



(version 11 December 2018)

Biodiversity has been in the focus of ecological research for a long time but we still have just a vague understanding about why some species are present and others are absent in a locality. Biodiversity varies both along natural and anthropogenic gradients, but disentangling the actual factors that influence it has been limited. Coordinated global sampling has offered very promising avenues to advance our understanding on biodiversity. **DarkDivNet** is a research network which aims to explore the dark diversity of plant communities—the set of species present in a region but absent in a site, albeit the site conditions are potentially suitable (Pärtel *et al.* 2011). We plan vegetation surveys replicated on the global scale to advance dark diversity studies.

Only a subset of the species present in a region can live in the local ecological conditions of a given site (the site-specific species pool, to which we will refer to as “species pool” from here on; Zobel 2016). Moreover, not all the species from the species pool are realised into local communities. Typically, a large proportion of the species from the species pool are not present; this unobserved part of the species pool represents the dark diversity of a community. Through knowledge of dark diversity, it is possible to estimate the completeness of communities and, thus, quantifying the proportion of the pool that is locally realized. This estimate enables a measurement of biodiversity at the relative scale by considering the natural variation in species pool sizes, often providing a more appropriate estimate of biodiversity than the absolute richness for comparisons across habitat types or regions. In particular, large-scale patterns of completeness remain mostly unknown.

Considering dark diversity has enabled to analyse processes shaping local diversity across different regions and habitat types. With dark diversity measures, we can also detect and disentangle the relative importance of biogeographical history from local interactions like competition or facilitation, or the effects of mutualistic and pathogenic organisms (e.g. soil microbiota). When the functional traits of observed and dark diversity are examined, we can get ideas about why species are absent from local communities. Ultimately, knowledge on dark diversity allows to evaluate the conservation value of natural sites and restoration potential of degraded habitats (Lewis *et al.* 2017).

Since dark diversity is, by definition, unobservable, it can only be estimated. Ecologists have faced this challenge by developing and using tools that rely on different sources of information, such as species’ ecological preferences or the opinion of experts. One particularly effective way to estimate dark diversity is to consider species co-occurrence patterns, which can be retrieved from different sources, such as species distribution maps or vegetation surveys. Ideally, if the realized community contains species that tend to be found together with a focal species, then one would expect that the probability of it being a member of the local dark diversity is high, and vice versa (Lewis *et al.* 2016). Vegetation databases, which are increasingly available, could be potentially very useful to characterize these co-occurrence patterns, but very little is known as to how they can be used under a wide range of ecological conditions, including different habitats and regions. Finally, species pool and dark diversity studies have mostly considered only taxonomic diversity, but techniques are needed to estimate the same parameters for functional and phylogenetic diversity as well.

*DarkDivNet* is coordinated by the Steering Committee of researchers in the University of Tartu (Meelis Pärtel, Carlos P. Carmona, Martin Zobel, Mari Moora, Kersti Riibak and Riin Tamme), and advised by a panel of international experts including Jon Bennett (Canada), Milan Chytrý (Czech Republic), Ove Eriksson (Sweden), Francesco de Bello (Spain), Susan Harrison (United States), Rob Lewis (Denmark), Angela Moles (Australia), Maarja Öpik (Estonia) and Jodi Price (Australia).

*DarkDivNet* can be contacted at [darkdivnet@ut.ee](mailto:darkdivnet@ut.ee).

Specifically, the main aims of *DarkDivNet* are:

- To bring together scientists around the world who are interested to contribute to the study of dark diversity of plants using a standardized approach.
- To determine observed and dark diversity, species pool and community completeness, and link them to geographical and environmental gradients, anthropogenic disturbance, and interactions with symbiotic soil biota (sampled with eDNA). This activity will enable to test the hypothesis that species pool size is driven by large scale factors like biogeographic history, while community completeness mirrors more recent changes in landscapes and ecosystems, most of which is of anthropogenic origin.
- To compare different methods to estimate dark diversity and to test their accuracy. We will compare the different estimations of dark diversity with the species composition recorded in the surroundings of sample plots, with species pools based on expert opinions, with plant taxa which have been in a site before (identified from eDNA preserved in the soil), and with the future occurrence of species through repeated sampling of permanent plots. The data collected in *DarkDivNet* will eventually become a benchmark that can be used to test new methods to estimate dark diversity in a wide range of conditions.
- To evaluate whether locally collected empirical information of species co-occurrences can be substituted by the use of available large vegetation databases (e.g. EVA, sPlot), and if yes, which filters have to be applied to databases to achieve satisfactory results (since databases most often do not provide data collected using standardized methods or plot sizes).
- To expand the dark diversity concept from purely taxonomy-based approaches to include the functional and phylogenetic aspects of diversity.

