturally changing world, with the associated global change challenges, large-scale data mining will allow us access to all kinds of new scientific research avenues, such as to the **detection of long-term evolutionary changes** over geological time-scales or **rapid, large modifications** at global scales.

From the taxonomic point of view, new opportunities will come from the **direct identification of species** with handheld devices built using **Deep learning** and **Artificial Intelligence** (AI) processes, linked to direct *in-situ* DNA sampling and open access **on-line knowledge libraries**. Global diversity also represents a reserve of solutions (**Biomimicry**) to discover and decrypt biological and evolutionary patterns that can potentially be used to meet new societal needs, but especially to responds to those that will appear in the face of global changes, and in a context of enhanced environmental conservation.

Research will greatly benefit from the creation of bridges (e-connections) between data that has been generated from within other communities. Such **interoperability** will be built based on shared and standardized ontologies and semantics, supporting a **global distributed network of resources** of primary data and of **interoperating tools** in both bioinformatics and diversity specific applications. New technologies that facilitate the exploration of specimens in different ways (chemical analysis, 3-D scanning) will allow for the potential of prospecting for chemical or biomechanical components, thus realising the innovation potential that may currently be hidden within collections, etc. Developments in technology will not only enable us to follow new paradigms but also to fulfil the old ones, like **completing the inventory of life on Earth**.

Obviously, the development of AI and machine learning will help to take over routine tasks in both collection management and access to data and we will be faced with such questions as, **How do institutions deal with the organisational consequences of technology? How will work with collections, and the physical relationship between a researcher and the material, change?**

QUESTION 3: Where do we want CETAF collective scientific research to be in 10 years?

The establishment of a stable classification, at least to the level of orders, upon which to build future research efforts will be essential to future progress. Achieving a stable phylogenetic classification, including the fossil record, and with the development of faster and more far-reaching techniques, should be within our sights as one of our major collaborative goals. Coupled with big-data and large-scale analysis approaches for phylogenetic and genomic work, the acquisition of a DNA barcode, or the future equivalent, for each species on Earth, is an important development towards documenting life on Earth.

In order to advance common research goals the organisation and establishment of an integrated **European Institute of Taxonomy** or **Virtual Institute of Taxonomy** initiated in Europe would be a relevant tool, as seen with the previous European Distributed Institute of Taxonomy (EDIT) project. The scientific challenges to be addressed in the context of global changes and the current '6th big extinction' cannot be restricted to local actions. On the contrary, they will only be successful if they are placed in a global perspective so the aim to create a virtual taxonomy institute should not be restricted just to the political boundaries of Europe. An institute dedicated to taxonomy and integrated research, including biodiversity data applications, would centralize information promoting efficient time and money usage, as well as exchange within the community at a number of different scales and centred-around different questions or expertise. Such an institute would be a natural hub for training activities and could also be used as a tool to convince decision makers to support both taxonomy and collections. This institute would also be a great supporting mechanism for **integrated research** with a synthesis of taxon data made available via **the provision of integrated and linked data**.

However, conceiving of and creating a **virtual institute of taxonomy** is not in itself research, but a rather a means to facilitate it. A number of concrete steps could be taken to assist research (for example, a list of all taxonomic names with the original descriptions linked to the original names, digitisation and online access to all types, a gap analysis of staff, collections and specialists) along with more research-orientated activities, such as the development of collaborations in biodiversity exploration or the systematic integration of fossils into studies on biodiversity. Generally, a global vision of what research the community wants to do, the ambitions that it may have and the research directions that it wants to follow require a more detailed and profound analysis that seeks to develop a collective set of aims and targets for research, and to explore some big-thinking ideas and new research paradigms.

