

Ecological signals of plant-microbe associations are consistent across eDNA and vegetation surveys in northeast Greenland

Supplementary information 1

Parisy B^{a†}., Schmidt N.M^{b+}., Wirta H^a., Stewart L^c., Pellissier Ld,^e., Holben W.E.^f., Pannoni S^f., Somervuo P^g., Jones M.M.^{g,h}., Siren J^h., Eero Vesterinenⁱ, Ovaskainen O^{g,j,k}., Roslin T^{a,l}.

a: Department of Agricultural Sciences, University of Helsinki, Helsinki, Finland. b: Department of Ecoscience, Aarhus University, Roskilde, Denmark. +: Arctic Research Centre, Aarhus University, Aarhus, Denmark c: Department of Natural Sciences and Environmental Health, University of South-Eastern Norway, Bø, Norway. d: Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland. e: Swiss Federal Research Institute WSL, Birmensdorf, Switzerland. f: Division of Biological Sciences, University of Montana. g: Organismal and Evolutionary Biology Research Programme, University of Helsinki, Helsinki, Finland. h: Institute of Biotechnology, HiLIFE Helsinki Institute for Life Science, University of Helsinki, Helsinki, Finland. i: Department of Biology, University of Turku, Finland. j: Department of Biological and Environmental Science, University of Jyväskylä, Jyväskylä, Finland. k: Department of Biology, Centre for Biodiversity Dynamics, Norwegian University of Science and Technology, Trondheim, Norway. l: Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden

† Corresponding author: bastien.parisy@helsinki.fi

Table of contents

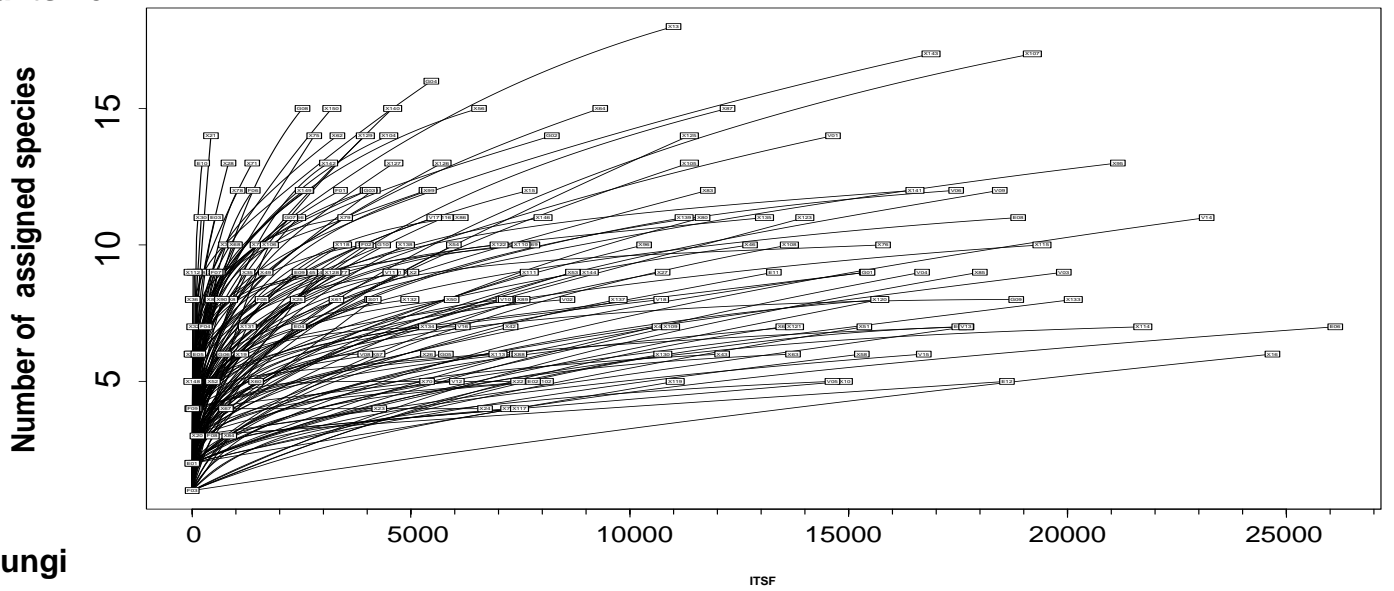
Table S1: Summary of the taxonomic assignment of different loci	1
Fig. S1: Rarefaction curves for OTU accumulation within and across samples.....	2
Supplementary text S1: Choice of specific priors for our HMSC model	3
Fig. S2: Venn diagram showing the shared and unique plant taxa detected by different methods of identification	6
Fig. S3: Spatial patterns in the distribution of plant taxa across the two different methods of identification.....	7
Fig. S4: Frequency of detection of a species by eDNA as a function of the semi-quantitative scoring of the relative coverage of the species by direct observation.....	69
Fig. S5: Boxplot of plant species richness per plot for two methods of identification.....	70
Fig. S6: Model convergence and discrimination success achieved for the HMSC model.....	71
Fig. S7: Differences in species responses among organism groups and scoring methods.....	72
Fig. S8: Summary of taxon-specific responses to environmental covariates	73
Fig. S9: Posterior predictions of the mean probabilities of occurrence of 19 plant species based on observational vs eDNA data along gradients in soil temperature, pH and moisture	74
Fig. S10: Numerical summary of associations detected between taxa.....	80
Fig. S11A: Estimated pairwise residual associations among plants and different functional groups of fungi	81
Fig. S11B: Estimated pairwise residual associations among plants and different functional groups of bacteria	82

Table S1. Summary of the taxonomic assignment for different loci. Each entry identifies the number of sequences achieving a plausible (probability of correct assignment >0.5; top) or reliable (probability of correct assignment >0.9; bottom) assignment at the respective taxonomic level for the locus in question. Column “% reads assigned” represents the percentage of sequences identified to a given taxonomic rank, as a proportion of the original, “raw” number of reads. ITS2 and rbcLa correspond to the loci used to identify plants, ITSF the locus for identifying fungi and 16S bacteria.

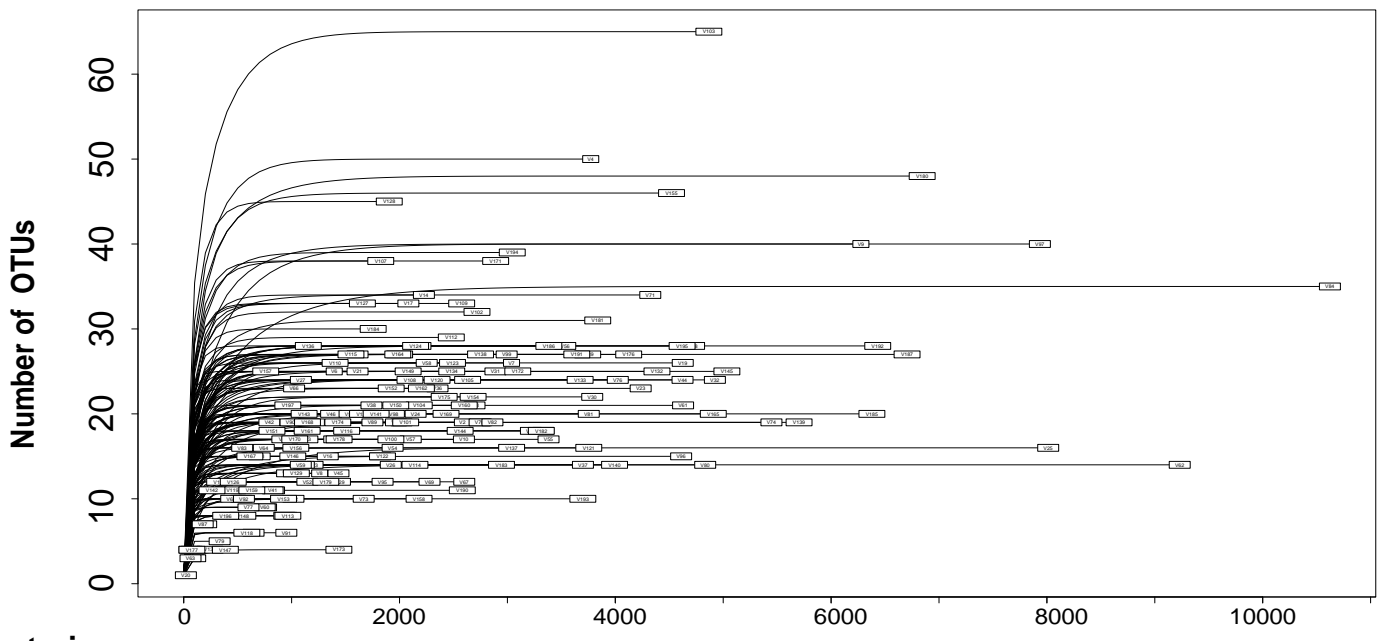
Plant						
Plausible (P>0.5)	ITS2			RBCLA		
	Total reads	% reads assigned	Number of taxa	Total reads	% reads assigned	Number of taxa
Raw	2.8M			5.1M		
Phylum						
Class	2.7M	96.4	2	3.34M	65.5	2
Order	2.6M	92.9	15	3.29M	64.5	16
Family	2.6M	92.9	21	3.21M	62.9	25
Genus	2.5M	89.3	52	705K	13.8	52
Species	2M	71.4	102	35K	0.7	37

Plant						
Reliable (P>0.9)	ITS2			RBCLA		
	Total reads	% reads assigned	Number of taxa	Total reads	% reads assigned	Number of taxa
Raw	2.8M			5.1M		
Phylum						
Class	2.5M	89.3	2	3.34M	65.5	2
Order	2.5M	89.3	13	3.28M	64.3	16
Family	2.5M	89.3	25	3.21M	62.9	25
Genus	2.28M	81.4	46	656K	12.9	23
Species	1.3M	46.4	82	5K	0.1	25

A) Plants - eDNA



B) Fungi



C) Bacteria

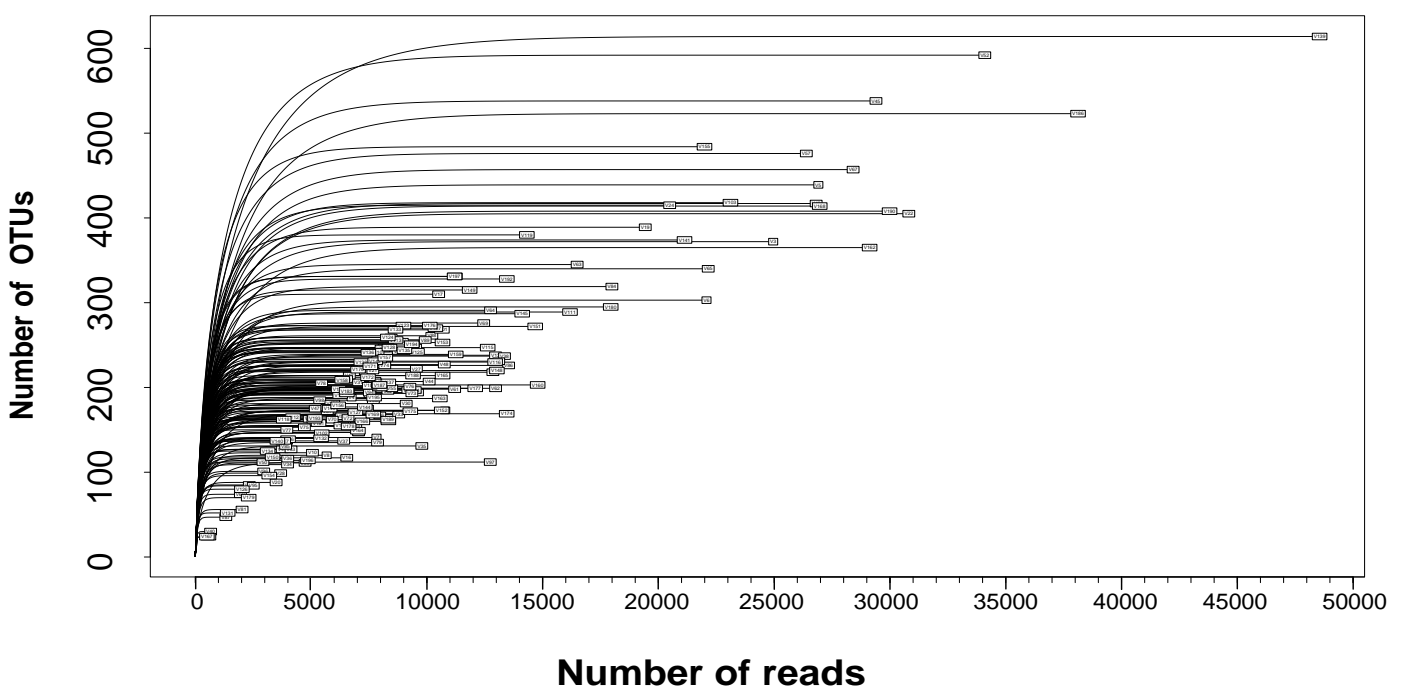


Fig. S1. Rarefaction curves of the number of OTUs per sample as a function of read numbers. Panel (A) shows the number of plant species accumulated per site, panel (B) shows the number of fungal OTUs, and panel (C) shows the number of bacterial OTUs. Curves produced with the vegan package in R (Oksanen, 2010).

Supplementary text S1: Choice of specific priors for our HMSC model.

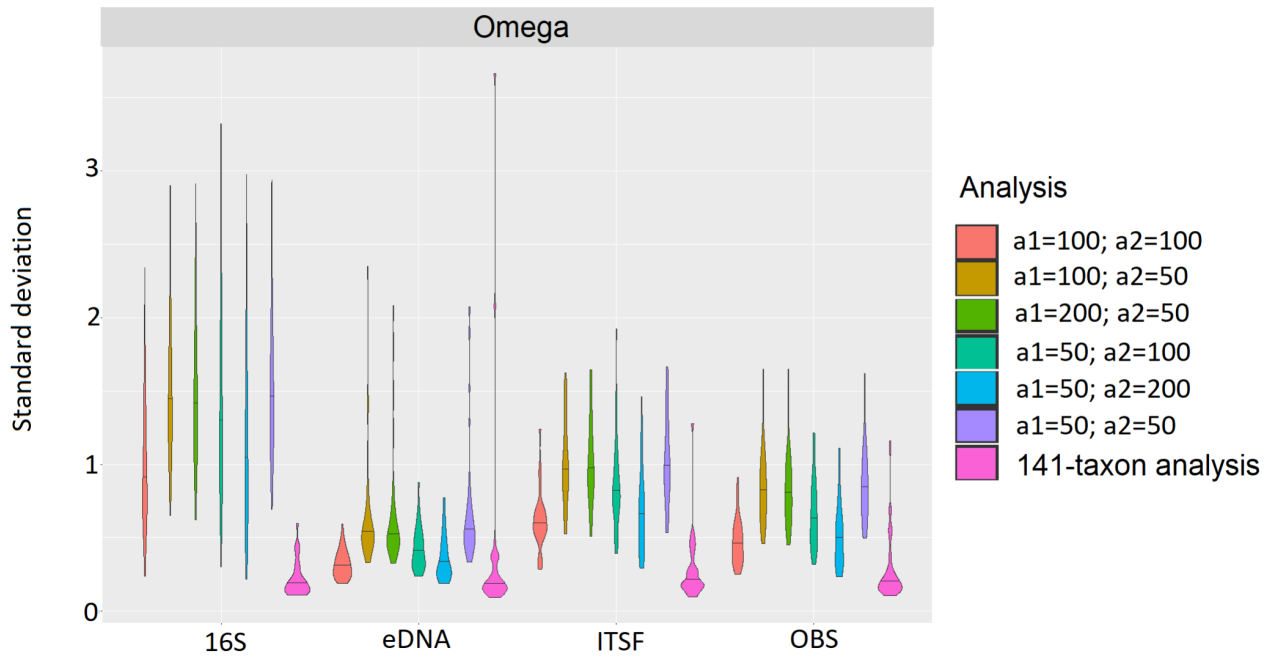
HMSC analyses of two data sets with 141 and 333 taxa, respectively, resulted in some highly different parameter estimates for the plant species, common to both data sets. For the smaller analysis (i.e., including 141 taxa), most of the marginal standard deviation estimates of the site random effect were close to zero, whereas they were much higher in the larger analysis framed on 333 taxa (Subfig S1, S2). The HMSC model defines the random effects jointly over the species, and consequently the estimates for a single species are affected by the other species in the data set (Ovaskainen & Abrego, 2020). In the smaller data set, there were no clear site effects for almost any of the species, and the estimates were almost uniformly smaller. In the larger data, there was signal of a site effect for many of the microbial species and the variances of the random effects were larger therefore estimated to be larger. As a result, the latter model did not penalize as strongly against higher site variances for the plant species as for the small data set, and their site effect estimates were consequently significantly higher.

The Multiplicative Gamma Process Shrinking Prior in the HMSC model penalizes against overly strong random effects. Without penalization the random effects would not be identifiable with occupancy data and the model might overfit. The prior is defined using several parameters that can be changed to modify its behaviour. Parameters a_1 and a_2 control the level of shrinkage for the species association matrix Ω , and their values can have a significant impact on the results (Ovaskainen & Abrego, 2020; Chapter 8.4.2). The first parameter a_1 controls the overall strength of the random effect, while the second parameter a_2 controls the structure of the association matrix via the effective number of latent factors.

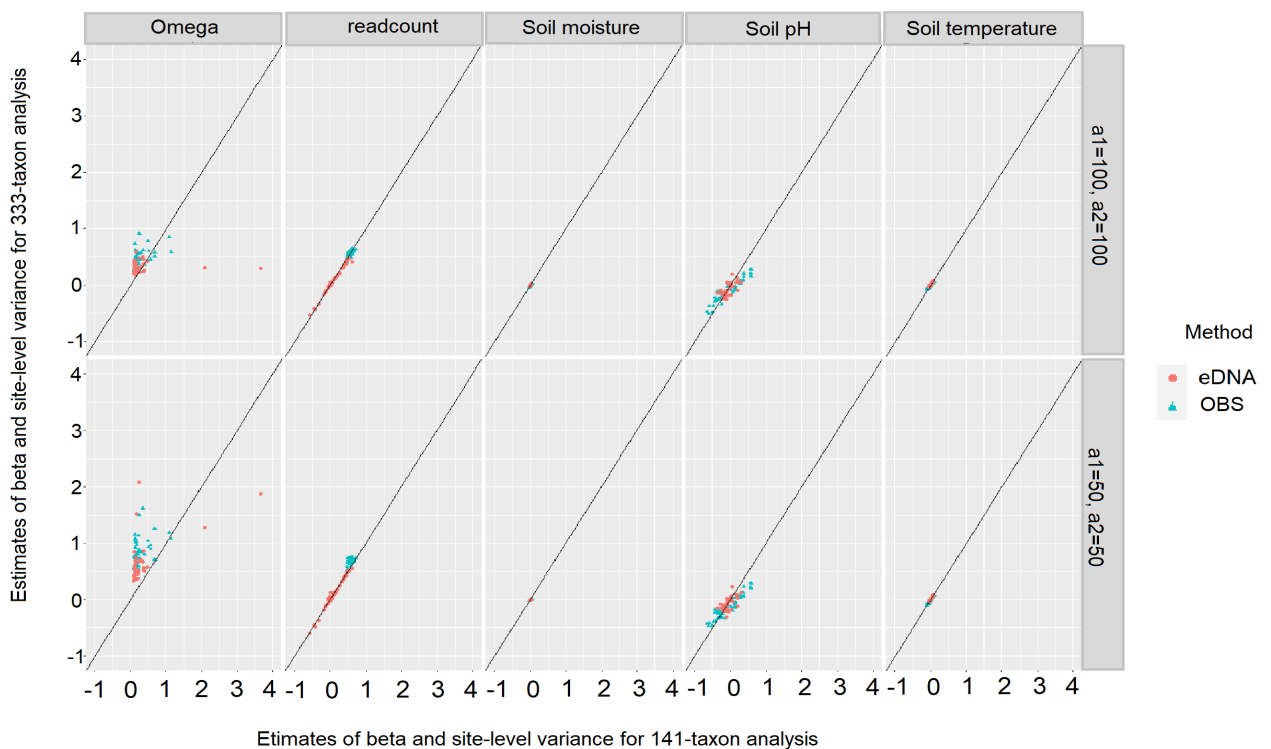
We tested whether using stronger shrinkage resulted in random effect estimates for plants in the larger model that were more similar to those in the original model of the smaller dataset. The default parameter values in the Hmsc R-package are ($a_1=50, a_2=50$). We tested five additional prior combinations ($a_1=100, a_2=50$), ($a_1=50, a_2=100$), ($a_1=100, a_2=100$), ($a_1=200, a_2=50$) and ($a_1=50, a_2=100$).

Systematically changing the values of parameters a_1 and a_2 had a clear impact on the distributions of the marginal standard deviation estimates of the site random effect over species (Subfig S1). Increasing only the value of a_1 resulted in highly similar estimates to those obtained with the default prior. The strongest shrinkage and the most similar estimates for plant species compared to those of the 141 taxa analysis were obtained with the priors ($a_1=100, a_2=100$) and ($a_1=50, a_2=100$). The prior ($a_1=50, a_2=200$) also strongly affected the beta parameter estimates for soil pH for some of the species (Fig S3), while they remained mostly similar with the prior ($a_1=100, a_2=100$).

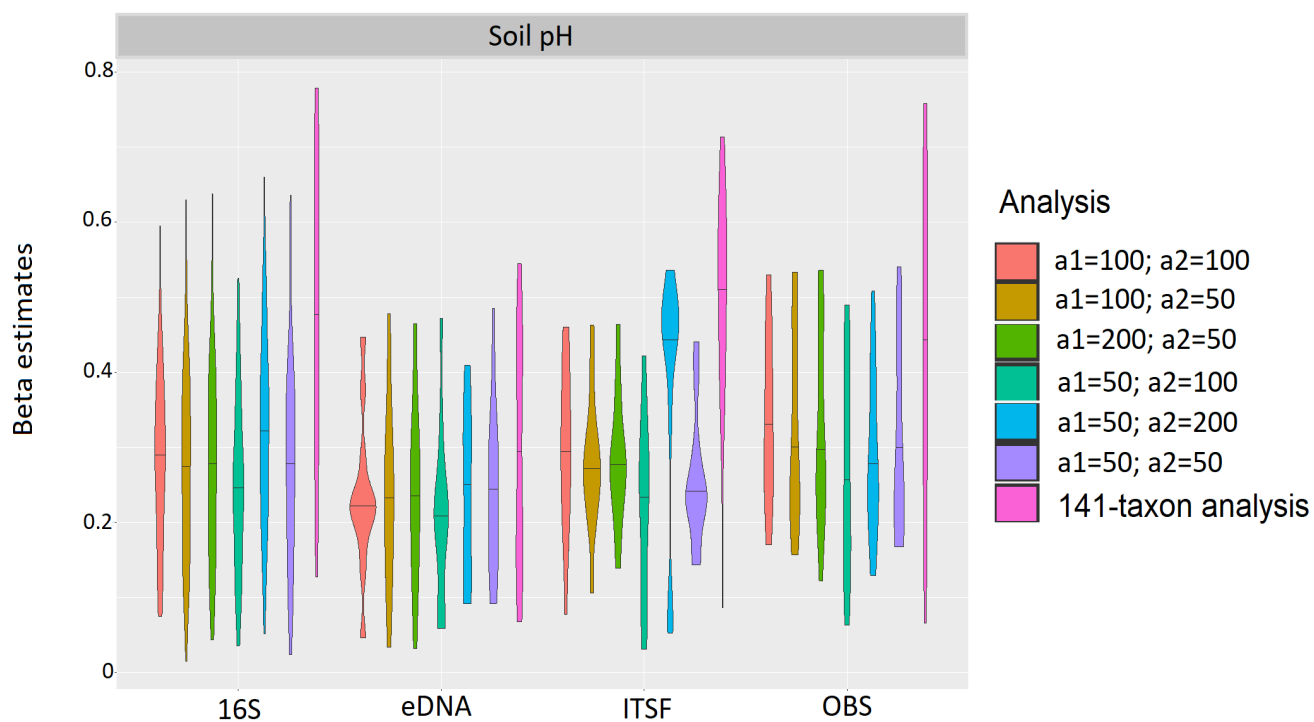
Based on these trials, we chose to use the priors of ($a_1=100, a_2=100$) in our analyses. This choice successfully decreased the discrepancy of the random effect parameter estimates for the plant species between the 141 and 333 taxa analyses, without having a strong impact on the other model parameters. Direct comparison of the estimates between these two analyses show that for most of the species, the parameter estimates are more similar with the ($a_1=100, a_2=100$) prior than with the default prior (Subfig S2). The stronger shrinkage prior also results in some decreases in the marginal standard deviations for the 16S and ITSF microbial and fungal taxa that are present only in the larger dataset. However, underestimating random effects is usually preferable to overestimating them, because overly high random effect variances can lead to overfitting with this kind of data. The HMSC model used here is complex for the data, and without regularization with the shrinkage prior it would not be identifiable.



Subfigure S1. The distributions of marginal standard deviation estimates of the site random effect over species separately for each taxon group. Each color represents the estimates for a single HMSC analysis: The pink color (i.e., PLMI) refers to an analysis including 141 taxa, whereas the other colors show the results for analyses including 333 taxa with different prior distributions as indicated in the legend. The eDNA and OBS plant species are the same in 141 and 333 taxa analyses, while the 16S and ITSF taxa are not shared between the two analyses.

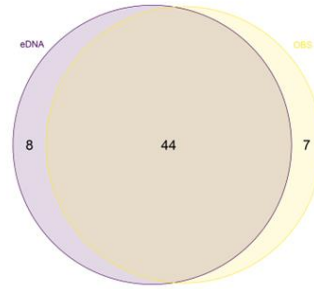
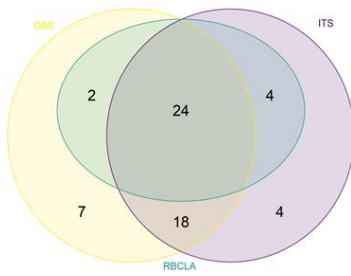


SubFigure S2. Scatterplots of beta parameter estimates and marginal site random effect variances for the plant species under two analyses. The x-axis shows the parameter estimate in the 141 taxa analysis and the y-axis shows the corresponding estimate in the 333 taxa analysis. The rows show estimates based on different priors in the 333 taxa analysis: top stronger shrinkage prior ($a1=100, a2=100$), bottom default prior ($a1=50, a2=50$), and the columns show different parameters with Omega referring to the marginal standard deviation of the site random effect. The different colors indicate whether the plant species was directly observed (OBS, blue) or identified by environmental DNA (eDNA, red). Identity lines are included to facilitate comparison.



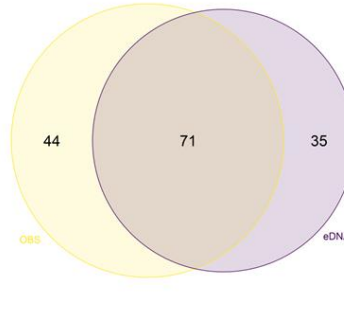
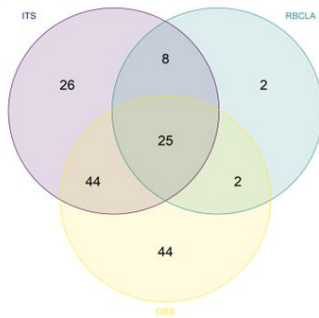
Subfigure S3. The distributions of beta parameter estimates for soil pH over species separately for each taxon type. Each color represents the estimates for a single HMSC analysis: PLMI corresponds to an analysis including 141 taxa, whereas the other colors show the results for analyses including 333 taxa with different prior distributions as indicated in the legend. The eDNA and OBS plant species are the same in 141 and 333 taxa analyses, while 16S and ITSF species are not shared between the two analyses.