We devise that these goals will result in several major publications, and we encourage members of the network to submit new ideas for further manuscripts using the information available.

## ***DarkDivNet* PROTOCOL**

### **1. Study area and target habitat selection**

Each participant or group of participants in *DarkDivNet* will contribute at least one study area, which will be defined by a circle of 10 km radius. Ideally, different study areas should be spread geographically (at least 50 km from each other) and ecologically as much as possible. To ensure this and avoid redundancies, the coordinators of *DarkDivNet* must be informed about the location and characteristics of potential study areas. Please do not perform any fieldwork until the coordinators have confirmed that the study area is valid to be included in the network. Within each study area, the contributors will select 1-3 typical natural or semi-natural target habitat types (e.g. tropical rain forest, dry forest and shrubland; temperate coniferous forest, deciduous forest and semi-natural grassland). The selected habitat types should be relatively frequent in the study area (i.e., avoid very rare and extreme habitats). Each of these 1-3 typical target habitats will be sampled by pairs of permanent sites (localities) each one under different anthropogenic disturbance conditions (relatively natural vs. disturbed). By establishing permanent plots, we will be able to estimate the predictive ability of methods to estimate dark diversity, and will allow the network to be used for testing potential new ideas proposed by its members. In addition to permanent plots, several plots within the 10 km radius study area will be sampled just once to estimate the species co-occurrence matrix (see below for further explanations).

## 2. Selection and marking of permanent sites

Each of the 1-3 target habitat types will include a pair of permanent sites. One permanent site represents the relatively typical state of the target habitat type (“natural site”) while the other represents a site which was initially similar to the natural site but has recently been affected by anthropogenic disturbances (“anthropogenic site”, e.g. recent clear cut of a forest, overgrazing by domestic animals, abandonment of grazing in otherwise grazed semi-natural grassland etc.). Each permanent site should be a good representative of the considered habitat type (please avoid sites with particular vegetation or geomorphological features). Ideally the participant should have previous experience with the habitats considered. There are no restrictions regarding the location of the sites of 1-3 target habitats other than they should be within the 10 km radius study area (i.e. no more than 20 km apart from each other).

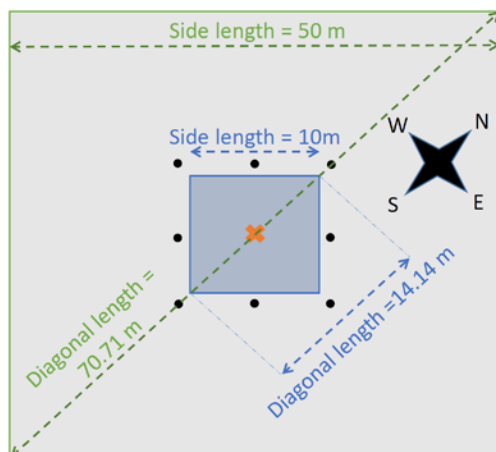
Each permanent site will have a nested structure, so that sampling will be done at two scales: a core plot, where more detailed measurements will be made, and a surrounding plot. The nested design will be used to estimate accuracy of dark diversity methods: species estimated for dark diversity for the core plot should be more likely found in the surrounding plot (Lewis *et al.* 2016). To make sampling uniform across regions and habitats, *DarkDivNet* plot sizes are fixed to 10 x 10 m for the core plot and 50 x 50 m for the surrounding plot.

The core plot must be centered within the surrounding plot (see Fig. 1 for a guide on how to set the plots). Participants should make sure that each site is reasonably uniform in terms of vegetation (e.g. no forest patches are included within the plots in a grassland site) and environmental conditions (e.g. slope, aspect, elevation, and soil types are relatively uniform within and among the sites in a study area). Please avoid sites that you suspect could suffer drastic changes in vegetation composition in the near future (e.g. forest plots in areas that are going to be clear-cut soon). Make sure that your access to both the natural site and the anthropogenic site is granted also in the future.