- **Taxonomy**: To produce a fully illustrated on-line catalogue of species in Europe with their distributions (GBIF occurrences) and features / illustrations / photos, with the aim of producing an on-line identification key to all of them. We could ask to following questions: **How many genera and species do we have in Europe across all the major lineages? Do we have a sampling or documentation bias? What is the rate of species discovery versus species extinction in Europe?** More long-term aims could be to identify / document all major lineages (down to tribes) in Europe and build an on-line key (toward a global Keys to Life project), led by Europe and with strong links to the Tree of Life project and CoL+. Exploration of AI recognition of species and leading innovation by, for example, the creation of smartphone identification and observation recording tools, would help to achieve the expected goals.

- **Evolution:** To place a special focus on producing a phylogeny of the major lineages present in Europe. Questions relevant to this are those such as **Where are the hot-spot of species diversity and of phylogenetic diversity in Europe? How can CETAF institutions contribute to such a collective effort?**

- **Global changes**: To explore documented models of evolution of fauna and flora in the context of global warming. Relevant questions are **What will the future of European diversity and phylogenetic hot-spots look like, and where would they be?** or **How does biodiversity in northern Europe differ from that in the Mediterranean? What is characteristic of each (adaptation to colder climates, recent or old adapted lineages?** It will be equally important to **explore the scenarios for latitudinal migrations to the north, or elsewhere, under the effects global warming**, and keep questioning issues such as **Can mountains become barriers or 'dead ends' for such migrations?** or **Did the Eocene-Oligocene transitional cooling push fauna and flora to the south but will global warming will push it to north?, and if so, how fast?** or **What can we learn from the past for predicting what might be before us?**

- **Conservation:** To produce a harmonized list of threatened and endangered species in Europe. An important effort should be devoted towards maintaining a fully documented list of invasive or alien species present in Europe (description, occurrence, audio-video documents, photos, DNA characterisation,…). In a changing world, **Are the same criteria of evaluation for conservation needed in all parts of the world? For all taxa? Does a threatened/endangered species in Europe belong to the same class of risk as species that are threatened or endangered in other parts of the world, and especially in global hotspots?** By answering these questions a harmonized list of threatened and endangered species in Europe, including possible conservation measures, could be elaborated. A collaborative project, within inputs from different stakeholders, would allow us to develop a European-wide biodiversity monitoring system, with common protocols and common standards. Such a system would also allow us to map in much more precise detail the past, current and changing distribution patterns, as well as species frequency / abundance over time.

QUESTION 4: What do our collections contribute to scientific development that is unique?

Collections of specimens are unique assets that represent "time capsules" giving insights into individual parts of a time series of evolutionary history as well as the recent ‘anthropisation’ of Earth. Each individual specimen contains rich biological, cheminal or genetic data as well as information at the suitable geographical resolution that is necessary for research and conservation purposes. Reproducibility of research results in biodiversity and geodiversity sciences is ensured by the collections themselves (voucher specimens and general specimens) that can be used to study and establish concepts, to test hypotheses and to continually re-evaluate our knowledge base. The physical specimens also represent a vast source of ecological metadata via the analysis of pollen and contaminants, or pests and diseases. This almost endless – and as yet largely untapped – trove of information gives access to data from within and across evolutionary and geological time. These time series are unique and provide a window into the past that can be exploited to complement a range of different studies (dating of taxa, character chronology, etc.). Historical data that can be obtained from collections, and especially from fossils, provides a rich and informative dataset from which to approach evolutionary and well as climate change-response questions. The genetic information stored in the collection allows not only to explore evolution of life on Earth, but also to date its major events (molecular calibration by fossils) and to give account of its current diversity and species community variability. These reservoirs of genomic variation are of interest for industry as well as for conservation efforts.

The collections themselves hold potential for use in a broad range of issues, such as the genomic recovery of populations by providing information of past and present genetic diversity and its origins, in habitat restoration for establishing which plants/animals may have been present in a particular place at a certain time, and in discovery of novel genetic solutions for food sustainability or productivity issues as well as to overcome artificial genetic ‘bottle-necking’, if they house the wild-relatives of domesticated breeds. The material preserved in collections can also be used to provide ‘bio-models’ for industry, biomechanics and bioengineering.

