biodiversa

The ERA-Net promoting pan-European research on biodiversity and ecosystem services

Mapping the biodiversity research landscape with the

BiodivERsA database

TEMPORAL AND GEOGRAPHICAL TRENDS IN THE TYPE OF BIODIVERSITY RESEARCH FUNDED ON A COMPETITIVE BASIS IN EUROPEAN COUNTRIES

Mapping the biodiversity research landscape with the BiodivERsA database: temporal and geographical trends in the type of biodiversity research funded on a competitive basis in European countries.

To cite this report

Gambette P., Eggermont H. & Le Roux X. (2014) Temporal and geographical trends in the type of biodiversity research funded on a competitive basis in European countries. BiodivERsA report, 33pp.

Photography credits

Cover:
Paul Kempf

Other pictures, order of appearence:
Inside front cover: Adrien Regnier-Laurent - The BiodivERsA partners: Adrien Régnier-Laurent - Content: Frédéric Lemaitre - Page 7: Edwin Brosens - Pages 8 and 9: Paul Kempf - Pages 11 and 12: Adrien Régnier-Laurent - Pages 14 and 15: Adrien Regnier-Laurent - Page 18: Adrien Régnier-Laurent - Page 22: CNRS Photothèque/ VIGNAUD Thomas - Pages 28 and 29: Adrien Regnier-Laurent - Page 30: Paul Kempf - Back cover: Paul Kempf.

Acknowledgements

We would like to thank the BiodivERsA partners who provided information on national programmes and projects, and reviewed a previous version of this report. We also gratefully acknowledge the assistance from André Heughebaert (Belgian Biodiversity Platform) for developing and co-managing the BiodivERsA database.

Layout

Thibaut Lochu

To contact BiodivERsA

BiodivERsA Coordination

Xavier Le Roux, Coordinator and CEO
xavierleroux@hotmail.fr
Tel: +33 (0)6 31 80 38 20

BiodivERsA Secretariat

Frédéric Lemaître, Executive Manager
frederic.lemaitre@fondationbiodiversite.fr
Tel: +33 (0)1 80 05 89 37

Fondation pour la Recherche sur la Biodiversité
195 rue St Jacques, 75005 Paris France
www.biodiversa.org

To access BiodivERsA database

www.biodiversa.org/database

The BiodivERsA partners

Fondation pour la Recherche sur la Biodiversité (FRB), FRANCE [coordinator]

Fonds zur Forderung der Wissenschaftlichen Forschung (FWF), AUSTRIA

Belgian Science Policy Office (BELSPO), BELGIUM

Bulgarian National Science Fund (BNSF), BULGARIA

Estonian Research Council (ETAG), ESTONIA

Agence National de la Recherche (ANR), FRANCE

Ministère de l'Ecologie, du Développement Durable et de l'Énergie (MEDDE), FRANCE

Bundesministerium für Bildung und Forschung (BMBF)/ Projektträger im Deutschen Zentrum für Luft-und Raumfahrt e.V. (PT-DLR), GERMANY

Deutsche Forschungsgemeinschaft (DFG), GERMANY

Ministry of Rural Development (VM), HUNGARY

Research Council of Lithuania (RCL), LITHUANIA

Nederlandse organisatie voor Wetenschappelijk Onderzoek (NWO), THE NETHERLANDS

Research Council of Norway (RCN), NORWAY

Fundação para a Ciência e a Tecnologia (FCT), PORTUGAL

Ministerio de Economia y Competitividad (MINECO), SPAIN

Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning (FORMAS), SWEDEN

Swedish Environmental Protection Agency (SEPA), SWEDEN

Ministry of Food, Agriculture and Livestock (MFAL), TURKEY

Department for Environment, Food and Rural Affairs (DEFRA), UNITED KINGDOM

Joint Nature Conservation Committee (JNCC), UNITED KINGDOM

Natural Environment Research Council (NERC), UNITED KINGDOM

Natural Environment Research Council (NERC), UNITED KINGDOM

CONTENTS

INTRODUCTION

Research on biodiversity and associated ecosystem services in Europe: a fragmented landscape

Biodiversity – the variety of living organisms, their habitats and their genes - on which so much of human life depends, is under pressure and its degradation is one of the world's most pressing crises (Rockström et al., 2009). The current species extinction rate is estimated to be between 100 and 10000 times higher than it would naturally be, raising issues for the preservation of living forms on Earth and putting at risk the ecosystem services they provide to humanity (Chivian and Bernstein, 2008). If this trend continues, we could reach tipping points where these vital services are no longer sufficiently delivered. Hence, there is a strong and urgent need for more research on biodiversity and associated ecosystem services to better identify and understand the opportunities and risks associated with biodiversity protection, management and use. New knowledge has to be gained at all scales, from national to regional and global, with a major challenge to promote approaches that cross national borders and gather different disciplines and types of actors.

Biodiversity research can be promoted through devoted programmes, but more often it is supported through non-specific programmes. Indeed, biodiversity appears in various research themes such as sustainable management and use of natural resources, ecotoxicology and environmental health, ecological engineering and green economy, management of protected areas, and global change impacts: this is true for many research strategies and priorities both at the international and national levels, and for most research funding schemes.

Since there is not a unique and delimited entrance to fund biodiversity research, the total amount of funding allocated to biodiversity projects and the type of research funded can hardly be quantified at the European level. It is particularly challenging to profile the type of research funded and analyse complementarities between countries or funding programmes. This requires to first sort and gather information from a range of sources. Nice examples of such an exercise are given by Matei et al. (2011), analysing the funding of the "biodiversity" topic within the 7th Framework Programme (FP7) "Environment" theme; and Chaveriat et al. (2011) zooming in on the landscape of biodiversity research in France.

The BiodivERsA ERA-Net and its activities to map the research landscape in Europe

This context is of major importance for BiodivERsA, a tool part of the European Research Area Network (ERA-NET) scheme of the European Commission (EC). Now at the end of its second four-year funding phase (2010-2014), BiodivERsA brings together a network of 21 funding organisations from 15 European countries that aims at building a dynamic platform for encouraging excellent and policy-relevant research on biodiversity and associated ecosystem services at pan-European scale. The BiodivERsA network is now launching annual calls for proposals on topics that correspond to the most pressing issues that biodiversity and ecosystem services currently face. Every year, the network updates its research agenda based on existing national, European and international agendas, ensuring that the most relevant topics are prioritized. This shared agenda avoids duplications and inefficiencies resulting from a fragmented approach. During the 2008-2014 period, BiodivERsA launched 5 joint calls for research proposals for a total amount of ca. €100M including €50M of new money raised by the participating agencies.

In order to support these mapping activities, BiodivERsA has developed a regularly updated database (Eggermont et al. 2013) holding information about:

- funding programmes and associated calls for research proposals on biodiversity and associated ecosystem services in Europe; this includes

thematic programmes devoted to biodiversity, thematic programmes including biodiversity issues, and blue sky programmes where biodiversity research applications are eligible.

- research projects on biodiversity and associated ecosystem services funded through these programmes
- research institutes and other organisations (including stakeholders) involved in the projects funded, and researchers leading the projects.

This database ultimately aims at including the different funding schemes that can fund biodiversity research projects. At the European level, these instruments include the Framework Programme for research and development, European Research Council (ERC) grants, LIFE+, the European regional development fund, the European agricultural fund for rural development, and ERA-nets (the latters can mix EC and national funding sources). Funding opportunities also emerge under, and together with Horizon 2020 that is the new EU framework programme for research and innovation. Research programs from ERA-nets are also taken into account, identifying resources allocated from each national funder for each project.

The BiodivERsA database focuses on competitive allocation of funds to research only, either at national or European level, excluding e.g. funding by national institutes only targeting the scientific community they are in charge of, or schemes that fund research at a local level.

The BiodivERsA database is accessible at www.biodiversa.org/database. Upon registration and acceptance of the data user agreement, the information is available to anyone interested in biodiversity research funding in Europe. Besides allowing the analysis of the funding landscape for biodiversity research (Eggermont et al. 2013), the profile of funded research and the possible complementarities of biodiversity research priorities among countries and agencies in Europe (this brochure), the database will also help scientists to identify potential resources and network opportunities to further develop their research. In addition, it will help at finding scientific expertise for specific policy questions. As such, the BiodivERsA database can be considered a valuable tool for strategic cooperation and expertise in the large, fragmented domain of research on biodiversity and associated ecosystem services.

A database to analyse the type of biodiversity research funded in Europe

A first brochure published by BiodivERsA (Eggermont et al., 2013) presented the BiodivERsA database, and its use to analyse selected features of the European funding landscape for biodiversity research (i.e. level of funding encompassing a multitude of funding schemes; temporal funding trends for the 2005-2011 period; and the level of funding by national agencies versus Framework Programme at the European Scale).

Based on a semantic analysis of the texts of project abstracts, the main goals of this second brochure are to:

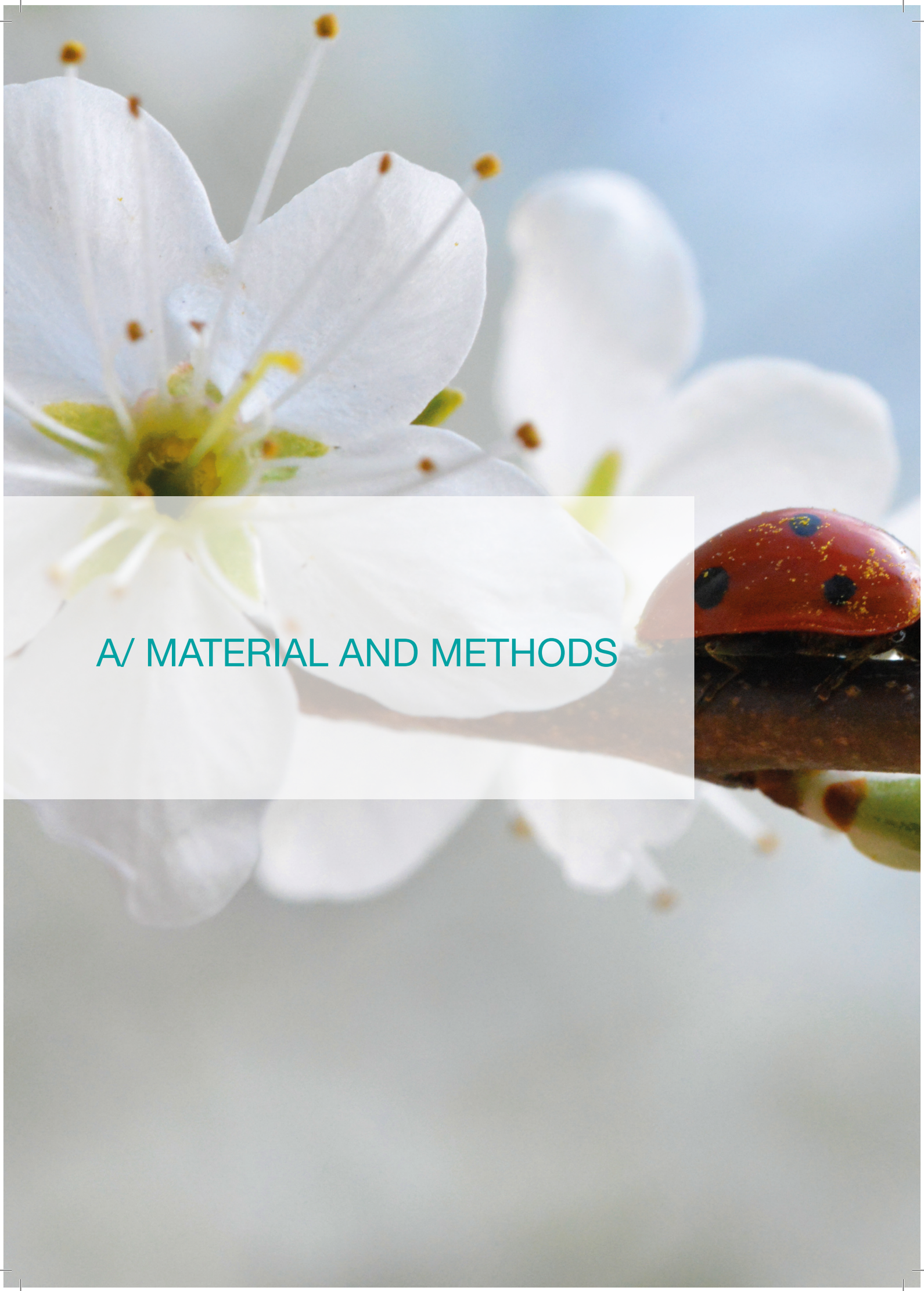
- Test the existence of any **coherent temporal trends (over the 2004-2011 period) in the type of research funded** at the European level in this domain, i.e. when aggregating information from many national funding agencies across Europe;
- Analyse the changes in the type of biodiversity research funded, identifying the **main research topics that are decreasingly and increasingly supported over the 2004-2011 period;**
- Compare the **type of biodiversity research funded between different national agencies** that program and fund research in European countries.

The BiodivERsA consortium of national programmers and funders of research on biodiversity and ecosystem services includes agencies as well as Ministries, and is tightly linked to research stakeholders (from Ministries to scientific community and NGOs, practitioners and businesses). BiodivERsA is thus well placed to identify and provide explanations for the temporal and geographical trends that

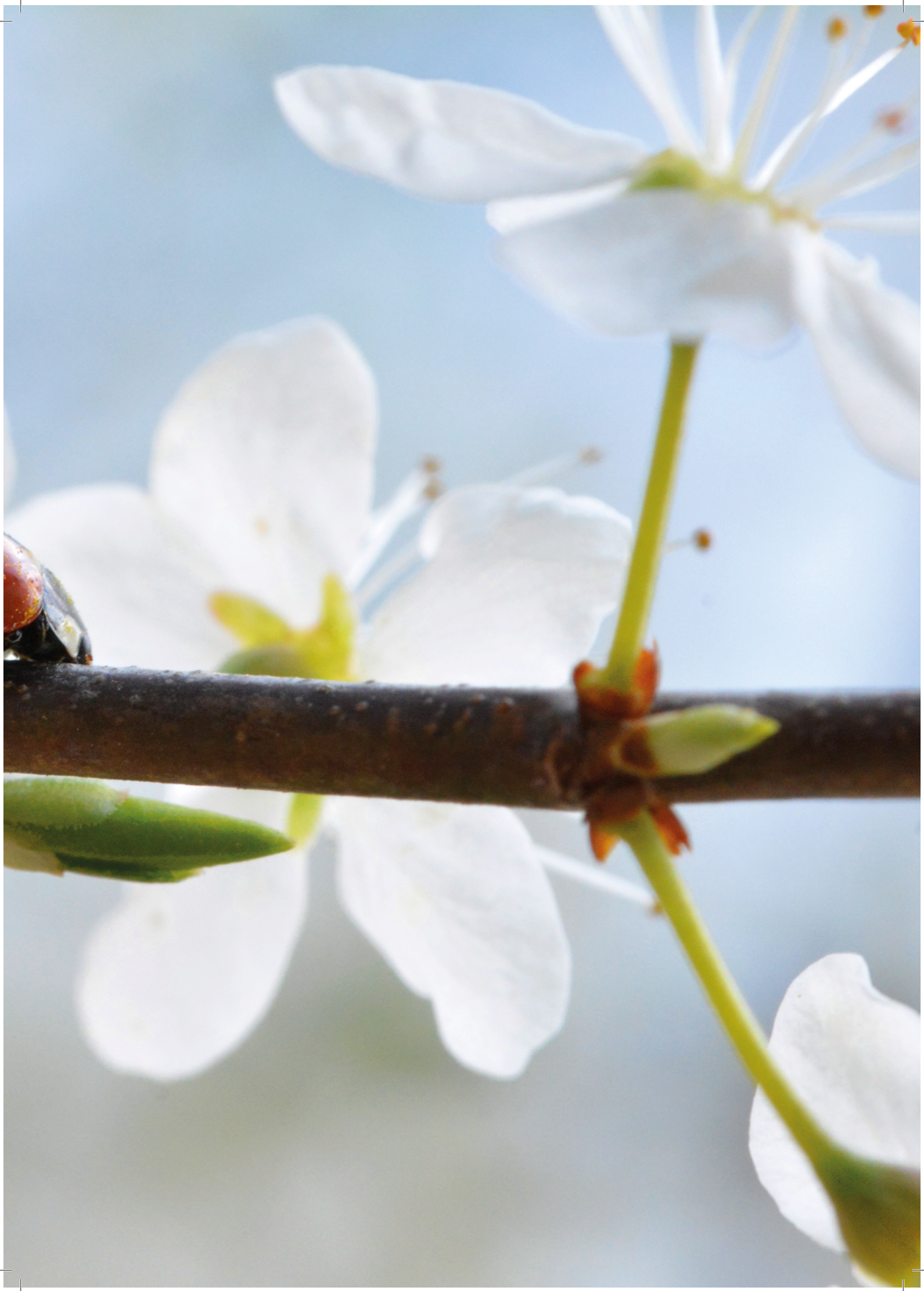
are observed in the field. Still, as mentioned above, obtaining extensive information on all the biodiversity research projects funded in a given country is hardly reachable. Here, we studied 4159 abstracts of projects funded by 15 national funding agencies that are main competitive funding sources for research on biodiversity and ecosystem services at the national scale in Europe (Eggermont et al., 2013), or by the European Commission. We explored this corpus with statistical textual analysis. Though this analysis does not allow for a complete between-country comparison, it can elucidate some major trends in the type of biodiversity research funded in European countries. In addition, we hypothesized that changes in the vocabulary used in project summaries provide good information on changes in the type of research itself. Although this has limits (for instance, researchers can adapt summaries and the vocabulary used to expectations from funders), our results of comparisons between years and between countries seem to be very consistent and robust. Hence, they can be reliably interpreted in terms the type of funded research.

We hope this mapping activity will help a large range of stakeholders (researchers, institutes, funding agencies and policy makers) in getting a more comprehensive view of the current geographical, institutional and temporal trends of the type of research on biodiversity and ecosystem services funded in Europe.

Field sampling at the Col du Lautaret



A/ MATERIAL AND METHODS



A/ MATERIAL AND METHODS

A1. CORPUS OF PROJECT ABSTRACTS

Our corpus was extracted from the BiodivERsA database on June 2013. The database was built by importing and gathering data from around 50 national bodies that fund biodiversity research (mainly national funding agencies and ministries), also including data from major European calls from the European Commission.

The BiodivERsA database presently includes **605 annual calls** (including the biodiversity-relevant European FP6 and FP7 programmes, and BiodivERsA ERA-net calls) including 100 biodiversity-specific calls funding **6546 research projects**. Those data span the **2000-2014 time window**, and currently cover **17 countries**: Austria, Belgium, Bulgaria, Estonia, France, Germany, Hungary, Ireland, Italy, Lithuania, the Netherlands, Norway, Portugal, Spain, Sweden, Turkey, and the UK. For a detail of the database, see Eggermont et al. (2013).

Projects were carefully screened to check whether they could unambiguously be defined as projects on biodiversity and associated ecosystem services. Biodiversity is defined here according to the United Nations Convention on Biological Diversity, as "the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems". This means that the projects, at least partly, explicitly analyse and account for biological diversity; projects focusing on services without any link to biodiversity (e.g. a project focused on C fluxes and sequestration without any focus on e.g. the diversity of soil micro-organisms, soil fauna, plants or ecosystems) are not selected. The information was always validated by the partners using these criteria; as such, data quality remains their responsibility. However, the authors of the present report checked data quality on a sample of the corpus from individual agencies, ensuring that the criteria have been correctly used.

For the present analysis, we focused on projects with an English abstract, and on years with a total of at least 399 of such projects, within the period 2004-2011 (Figure 1)

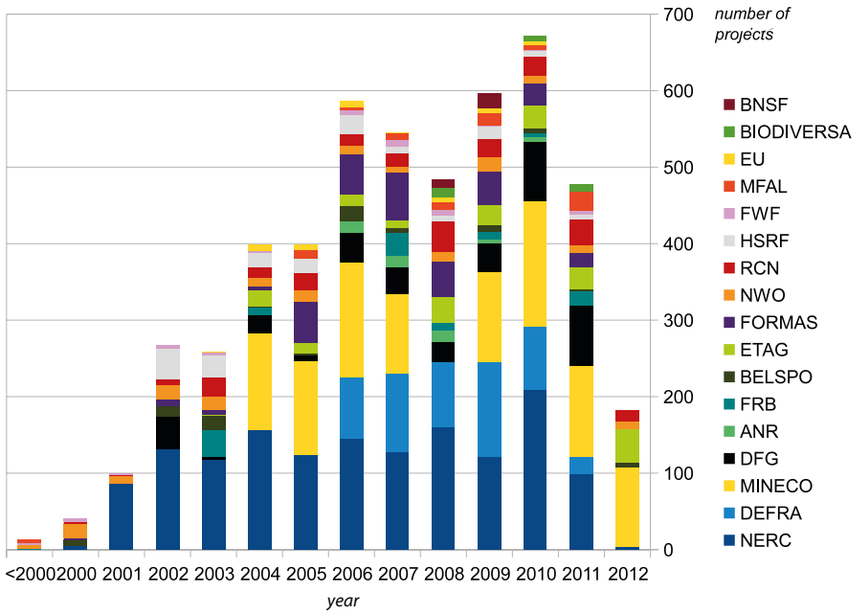


Figure 1: Distribution of the total number of projects referenced in the BiodivERsA database per year and per funding agency over the 2000-2012 period. The present study focused on the 4159 projects with abstracts in English within the 2004-2011 period.