Mark corners of both the core plot and the surrounding plot with metal or plastic poles so that you can find them in the future. It is useful to have metal poles hammered almost completely to the soil (you can use a metal detector to find them), as well as taller poles than can be found visually, or some other method that allows finding the corners in the future in case highly visible poles increase risk of disturbance in the plot (e.g. by curious domestic animals, vandalism). If possible, orient diagonal in north-south direction and label poles with coordinates N, E, S, W (then it is easier to find other corners once you have detected one). Practical suggestions to mark plots: (1) mark the center of the core plot; (2) measure the half diagonal in N direction and mark N corner; (3) having the direction from N to the central mark measure the S corner; (4) by using two tapes measure both from N and S corners plot side lengths and find the point where two measuring tapes meet – these will be E and W corners, respectively (see Fig. 1 for a graphical example).

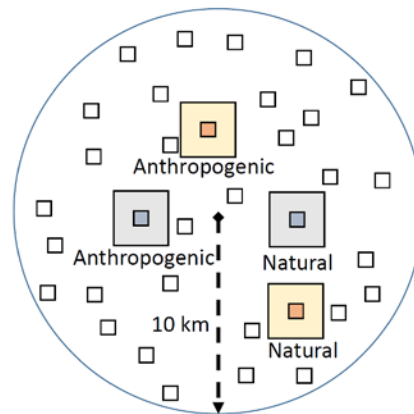
## 3. Plots for co-occurrences

In addition to the nested plots at permanent sites described above, the participants will sample vascular plants on 30-90 co-occurrence plots of 10 x 10 m. Species co-occurrences will be used to predict the probability of species being in the dark diversity of core plots. Permanently marking these plots is not needed. Please consider that 30 is an absolute minimum requirement, and we strongly encourage participants to sample up to 90 plots, particularly in the cases when two or three habitat types are sampled by permanent plots. This sampling scheme allows to estimate also gamma diversity and find the difference between gamma diversity and species pool. Co-occurrence plots will be randomly selected from an area of 10 km radius around the center of the study area (defined as the midpoint of the permanent sites), and should include relatively undisturbed natural or semi-natural vegetation of any kind, with plots located in different habitats. The number of plots per habitat should approximately reflect the proportion of each habitat in the study area. Focusing on undisturbed vegetation allows to identify natural species pool variation. Once more, note that co-occurrence plot sampling should not be restricted to the same habitat type(s) as the one(s) in the nested permanent plots. Since permanent plots at natural sites represent typical vegetation in the area, several co-occurrence plots will likely share similar vegetation as well. However, the participants that have sampled only woody species in their permanent plots (see below), must ensure that all their co-occurrence plots include woody species.



#### PERMANENT SITES

- Setting up the plots in the permanent sites:
  1. Mark the common center for the core plot and the surrounding plot (orange x in the figure)
  2. Make a diagonal (South to North) to find the first two corners of the core plot and the surrounding plot.
  3. Repeat the procedure with the other diagonal (West to East).
  4. The 8 points for soil analyses are placed 10 cm outside of the small subplot, in the relative positions shown by the black dots in the figure.



#### GENERAL STRUCTURE OF THE STUDY AREA

- For each habitat type (up to three; two habitats shown in the example), a pair of sites: core plot nested within surrounding plot. All sites must be less than 20 km apart from each other.
- Each permanent plot in a different level of disturbance (natural site and anthropogenic site).
- 30 (preferably more) plots for co-occurrences within 10 km radius from the center of the permanent sites. Same size as the core plot (10 x 10 m).

**Fig 1.** Schematic representation of the sampling design.

#### 4. Core measurements

The webpage <http://bit.do/DarkDivNet> will contain protocol updates and data-sheets that should be filled and submitted to [darkdivnet@ut.ee](mailto:darkdivnet@ut.ee). The following protocols should be followed carefully. In case of doubt, do not hesitate to contact the coordinators.

##### a. Measurements in permanent plots

For each site, the participants will record some general characteristics:

- Geographic coordinates (latitude, longitude) of the center of the permanent plot.
- ID of the site (codes will be assigned by the coordinators).
- Date of sampling.
- Broad vegetation type (chosen from a series of fixed options)
- Vegetation type specification (a more detailed description of vegetation type).
- Slope, aspect, elevation.
- Position of the site along the disturbance/modification gradient (natural/anthropogenic), including a description of the disturbance (type, history, intensity).
- Up to 10 photos of the site (minimum one photo per site). These photos will be used to have a general overview of the characteristics of each site. Add the name and contact information of the photographer (in case we want to use the photo in media).