A number of cultural connections with collections are also apparent, namely the cultural heritage that they represent, and especially their importance as a source for documenting scientific thinking, as it evolves over time. Collections were also seen as an important asset in generating ideas and raising curiosity in society (development of question-asking and scientific literacy).

Annex 1. Workshop participants

Austria
Biology Centre of Upper Austria State Museums, Linz
BERNING, Björn
Natural History Museum Vienna
KROH, Andreas
WÖLFEL, Katharina
Belgium
Botanic Garden Meise
DESSEIN, Steven
GROOM, Quentin
Royal Belgian Institute of Natural Sciences, Brussels
VERHEYEN, Erik
PALECO, Carole
MARTENS, Koen,
Royal Museum for Central Africa, Tervuren
MERGEN, Patricia
GERARD, Isabelle
Czech Republic
Czech Consortium: National Museum, Prague
KVACEK, Jiří
FRANK, Jiří
Denmark
Natural History Museum of Denmark, Copenhagen
SEBERG, Ole
Estonia
Estonian Consortium: Estonian Academy of Sciences, University of Tartu
KõLJALG, Urmas
Finland
Finnish Museum of Natural History-LUMOS, Helsinki
SCHULMAN, Leif
France
National Museum of Natural History, Paris
BOURGOIN, Thierry
BENICHOU, Laurence
DEMANOFF, Vanessa
VIGNE, Jean-Denis
ILLIEN, Gildas
GUIRAUD, Michel
Germany
Bavarian Natural History Collections, Munich
NATZER, Eva Maria
Botanic Garden and Botanical Museum Berlin
GÜNTSCH, Anton
HÄFFNER, Eva
Natural History Museum Berlin
HÄUSER, Christoph
QUAISSER, Christiane
German consortium: NORe, Zoological Museum Hamburg
HUSEMANN, Martin
Senckenberg Society for Nature Research, Frankfurt am Main
HÖRNSCHEMEYER, Thomas
Stuttgart State Museum of Natural History
EDER, Johanna
TILLEY, Laura
Zoological Research Museum Alexander Koenig, Bonn
GROBE, Peter
Greece
Natural History Museum of Crete, Heraklion
VOREADOU, Catherina
Hungary
Hungarian Natural History Museum, Budapest
PAPP, Beáta
BABOCSAY, Gergely
Italy
Italian Consortium, Natural History Museum, University of Florence
BARTOLOZZI, Luca
Israel
Steinhardt Museum of Natural History, Jerusalem
MEIRI, Shai
NHCOLL, Jerusalem
GAL, Gila Kahila-Bar
RABINOVICH, Rivka
Netherlands
Naturalis Biodiversity Center, Leiden
SMETS, Erik
KOUREAS, Dimitris
Norway
Natural History Museum - University of Oslo
MYHR, May Britt
MEHLUM, Fridtjof
Poland
Museum and Institute of Zoology - Polish Academy of Sciences, Warsaw
BOGDANOWICZ, Wieslaw
Slovakia
Slovakian consortium: Slovak Academy of Sciences - Institute of Botany, Bratislava
BÉREšOVÁ, Anna
Spain
Spanish consortium: CSIC, National Museum of Natural Sciences, Madrid
REY, Isabel
SANTOS, Celia
Spanish consortium: CSIC Royal Botanic Garden of Madrid
MUÑOZ, Jesús
Sweden
Swedish Museum of Natural History, Stockholm
JOHANSSON, Kjell-Arne
BISANG, Irene
Gothenburg Consortium, Department of Biology and Environmental Sciences, Gothenburg University
PERSSON, Claes
Switzerland
Geneva Consortium, Conservatory and Botanical Garden of the City of Geneva
PRICE, Michelle
United Kingdom
London Natural History Museum
SMITH, Vince
Royal Botanic Garden Edinburgh
HASTON, Elspeth
BHL
RINALDO, Connie (USA)
Species2000
BANKI, Olaf
CETAF General Secretariat
CASINO, Ana & GÖDDERZ, Karsten

Annex 2. Workshop results.