This amounts to 4159 project summaries, corresponding to a total of 1 218 741 words. We decided not to translate abstracts not written in English because the choice of words by applicants is at the heart of the present analysis. The distribution of projects according to their funding agency over the 2004-2011 period is given in Figure 2.

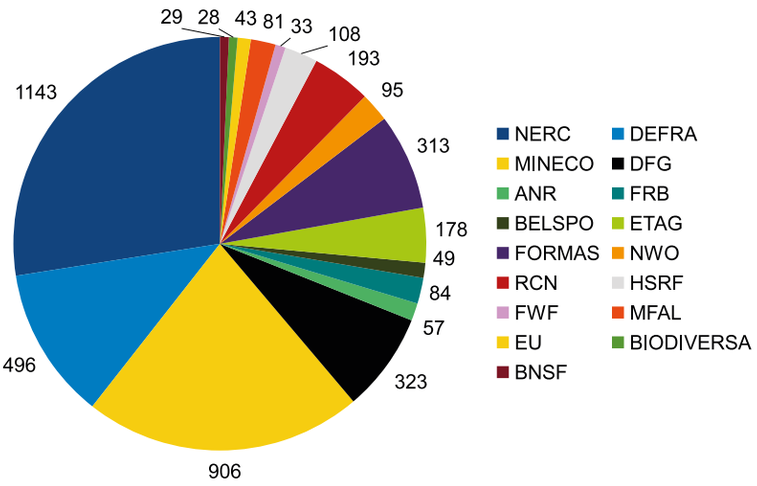


Figure 2: Distribution of the 4159 projects with English abstracts among funding agencies for the 2004-2011 period.

A2. ANALYSIS OF TEMPORAL AND GEOGRAPHICAL TRENDS, AND IDENTIFICATION OF OVER- OR UNDER-REPRESENTED VOCABULARY

Each project summary was characterized with an identifier, the year of project start, and the agency funding the project. The whole corpus was converted to lowercase in order to normalize the words, especially those at the beginning of a sentence, as well as those appearing in project summaries entirely written in uppercase letters.

Lexico 3 (Salem et al., 2003) was used to generate correspondence analyses of the corpus, focusing on the funding year for each project (for temporal trend analysis) or on the funding agency (for 'geographical' trend analysis).

Lexico 3 was also used to compute the under- or over-represented vocabulary in a sub-corpus compared to the whole corpus. The software provides a specificity score indicating whether or not the over-/under-representation is statistically significant, based on an assumption of hypergeometric distribution of words (Lafon, 1980 ; Lebart et al. 1998). This computation was applied to single words or coherent groups of words.

A3. TREE CLOUD VISUALIZATIONS

We generated trees of words with TreeCloud (Gambette & Véronis, 2009). This program visualizes the most frequent words of a text by arranging them according to a co-occurrence principle: the words appear close to each other in the tree if they frequently appear together in a 10-word window (right or left) in the same project summary. TreeCloud was used with the Liddell co-occurrence distance (Gambette, 2010), ignoring the words in a customized stoplist specially built for this study. The latter contains the words from the default stoplist provided by TreeCloud for the English language, plus some frequent words which carry no information about research topics, like investigate, determine, project, etc. Starting from the co-occurrence distances between the words, the tree was built with the Neighbor-Joining algorithm (Saitou & Nei, 1987) and visualized by SplitsTree 4 (Huson & Bryant, 2006).

Tree clouds were constructed for the 100 most frequent words used: in this case, the sizes of the words in each tree cloud logarithmically increase with their frequency. Tree clouds were also constructed for the words having the highest specificity scores: in this case, the sizes of the words in each tree cloud are proportional to their specificity score.

Branch lengths were set to unit values to improve the readability of each tree. Hence, only the topology of the tree and word size are significant (interpretable) in the visualization, not the distance between the words.

We compared the tree clouds obtained for the first and second half of the 2004-2011 period, which correspond to periods with sufficient numbers of projects and words for the analysis, i.e. :

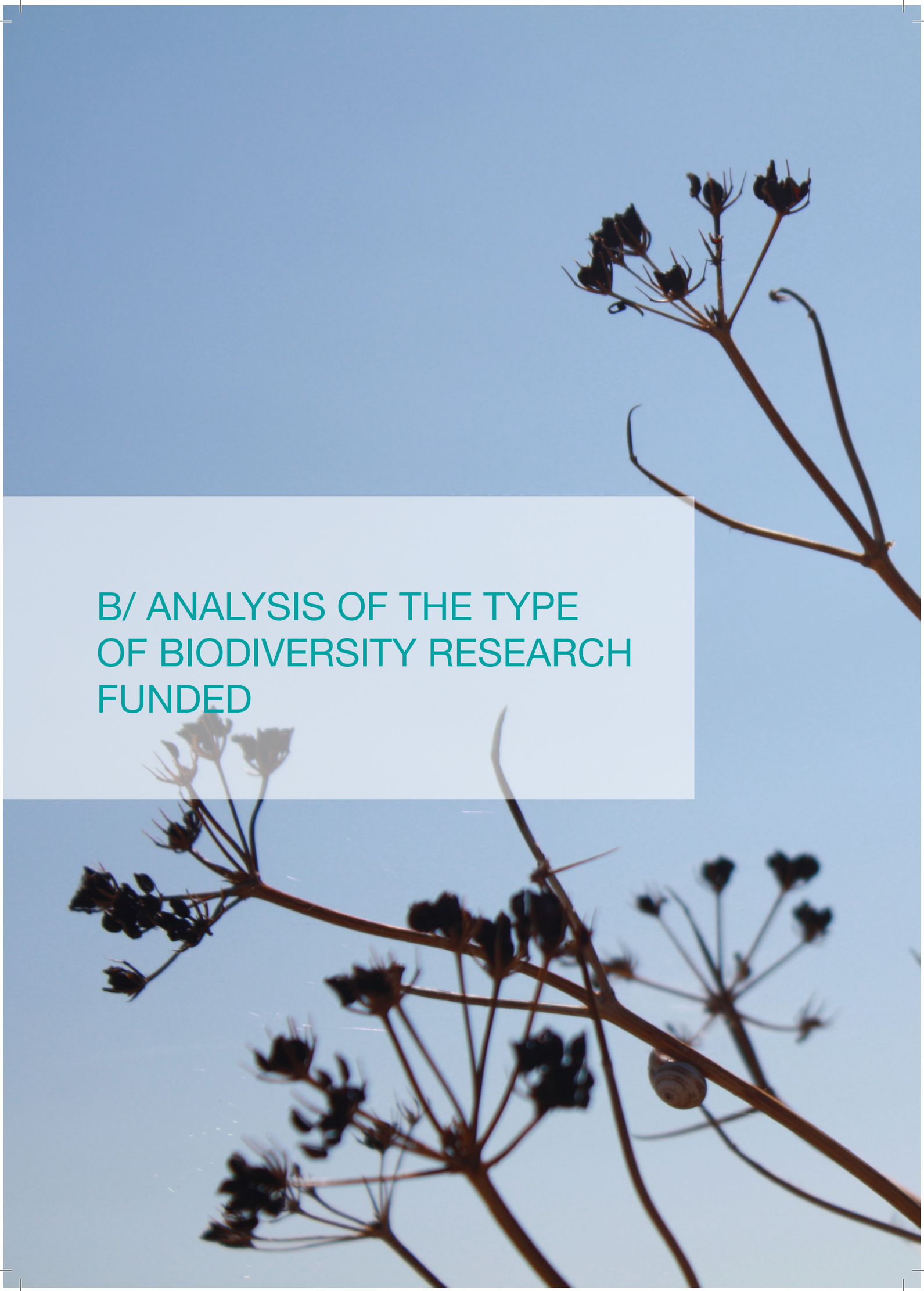
- 1931 projects for 2004-2007, with a total of 527 165 words
- 2228 projects for 2008-2011, with a total of 691 576 words

A4. ASSESSING THE RESULTS' ROBUSTNESS AND POSSIBLE BIASES IN THE ANALYSIS

We tested the robustness of the results and assessed possible biases in the analysis by two types of analyses.

Firstly, we normalized the size of the abstracts in terms of word number by prescribing similar weight of each abstract: abstracts can indeed differ in term of lengths, in particular among funders and funding schemes. We tested whether the results of the semantic analysis were modified by the normalization.

Secondly, we tested whether our results were affected by the funding level of projects. We focused on projects with medium funding levels (between 200 and 800 k€ per project, corresponding to projects from ANR, BELSPO, DEFRA, DFG, FORMAS, FWF, NWO, RCL & RCN) by excluding projects with low funding levels (between 40 and 190 k€ per project, corresponding to projects from BNSF, ETAG, FRB, MFAL, MINECO & VM) and the projects funded by the EC (with funding typically above 4 M€ per project). The temporal trend in the type of research funded for this group of projects of medium financial size was compared to that obtained with the complete corpus.



B/ ANALYSIS OF THE TYPE OF BIODIVERSITY RESEARCH FUNDED

B/ ANALYSIS OF THE TYPE OF BIODIVERSITY RESEARCH FUNDED

B1. TEMPORAL VARIATION OF ANNUAL SUB-CORPUSES OF WORDS USED IN PROJECT ABSTRACTS

The correspondence analysis comparing the 8 annual sub-corpuses of words clearly shows that the vocabulary used in project abstracts changed during the 2004-2011 period (Figure 3). Moreover, a clear temporal trajectory was observed, with sub-corpuses of words progressively shifting along the first axis during the 2004-2010 period, and with the 2011 sub-corpus displaying a backward trend. This demonstrates a clear, overall evolution in the vocabulary used in the projects through time.