- Total percent cover of vegetation (%), as well as the total cover of bare ground and distinct layers (litter, herbaceous vegetation, bryophytes, lichens, shrubs, trees). Visual estimation of the average height of each of the vegetation layers.
- Vascular plant species lists at the two considered scales (core plot and surrounding plot). Nomenclature will follow The Plant List (<http://www.theplantlist.org>). At both scales, include all the vascular plant species rooted in the corresponding plot. The percent cover of each species present will be characterized. Species cover will be estimated in both the core and the surrounding plot (with separate lists for each of them). In principle, all vascular plants should be included in the survey. However, after agreement with the coordinators, exceptions could be made in some cases, particularly in hyperdiverse tropical forests. In this case, it will be allowed to sample only woody species using a DBH threshold and to use a rough estimation of total basal area for each species instead of species covers.
- To estimate functional dark diversity, the following traits will be measured for all the species found in each site:
  - Maximum vegetative height (cm): distance between the ground and the highest vegetative leaves of the plant (i.e. excluding inflorescences, Cornelissen *et al.* 2003). For this, the tallest individual of each species will be selected and measured. Plants will be measured in its normal state, i.e. without stretching them.
  - Leaf width and length (cm): measured in the same individuals in which height is estimated, the participants will choose an average sized mature leaf, preferably from a part of the canopy exposed to light and without signs of herbivory, and measure its length (excluding the petiole) and width. Measurements will be given with a precision of 1 cm, except for leaves with dimensions smaller than 1 cm, in which case a level of precision of 0.1 cm will be used. In the case of compound leaves the whole leaf should be measured (not individual leaflets). For graminoids measure only the lamina, not leaf sheath. Record leaf size as zero for leafless species.
  - Growth form of all species. The participants will follow Cornelissen *et al.* (2003). The considered categories are:
    1. **short basal**: leaves <0.5 m long concentrated very close to the soil surface, e.g. rosette plants or prostrate growth forms (compare with 5, 6 and 11);
    2. **long basal**: large leaves (petioles) >0.5 m long emerging from the soil surface (e.g. bracken *Pteridium aquilinum* or certain agaves), but not forming tussocks (cf. 6);
    3. **semi-basal**: significant leaf area deployed both close to the soil surface and higher up the plant;
    4. **erect leafy**: plant essentially erect, leaves concentrated in middle and/or top parts
    5. **cushions** (=pulvinate): tightly packed foliage held close to soil surface, with relatively even and rounded canopy boundary;
    6. **tussocks**: many leaves from basal meristem forming prominent tufts;
    7. **dwarf shrubs**: woody plants up to 0.8 m tall;
    8. **shrubs**: woody plants taller than 0.8 m with main canopy deployed relatively close to the soil surface on one or more relatively short trunks;
    9. **trees**: woody plants with main canopy elevated on a substantial trunk;
    10. **leafless shrubs or trees**: with green, non-succulent stems as the main photosynthetic structures;
    11. **short succulents** (plant height <0.5 m): green globular or prostrate ‘stems’ with minor or no leaves;
    12. **tall succulents** (>0.5 m): green columnar ‘stems’ with minor or no leaves;



13. **palmoids**: plants with a rosette of leaves at the top of a stem (e.g. palm trees and other monocotyledons, certain alpine Asteraceae such as Espeletia);
  14. **epiphytes**: plants growing on the trunk or in the canopy of shrubs or trees (or telegraph wires);
  15. **climbers and scramblers**: plants that root in the soil and use external support for growth and leaf positioning; this group includes lianas;
  16. **hemi-epiphytes**: plants that germinate on other plants and then establish their roots in the ground, or plants that germinate on the ground, grow up the tree and disconnect their soil contact. This group also includes tropical ‘stranglers’ (e.g. some figs);
  17. **hemiparasites or holoparasites**: with haustoria tapping into branches of shrubs or trees, to support green foliage (mistletoes, e.g. Loranthaceae, Viscaceae; also Cuscutaceae);
  18. **aquatic submerged**: all leaves submerged in water;
  19. **aquatic floating**: most of the leaves floating on water;
  20. **other growth forms**: give a brief description, and send a picture to the coordinators.
- NOTE 1: The traits of those species found both at the natural and anthropogenic sites will be measured in both sites.
  - NOTE 2: Please, if you have any questions regarding the trait measurement protocols, or particular difficulties, contact coordinators to discuss potential solutions.
- **Soil sampling protocol**
    - It is extremely important to follow the protocol carefully to avoid contaminating eDNA samples.
    - Never leave silica gel bags open for longer than necessary! Dry silica gel is **orange**!