Comments from each of the workshop breakout groups

1 What are the relevant scientific questions that collections data can answer?

Convenor: Thierry Bourgoin

- Integrated virtual institute of taxonomy
- Synthesis of taxon data
- Integrated and linked data (DISSCO)
- Biology and earth sciences – integrated research combining these fields
- Researchers are trusted - repeated measurements
- Collection based research (joined with many NHMs all fields)
- Combine data from different NH areas to achieve new results – linking data
- Stable classification
- Synthesys to promote taxonomy
- Database of all types of all genera (tools)
- Names of collectors
- Registration of all names including fossils (plants particularly)
- Collection management system
- Working classification above order level
- DNA barcoding of all genera (taxonomists first) – new results combination - BIG data
- European monitoring system (common protocols)
- Registration – names on-line
- Names of fossil plants on-line
- Mechanism in place to help promoting alpha taxonomy, institutions provide less and less resources
- European institute of taxonomy – we do not have specialists – losing taxonomists – need positions, people to fill them
- Providing type specimens from EU museums (and other services – digitisation) to tropical countries
- Original descriptions associated with specimens (to EOL)
- Societal relevance of science
- On European level – convince decision makers (directors) – gap analysis (expertise, collections)
- Optimizing resources of NH knowledge in Europe Designation of
- Actual taxonomic work is not well supported
- Gap analysis - staff, expertise, collections

Topics covered in EDIT: Species discovery & identification / Evolutionary relationships and patterns / Micro- and macroevolution / Evo-devo / Biological systems (molecular, organisms, ecosystems) / Molecular ecology.

2 What are the new paradigms of research that new technologies will allow us to answer?

Convener: Vince Smith

- DNA sequence data from NGS platforms
- Applying technology to see how, in quality and quantity, we can go with starting DNA in standard set-ups
- Adaptive traits over time (in response to the environment) through next generation techniques - hybrid baits
- Rapid automated identification of existing and new species, and linking this automatically to specimen data
- Image and data capture > big data and big analyses linked to large scale technology facilitated new field data (drones, satellites, eDNA, meta-barcoding)
- Biomechanics and material science - 3D or CT scanned organisms provide detailed models on structural solutions
- Build / expand better image recognition tools so that images in digital databases can be extracted, identified and linked
- Terra-forming Mars
- Formalized ontology based descriptions
- Chemical profiles via various methods
- Imaging of the determination of chemical ingredients of medicinal or aromatic plants (which could be useful in the pharmaceutical industry)
- Ensure that existing networks communicate and are integrated
- Landscape genomics combining distribution modelling with genomic data
- Genetic diversity in time and space - when all collections are connected they start forming population samples
- Better integration and standardisation of different types of data
- Nano-publications
- Field identification devices (citizen science) via e-DNA applications, AI to speed up human work (but not replacing them completely), modelling, data capture
- Semantic techniques (LOD, graph based analyses)
- All taxa inventories - metagenomics (collections provide a reference)
- New tools to integrate molecular phylogenies, including fossil and morphological time-series as well as ecological information to solve macro-evolutionary problems
- By linking databases we can detect long-term evolution of biological & geological diversity from the past geological times until today
- Autonomous sentinels discovering new species in the wild
- Whole globe biodiversity map -when all collections with data are digitised it will be possible to link species occurrences with abiotic factors correlates and other species - model of BD map
- Drop relational databases - data sink (hello to graph bases)
- Key is bringing data together (SEWEB) - occurrence, geology, climate, molecular, traits, social, policy combined with new methods of analysis (mining, correlation etc.).
- More interlinking / integration with specimens linked to taxon descriptions in the literature (type specimens have links to collection, to literature, to archival field notes and ledgers and molecular data)
- Take advantage of the new generation of infrastructure (chemical imaging) to discover new potential in specific collections.
- Open Access question - publication of articles and the publication of research datasets
- Description / discovery of all species on Earth
- Shorten "time to market" - one authority for taxon names with published names immediately available
- Seamlessly zoom in and out between micro and macro levels - - from subatomic to full specimens, and even groups of specimens