The backward trend observed for the 2011 sub-corpus could reveal an actual trend but may also be due to incomplete inventarisation of the 2011 projects. In particular, the number of DEFRA projects referenced in the database for 2011 is lower than for the previous years, which may partly explain this backward trend. Improving the completion of the database for the 2004-2011 period and extending the data set to one or two additional years (which is part of BiodivERsA plans) will allow us to test this and more reliably explain the observed trend.

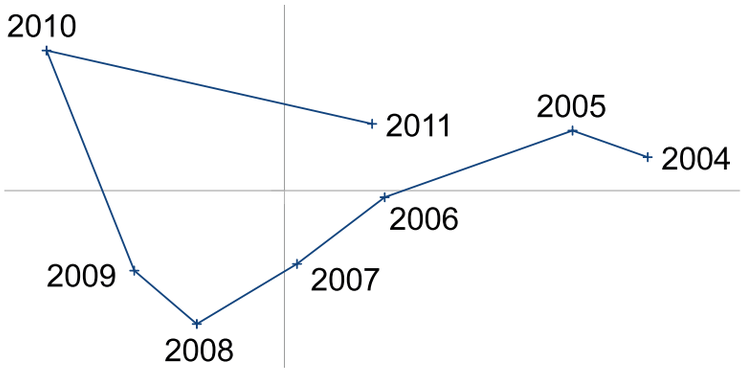


Figure 3: Correspondence analysis of the annual sub-corpuses of words used in project abstracts over the 2004-2011 period. The closer two points, the more similar the funded research for these two years.

The temporal trend observed here was robust and only weakly influenced by the size of project abstracts (see Appendix 1 Top). Similarly, the temporal trend observed when focusing on projects with funding levels between 200 and 800 k€ per project (which corresponds to projects from ANR, BELSPO, DEFRA, DFG, FORMAS, FWF, NWO, RCL & RCN) was very close to the one presented in Figure 3 (see Appendix 1 Bottom).

B2. TOPICS FREQUENTLY ADDRESSED BY FUNDED PROJECTS OVER THE 2004-2007 AND 2008-2011 PERIODS

We compared the tree clouds based on the most frequent words used for the first and second half of the 2004-2011 period (Figure 4).

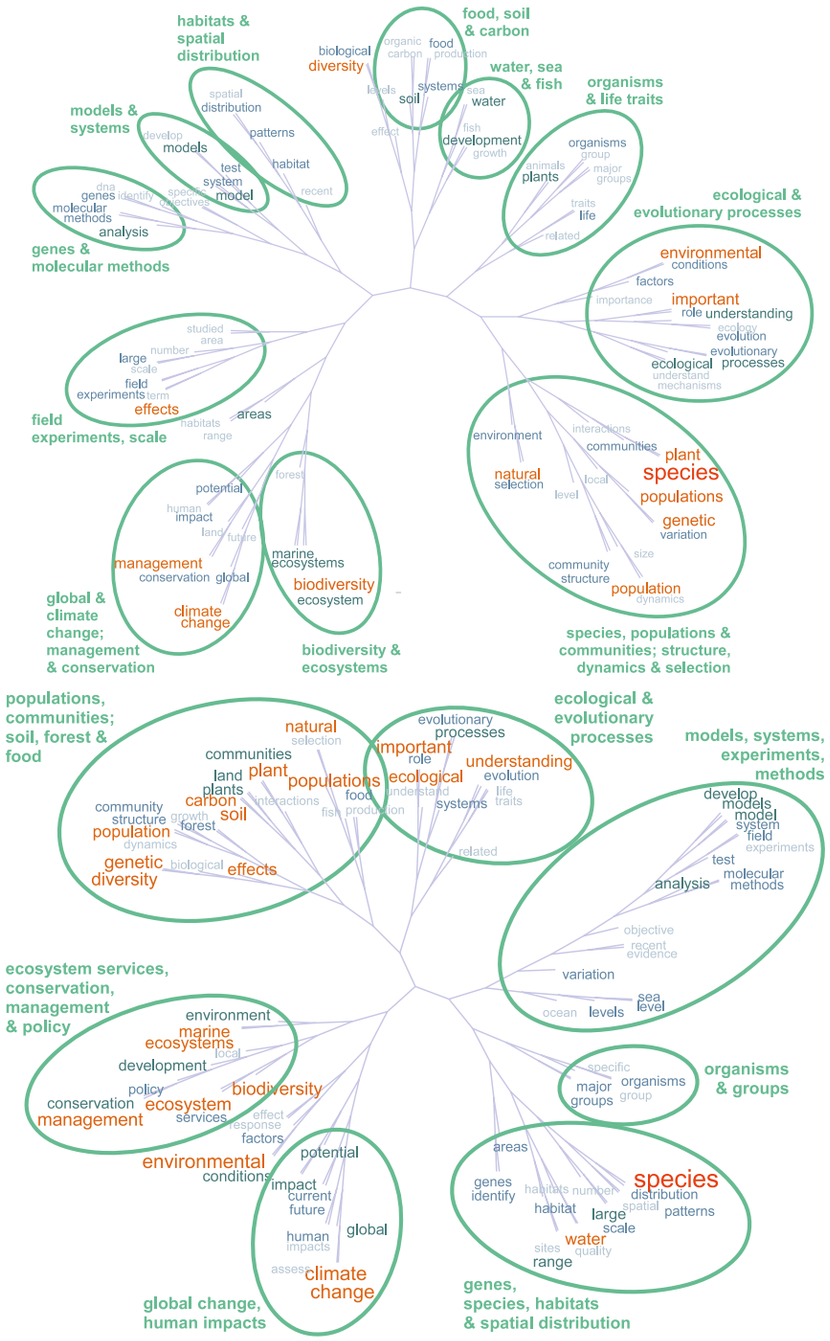


Figure 4: Tree cloud of the 100 most frequent words in abstracts of all the funded projects analysed, revealing the main coherent clusters of words used by researchers over (Top) the 2004-2007 period, and (Bottom) the 2008-2011 period. The size of a word is proportional to the frequency of its use.

When comparing the trees built from the most frequent words used for the 2004-2007 and 2008-2011 periods (Figure 4), the following conclusions arise:

- Some clusters are always observed with a similar pre-eminence in each tree, like the word clusters 'Ecological and evolutionary processes', and 'organisms and groups'. These two clusters likely correspond to basic bricks for biodiversity research.
- The three closely related clusters corresponding to methods and approaches observed over the 2004-2007 period ('Models & systems', 'Genes and molecular methods', and 'Experiment and scales') form a single cluster over the 2008-2011 period. Again, this corresponds to basic bricks for biodiversity research.
- For the 2004-2007 period, we detected a cluster 'Species, populations & communities; structure, dynamics and selection', a cluster 'Habitats and spatial distribution', and a cluster 'Food, soil and water'. Over the 2008-2011 period, these clusters are partly reorganised, mainly into the two clusters 'Genes, species, habitat and spatial distribution' on the one hand, and 'Populations, communities, soil, forest and food' on the other hand. This could underline a progressive separation of studies focusing on genes and species and their spatial distribution, as compared to studies focusing on higher levels of biodiversity, i.e. populations to communities. This trend is consistent with other analyses presented below (see B4.).
- For the 2004-2007 period, we observed two related branches: one corresponding to a 'Biodiversity and ecosystems' cluster, and the other one to a 'Climate change, management and conservation' cluster. These branches correspond to research issues which were far more addressed in projects during the 2008-2011, forming more visible and restructured clusters. The first cluster evolved into a cluster 'Ecosystem services, biodiversity, conservation, management & policy', whereas the second one evolved into a 'Global and climate change, human impacts' cluster that includes the terms current and future. This noticeable trend is linked to the increasing importance of the notions of ecosystem services and valuation of biodiversity, and its accounting for in management, development and policy actions. This demonstrates that biodiversity research increasingly focuses on socio-ecosystems. In parallel, this also corresponds to increasing research efforts aiming to characterize the impacts of human activities and global change on biodiversity and ecosystem services. The same trend is observed when analysing specificity scores of the main words used by researchers through time (see B3.)

B3. TOPICS SPECIFICALLY ADDRESSED BY FUNDED PROJECTS IN THE 2004-2007 VERSUS 2008-2011 PERIODS

The profile of funded research between the two periods can also be compared using the tree clouds based on the most specific (rather than most frequent) words for each period (Figure 5).