A) Genus level

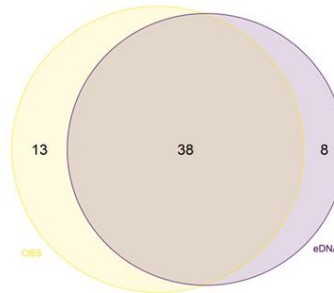
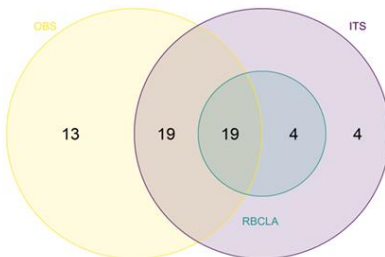


**Plausible
($P > 0.5$)**

B) Species level



C) Genus level



**Reliable
($P > 0.9$)**

D) Species level

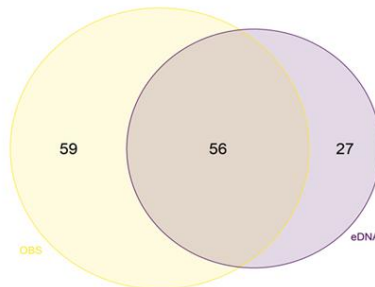
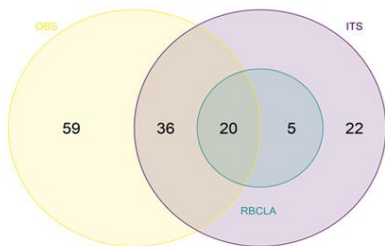
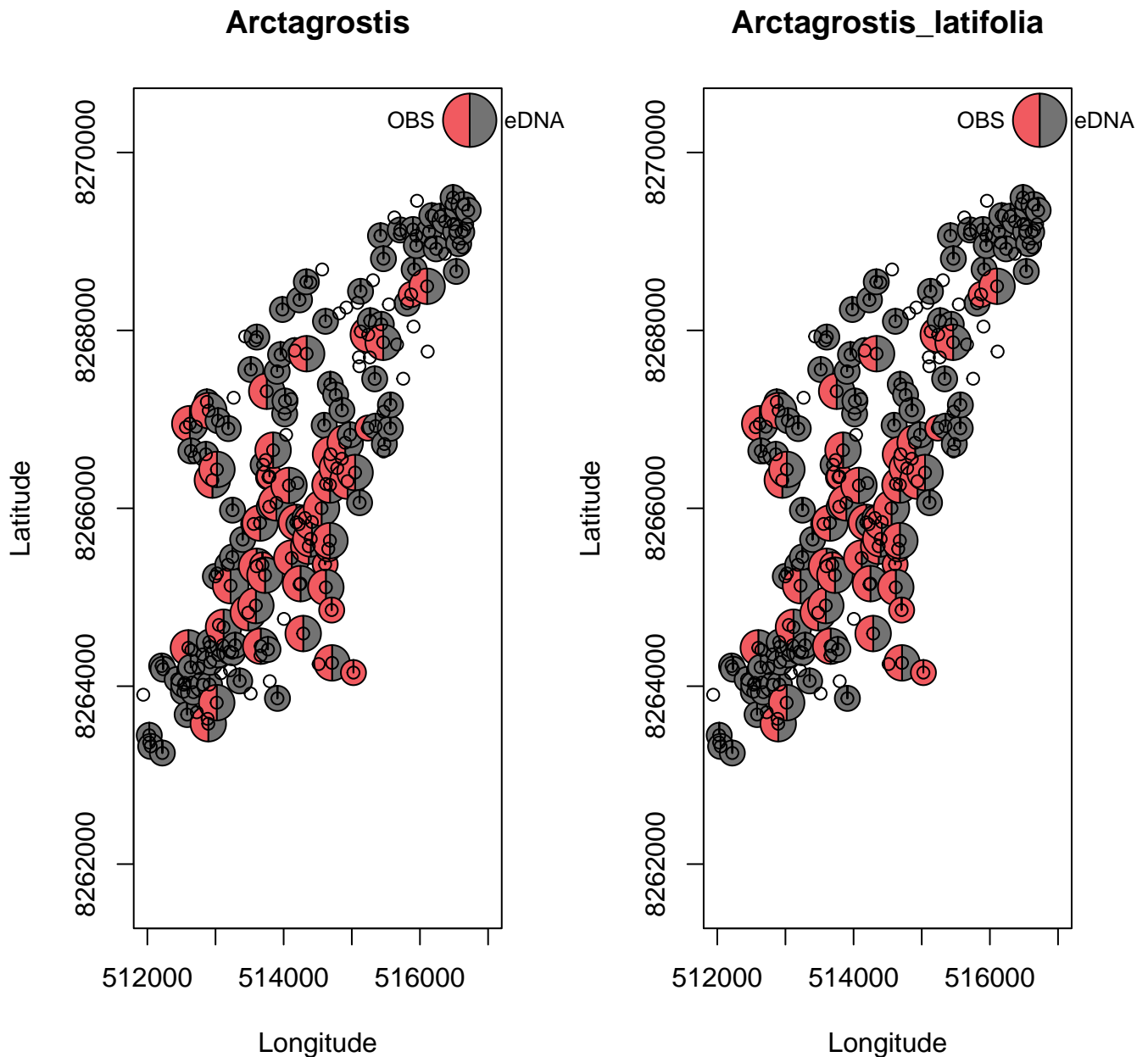
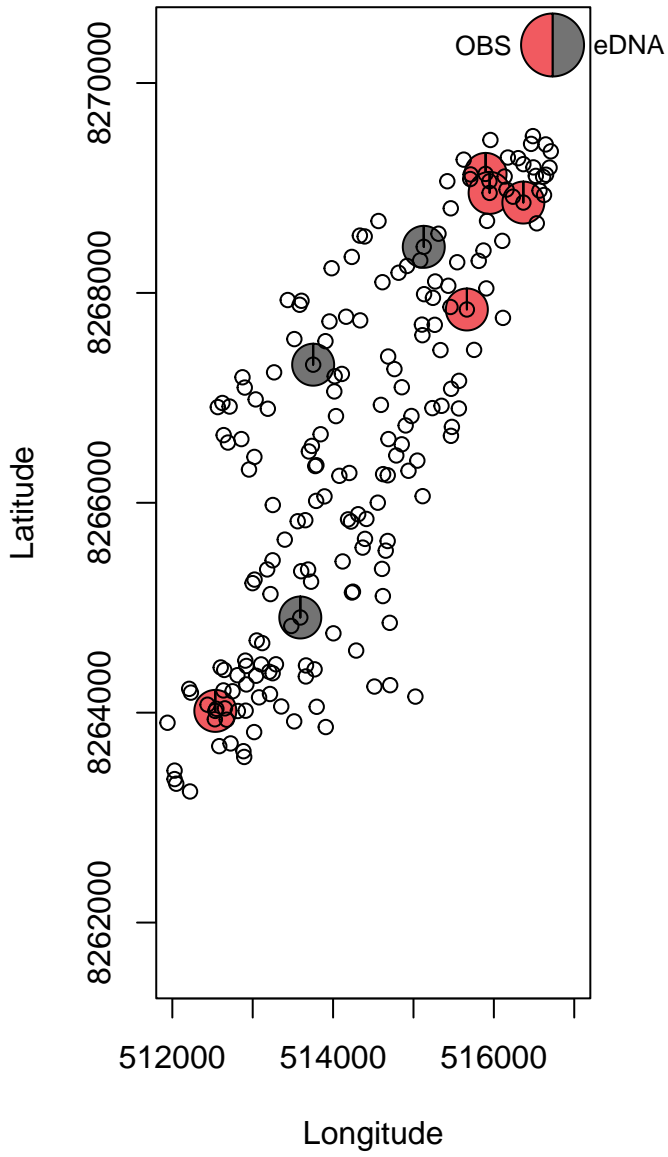


Fig. S2. Venn diagram showing the shared and unique plant taxa detected by different methods of identification. The sizes of the circles represent the total number of plant taxa detected per method, with specific numbers given for each method. Numbers within intersections identify the number of plant taxa detected by all of the respective methods. ITS refers to plants detected using gene region ITS2; rbcLa to plants detected using gene region rbcLa; OBS to plants detected by observation, and eDNA to all plants detected by metabarcoding (as combining evidence from ITS and rbcLa). Results are shown separately at different level of taxonomy (i.e genus vs species) for the two taxonomic assignment threshold used. (i.e., Plausible, $Pr > 0.5$ and reliable $Pr > 0.9$), as detailed in Table S1.

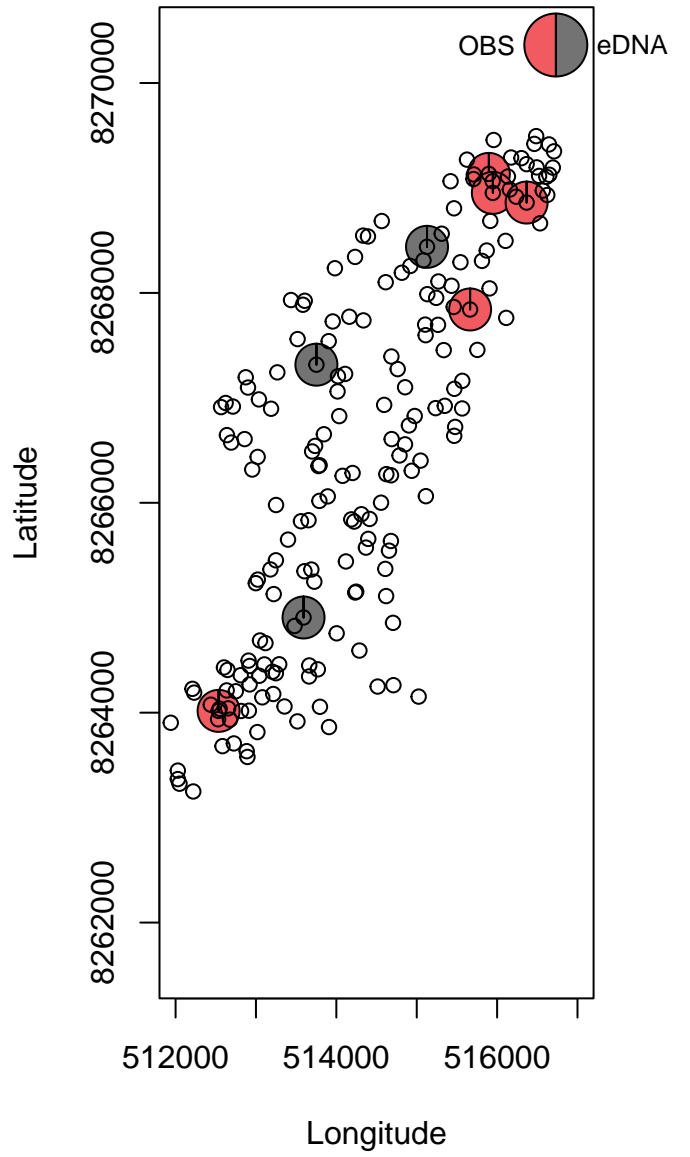
Supplementary Fig. S3. Pie charts showing spatial patterns in the distribution of plant taxa across the two different methods of scoring. To illustrate differences in detection at different levels of taxonomy we show patterns at two levels: genus (left-hand plot) and species (right-hand plot). The size of the individual pie charts indicates whether the taxon in question was locally detected by both methods (large bicolored circles), by a single method (small unicolored circles) or no method (empty circles). Note that for genera with a single species in the Zackenberg species pool, the two plots will obviously be identical.