- Virtual access on demand through robots looking up and digitizing specimens (when requested) in various levels of details (e.g. 3D)
- Automated biodiversity mapping and monitoring (images and sound recordings, AI) - collections form baseline references
- Data extraction from combined datasets (literature plus museum databases plus molecular databases)
- See immediately what specimen has been studied for a taxonomic treatment and link specimen reference data easily to ecological studies
- Combining molecular and digital technologies to set up new paradigm on how we communicate species in research, nature protection, etc.
- Join all counting in Eurasia - forming a real possibility to have big data of biota (for example of invasive species). The sea realm is all missing.
- Massive data mining using data clusters (metadata from specimen databases; library, archives and catalogue; information from digitised texts or transcribed material; information from ? digital material (e.g. e-journals); bioinf. (DNA); information from biodiversity research / citizen science databases and protocols.
- Massive data mining using data elements (candidates) of interest. For example, authorities / entities (subjects - taxa, agents - authors names, dates - events, places – georeferencing) with links between those elements
- Massive data-mining using usage scenarios (potential functionalities). For example, algorithmic computing for data identification and consolidation, assessment, aggregation or interpretation or data visualisation for enhancing outcomes (link clusters, time-lines, tag clusters, maps)

3 Where do we want CETAF collective scientific research to be in 10 years?

Conveners: Erik Smets & Jiři Kvaček

- Collaboration in biodiversity exploration - exchange expertise to elaborate the whole material collected in expeditions
- Stable working classification of all organisms
- Combine and analyse large amounts of data from different realms routinely
- A DNA barcode of all genera
- Big data
- Shared portal of natural history collections - DiSSCo
- Have all original descriptions linked to scientific names in an online system
- A full register of collections
- Digitisation and online access to all types / Create a European and worldwide database identifying type material & type specimens
- A good CMS system
- Identification of all duplicate specimens
- Collaborative work on fossils in the catalogue of life
- A list of all taxonomic names
- A common data management plan
- A system of registration of plant names
- Our research results regarded by the public as honest and trustworthy because our collections are the documents and we are not dependant on the economy
- Truly united research community - divide European Staff along (among?) natural history collections
- European Institute of Taxonomy
- Support mechanisms for actual taxon-based work in Europe (Synthesys++)
- EU biodiversity monitoring system with common protocols
- Gap analysis of staff, collections and specialists
- Collaborative collection-based research for questions of general interest
- Apply fully integrative taxonomy (morphology, molecular,…)
- Fundamental and applied collection-based research
- Projects overarching biodiversity and Earth sciences (integrating both)

4 What do our collections contribute to scientific development that is unique?

Conveners: Nikolaj Scharff & Jesús Muñoz

- Digital preservation research
- Cultural heritage and documentation of scientific thinking
- Cultural historical information
- Reproducibility of research results (voucher specimens)
- Generating ideas and raising curiosity
- Endless source of (new) information (that we cannot predict)
- Untapped source of information of data above and beyond GBIF (specimens)
- Presence of the collection is a unique asset
- Genomic recovery of populations
- Reservoirs of genomic variation with links to the food industry
- Bio-models for industry, biomechanics and engineering
- Predicting future changes (sequence evolution distribution)
- Great spatial resolution an)d ecological metadata (contaminants, pollen
- Potential for interdisciplinary use of the data
- Specimens in the collections are physical time capsules
- Time series (e.g. genetic, for conservation)
- Keeping and updating conservation methods so collections will be as useful as possible in the future
- Testing the ground for new techniques
- Improvement of methodologies testing hypotheses (and compare to earlier hypotheses)