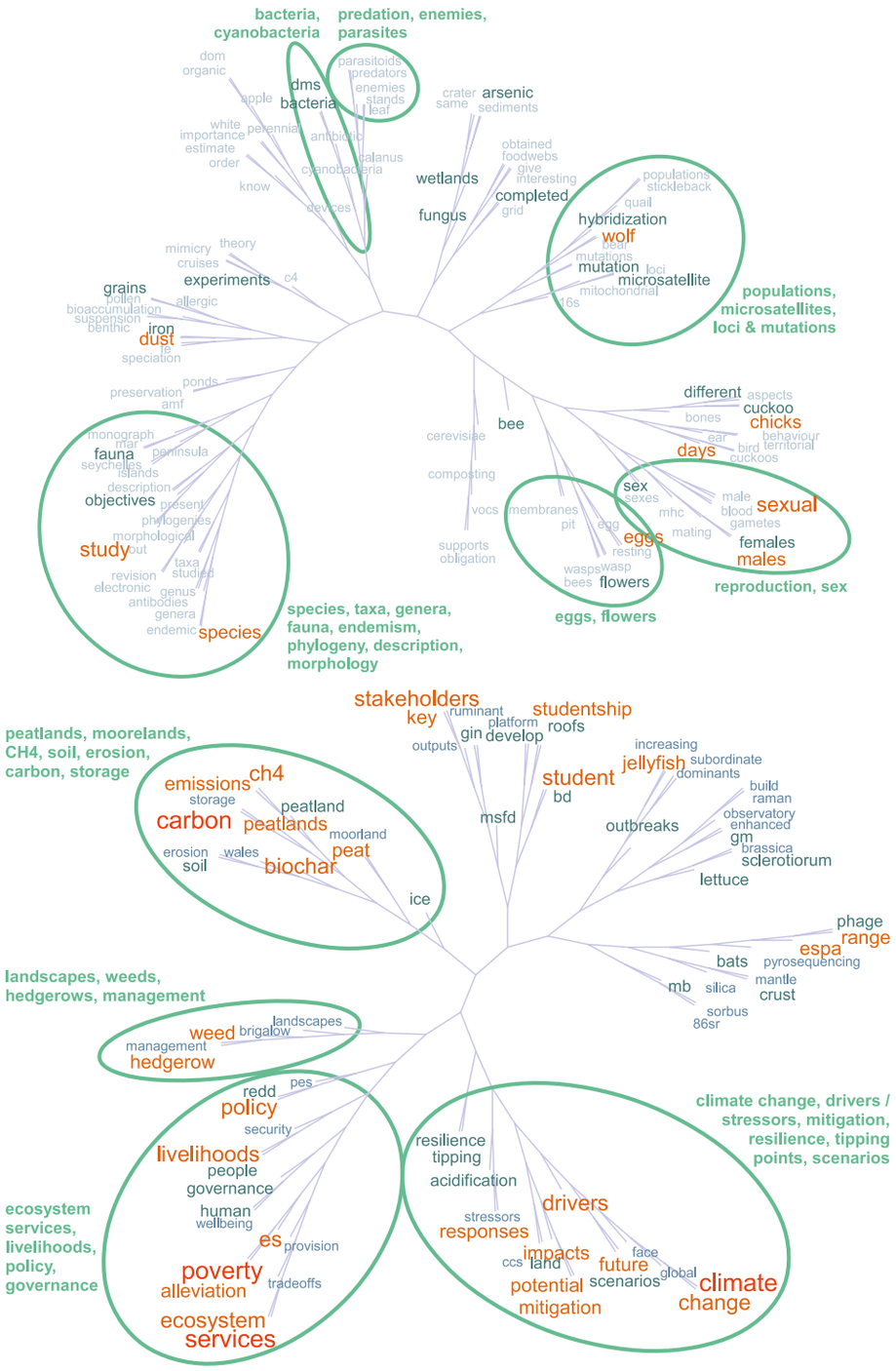


Figure 5: Tree cloud of the 100 most specific words in abstracts of funded projects over (Top) the 2004-2007 period, and (Bottom) the 2008-2011 period. The size of a word increases with its specificity score.

Figure 5 (Top) shows that the words with highest specificity scores for the 2004-2007 period mainly deal with biological entities or concepts. For example, there is a clear tree branch corresponding to words referring to reproduction (words sexual, sex, males, females, eggs, gametes). Several taxa also appear (wolf, bees, wasps, fungus, chicks, cuckoo); and there is a subtree typified by bacteria and cyanobacteria, which also includes the term antibiotic. The concept of food web appears both by the word itself, and by related concepts in a tree branch: predators, parasitoids, enemies. A tree branch corresponding to molecular aspects (micro-satellite, mutation, loci, 16s) and including the term populations is also present. Last, a large tree branch gathers the terms species, taxa, genera, fauna, endemism, phylogeny, description and morphology. This likely refers to the relative importance of species-, phylogeny- and taxonomy-based studies over the 2004-2007 period.

For the 2008-2011 period (Figure 5, Bottom), four tree branches clearly appear. The two major ones correspond to words related to climate change, drivers, stressors, mitigation, resilience, tipping points and scenarios; the second one to words related to ecosystem services, livelihoods, human wellbeing, poverty alleviation, people, policy and governance. This supports our previous conclusion based on the analysis of the most frequent words that biodiversity research increasingly focuses on socio-ecosystems. In addition, a smaller tree branch gathers words like landscapes, weeds, hedgerows and management, which is likely linked to the increasing importance of agro-ecology issues in biodiversity research. Yet, another branch gathers words like peatlands, moorelands, CH4, soil, erosion, carbon and storage, likely referring to the increasing importance of the links between biodi-versity and soil functioning, and the regulation of greenhouse gas emissions.

B4. TEMPORAL TRENDS OF TOPICS ADDRESSED DURING THE 2004-2011 PERIOD

To check whether under-representation of some words in the second half of the studied period really reflects a robust and significant trend, and not just a signal caused by special circumstances for a given year, we analysed the temporal change in the annual specificity scores of eight groups of terms identified from previous analyses (Figure 6 Top):

- Reproduction: egg(s), gamete(s), sex(es), sexual, mate, mating, female(s), male(s)
- Selection, speciation, phylogeny: selection, speciation, phylogeny
- Species and genera-based approaches: species, taxa, genus, genera
- Fauna and emblematic species: fauna, bear, bird, wolf
- Gene-based approaches: gene(s), mutation(s)
- Traditional molecular approaches: dna, microsatellite(s), molecular, 16S
- Population-based approaches: population(s), metapopulation(s)
- Foodweb-based approaches: foodweb(s), predator(s)

For each of those groups, the specificity score significantly decreased with time during the 8-year period, and a relatively monotonous decrease was observed in each case (Figure 6 Top). The main contrast was observed between years 2004-2005 on the one hand, and 2009-2010 on the other hand, which is consistent with the conclusions raised from the analysis presented in Figure 3.

Interestingly, the signal remained largely apparent when focusing on individual words of a given group (Appendix 2).

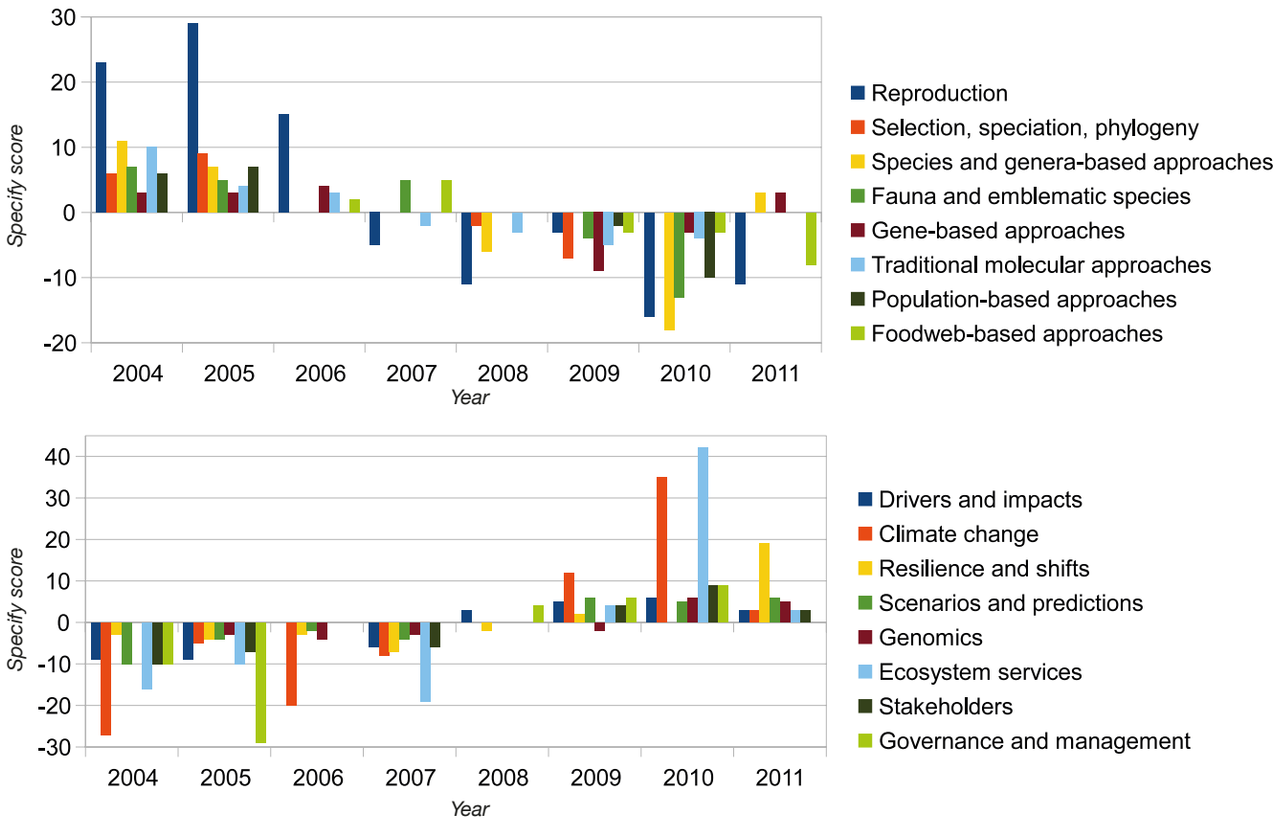


Figure 6: Temporal variations in the values of specificities for the major 8 groups of words that are (Top) decreasingly and (Bottom) increasingly used over the 2008-2011 period.

Similarly, we checked whether the over-representation of some words in the second half of the studied period really corresponds to a robust and significant trend by analysing the temporal change in the annual specificity scores of eight other groups of terms (Figure 6 Bottom):

- Drivers and impacts: drivers, stressors, impacts
- Climate change: CH4, N2O, CO2, emissions, climate, warming, change, changing, mitigation
- Resilience and shifts: resilience, shift(s), tipping
- Scenarios and predictions: future, prediction, scenario(s)
- Genomics: genome(s), genomic
- Ecosystem services: ecosystem(s), es, services
- Stakeholders: stakeholder(s), people, human, practitioner(s)
- Governance and management: governance, policy, management, planning, decision, tradeoffs

Again, for each of those groups, the specificity score significantly increased with time during the 8-year period, and a relatively monotonous increase was observed in each case (Figure 6 Bottom). Very positive signals were observed for ‘climate change’ and ‘ecosystem services’ in 2010. The increasing signal remained largely visible when focusing on individual words of a given group (Appendix 3).

The trends illuminated by this last analysis allowed better understanding of some aspects of the bulk analysis presented above (Figure 5). For instance

poverty alleviation was actually overrepresented only in 2010 (in NERC projects mainly). Similarly, the particular representation of some names of species was actually limited to one year, and linked to only a few projects. However, most of the results are consistent between Figures 5 and 6.

These results proved to be robust when the corpus was normalized for abstract length (see Appendix 4 Top). Similarly, the results were not significantly modified when focusing on projects with medium funding level (between 200 and 800 k€ per project, corresponding to projects from ANR, BELSPO, DEFRA, DFG, FORMAS, FWF, NOW, RCL & RCN) (see Appendix 4 Bottom).

All these results highlight **a decreasing trend for words - and we think funded research - focusing on "low organisation levels" of biodiversity, in particular individual genes, species and genera, and populations; or focusing on some emblematic/patrimonial species and fauna compartments.** This is also true for research focusing on associated mechanisms, i.e. reproduction, selection, speciation and for phylogenetic studies.

In contrast, there is **an increasing trend for words - and we think funded research - focusing on "higher organisation levels" of biodiversity, in particular (meta)genomes rather than genes, as well as (meta)communities and ecosystems rather than taxa and populations. This is also true for research on the drivers and mechanisms relevant at these higher organisation levels,** i.e. global change components, ecosystem management and environmental policies. More particularly, research focusing on the effects of climate change on biodiversity, and feedbacks to climate (CH4, N2O and CO2 emissions, mitigation options, etc.) is increasingly present. Similarly, research on the resilience, possible shifts and tipping points of biodiversity and ecosystems, and research developing scenarios and predictions of biodiversity and ecosystems, is also increasing.