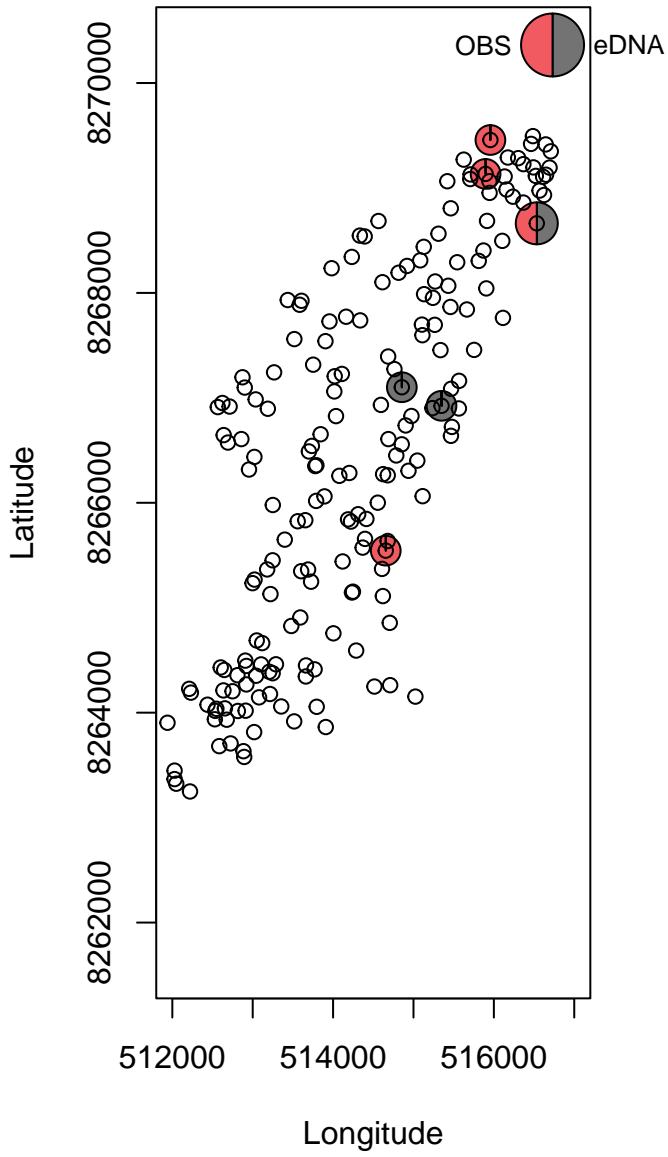
Arenaria



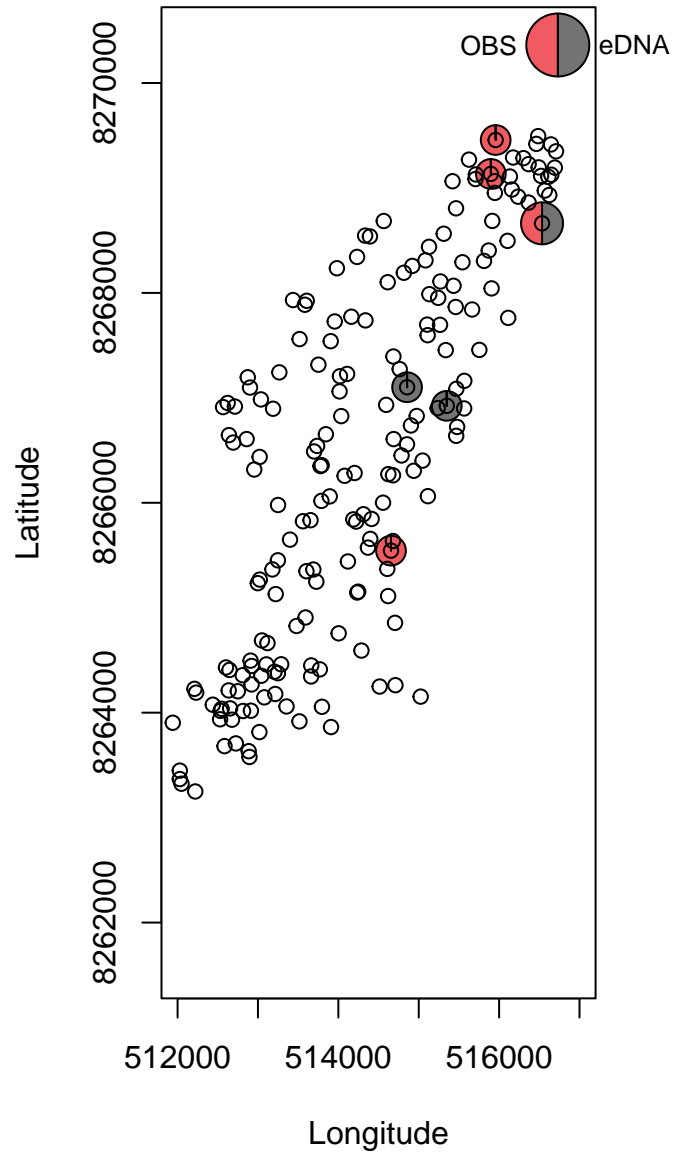
Arenaria_pseudofrigida



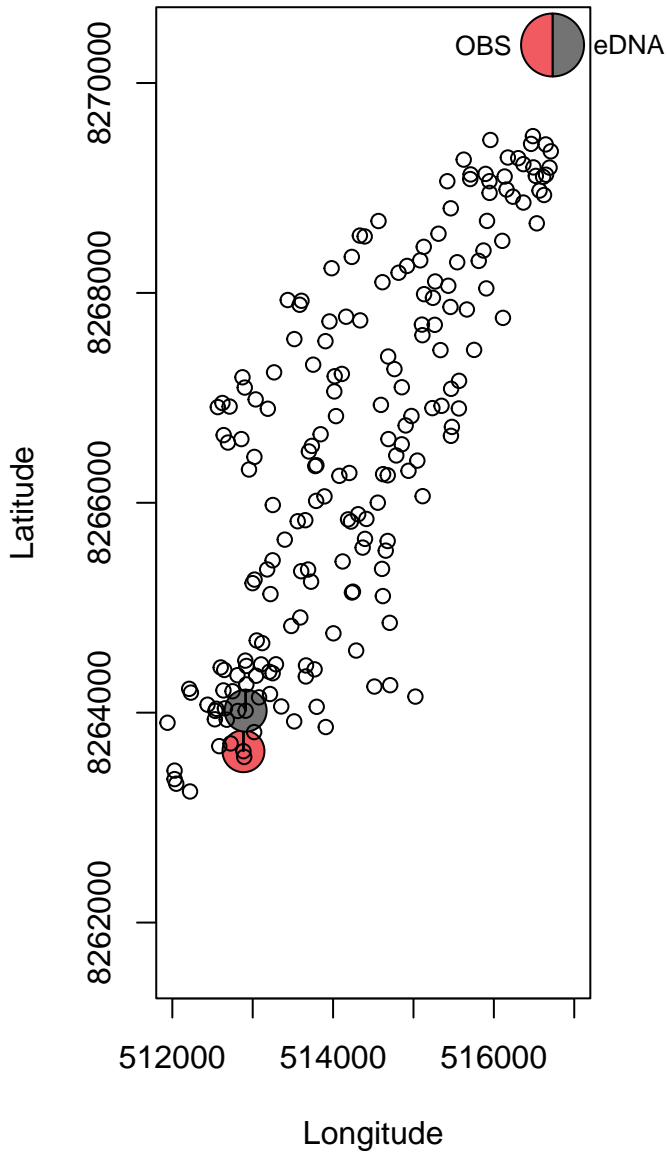
Arnica



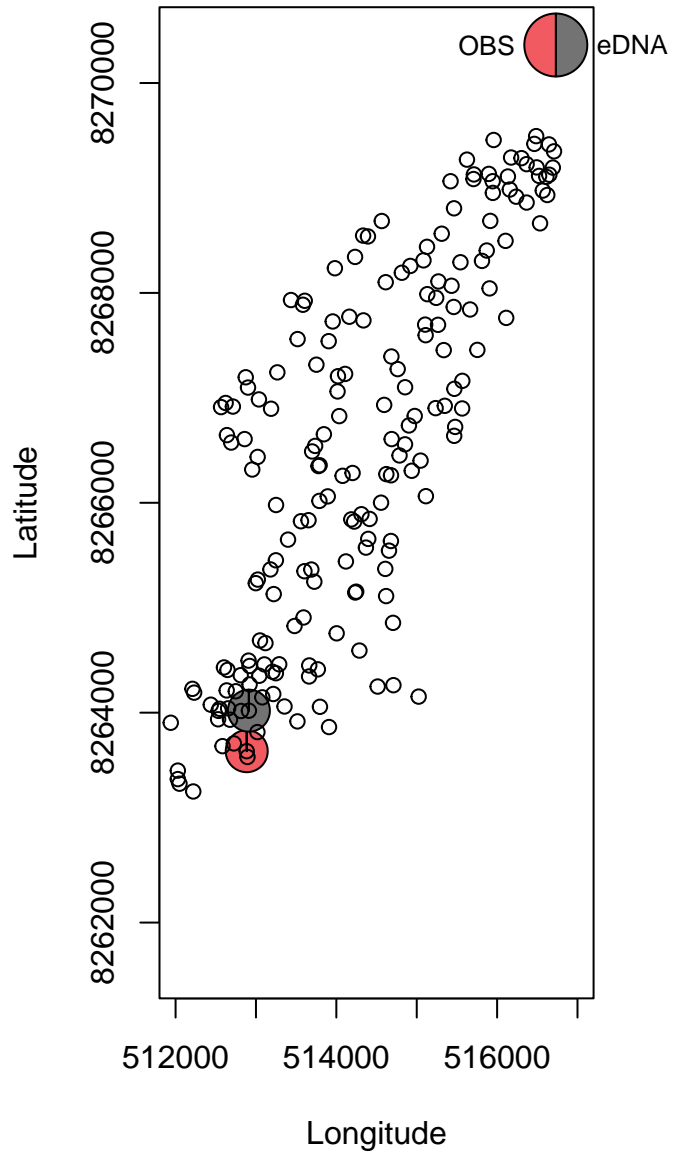
Arnica_angustifolia



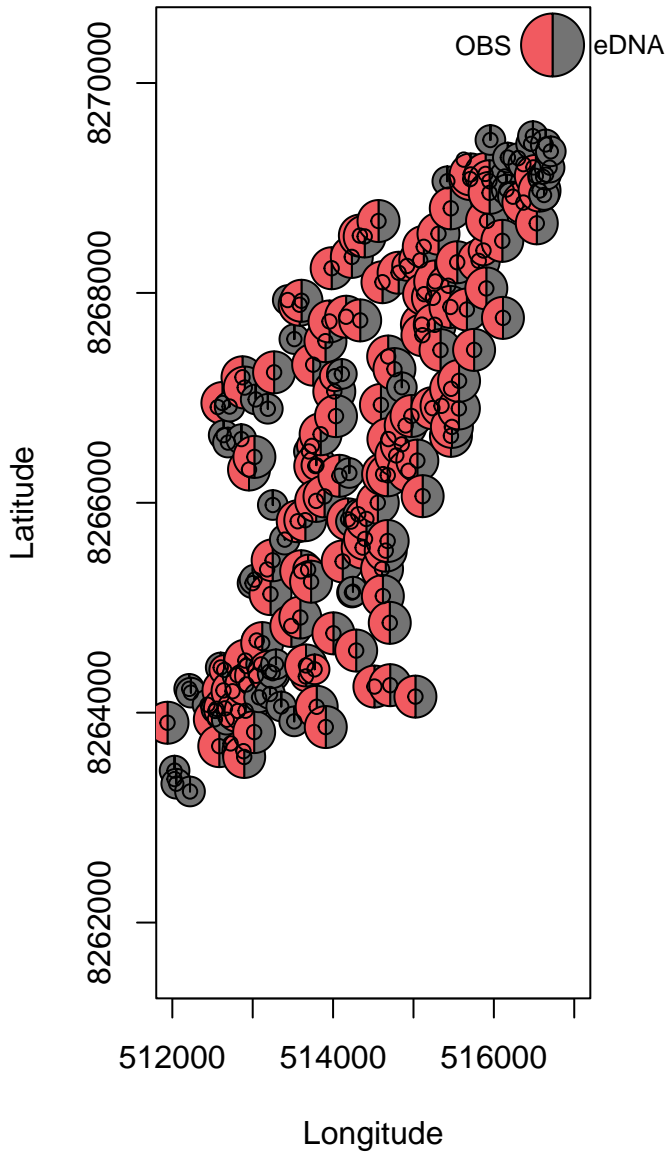
Betula



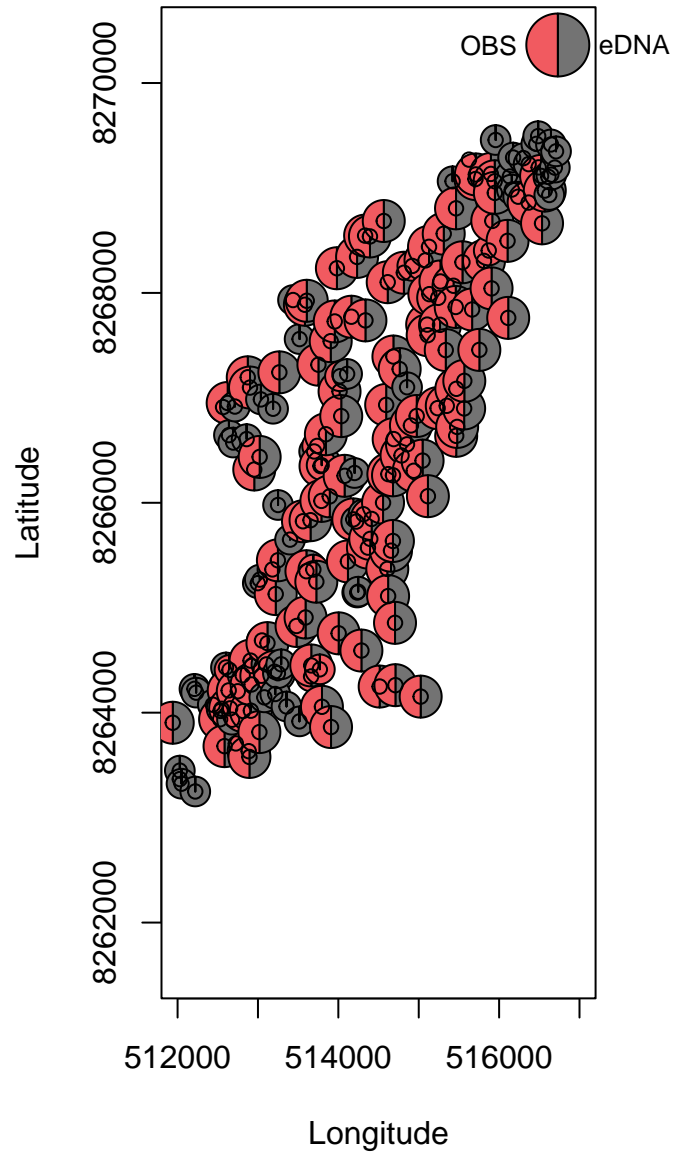
Betula_nana



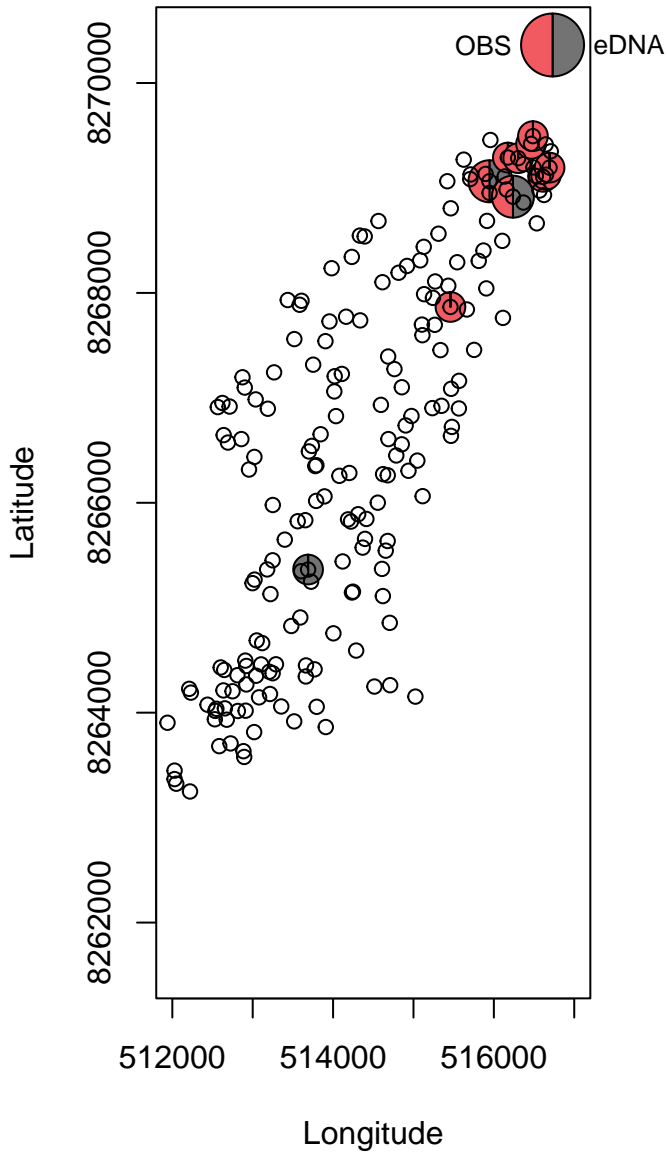
Bistorta



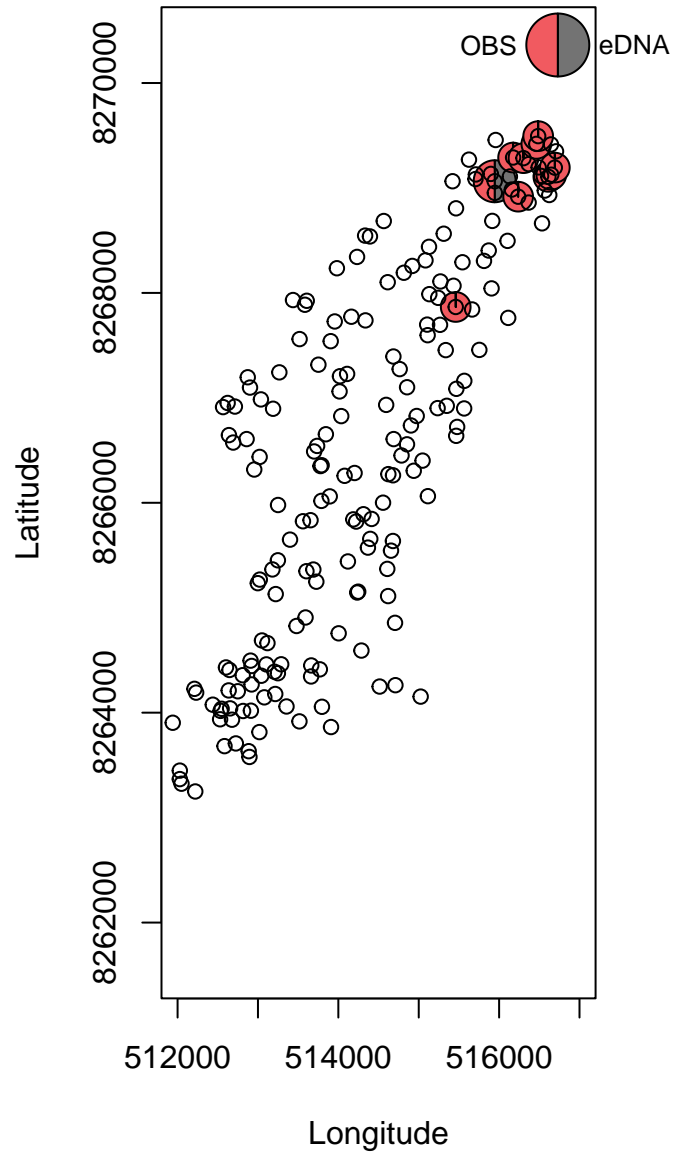
Bistorta_vivipara



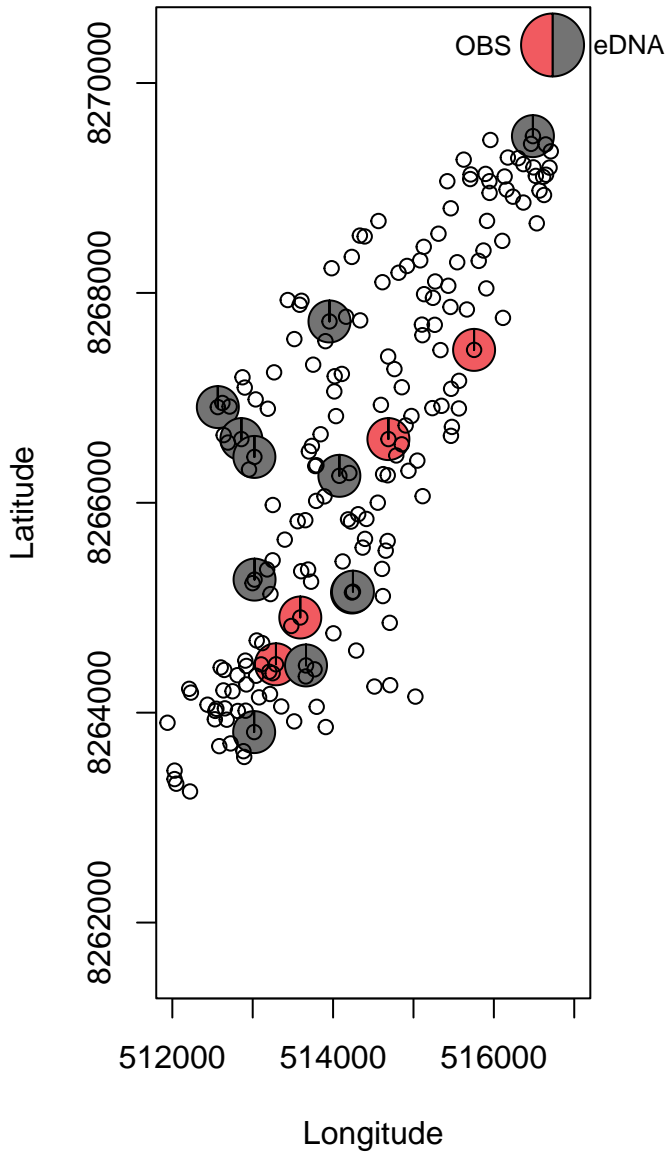
Campanula



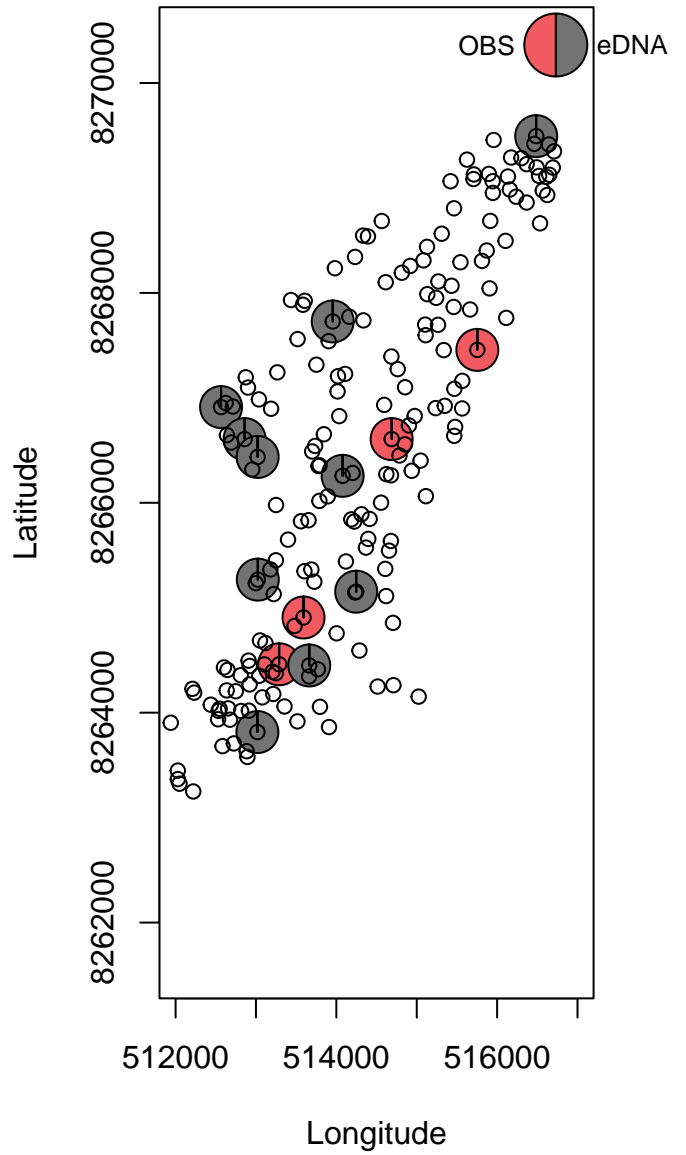
Campanula_uniflora



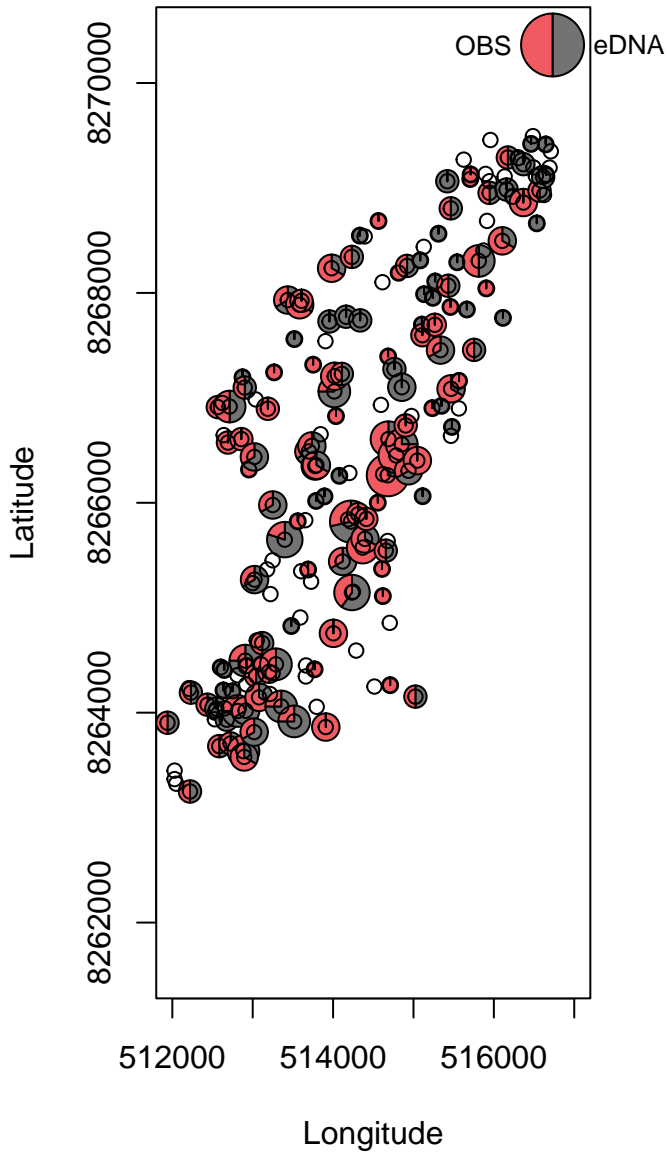
Cardamine



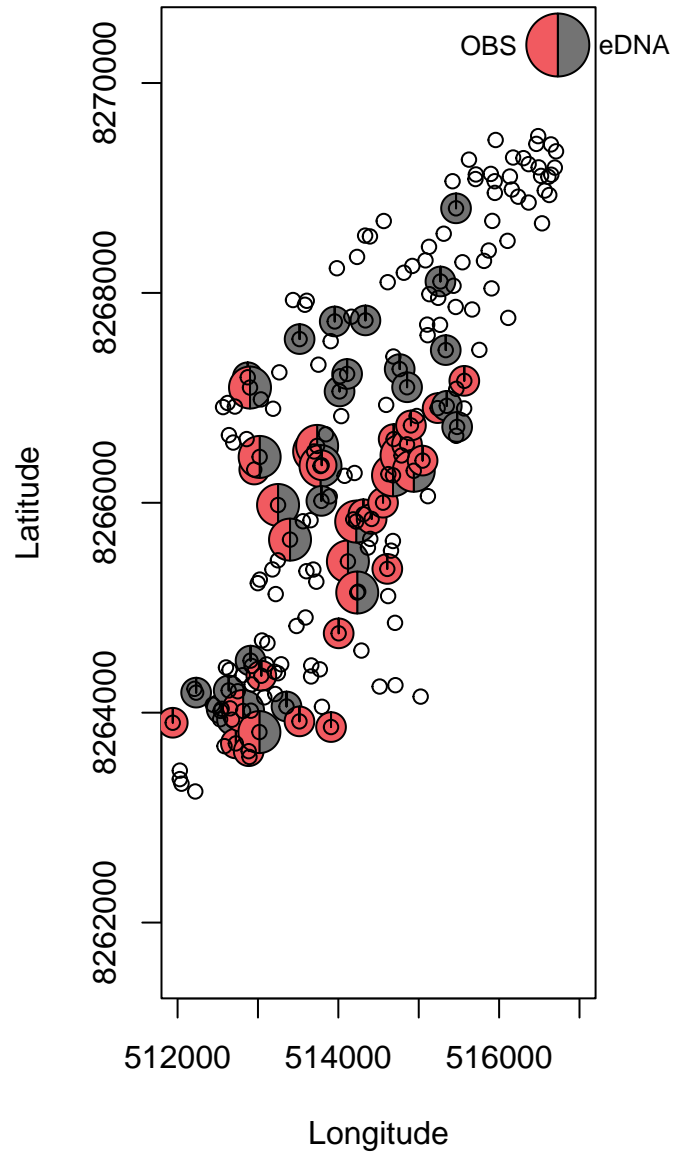
Cardamine_bellidifolia