*This soil sampling kit contains the following equipment: 3 pairs of rubber gloves, 6 teaspoons, 2 large plastic minigrip bags (includes the kit package itself; a unlabeled bag is for soil mixing, the other labelled bag is for sending the collected samples A and C), 3 labelled paper bags (to collect soil samples A, B and C), 2 small plastic minigrip bags and ca. 25+25=50 g of silica gel (to preserve eDNA samples A and B), 1 sticker (to close the bag of soil sample C).*

○ **In the field (at each site):**

1. Add the **full site ID** and **date** of sampling to all 3 *paper bags* and one labelled *large plastic minigrip bag*. The coding for full site ID code (6 characters long) is shown in the table:
2. Put on the *rubber gloves*.
3. Select 8 POINTS in the outer part of the core plot (ca. 10 cm distance from the margins of the core plot; the 8 points will correspond to the four corners of the core plot, and to the four central points of each side). Collect soil from each point using *a teaspoon* per plot (replace if it breaks). From each point, collect **3** heaped teaspoons of soil from the topsoil beneath the litter and put it into the *unlabeled large minigrip bag (24 teaspoons in total)*. Usually it is enough to remove vegetation and litter until you reach the first soil layer. Avoid stones, large roots, soil animals.
4. Close the *large minigrip bag* and mix the soil well by shaking and crushing through the plastic bag. Try to disintegrate large soil particles.

<b>D</b>	<b>0</b>	<b>0</b>	<b>1</b>	<b>N</b>	<b>A</b>	<b>1</b>
DarkDivNet	Study area ID (three digits assigned by DarkDivNet)			Select N - natural site or A - anthropogenic site	Habitat type number Assign 2 or 3 if additional habitat types are sampled	

5. Place **4 teaspoons** of mixed soil into the labelled *paper bag A*, fold the paper bag and place it into a *small minigrip bag*. Repeat the same with *paper bag B* and another *small minigrip bag* (this is backup sample).
  6. Place the rest of the mixed soil into the *paper bag C* for geochemical analyses.
  7. Pack **all** soil samples (*A* and *B* inside *small minigrip bags*) into the labelled *large plastic minigrip bag* which contained the sampling kit. Make sure that all 3 paper bags and the large minigrip bag are correctly labelled with the full *site ID code* and *sampling date*!
  8. Place used *gloves* and *teaspoons* into the *large plastic bag* which was used to mix the soil for disposal.
- **After field work indoors, within 12 hours:**
1. In case the soil samples are wet, take out *paper bags A* and *B* from the *minigrip bags* and let them dry at ambient temperature for 24-48 hours (make sure that paper bags are closed, kept on clean surface and NOT in contact with each other to avoid contamination). Once the soil is dry, put the *paper bags* back to their plastic minigrip bags. Before this procedure, make sure that your hands are clean; never put your fingers inside paper bags!
  2. Add ca. 25 g dry silica gel (half of the supplied amount) to each of the *small minigrip bags* containing the *paper bags A* and *B*, and **close them airtight!** Put both back to the labelled *large minigrip bag* and close it airtight.
  3. Dry the soil sample in the *paper bag C* at ambient temperature. Once dry, close the paper bag securely using the sticker, and put it into the *large minigrip bag* containing eDNA samples *A* and *B* and close it airtight.
  4. If the silica gel moisture indicator color changes (i.e. if it is not orange anymore), you should replace the silica gel (use the backup sampling kit). Make sure that all the minigrip bags containing samples are securely closed.
- **Sending samples:**
1. If you are outside the European Union, please contact the coordinators to obtain documentation for sample posting. Within EU no special documentation is needed.
  2. Take out the backup *sample B* from the *large minigrip bag* and store it in a separate *large minigrip bag* (use the backup sampling kit). Repeat the same with all backup samples. The backup samples *B* collected from all sites can be stored together in one large minigrip bag; make sure that the large minigrip bag is securely closed. Store the backup samples in a dry place at ambient temperature. You will be requested to send *sample B* in case something happens with *sample A*.
  3. Send **samples A and C** enclosed in the labelled *large minigrip bag* to:

**Kersti Riibak**  
**Inst. Ecol. Earth Sci., University of Tartu**  
**Lai 40, 51005 Tartu, Estonia**  
 Contact: +372 7376222, [darkdivnet@ut.ee](mailto:darkdivnet@ut.ee)

## **b. Measurements in plots for analyzing co-occurrences**

At each 10 x 10 m co-occurrence plot, the participants will record:

- Geographic coordinates (latitude, longitude) of the center of the plot.
- ID of the plot (assigned by the coordinators)
- Date of sampling.
- Slope, aspect, elevation.
- Broad vegetation type (chosen from a series of fixed options)
- Vegetation surveys, following the same procedure as in permanent plots (all species and their covers), but only at the 10 x 10 m scale.
- In all the co-occurrence plots the same traits (height, leaf width and length and growth form) will be measured in all species. This means that the traits of those species that are found in several co-occurrence plots will be sampled several times, once per plot.