This does not necessarily mean that the first type of research is not funded anymore, since this analysis relies on specificity scores. Indeed, as demonstrated by the analysis of the most frequent words, research focusing on "low organisation levels" of biodiversity (in particular genes, species, genera and (meta)populations) is still funded (see Figure 4, Top), but a new balance has clearly been achieved between studies on 'low' and 'higher' organisation levels during the 2004-2011 period. This seems to reflect a tendency in biodiversity research that is also visible at the level of individual national agencies. For instance, at NWO (Dutch agency), biodiversity research is shifting from the general ecological perspective to a more specific focus on socio-ecosystems, in particular agriculture.

B5. COMPARISON OF THE TYPE OF FUNDED BIODIVERSITY RESEARCH BETWEEN NATIONAL AGENCIES

The correspondence analysis comparing the 16 sub-corpuses of words, each corresponding to projects funded by a given funding source (15 national agencies and EC), clearly shows that the vocabulary used in project abstracts varies between agencies (Figure 7).

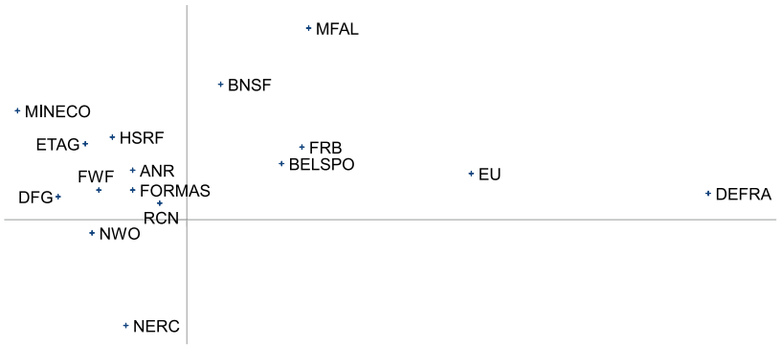


Figure 7: Result of the correspondence analysis for project abstracts from the 15 national funding agencies over 2004-2011. The profile of the research projects directly funded by EC (EU) is also indicated.

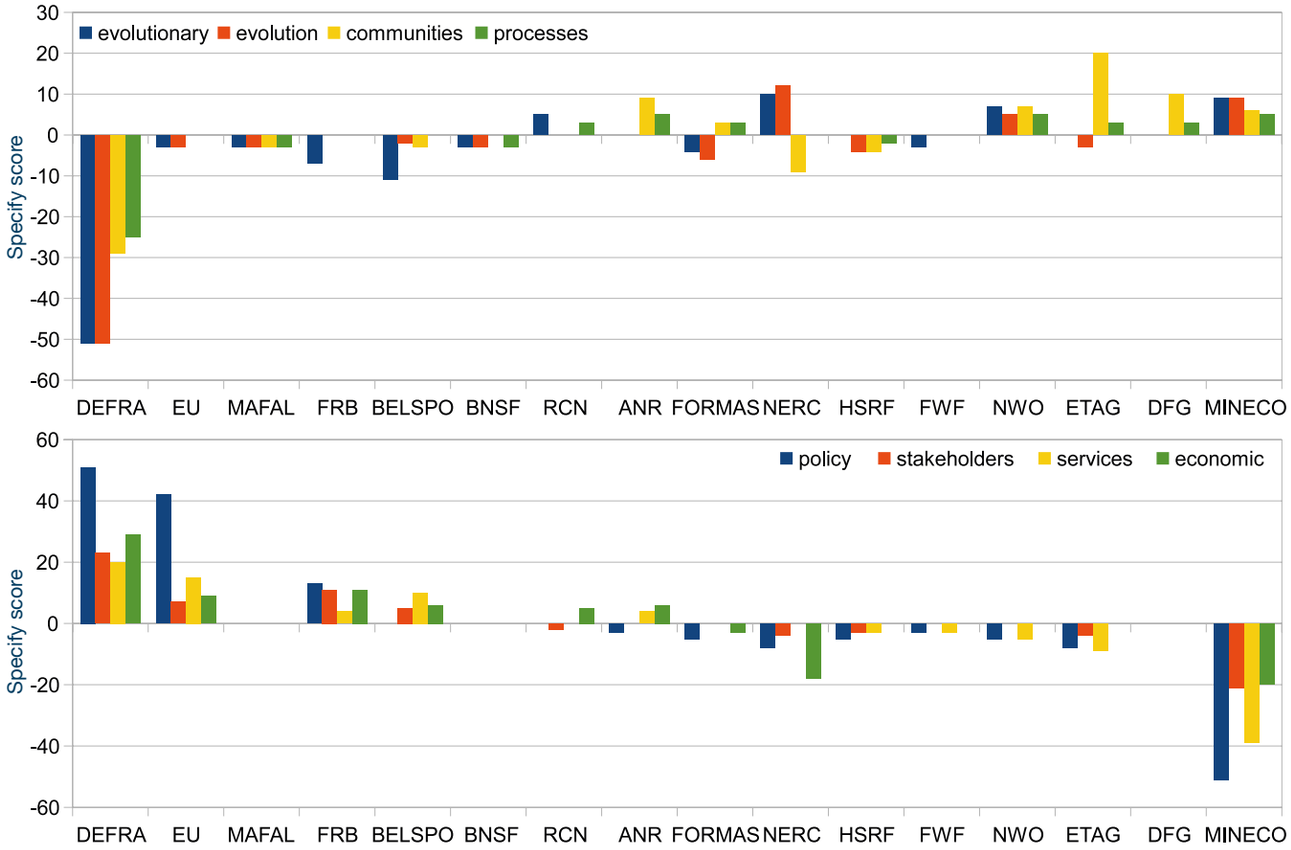


Figure 8: Major word specificity scores detected in project abstracts explaining the observed contrasts between national funding agencies in Figure 7. The values for abstracts of research projects directly funded by EC (EU) are also indicated for comparison. Funding sources are presented according to their location on the first axis of Figure 7.

The analysis of the specificity scores for words that discriminate the 16 funding sources supports our interpretation of the meaning of the main axis of Figure 7. Indeed, the summaries of projects funded by DEFRA and the EC are strongly characterised by specific words like services, economic, policy, and stakeholders that are under-represented in the summaries of projects funded by ETAG, NERC, and most importantly MINECO (Figure 8). On the reverse, the summaries of projects funded by MINECO and NWO are characterised by specific words like evolution/evolutionary, processes, and communities that are strongly under-represented in the summaries of projects funded by DEFRA, and to a lesser extent MFAL (Figure 8). It has been checked with BiodivERsA members that the agency profiles are logical. For instance, despite the shift towards a stronger support to agriculture, NWO mainly supports the basic ecological side of research as there is no specific programme influencing NWO's results in this period, and most projects stem from NWO blue sky programmes or other non-specific biodiversity programmes.

Because Figure 7 is largely influenced by the specific profile of research funded through a few funding sources, we excluded DEFRA, EU, MFAL and BNSF projects from the corpus, to better detail the differences that exist between the 12 remaining agencies.

Detailed comparison between national agencies

We analysed in more detail the type of research funded by 12 national agencies for the 2004-2011 period (ANR, BELSPO, DFG, ETAG, FRB, FORMAS, FWF, HSRF, MINECO, NERC, NWO, RCN).

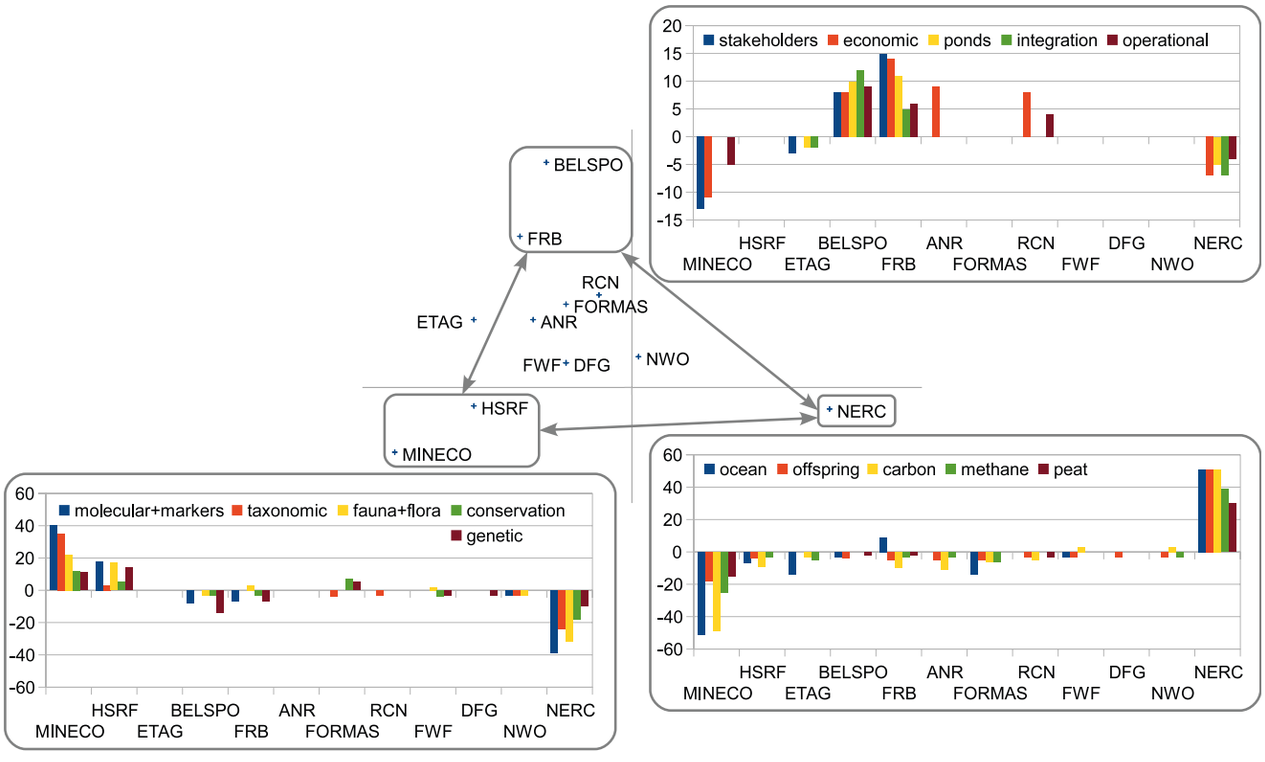


Figure 9: Result of the correspondence analysis of the sub-corpuses for project abstracts over the 2004-2011 period from the 12 national funding agencies. Variation of the specificity scores (Y axes) between agencies are presented for some words that are characteristic of projects for MINECO and HSRF (Bottom left panel), BELSPO and FRB (Upper panel) and NERC (Bottom right panel)

The correspondence analysis comparing the sub-corpuses of words, each corresponding to projects funded by one of these national funding agencies, confirms that the vocabulary used in project abstracts varies between agencies (Figure 9). Three agencies correspond to the three most specific profiles in term of vocabulary used, i.e. BELSPO (and to a lesser extent FRB), MINECO (and to a lesser extent HSRF), and NERC (Figure 9). The tree clouds presenting the most frequent words used in the project abstracts for BELSPO plus FRB, MINECO plus HSRF, and NERC, are presented in figures 10, 11 and 12, respectively.