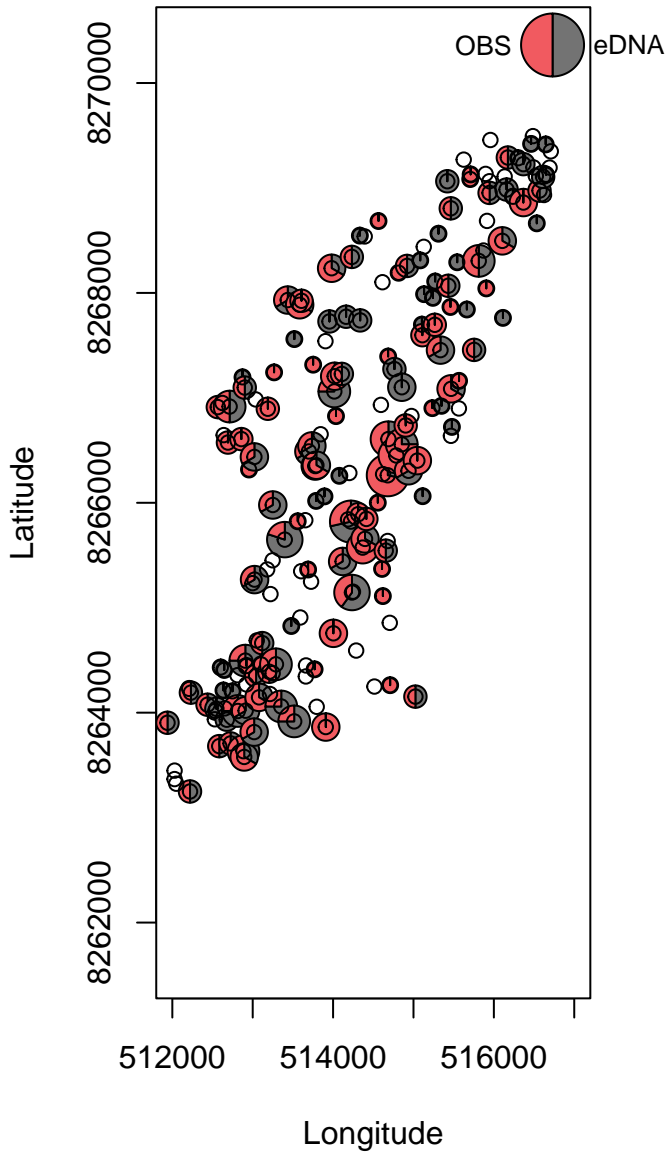
Carex



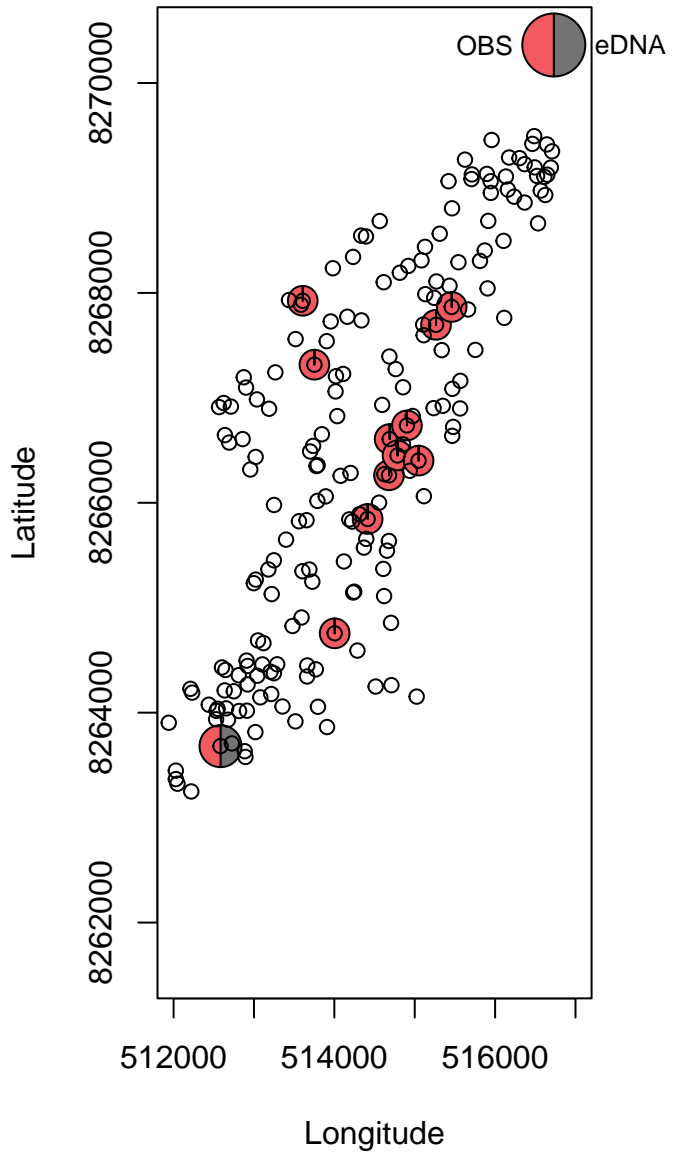
Carex_bigelowii



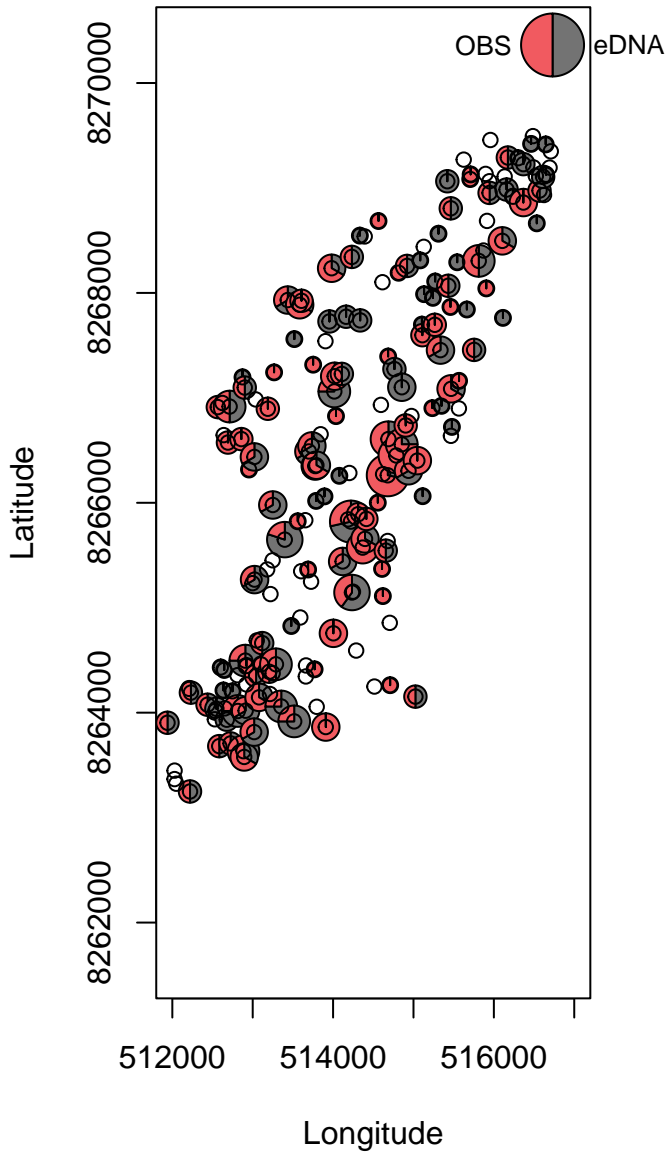
Carex



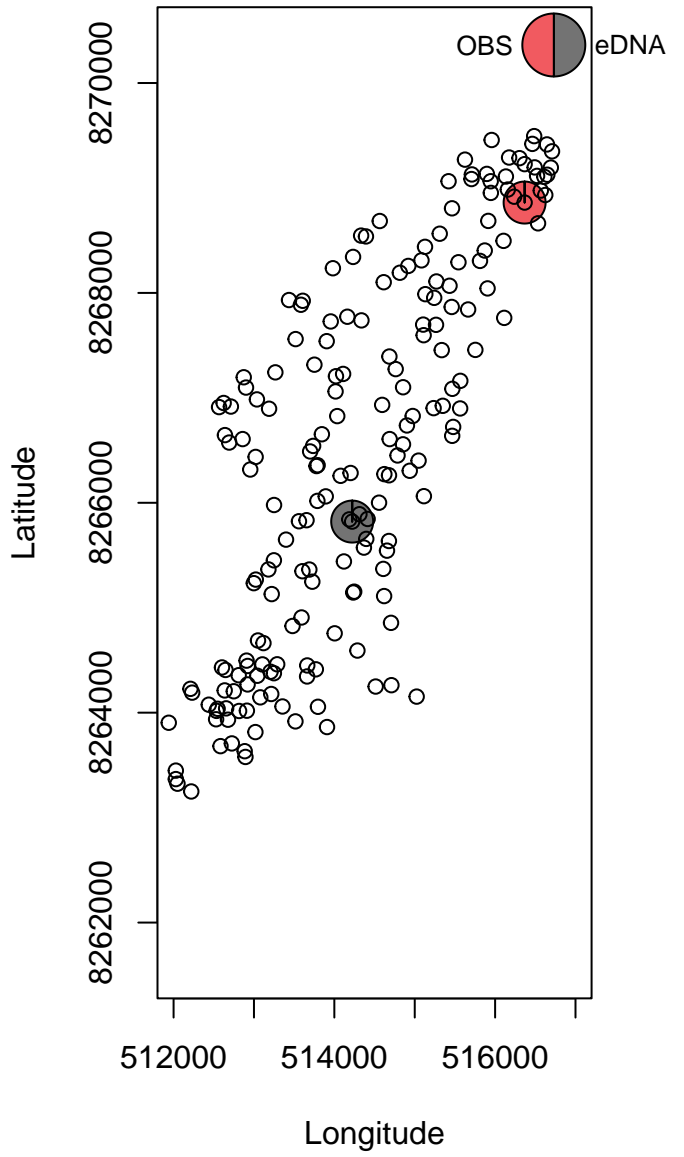
Carex_capillaris



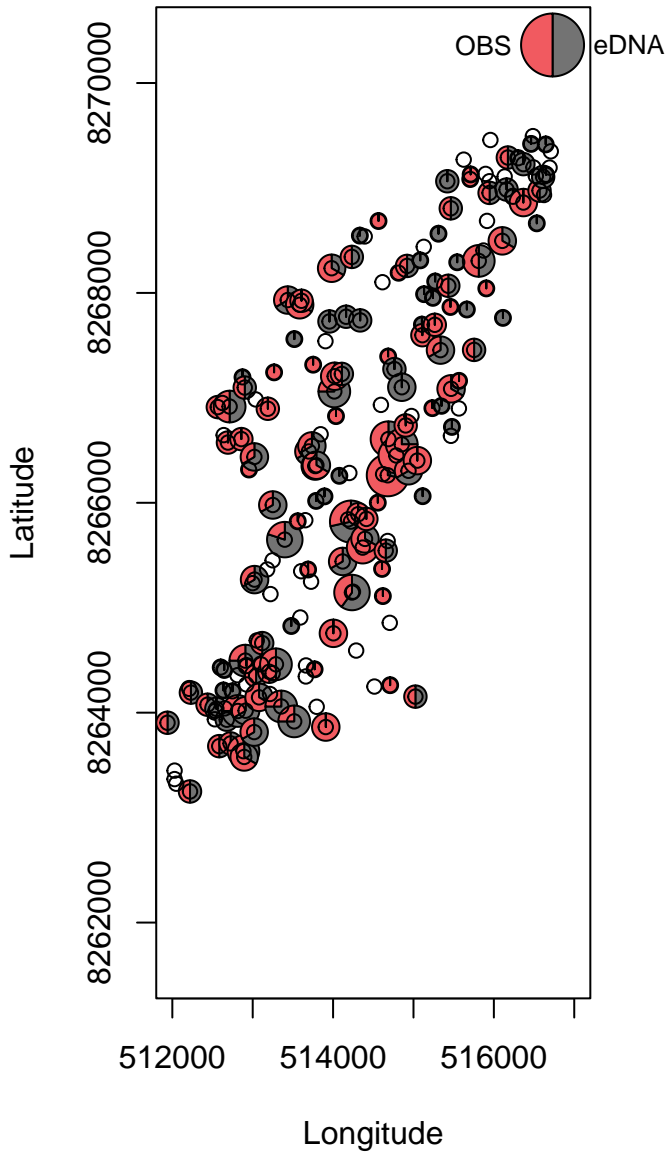
Carex



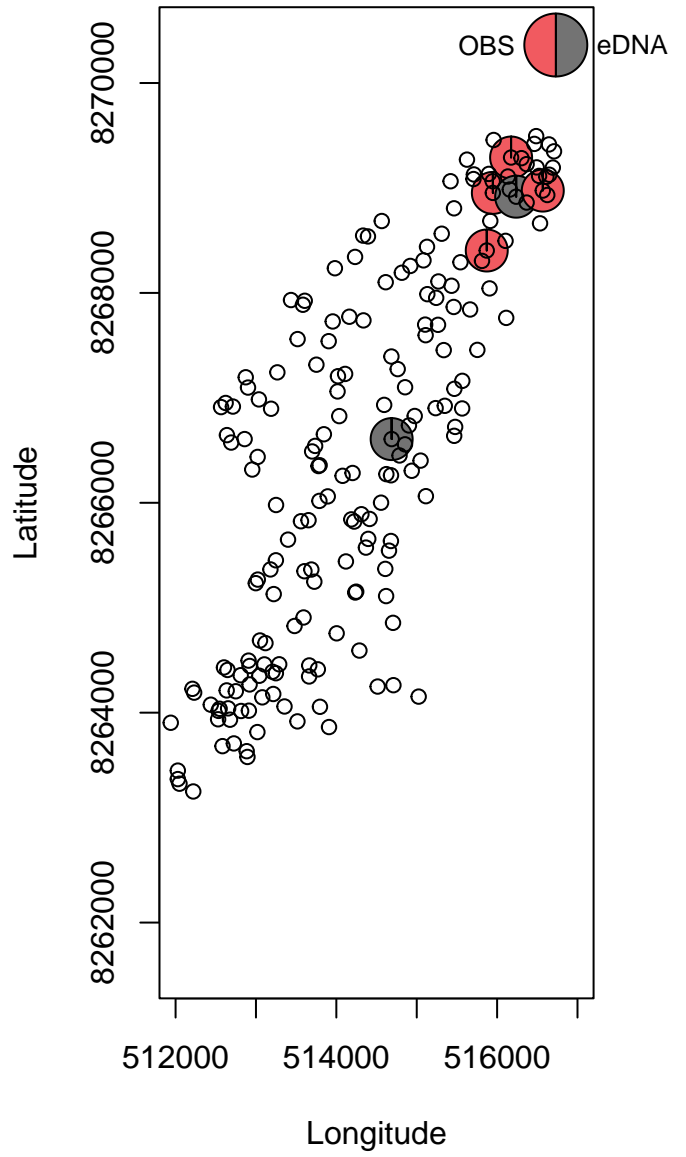
Carex_glaucosa



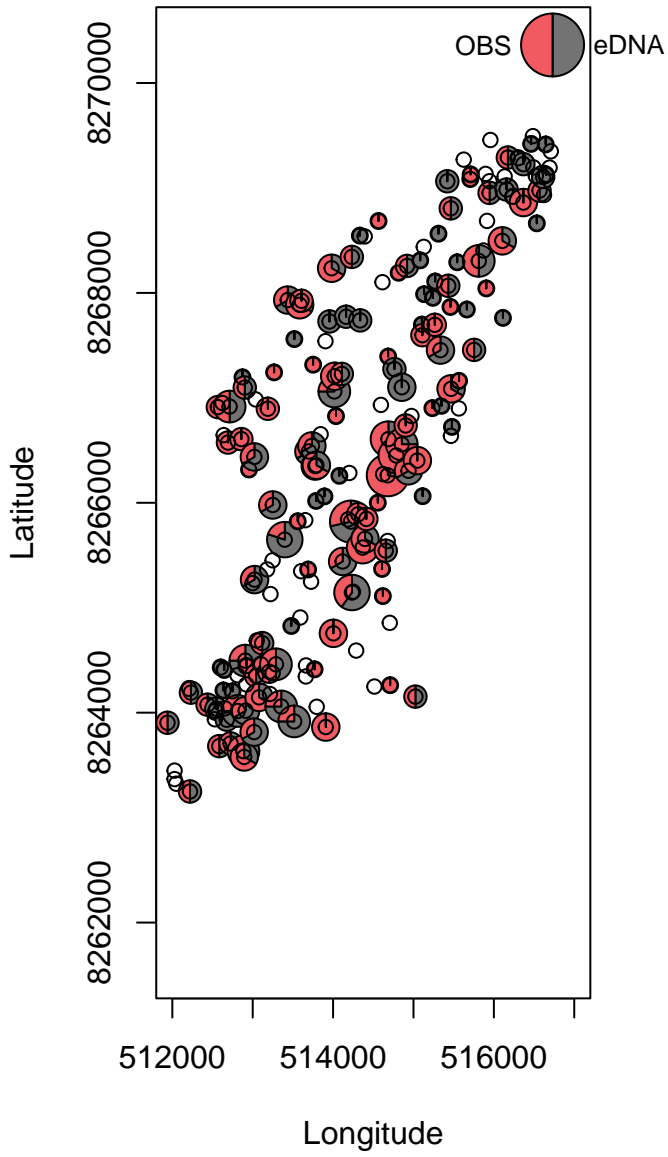
Carex



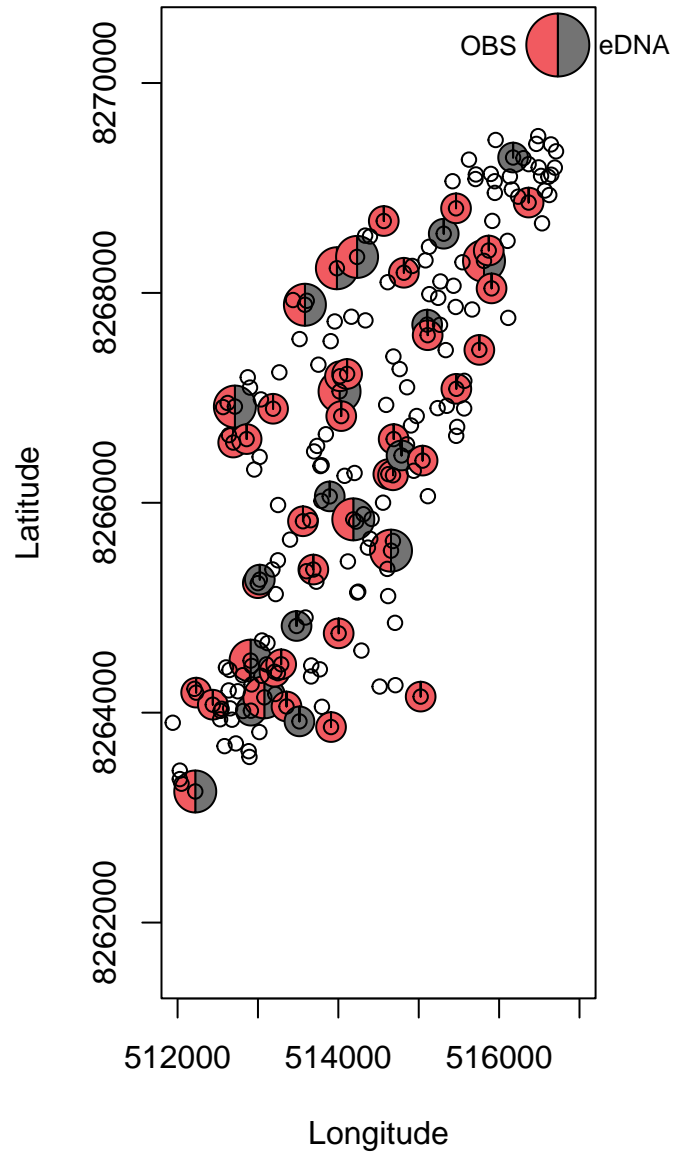
Carex_maritima



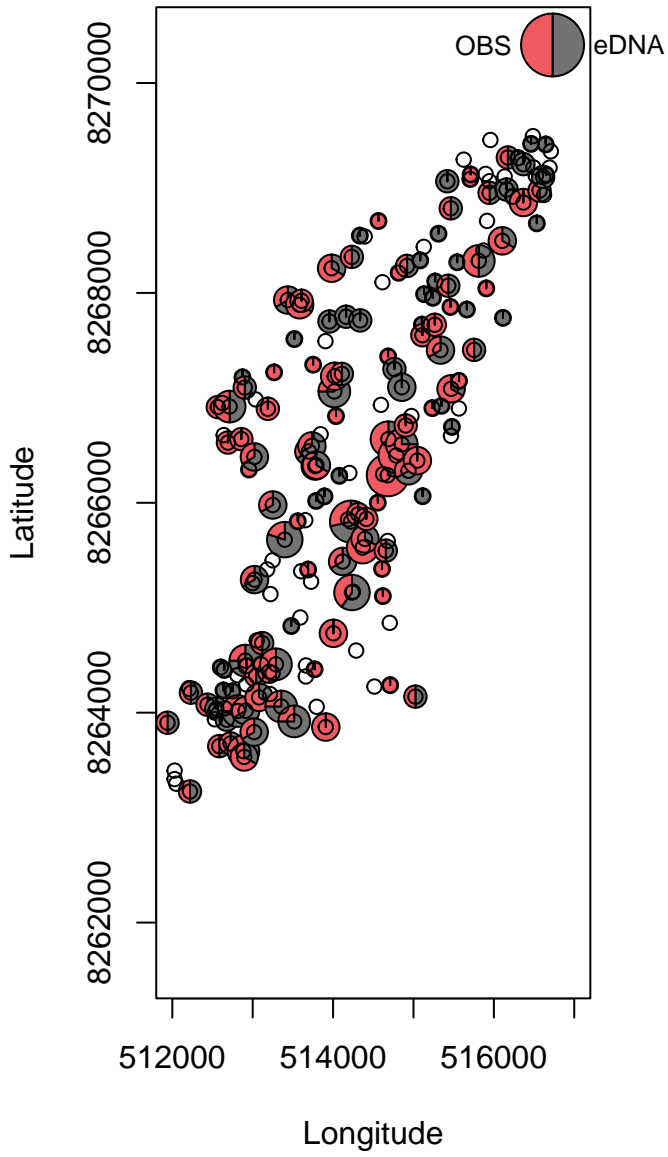
Carex



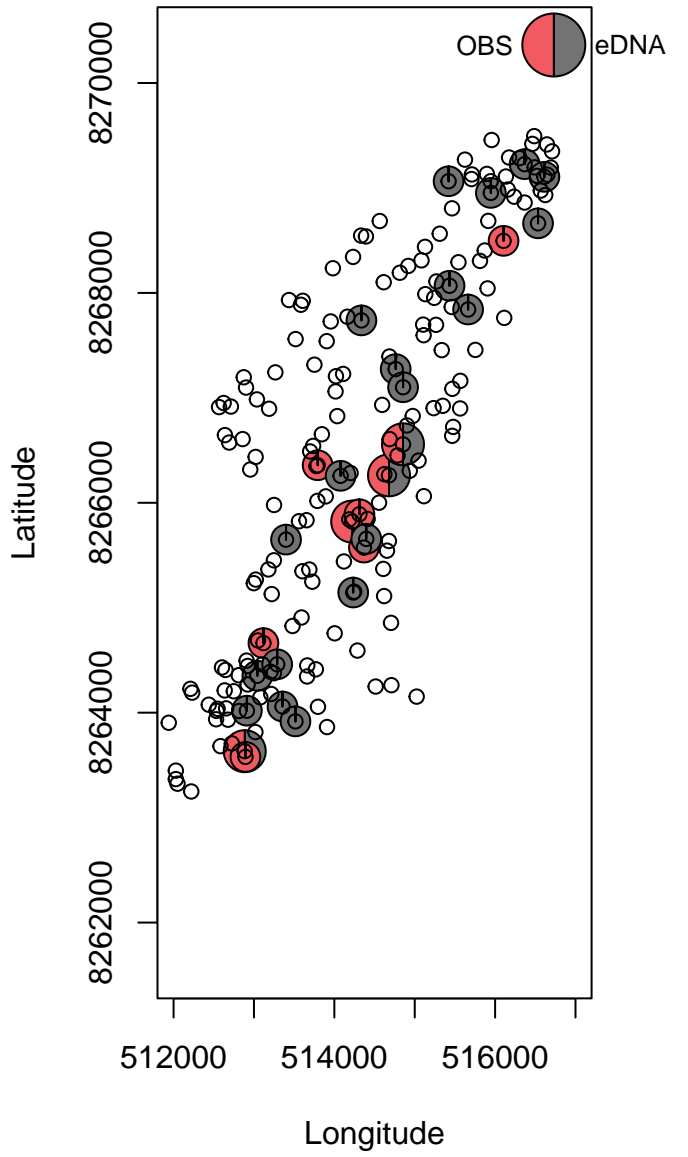
Carex_rupestris



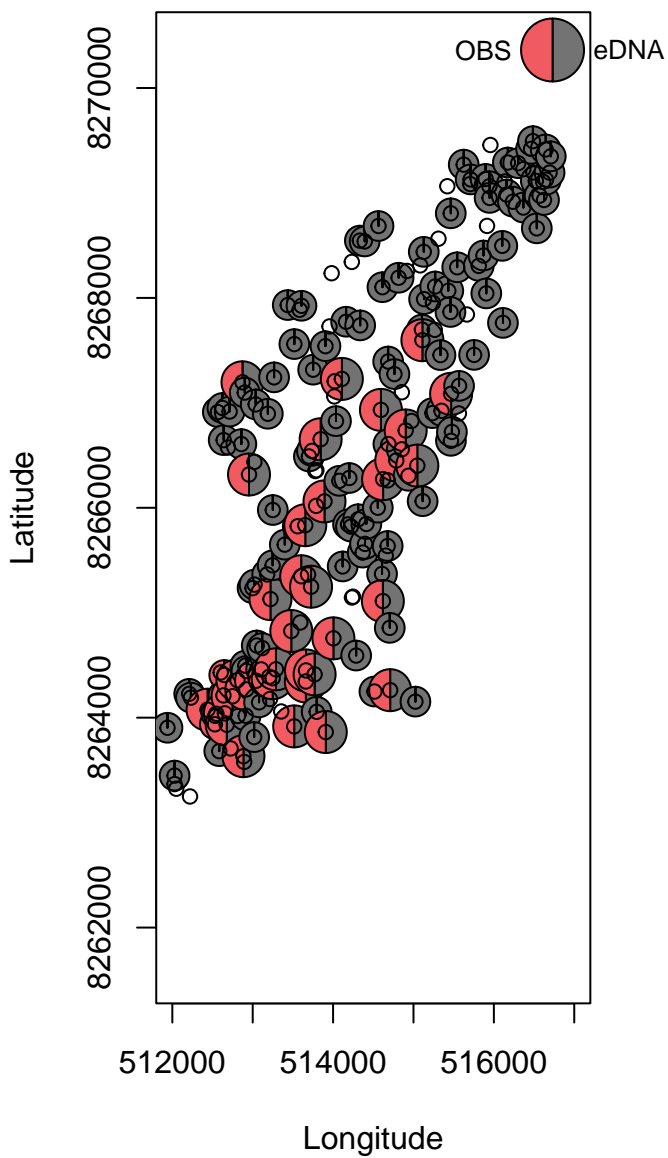
Carex



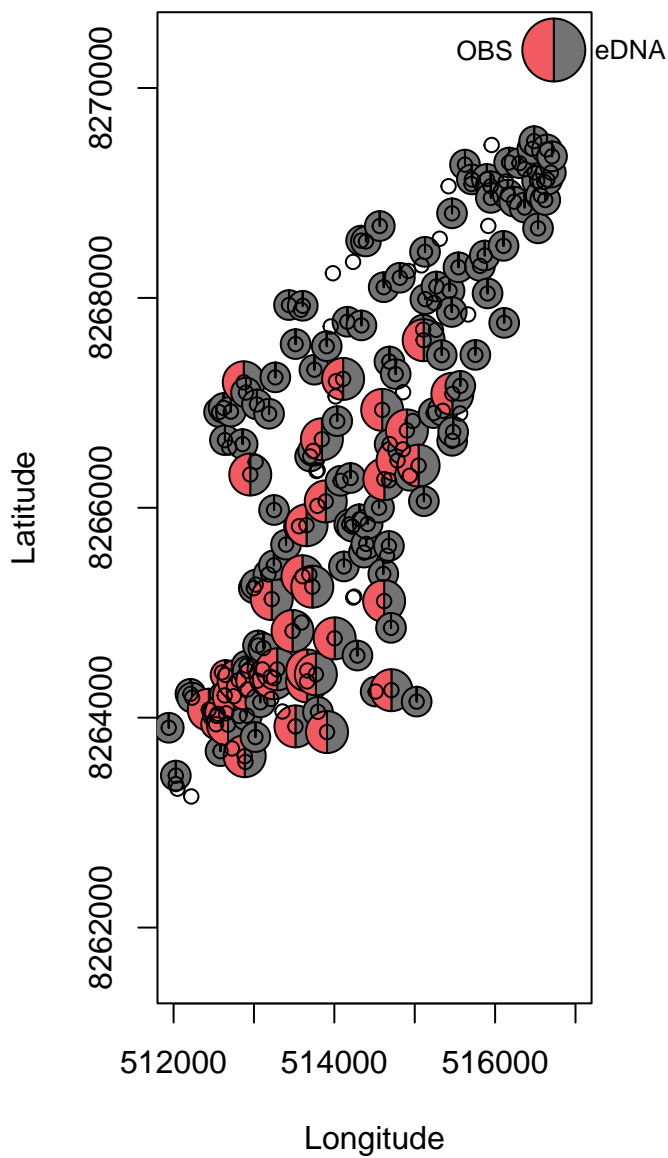
Carex_saxatilis



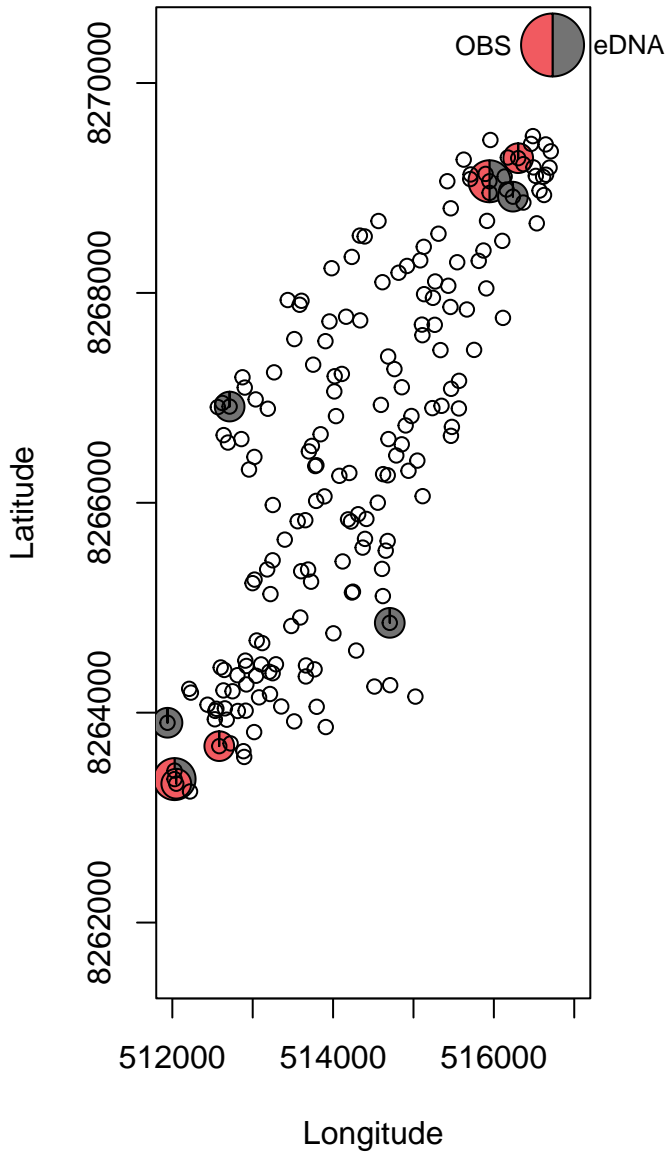
Cassiope



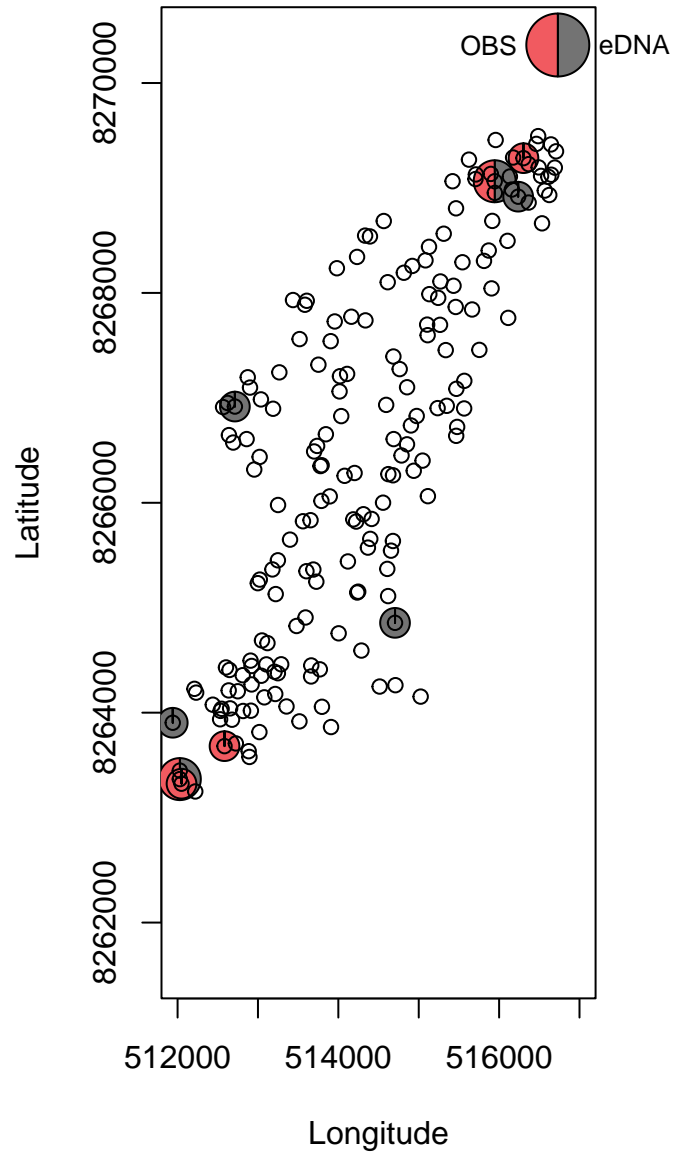