## **5. Optional information**

This information is not mandatory for a site to be included in the network, but it is highly recommended, and would allow the participants to take part in additional potential publications.

- When possible, the participants will prepare an expert-based habitat-specific species pool list for each of the sampled (semi)natural core habitat type within the study area. This will consist of assigning each species found in the co-occurrence plots to one of the four predefined levels: level 3 = definitely characteristic of the considered natural or semi-natural habitat; level 2 = often occurs in the considered habitat; level 1 = occurs, but not characteristic, 0 = not belonging to the species pool of that habitat. Note that when more than one habitat is considered (i.e. in study areas with more than one pair of sites), a different list will be done for each habitat type.
- After 2-3 years of the first sampling, the participants will be invited to resample vegetation in the core plots following the same protocols.

## **6. Guidelines for participation**

*DarkDivNet* is a network of researchers implementing local surveys replicated at a global scale. The expenses expected from participants are the time and resources invested in finding and sampling the different sites, and the costs of shipping soil samples to the coordinators. To facilitate communication with the coordinators, each study area will be led by a single person (participant). However, each participant can coordinate with colleagues to assemble a team to perform the sampling described above (see below for co-authorship guidelines).

Participants are expected to:

- Implement the sampling protocols exactly as explained above. Consult with the coordinators in case of questions about the protocols. Submit the corresponding information in the data-sheets provided on the project webpage.
- Share data associated with the project in a common repository, which will be centralized at the University of Tartu.
- Participate in writing of papers, follow deadlines, communicate with the coordinators.



Participants complying with the previous points are entitled to:

- Be included automatically as a co-author of those *DarkDivNet* papers that use data from their study area. To ensure co-authorship, the participants will have to communicate with the coordinators, follow deadlines, and revise/comment on the drafts of the papers.
- Each participant can offer co-authorship to one colleague who contributed considerably to field work. Further, in study areas including three natural/anthropogenic site pair and 90 co-occurrence plots (the maximum per area), an extra coauthor is justified.
- Propose to lead additional papers using the *DarkDivNet* data. These proposals will be sent to the coordinators, who will discuss them with the Advisory Panel, and suggest modifications when there are large overlaps with studies already in progress.
- Suggest add-on studies (additional measurements in study areas etc.). All communicating, data handling, publishing policy is solely responsibility of the initiator of an add-on study.
- Use the data obtained from their plots for publications, presentations, courses and other venues (e.g. blog posts) without the explicit approval of the coordinators.

## 7. *DarkDivNet* schedule

**October 2018 – March 2019:** Launching the *DarkDivNet*, introducing the project, planning study areas, sampling in the southern hemisphere.

**April 2019 – September 2019:** Sampling in the northern hemisphere, new sites planned.

**October 2019 – March 2020:** Sampling in the southern hemisphere; new sites planned.

**April 2020 – September 2020:** Sampling in the northern hemisphere; preliminary results reported in conferences.

**October 2020 onwards** starting writing first papers, new sites added continuously (will be included to following papers as soon as data becomes available), add-on studies.

## References

- Cornelissen, J. H. C., Lavorel, S., Garnier, E., Diaz, S., Buchmann, N., Gurevich, D. E., . . . Poorter, H. (2003). A handbook of protocols for standardised and easy measurement of plant functional traits worldwide. *Australian Journal of Botany*, 51, 335-380.
- Lewis, R. J., Szava-Kovats, R., & Pärtel, M. (2016). Estimating dark diversity and species pools: An empirical assessment of two methods. *Methods in Ecology and Evolution*, 7(1), 104-113.
- Lewis, R. J., de Bello, F., Bennett, J. A., Fibich, P., Finerty, G. E., Götzenberger, L., . . . Pärtel, M. (2017). Applying the dark diversity concept to nature conservation. *Conservation Biology*, 31(1), 40-47.
- Pärtel, M., Szava-Kovats, R., & Zobel, M. (2011). Dark diversity: shedding light on absent species. *Trends in Ecology and Evolution*, 26, 124-128.
- Zobel, M. (2016). The species pool concept as a framework for studying patterns of plant diversity. *Journal of Vegetation Science*, 27(1), 8-18.