The 3 main profiles observed (i.e. poles observed in figure 9, and figures 10-12) can be characterized as follows (we do not focus on words that are obviously specific, as Iberian or Balearic):

(1) Project abstracts for MINECO plus HSRF (Figure 10) over-used the following terms:

- Taxonomy-based approaches: species, taxa, genus, genera, tribe, families, taxonomic
- Fauna and flora: fauna, flora
- Biogeography: biogeography, biogeographic, geographical, distribution, endemic
- Phylogeny: phylogeny, phylogenetic
- Molecular approaches: molecular, nuclear, mitochondrial, markers

In parallel, the project abstracts for MINECO plus HSRF under-used the following terms:

- Marine: ocean(s), sea, marine, fish, fisheries
- Plankton: plankton, phytoplankton
- Lake: lake(s)
- Greenhouse gases: CO2, carbon, dioxide, emissions, atmosphere
- Ecosystem services: ecosystem, services
- Policy and stakeholders: stakeholder(s), policy

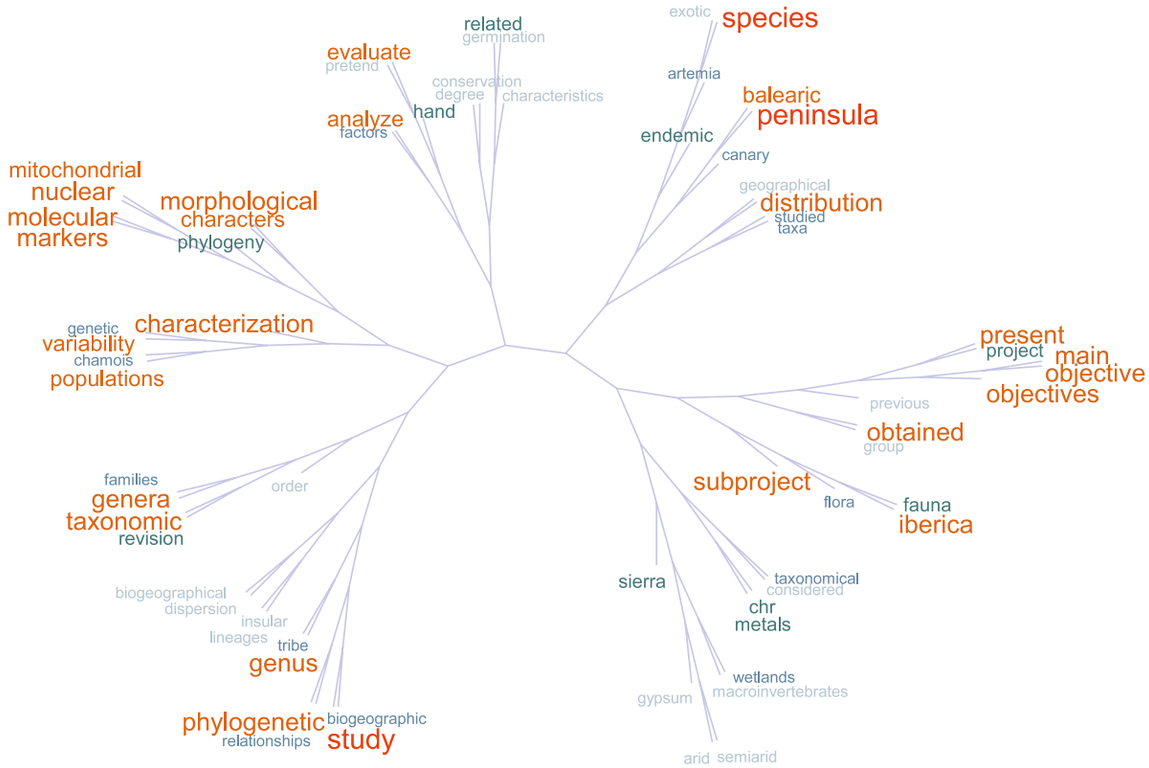


Figure 10: Tree cloud of the most specific words used in the project abstracts for MINECO (Spain), plus HSRF (Hungary).

This MINECO+HSRF profile (Figure 10) corresponds to a particular support of research projects focusing on ecological, and in particular taxonomy-based studies, biogeography and phylogenetic approaches. In contrast, it supports less than the mean of agencies research projects on biodiversity in marine and lake environments, on the link between biodiversity and greenhouse gas emissions, or on socio-ecosystems and ecosystem services including views on policy and stakeholders.

(2) Project abstracts for NERC over-used the following terms:

- Marine biodiversity: ocean(s), seabed, seawater, sediments
- Corals: coral(s), carbonate
- Greenhouse gases: CH4, methane, N2O, CO2, carbon, dioxide, greenhouse, gas(es), emission, atmosphere, atmospheric
- Microorganisms: bacteria, microbes
- Fossils: fossil(s), radiocarbon
- Livelihoods: livelihood, poverty, alleviation
- Reproduction: male(s), female(s), parents, offspring

In parallel, the project abstracts for NERC under-used the following terms:

- Taxonomy-based approaches: species, taxa, genus, genera
- Fauna and flora: fauna, flora
- Phylogeny: phylogeny, phylogenetic
- Invasions: dispersion, colonization, invasive, invasions
- Molecular approaches: molecular, nuclear, mito-chondrial, markers
- Spatial: landscape, spatial, fragmentation
- Conservation: conservation, endemic
- Function: functional, functioning
- Management: management

The NERC profile (Figure 11) thus corresponds to a particularly strong support of research projects on biodiversity in marine environments, including corals and seabeds, the link between biodiversity and greenhouse gas emissions, microbial biodiversity, and reproduction systems. In contrast, it supports less than the mean of agencies taxonomy-based studies, phylogenetic approaches, as research on biological invasions, and landscape ecology.

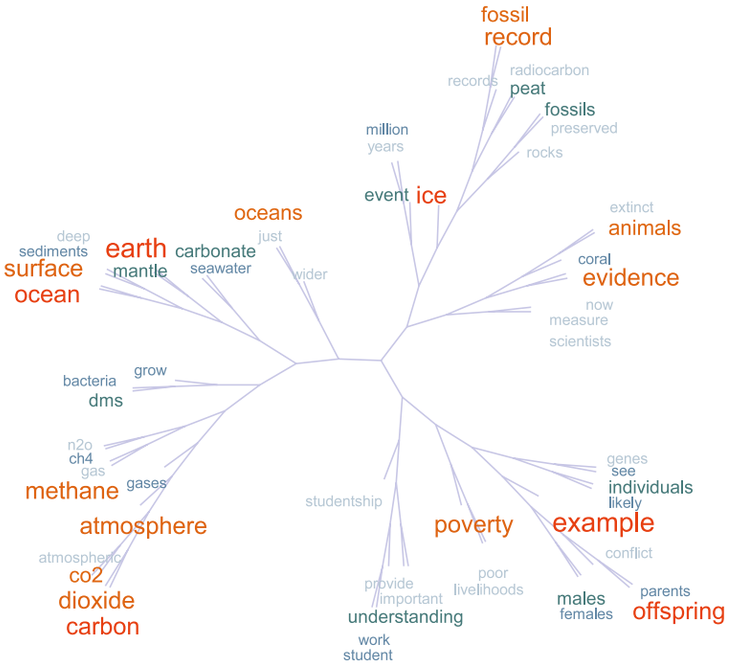


Figure 11: Tree cloud of the most specific words used in the project abstracts for NERC (UK).

(3) Project abstracts for BELSPO plus FRB (Figure 12) over-used the following terms:

- Ponds: pond(s)
- Agriculture and land: agricultural, farming, land
- Ecosystem services: ecosystem, services
- Management: management
- Valuation: valuation, economic, socio
- Stakeholders and practices: stakeholders, actors, actions, practices
- Policy: policy, policies
- Integration: integration, integrated
- Scenarios: scenario(s)
- Remote sensing: satellite, remote, sensing, hyperspectral

In parallel, the project abstracts for BELSPO plus FRB under-used the following terms:

- Species: species
- Genetic: gene(s), genetic
- Evolutionary: evolutionary

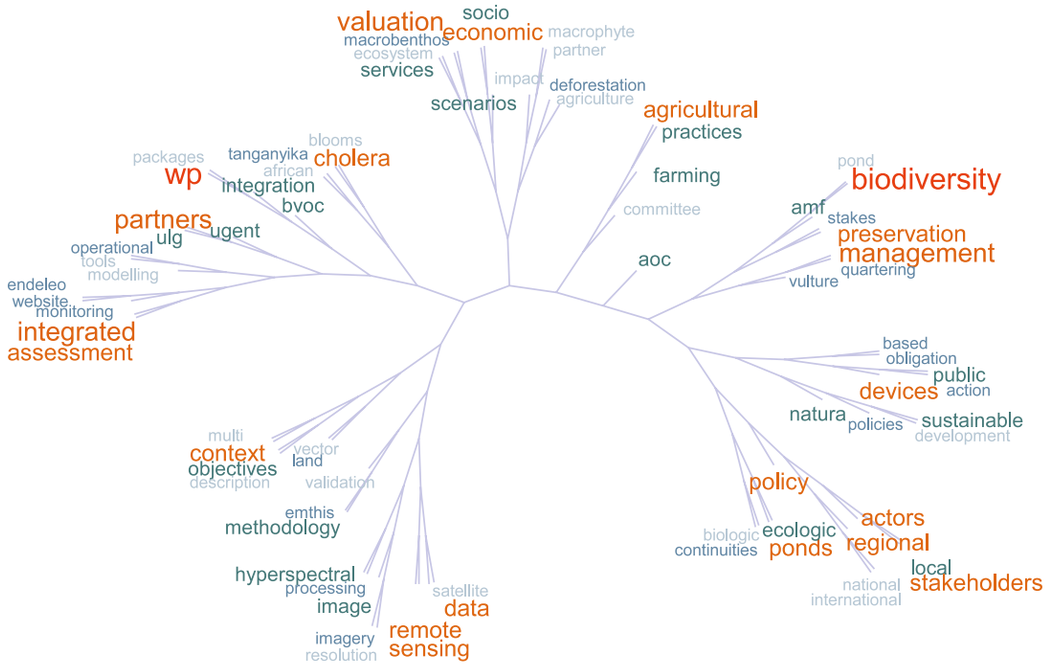
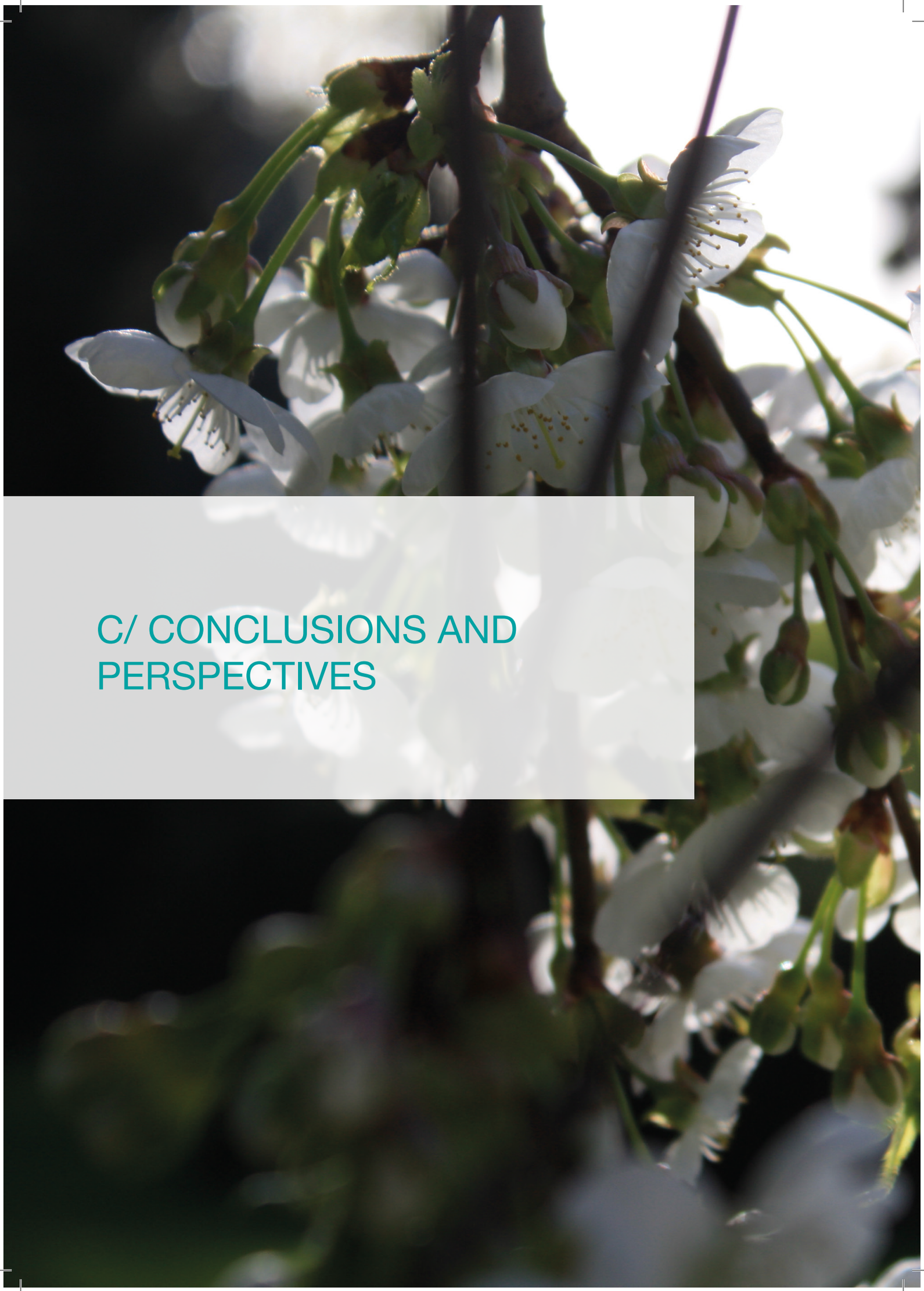


Figure 12: Tree cloud of the most specific words used in the project abstracts for BELSPO (Belgium), plus FRB (France).

This BELSPO+FRB profile (Figure 12) corresponds to a particular support of research projects focusing on socio-ecological systems, ecosystem services and their management and valuation, accounting for human activities and with a particular view on stakeholders, practices and policy. The specificity scores for the corresponding terms were particularly high. In contrast, BELSPO and FRB support less research projects on species-based and individual gene-based approaches and evolution, though the negative specificity scores for these terms are rather low.

For FRB, this profile is consistent with the nature of FRB, i.e. a science-society platform promoting research programmes that are co-designed by scientists and stakeholders, and address pressing issues for both science and society. For BELSPO, this profile is linked to a focus on multidisciplinary research programmes, supporting policy and scientific public services in view of sustainable development.

The detailed profile of the 7 other national agencies is presented in Appendix 5.



C/ CONCLUSIONS AND PERSPECTIVES

C/ CONCLUSIONS AND PERSPECTIVES

The results presented here show obvious changes in the type of research funded across Europe during the first decade of this century. **Over the 2004-2011 period, the observed changes shifted focus from research regarding organismal and population levels (taxa, populations, emblematic species, phylogeny, individual genes) and their conservation, to larger organisation levels (communities and ecosystems, landscapes, socio-ecological systems, whole genomes) and forecasts and scenarios of future biodiversity.** Emphasis also changed from acquiring basic knowledge on taxa and their dynamics to analyses of biodiversity protection costs and benefits and stakeholder's involvement.

The observed trends are strong, robust, and clearly established on a multi-annual perspective (see Figure 6). They correspond to an increasing number of funding programmes targeting studies on larger organisations levels, and concurrently to an important shift in the mindset of many scientists from the biodiversity research community. The observed trends are very consistent with the results of the Millennium Ecosystem Assessment (2005) and shifts in research policies to better address integrative and trans-sectorial issues (like fisheries, agriculture, etc.), but are not directly related to the goals set by the EU regarding the halt of biodiversity loss by 2010 first, and then by 2020.

It is important to remember that only programmes corresponding to competitive funding sources are considered here. It is expected that these programmes allow quicker changes of the type of research they aim at supporting. In parallel, other less/not competitive funding sources (e.g. direct funding through research institutes) likely allow the support of biodiversity research beyond the trends prescribed by competitive funding programmes. The relative importance of funding through competitive / not competitive sources is likely very different according to countries.

The results presented here show how quickly research on biodiversity and associated ecosystem services evolves, likely through 'internal' forces due to new and renewed topics addressed by the research community (e.g., see the role that the MEA had) and through the orientations given by research programmers and funders. Given the marked changes in the type of biodiversity research funded observed over the 2004-2011 period, **national and European research managers will have to clearly define the balance to be maintained between the different aspects of biodiversity research, and will have to carefully monitor this balance, so that biodiversity research support is based on an explicit, long-term and strategic approach.**

References

Chaveriat C., Ghitalla F., Pelegrin F., Fadil F. & Le Roux X. (2011). The French database for actors, structures and projects of biodiversity research: presentation and analysis of the research landscape (In French). Expertise & synthesis FRB, 36 pp.

Chivian E. & Bernstein A. (2008). Sustaining life: How human health depends on biodiversity. Center for Health and the Global Environment. Oxford University Press, New York.

Eggermont H., Le Roux X., Heughebaert A., Balian E. & BiodivERsA partners (2013). The BiodivERsA database: analysis of the competitive funding landscape for research on biodiversity and ecosystem services in Europe. *BiodivERsA report*, 2013; 33 pp

Gambette P. & Véronis J. (2009) Visualising a Text with a Tree Cloud, *Proceedings of the International Federation of Classification Societies 2009 Conference (IFCS'09), Studies in Classification, Data Analysis, and Knowledge Organization* 40:561-570.

Gambette P. (2010). User manual for TreeCloud. http://manual.treecloud.org

Huson D.H. & Bryant D. (2006) Application of Phylogenetic Networks in Evolutionary Studies, *Mol. Biol. Evol.* 23(2):254-267.

Lafon, P. (1981) Sur la variabilité de la fréquence des formes dans un corpus, *Mots* 1:127-165.

Lebart L., Salem A. & Berry L. (1998) *Exploring Textual Data*, Springer.

Matei S., Henkel L., Gauthier C.A. & Le Roux X. (2011) Biodiversity within the 'Environment' theme of the 7th framework programme (2007-2010): funding amounts, success rates, temporal trends & comparison between countries. Expertise and synthesis FRB, 32 pp.

Millennium Ecosystem Assessment (2005). Ecosystems and Human Well-being: Biodiversity Synthesis. World Resources Institute, Washington, DC., 100 pp.

Rockström J. et al. (2009) A safe operating space for humanity. Nature 461: 472-475.

Saitou N. & Nei M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees, *Mol. Biol. Evol.* 4(4):406-425.

Salem A., Lamaille C., Martinez W., Fleury S. (2003). *Manuel Lexico 3*. version 3.41. http://www.tal.univ-paris3.fr/lexico/manuels.htm

Electronic supplementary materials

Annexes are available online at: http://www.biodiversa.org/8

BiodivERsA database

www.biodiversa.org/database

For more information:

Coordinator and CEO
Xavier Le Roux
Fondation pour la Recherche sur la Biodiversité
xavierleroux@hotmail.fr
Ph.: +33 (0) 6 31 80 38 20

Secretariat
Frédéric Lemaître (Executive Manager)
frederic.lemaitre@fondationbiodiversite.fr
Ph.: +33(0) 1 80 05 89 37

Claire Blery (Management Assistant)
claire.blery@fondationbiodiversite.fr
Ph.: +33(0) 1 80 05 89 36

Fondation pour la Recherche sur la Biodiversite
195, rue Saint Jacques
75005 Paris, France

www.biodiversa.org

BiodivERsA is supported by the European Commission
within the 7th Framework Programme