Cassiope_tetragona



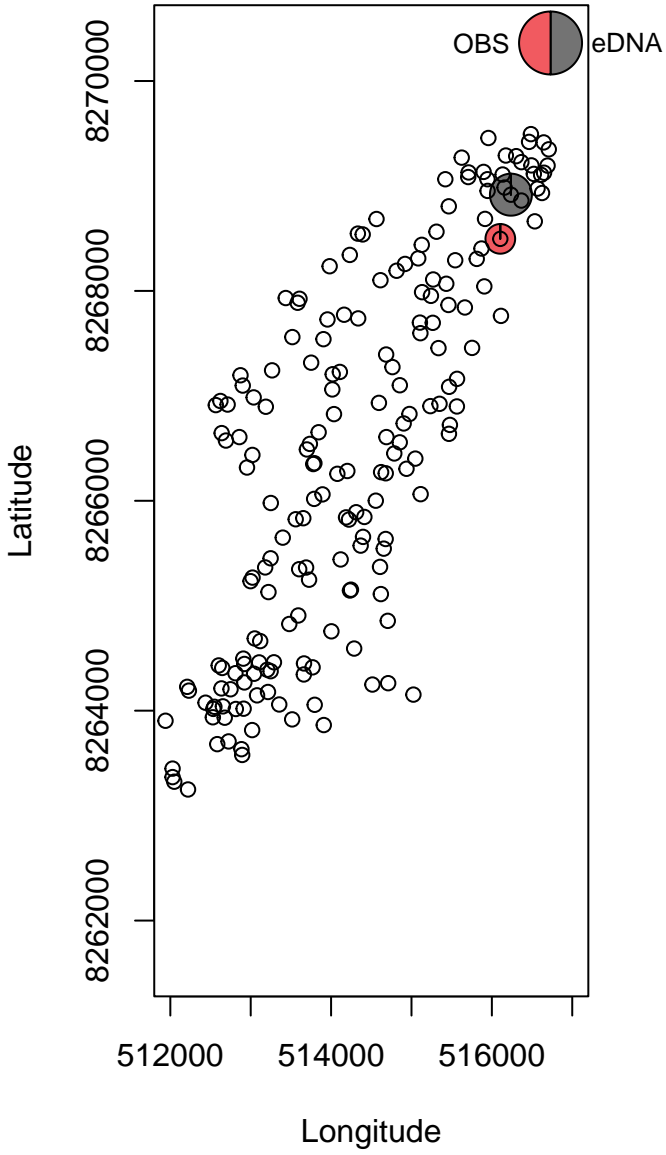
Chamerion



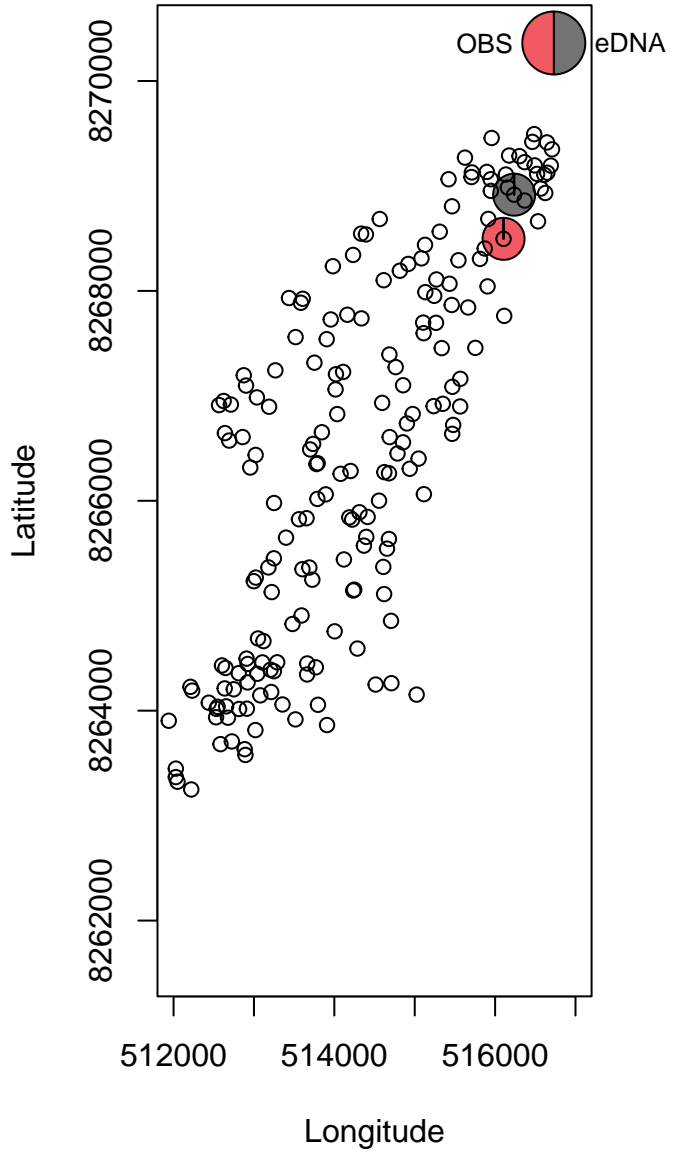
Chamerion_latifolium



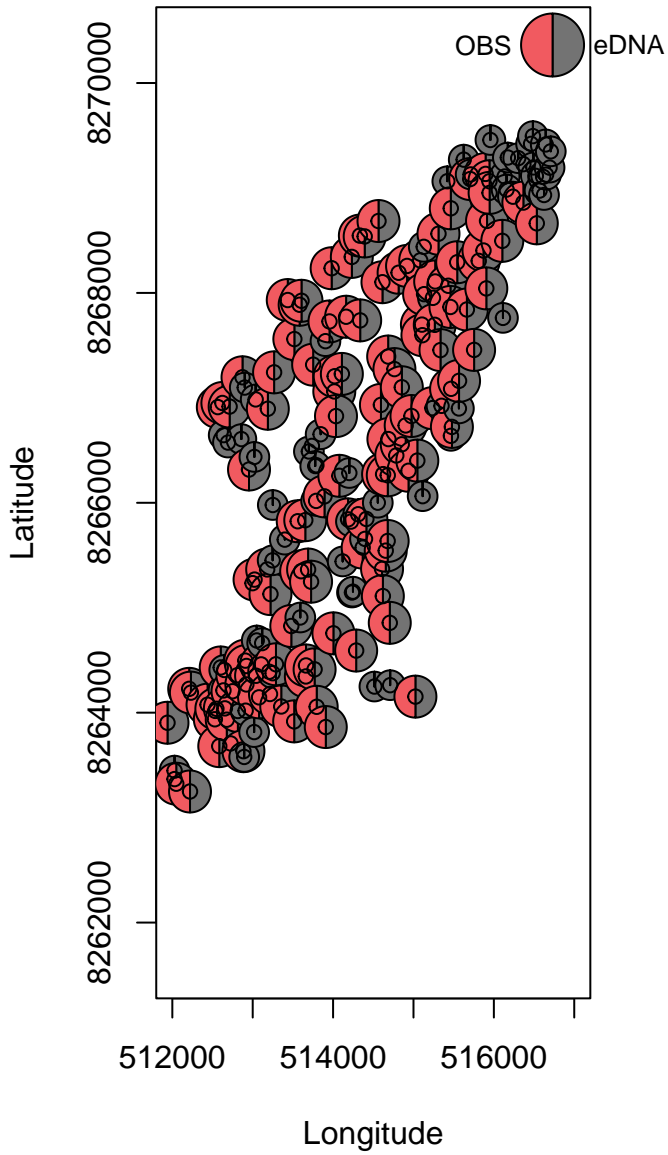
Deschampsia



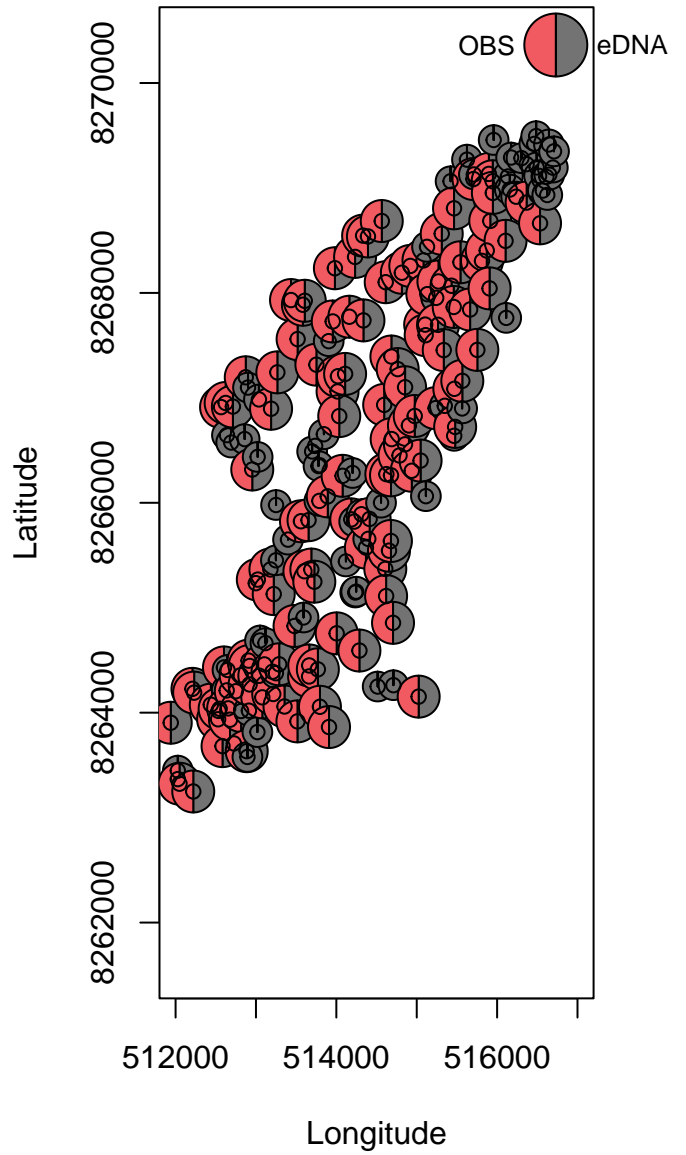
Deschampsia_brevifolia



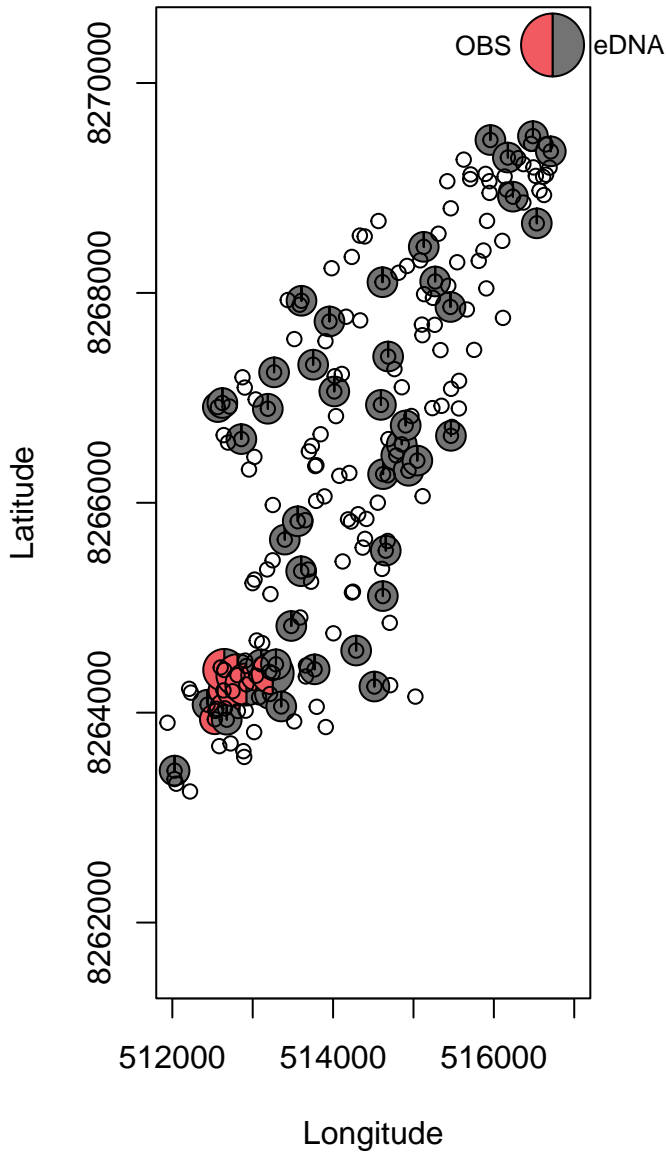
Dryas



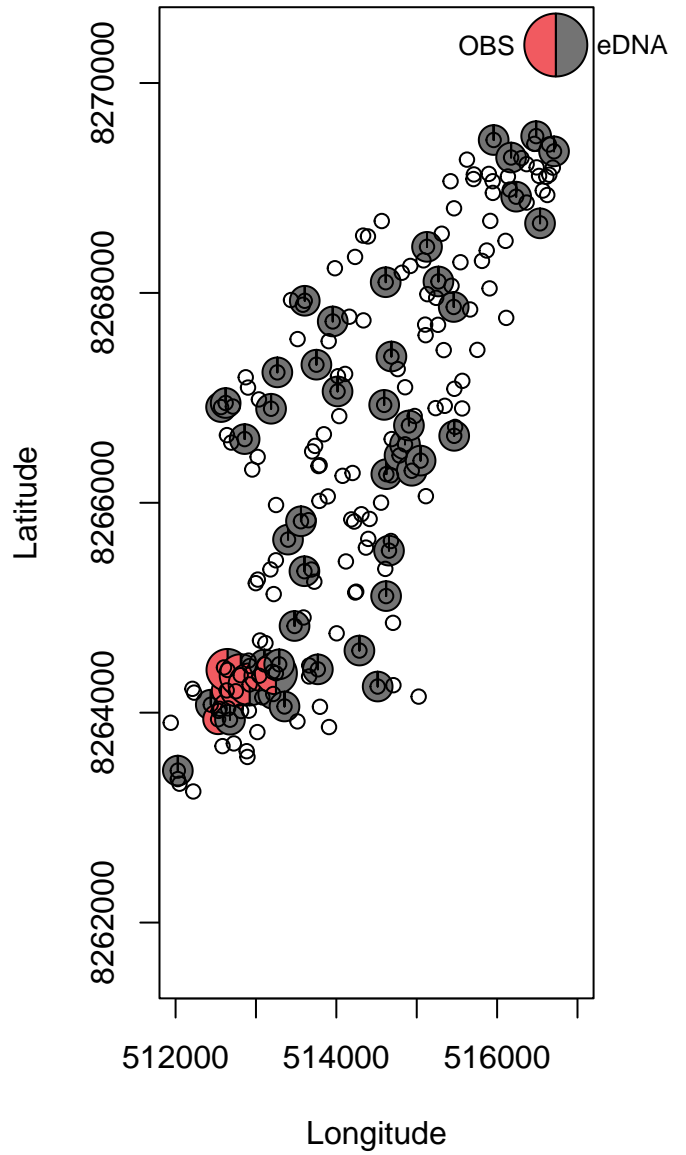
Dryas_octopetala



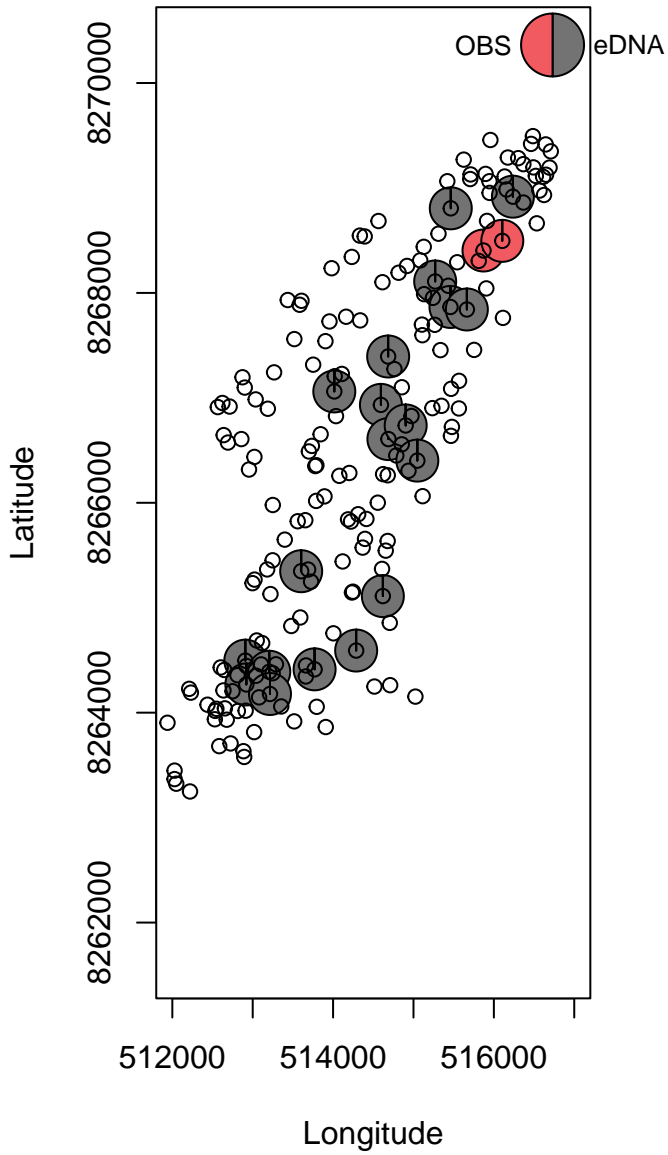
Empetrum



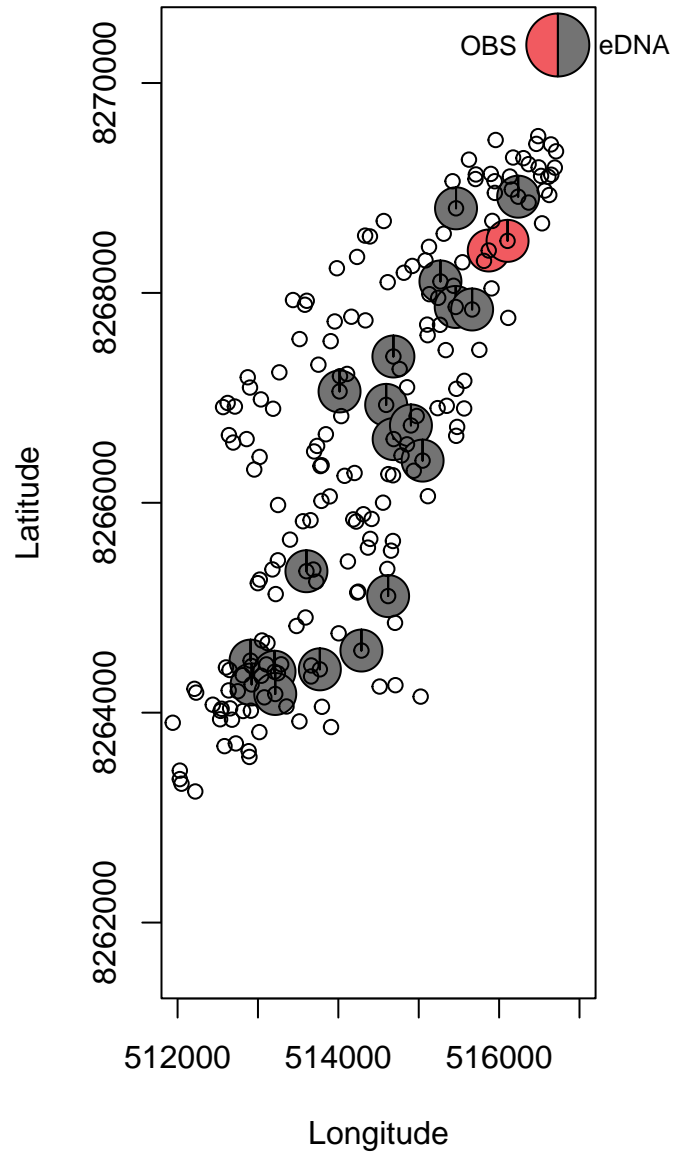
Empetrum_nigrum



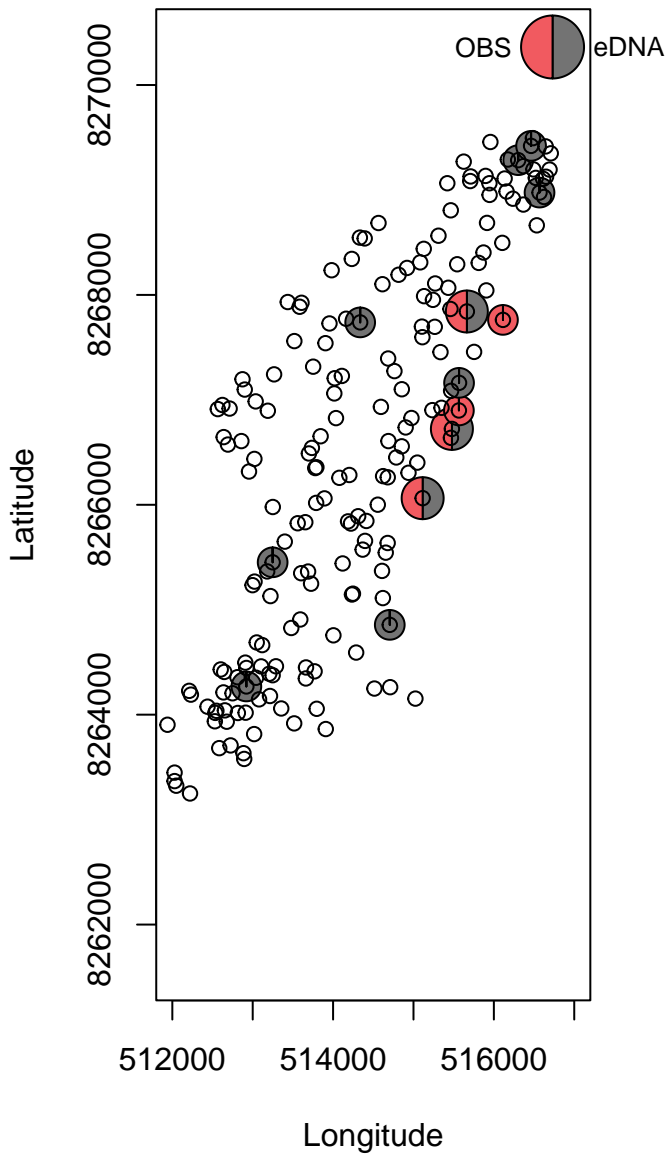
Epilobium



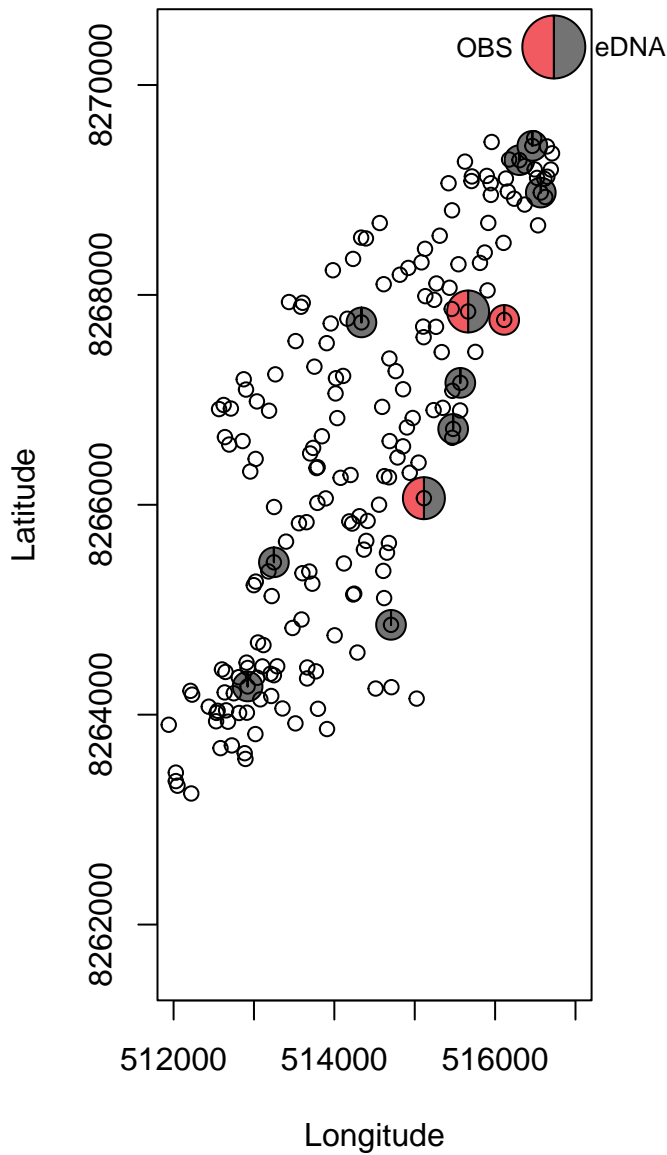
Epilobium_arcticum



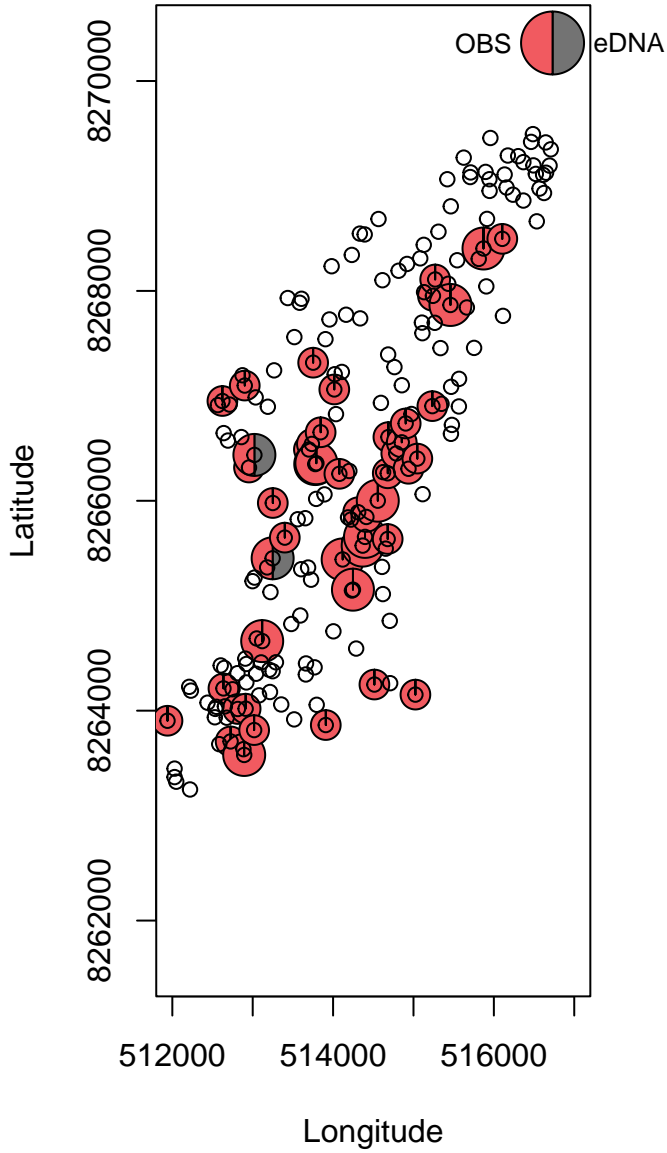
Erigeron



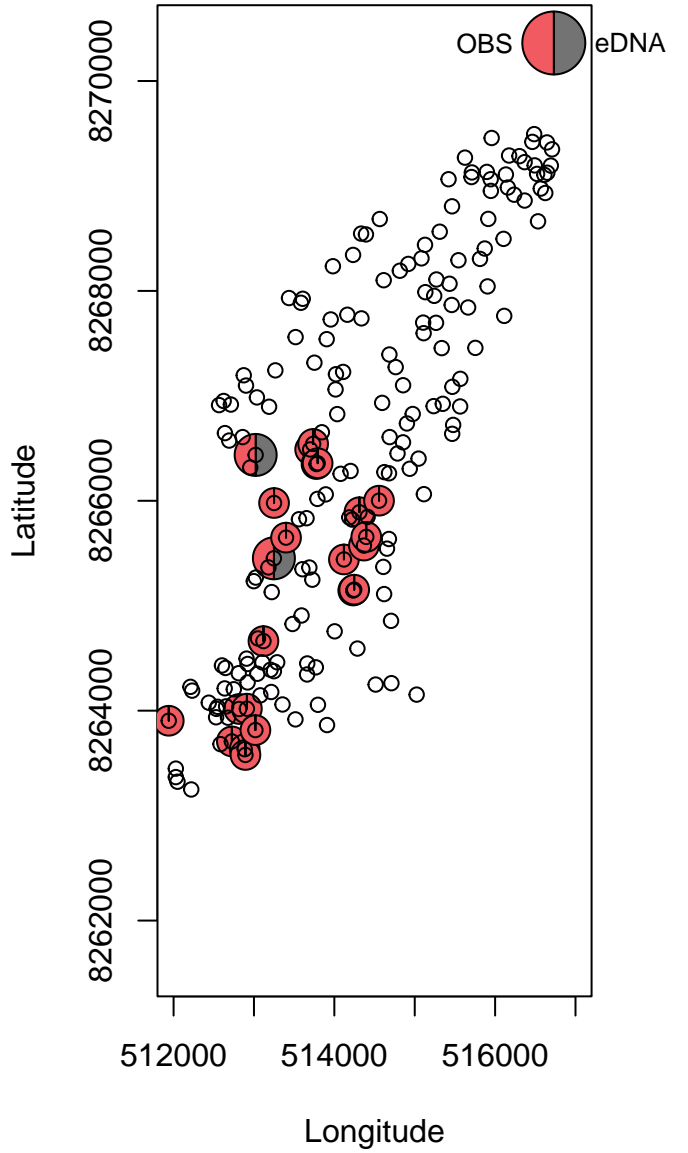
Erigeron_humilis



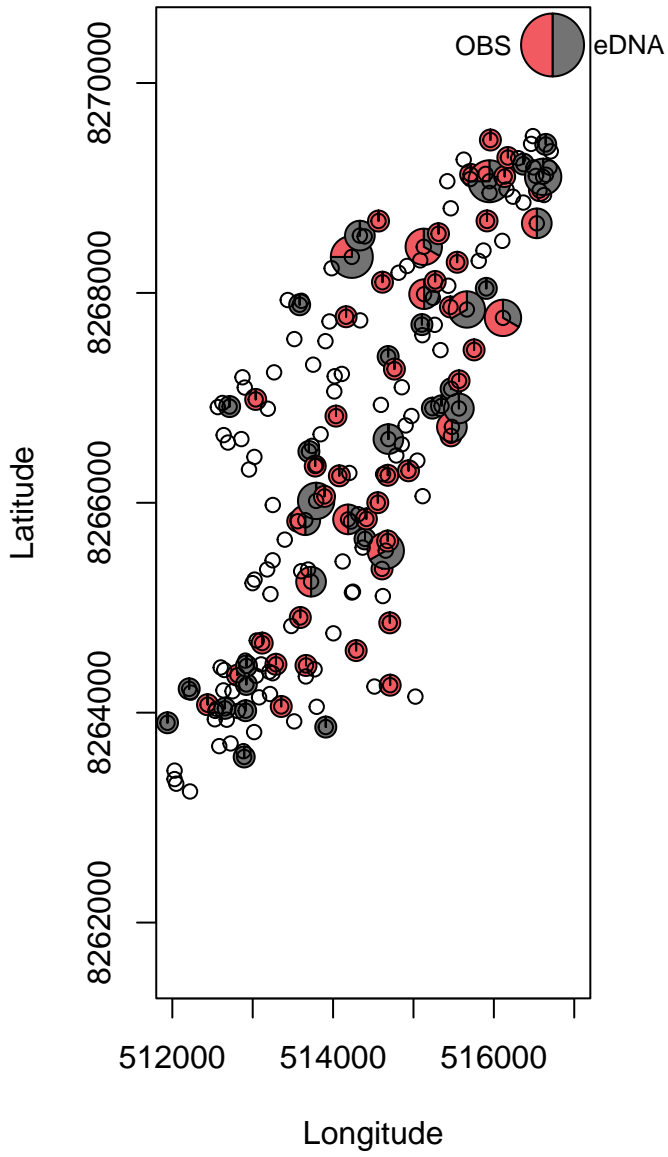
Eriophorum



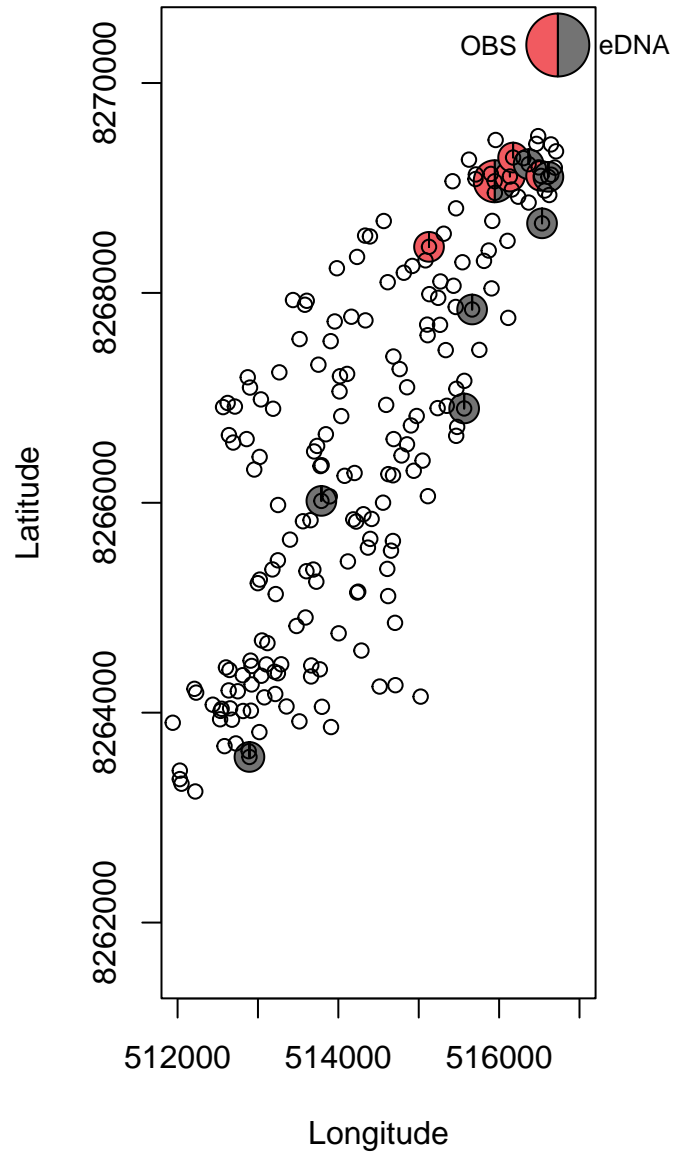
Eriophorum_scheuchzeri



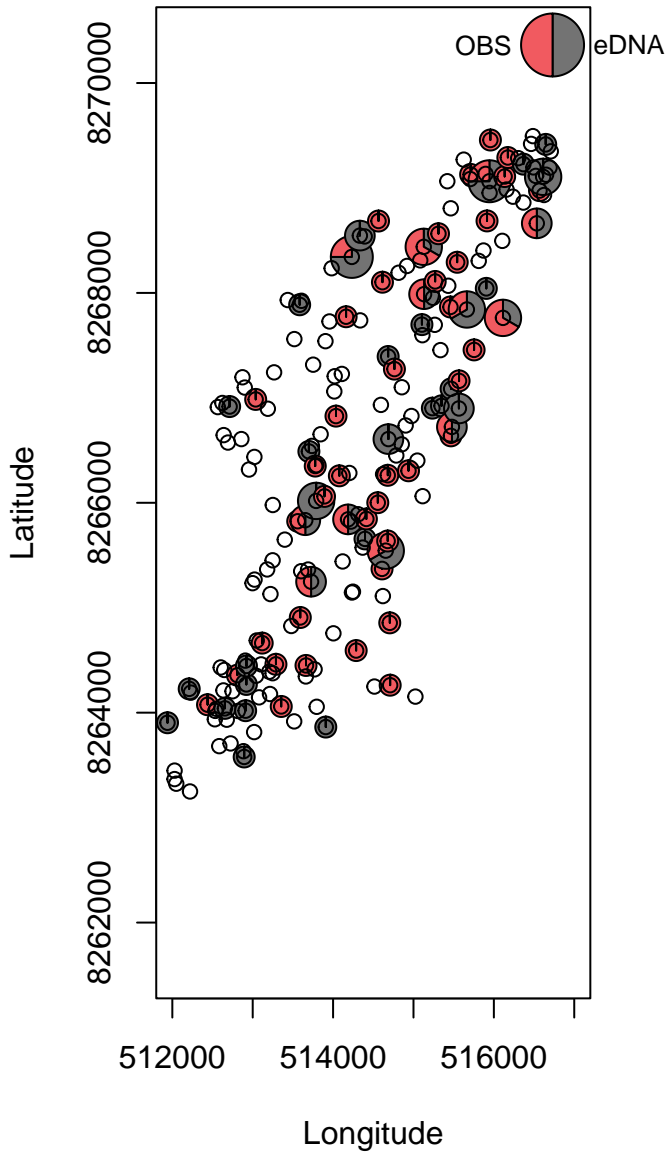
Festuca



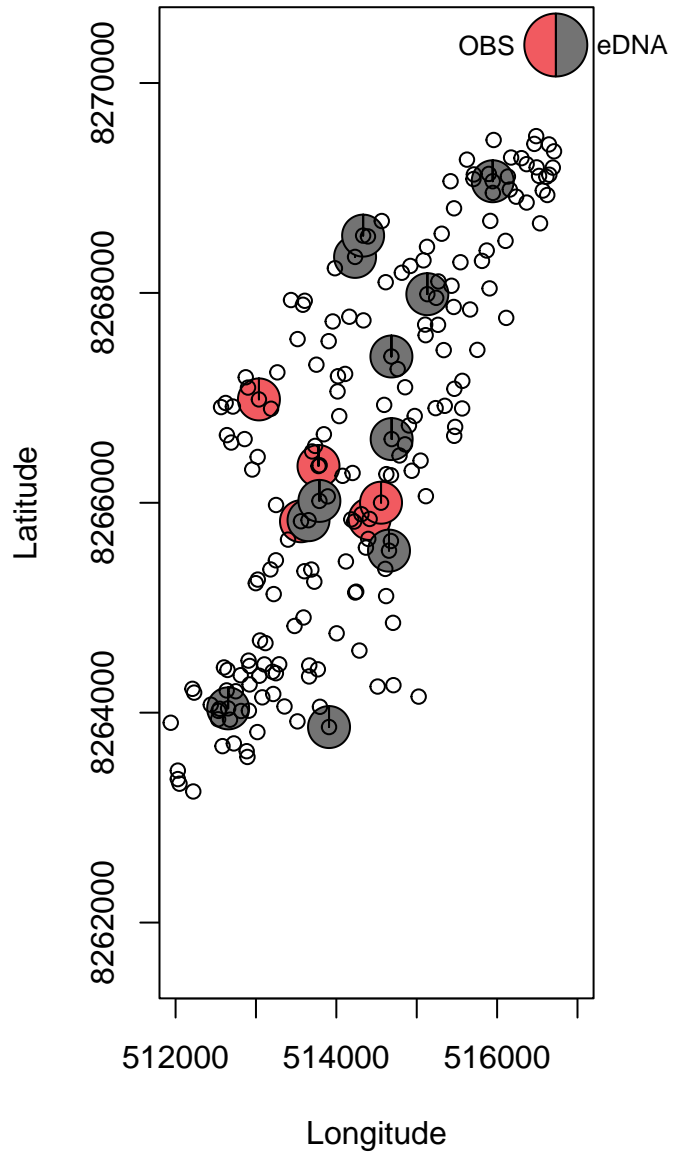
Festuca_baffinensis



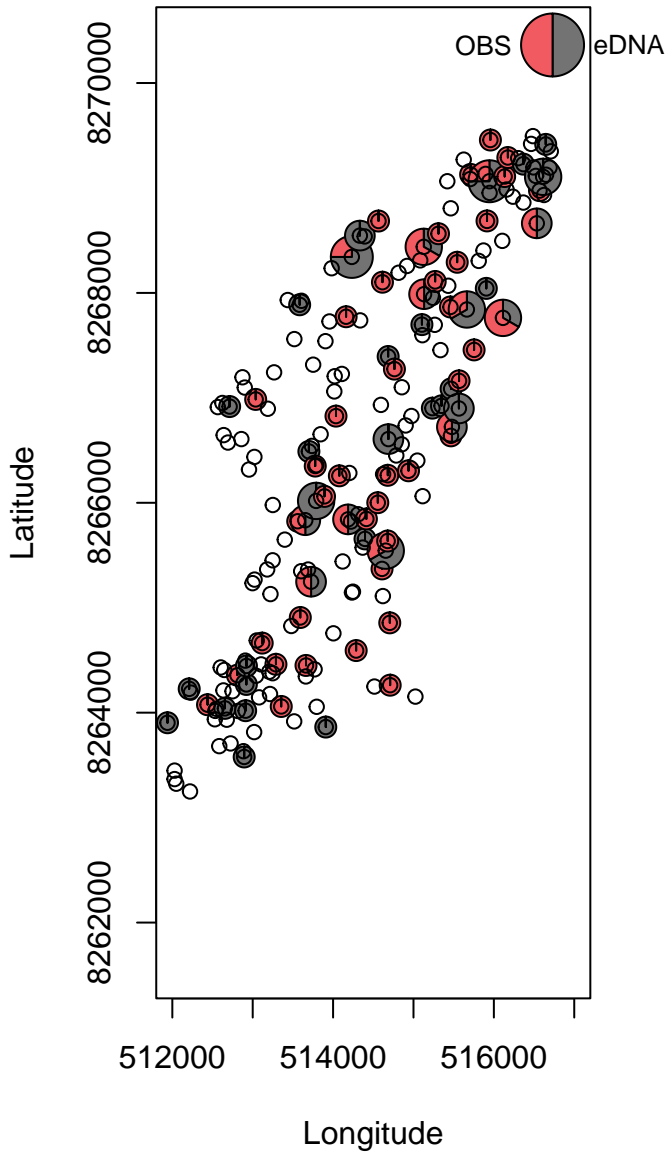
Festuca



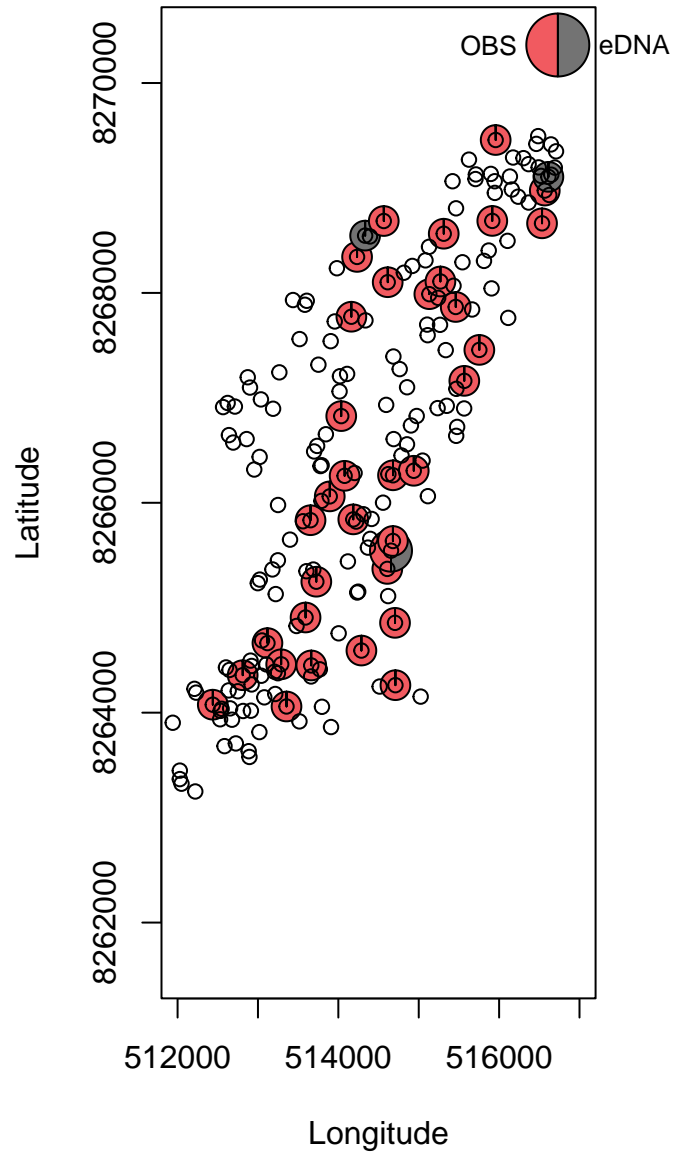
Festuca_brachyphylla



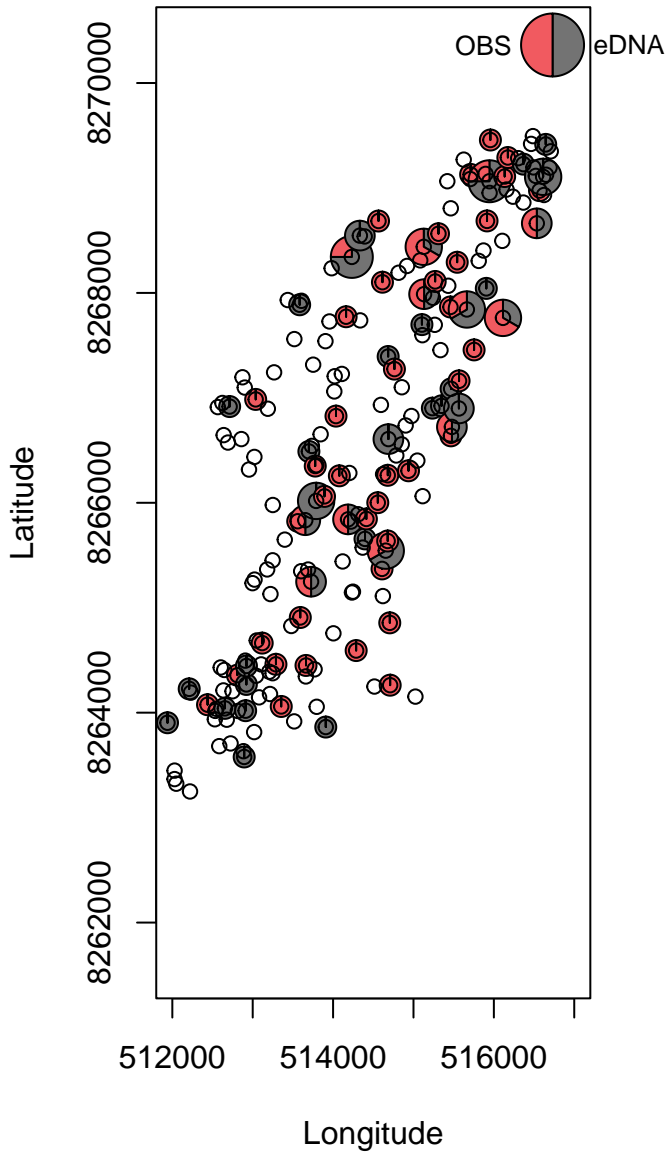
Festuca



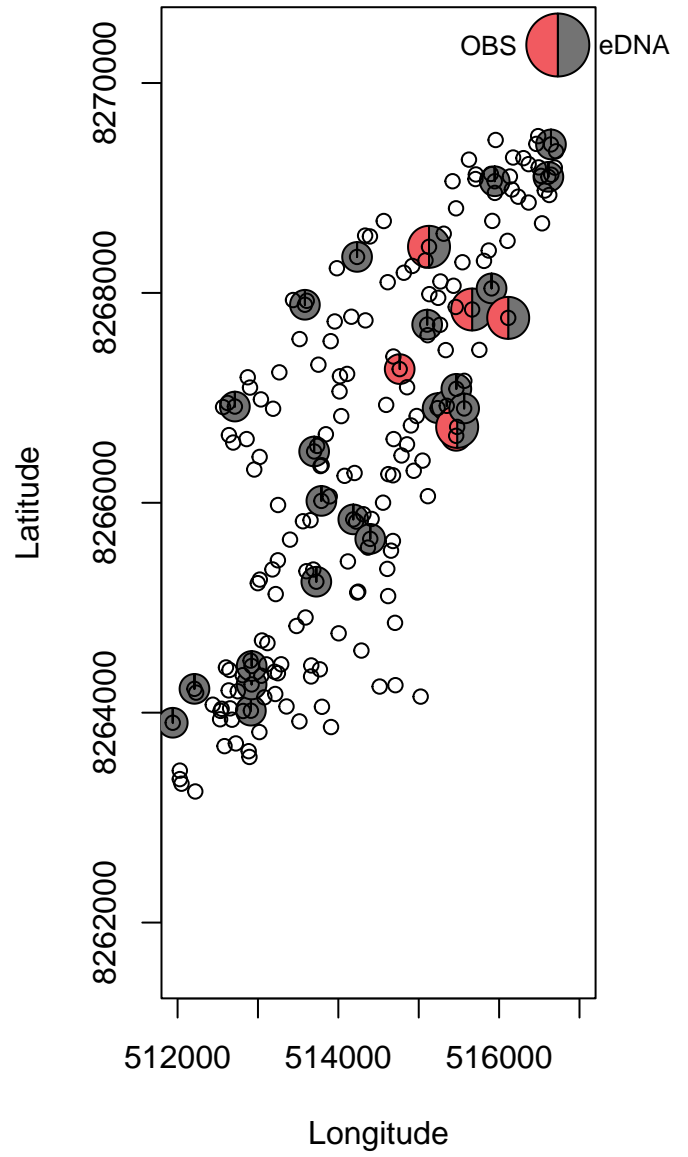
Festuca_hyperborea



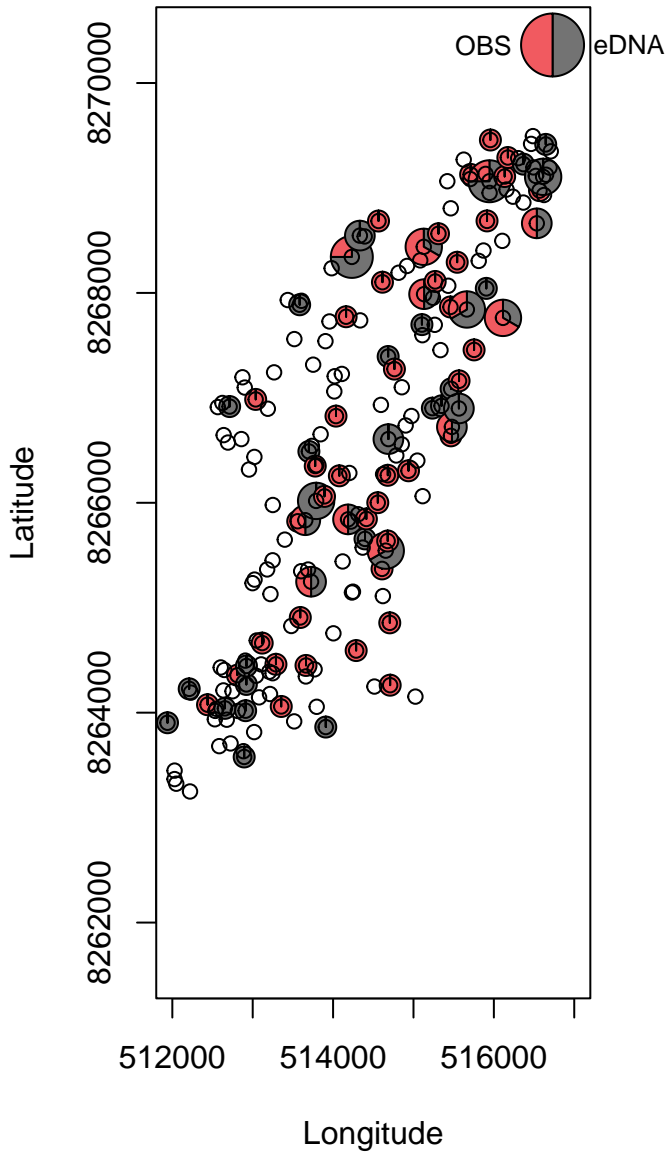
Festuca



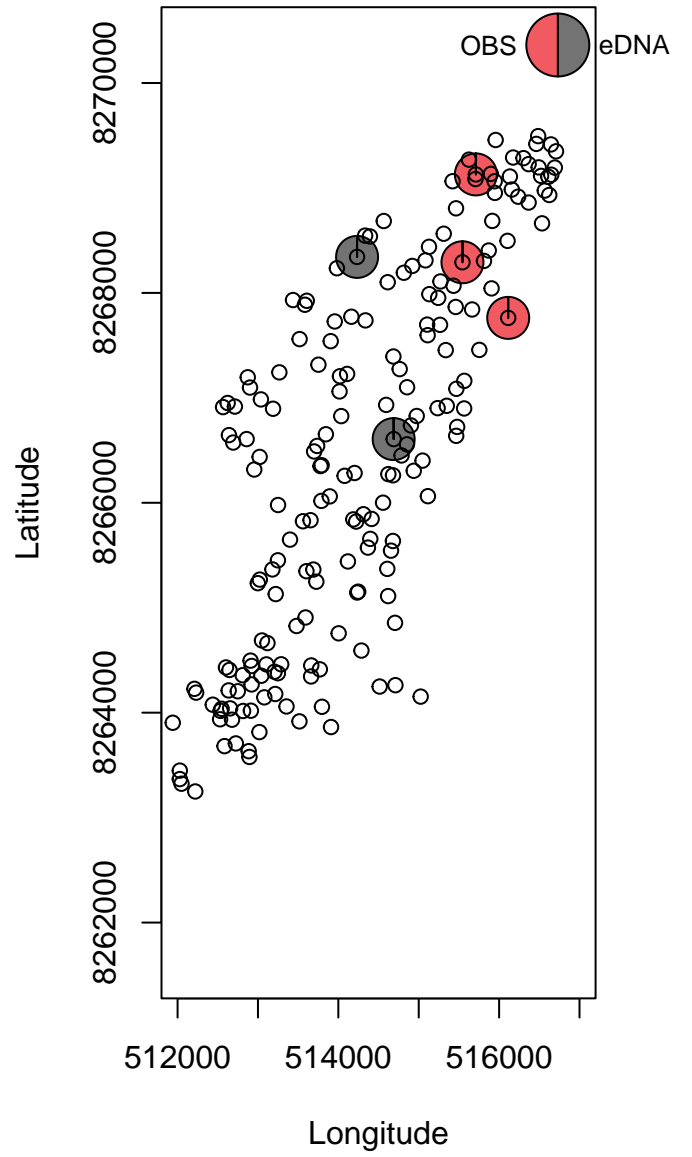
Festuca_rubra



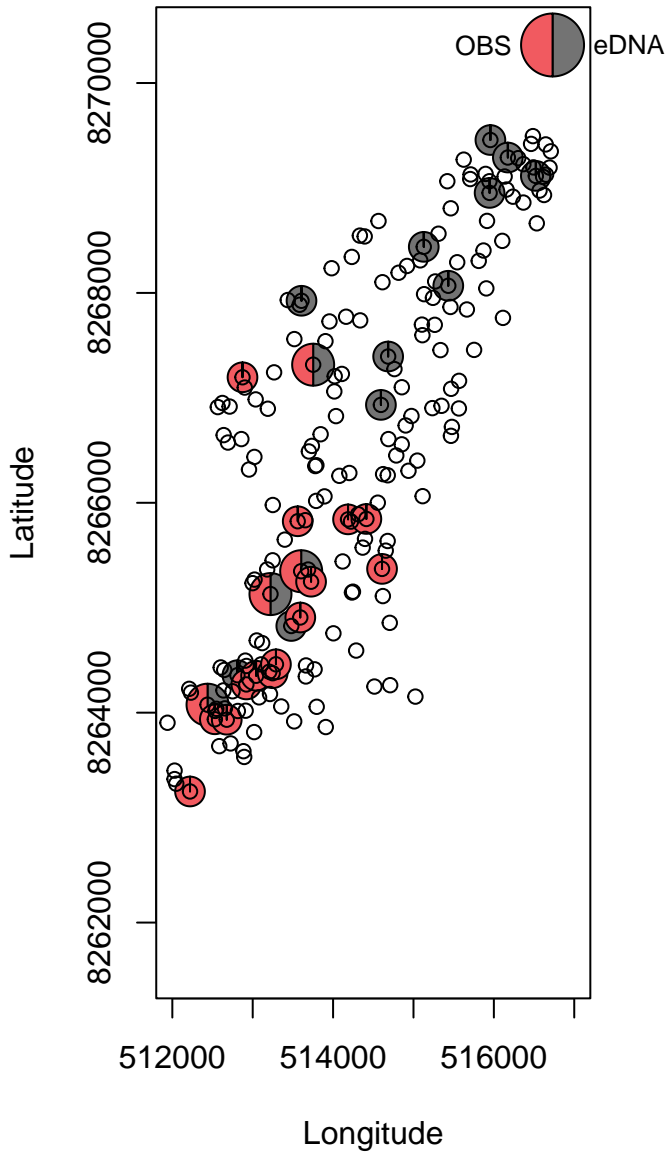
Festuca



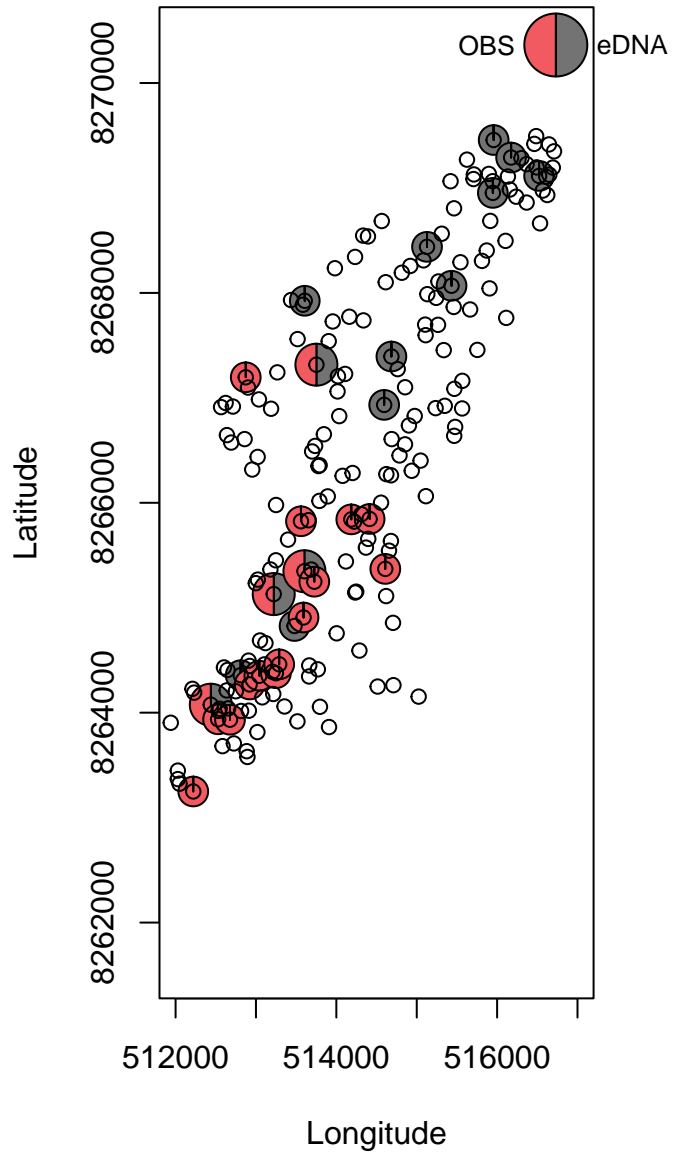
Festuca_vivipara



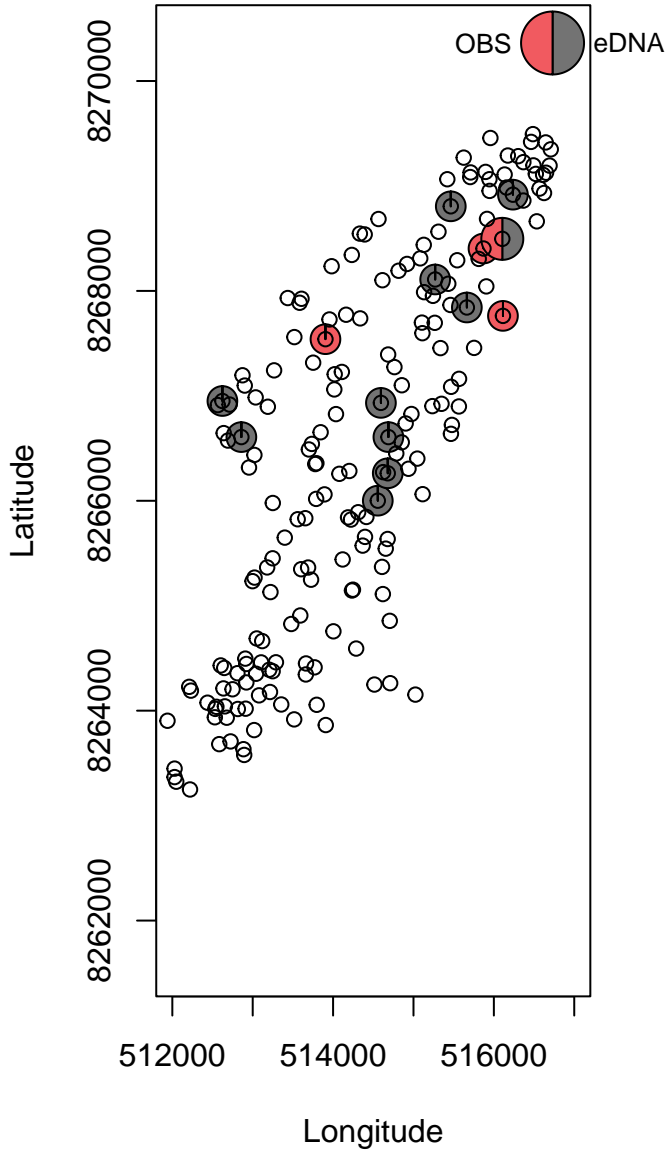
Hierochloe



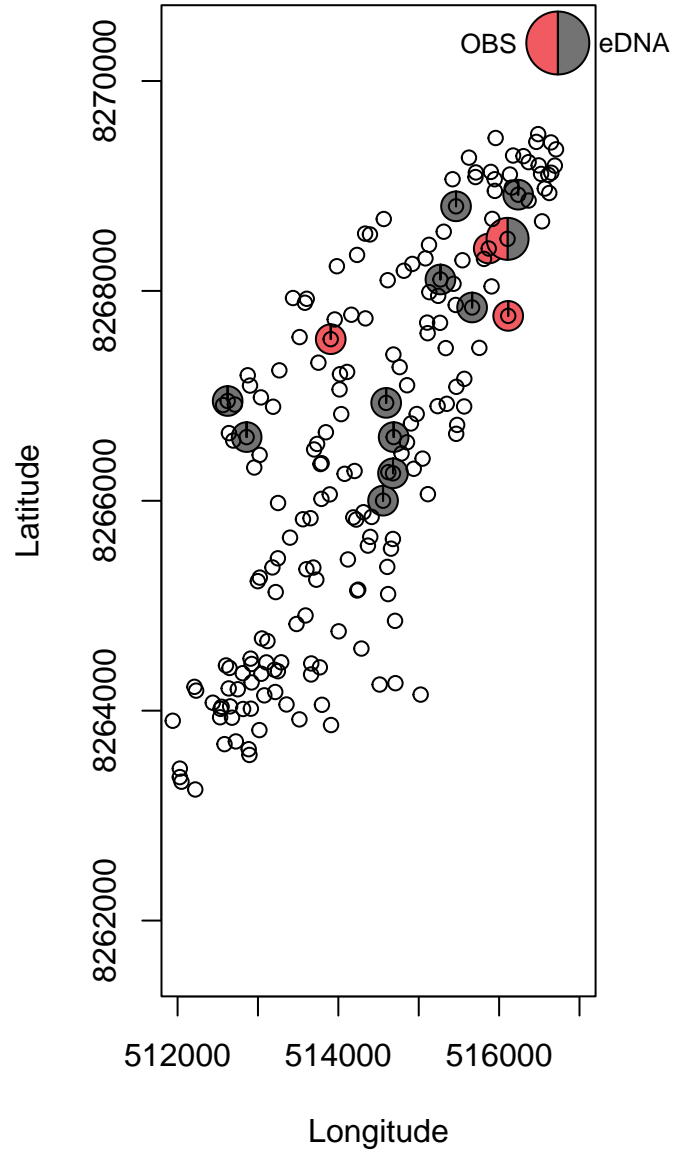
Hierochloe_alpina



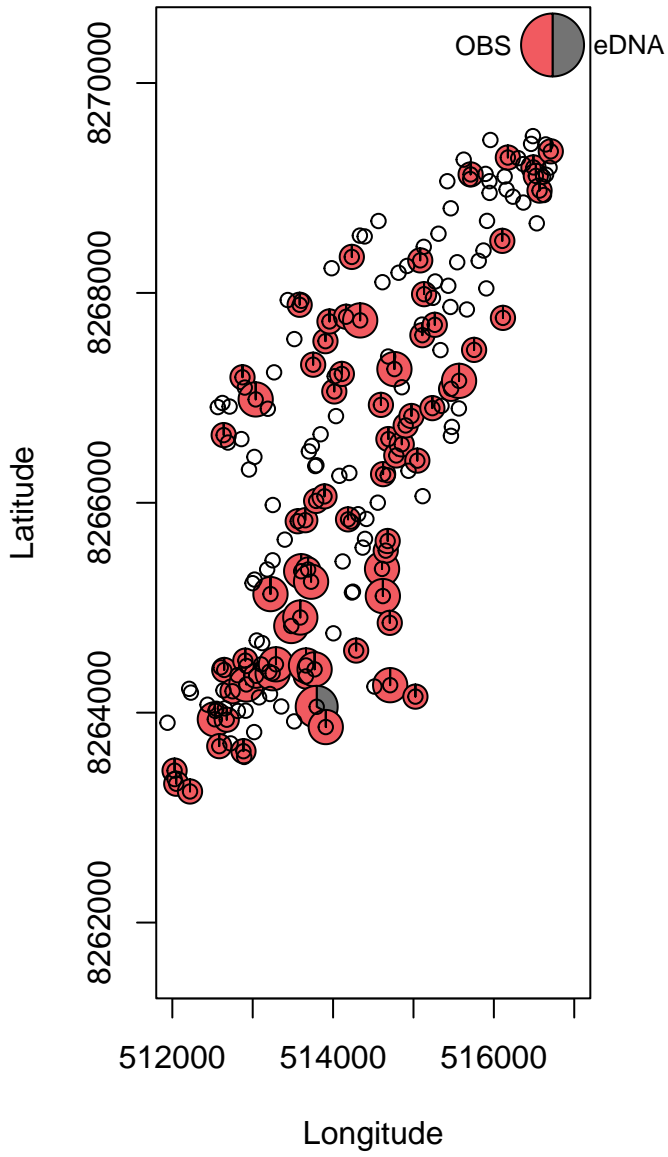
Koenigia



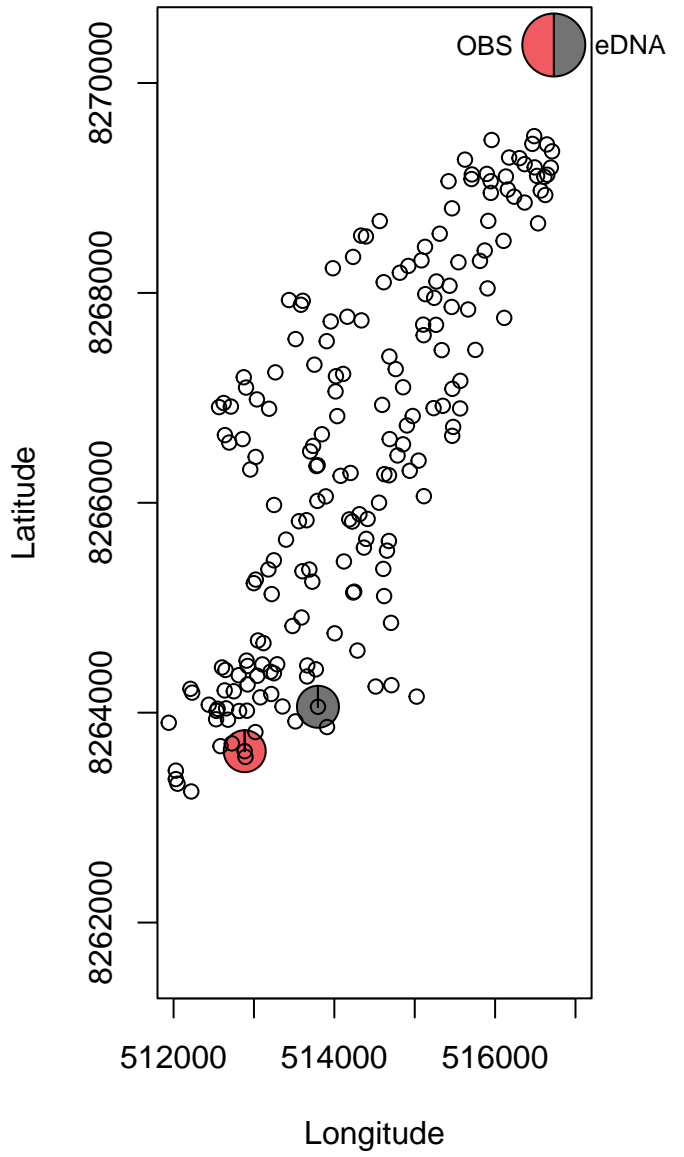
Koenigia_islandica



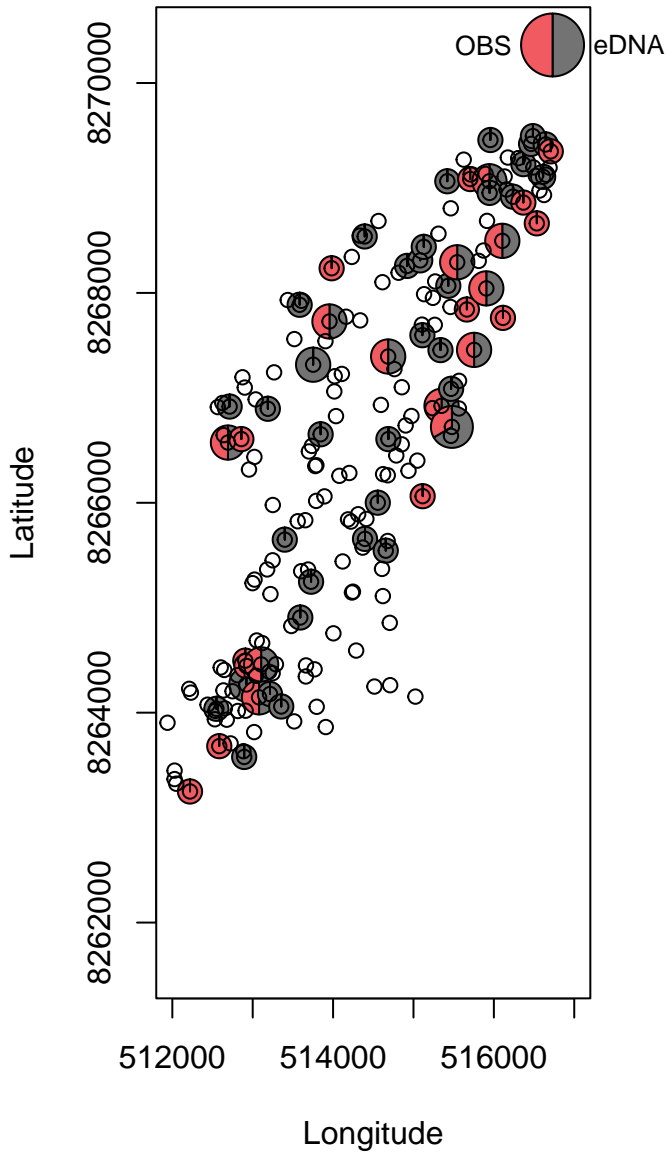
Luzula



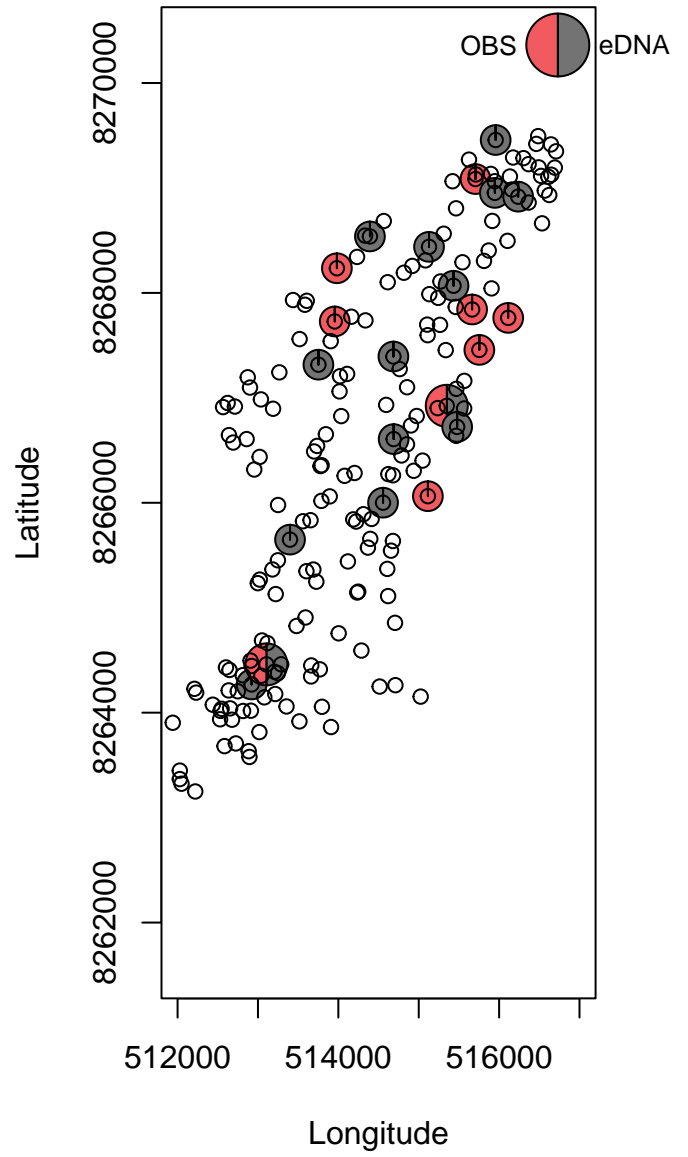
Luzula_wahlenbergii



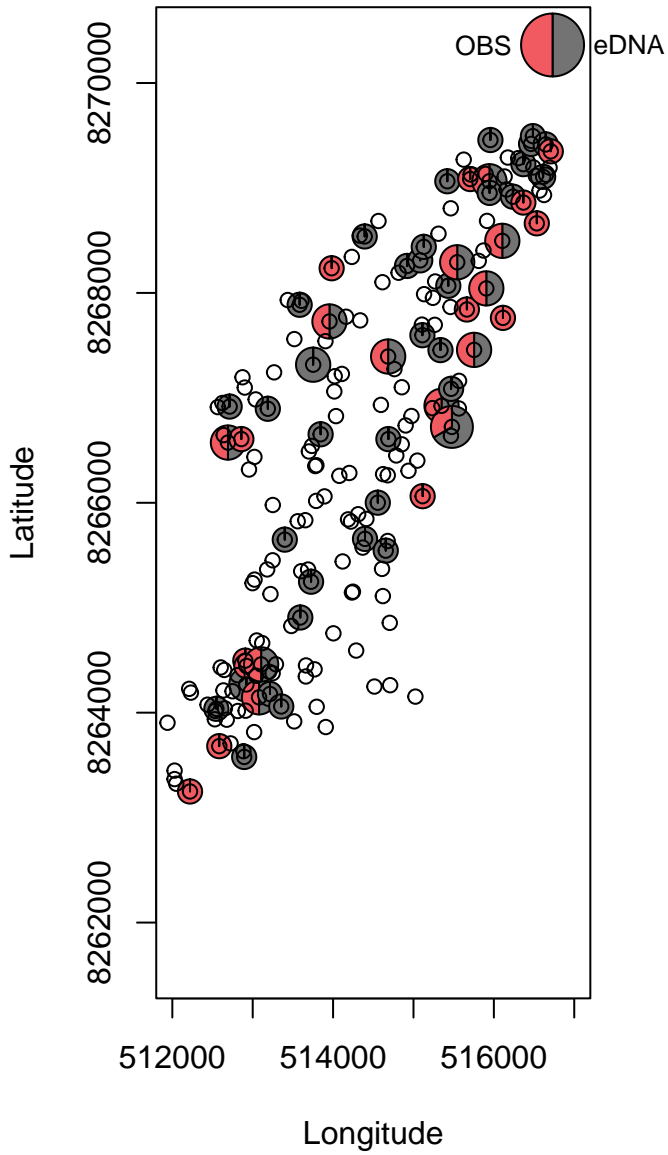
Minuartia



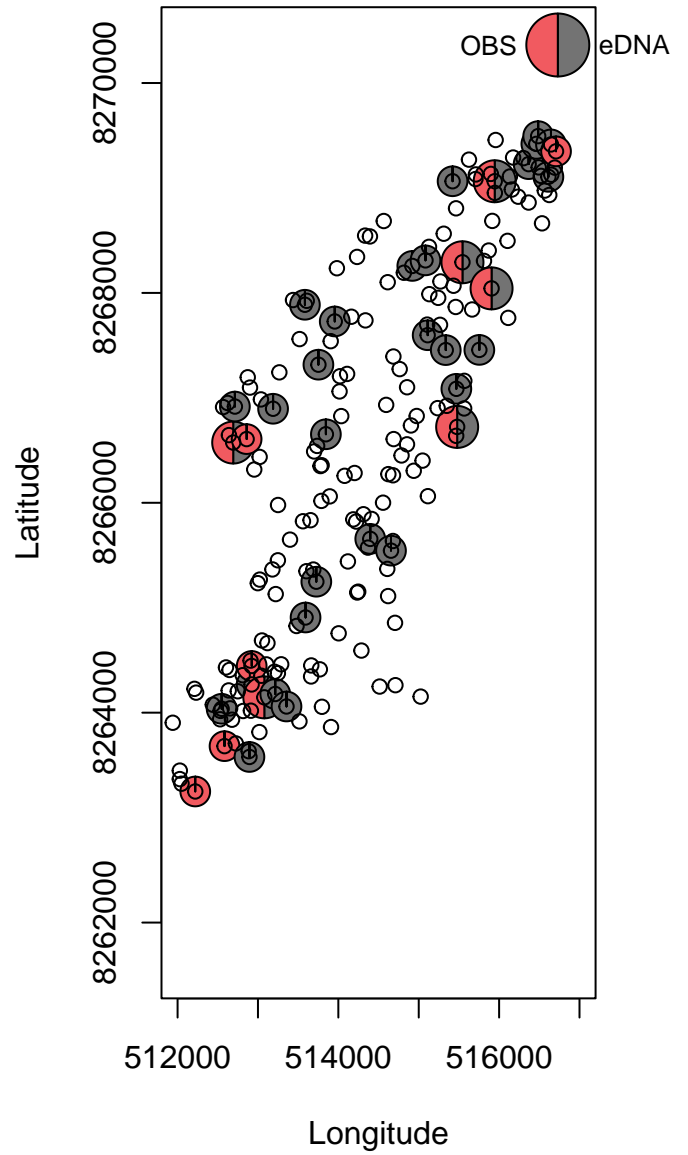
Minuartia_biflora



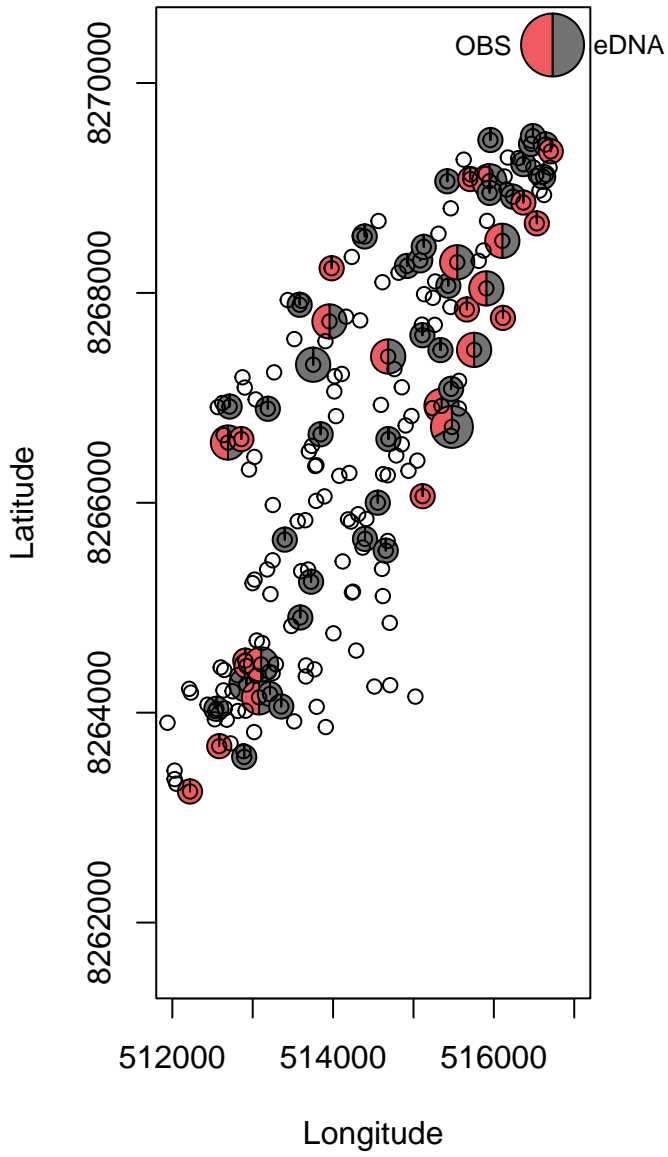
Minuartia



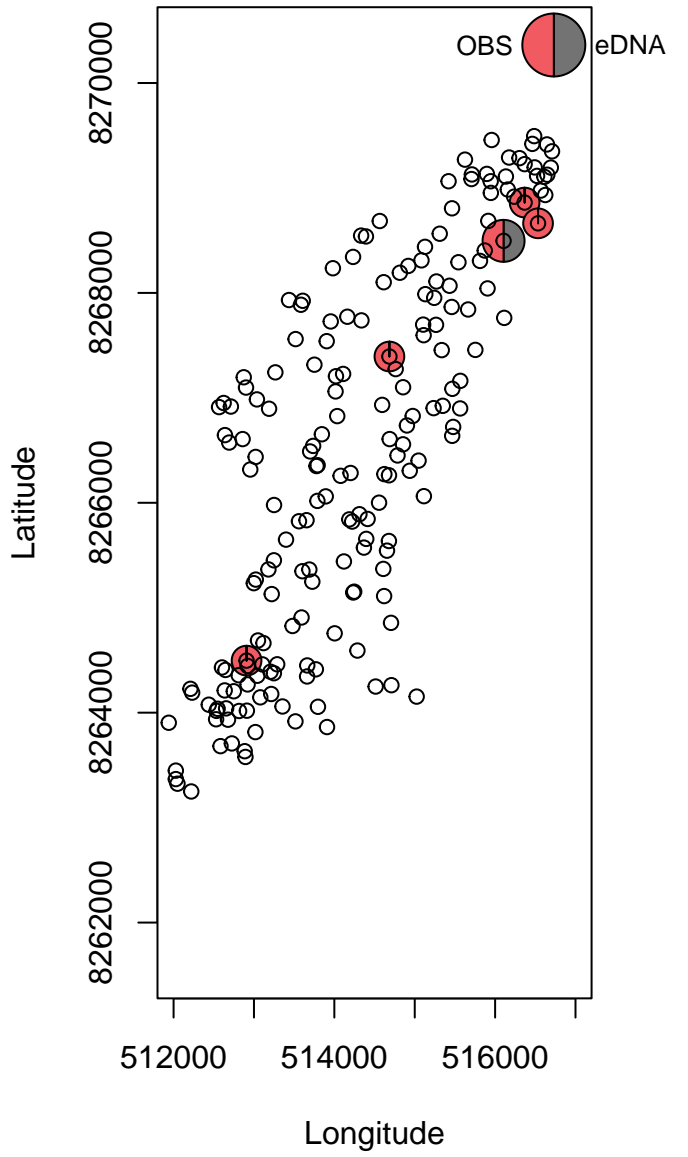
Minuartia_rubella



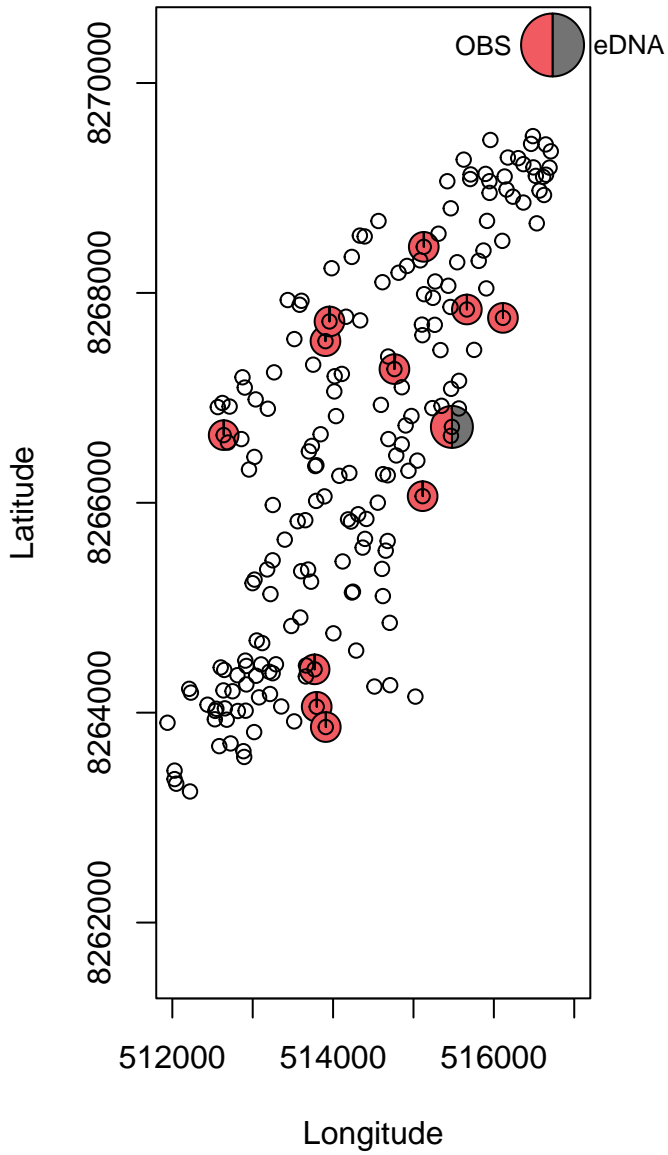
Minuartia



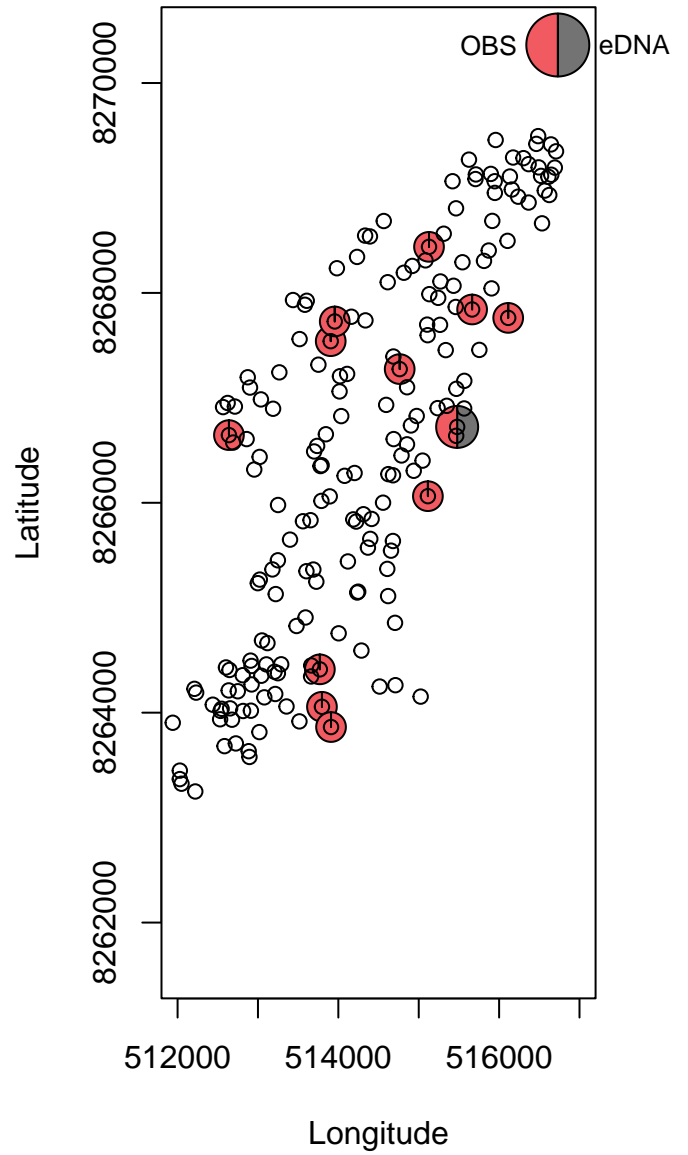
Minuartia_stricta



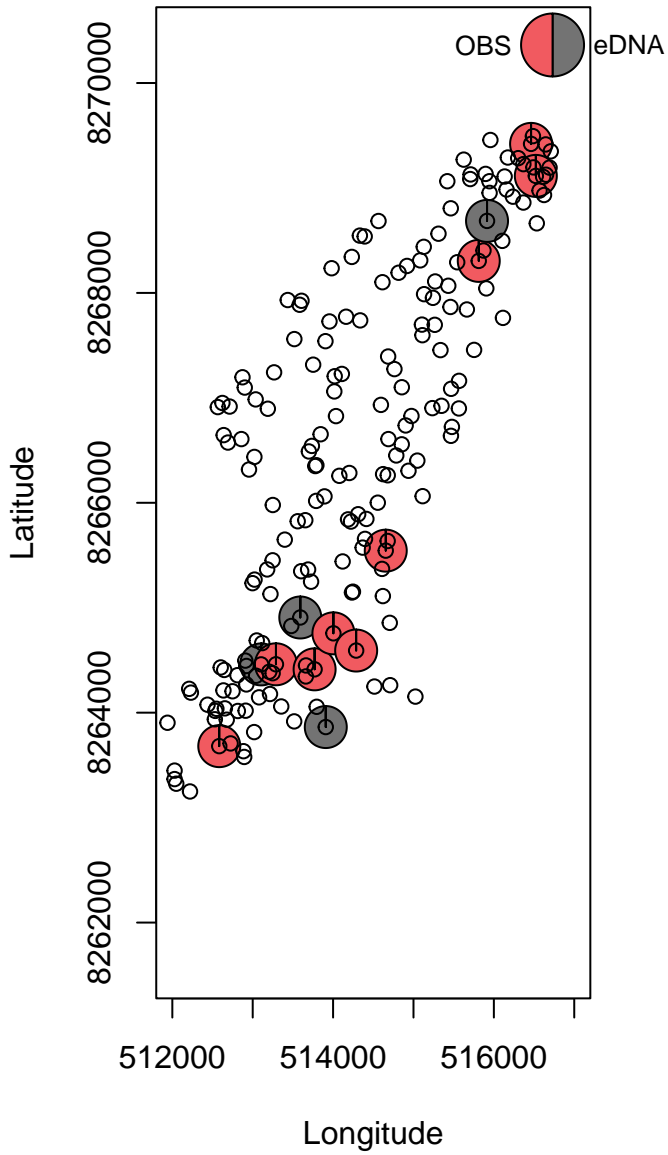
Oxyria



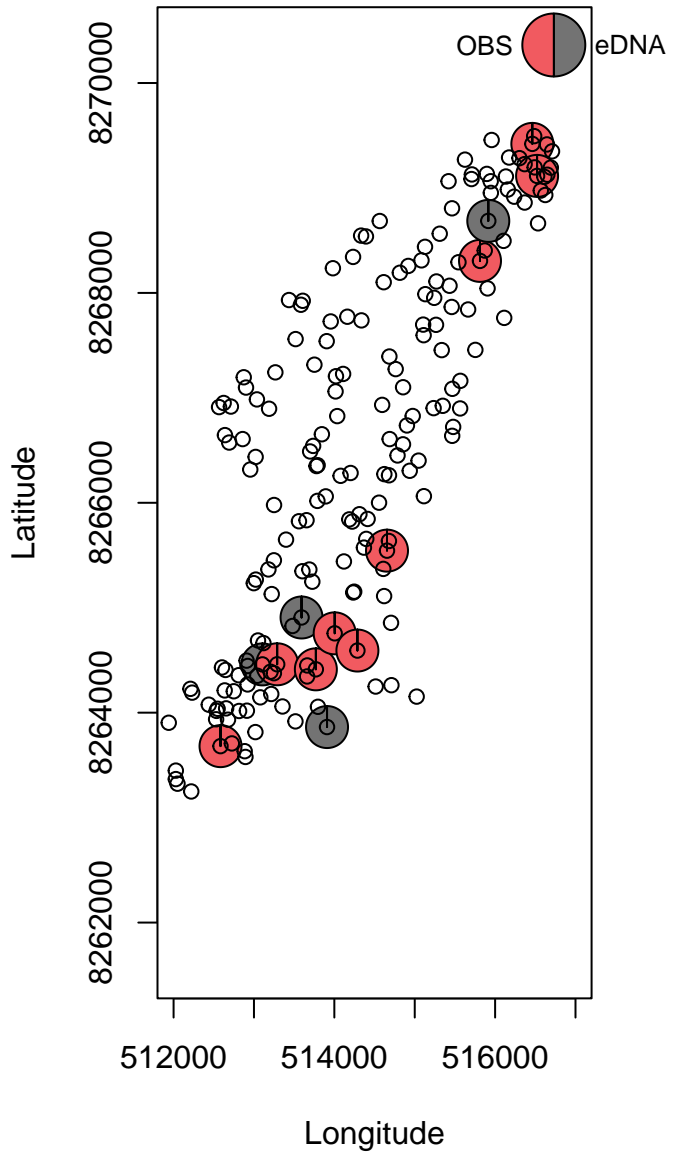
Oxyria_digyna



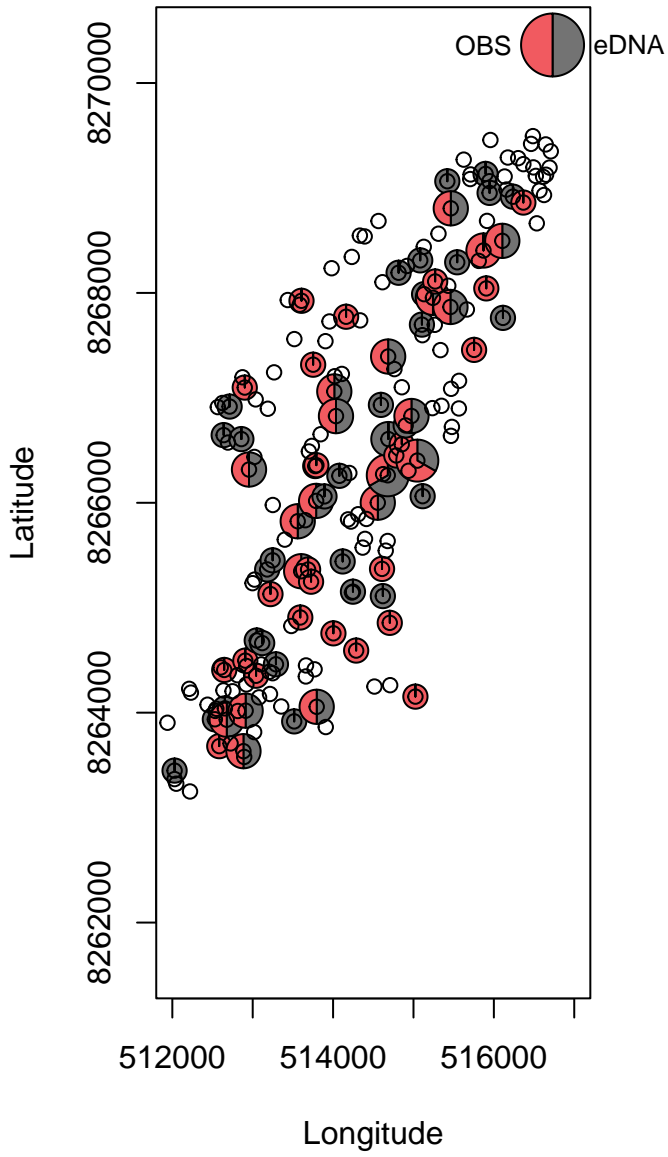
Papaver



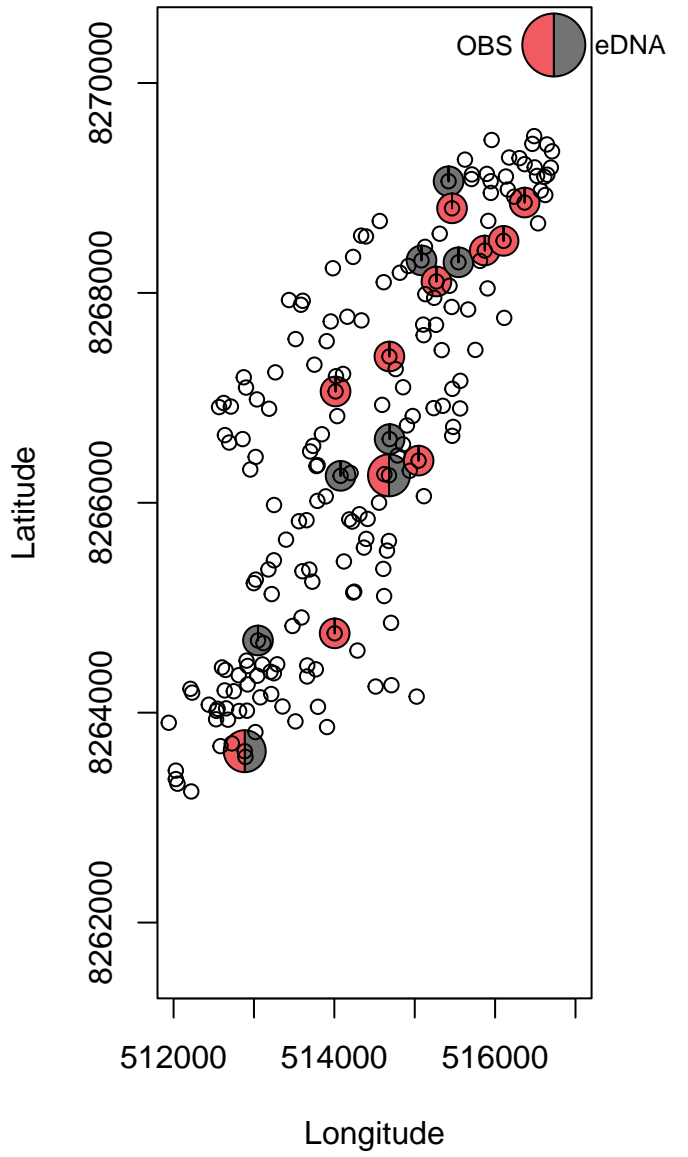
Papaver_radicatum



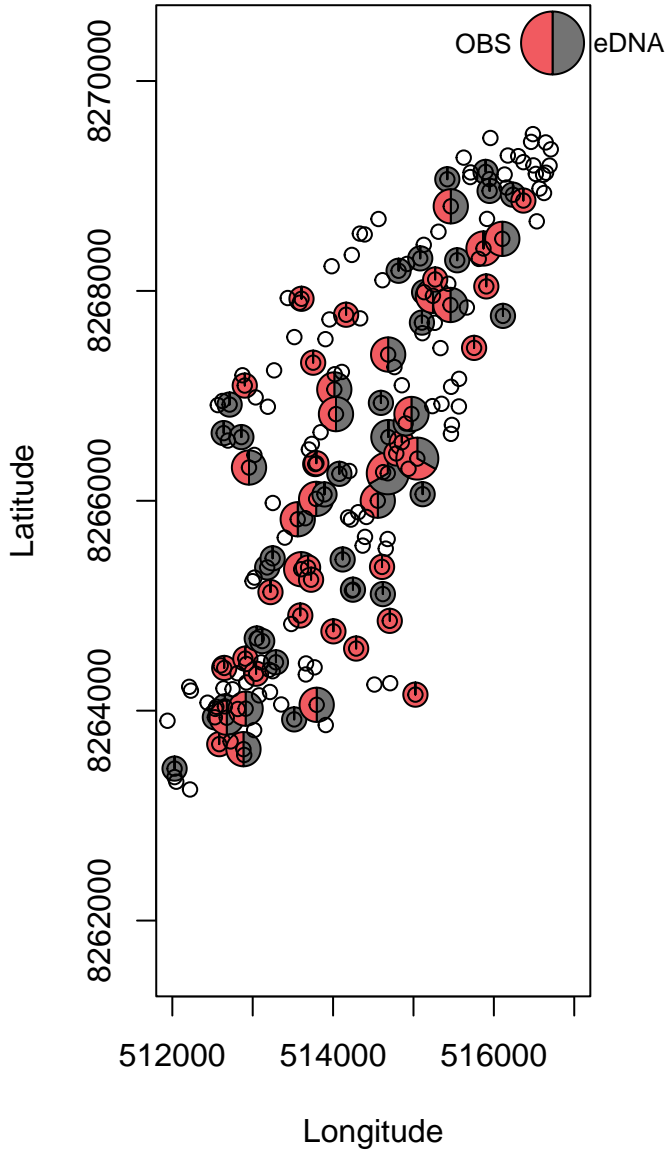
Pedicularis



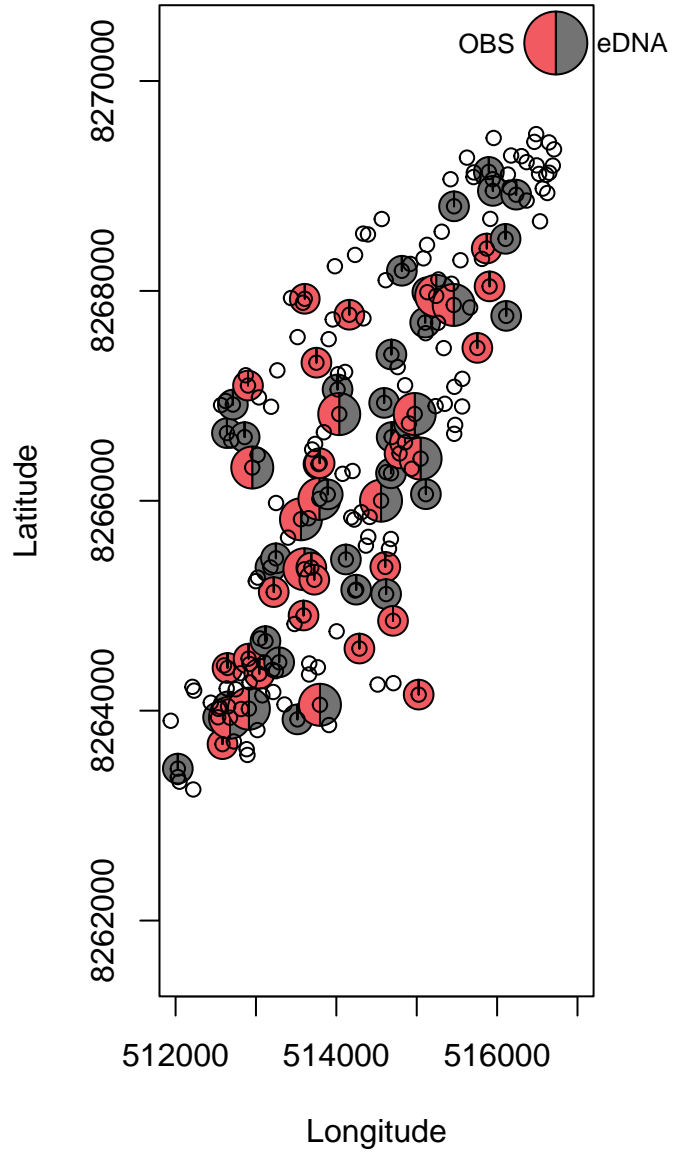
Pedicularis_flammea



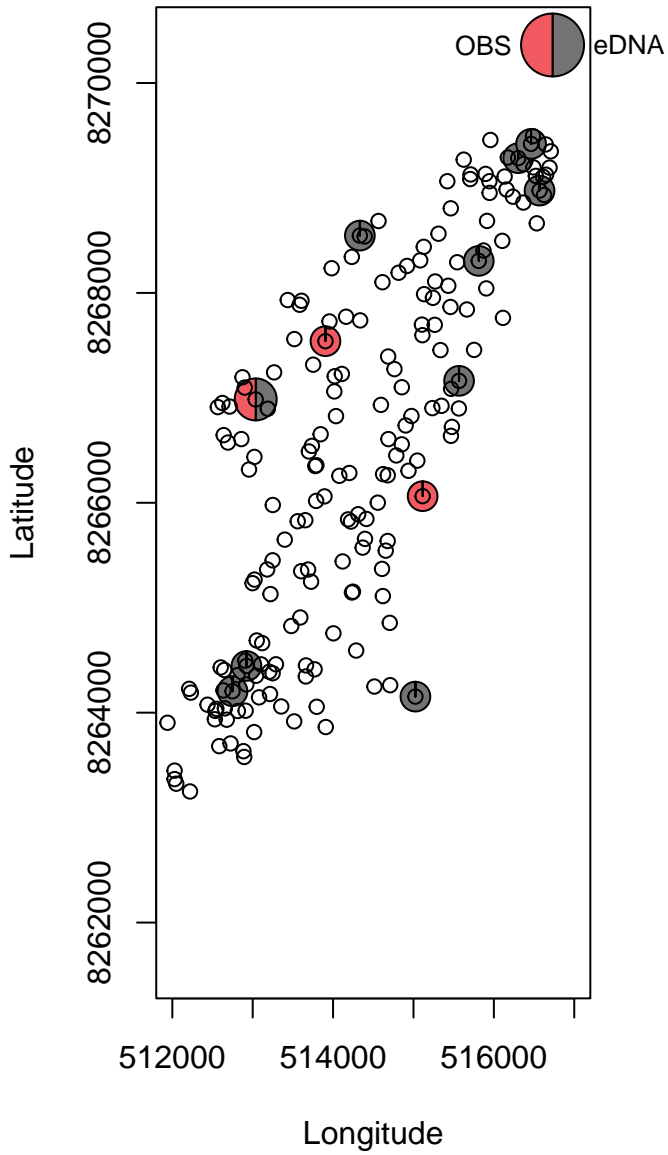
Pedicularis



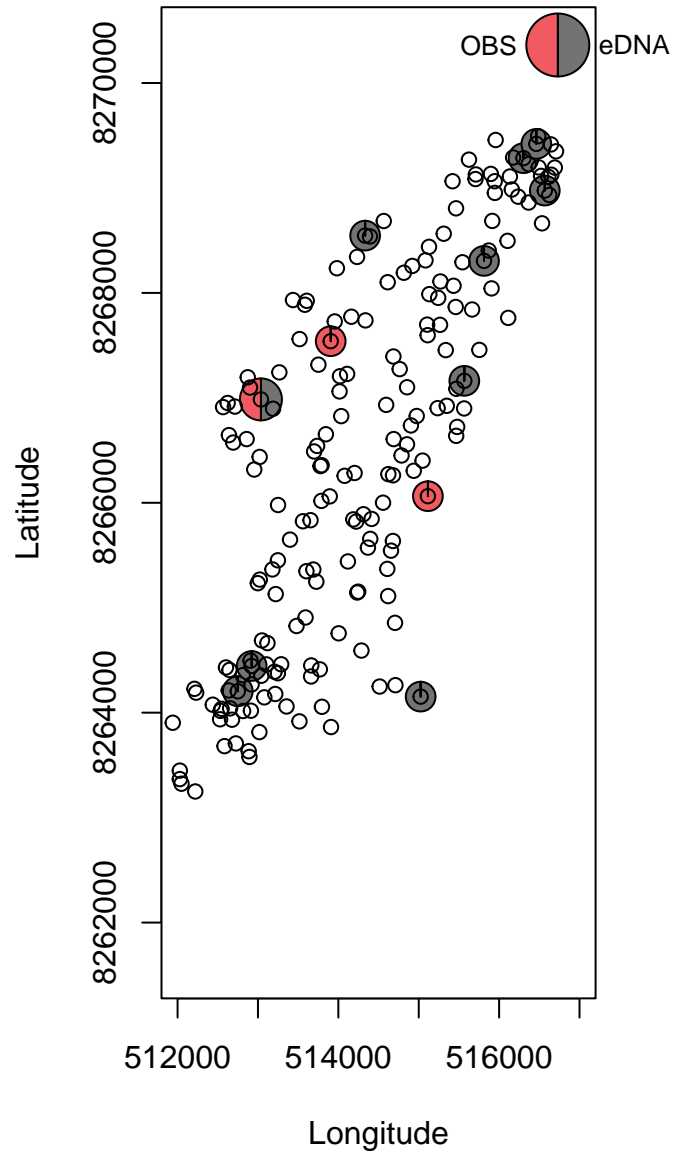
Pedicularis_hirsuta



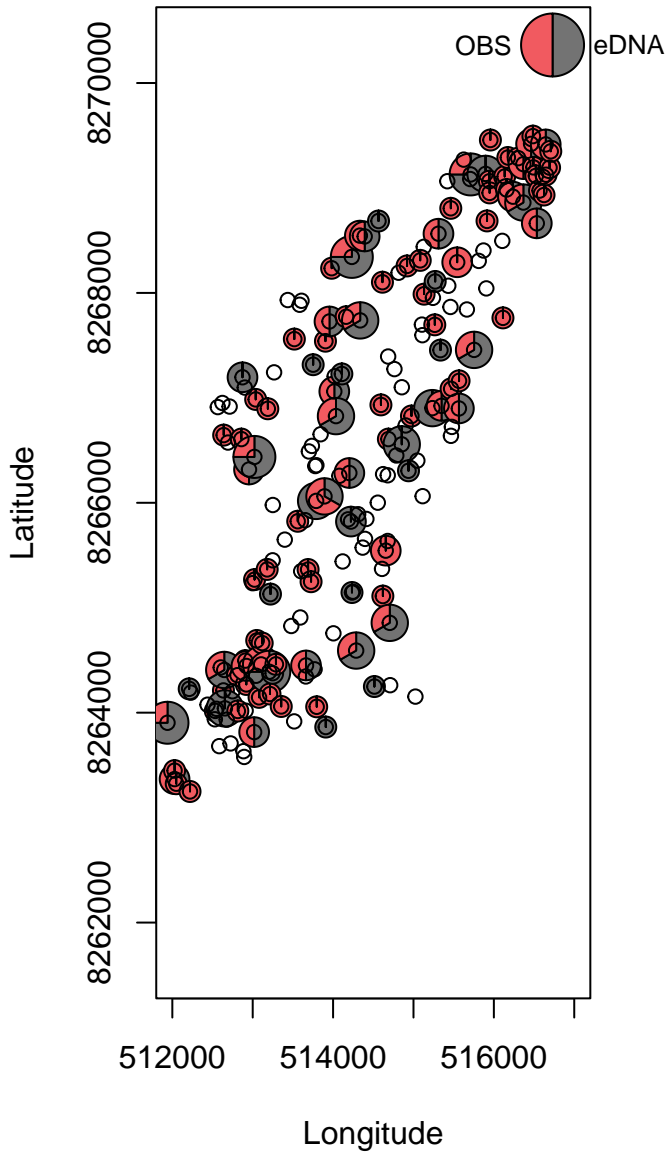
Phippsia



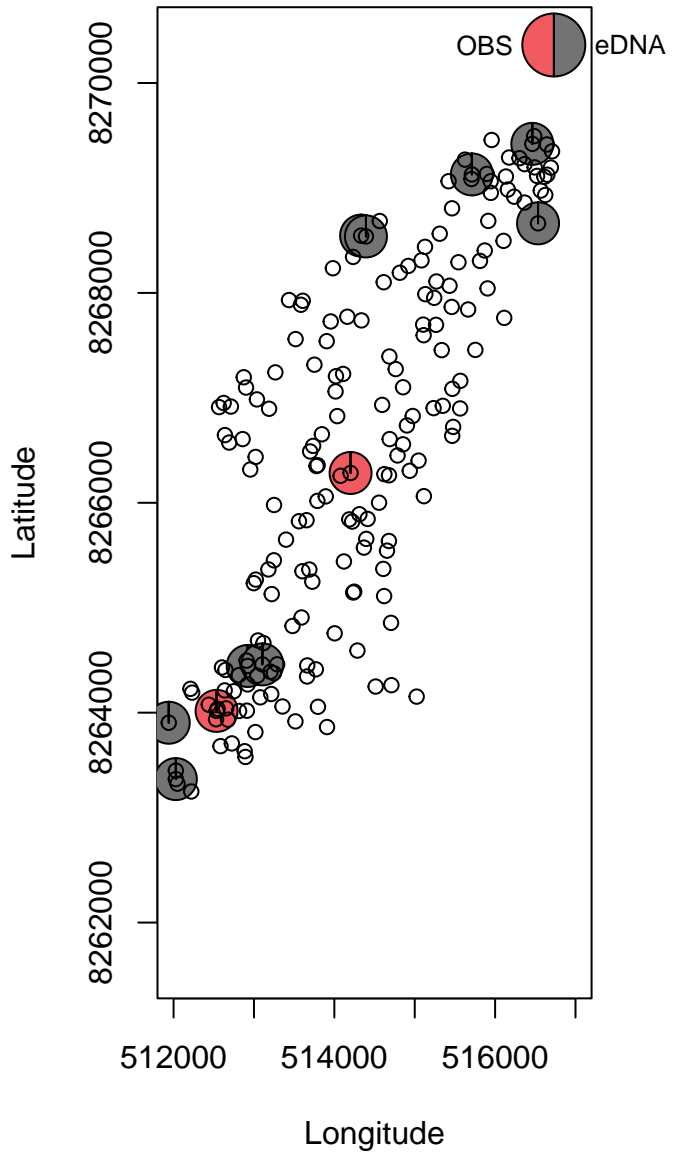
Phippsia_algida



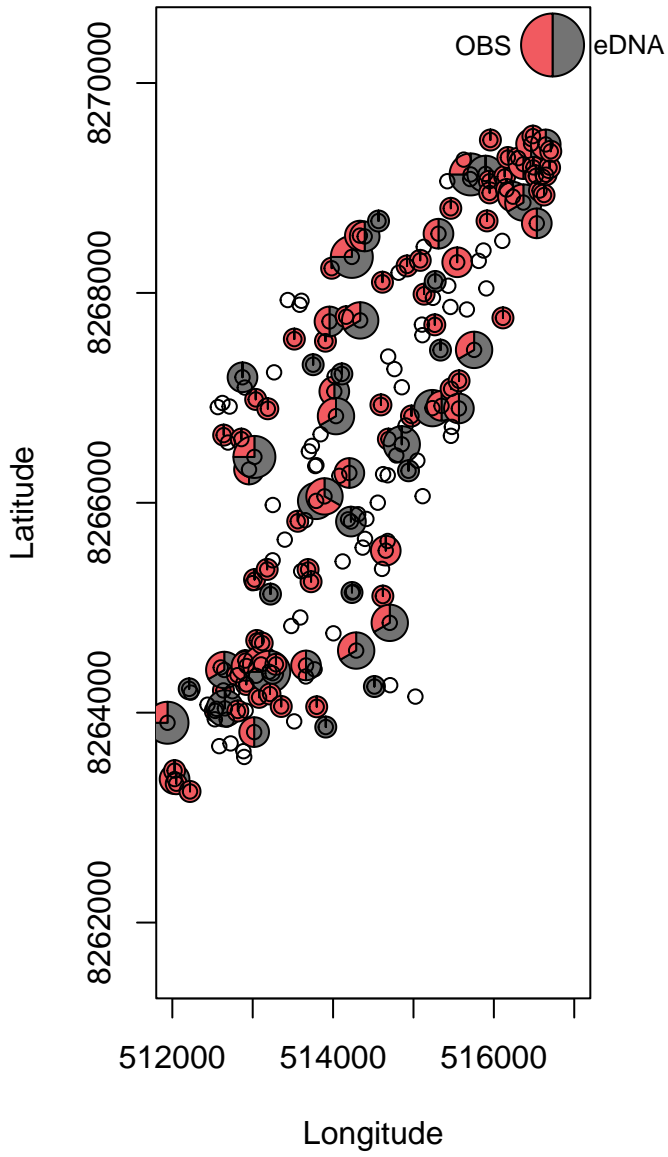
Poa



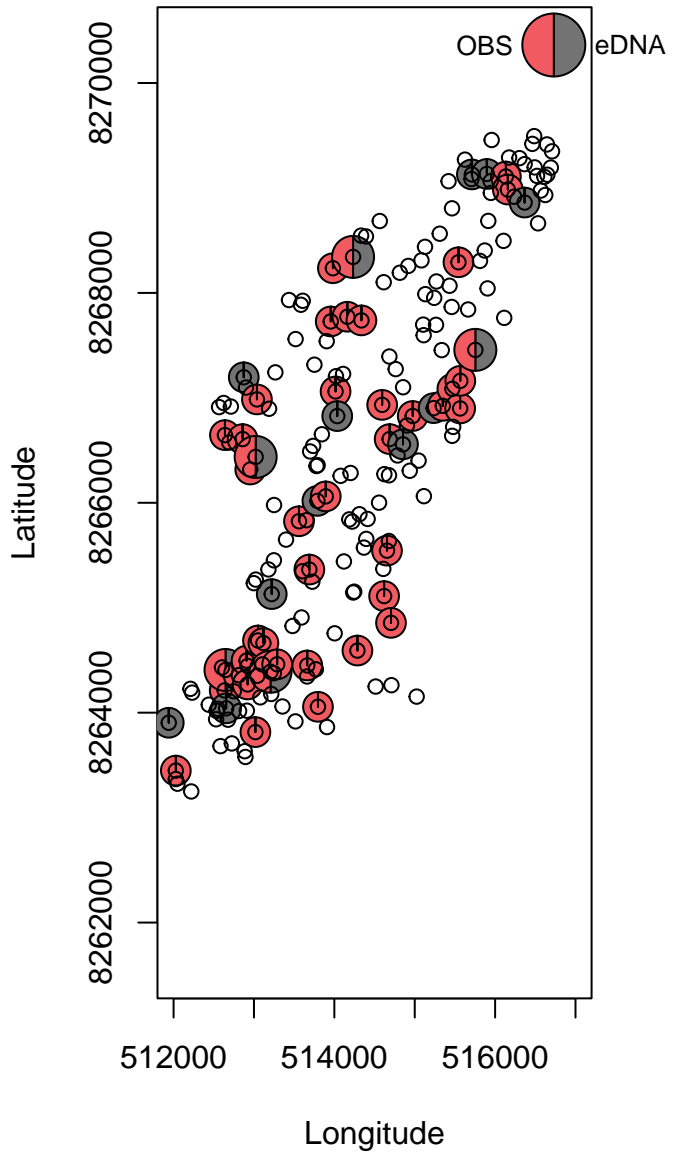
Poa_abbreviata



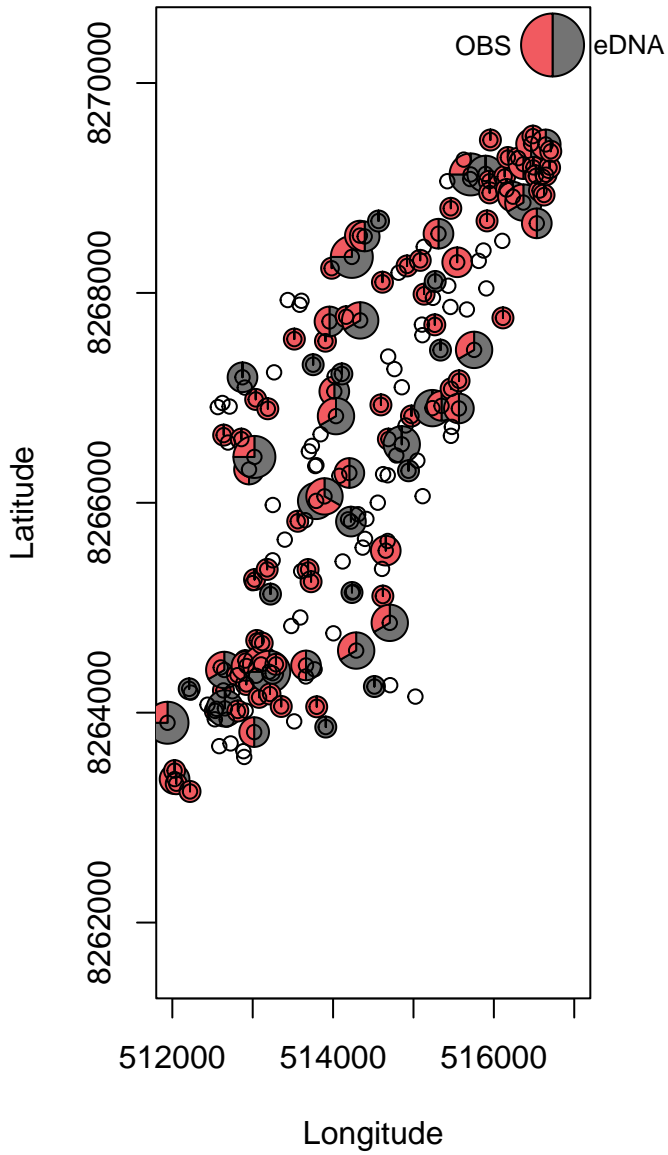
Poa



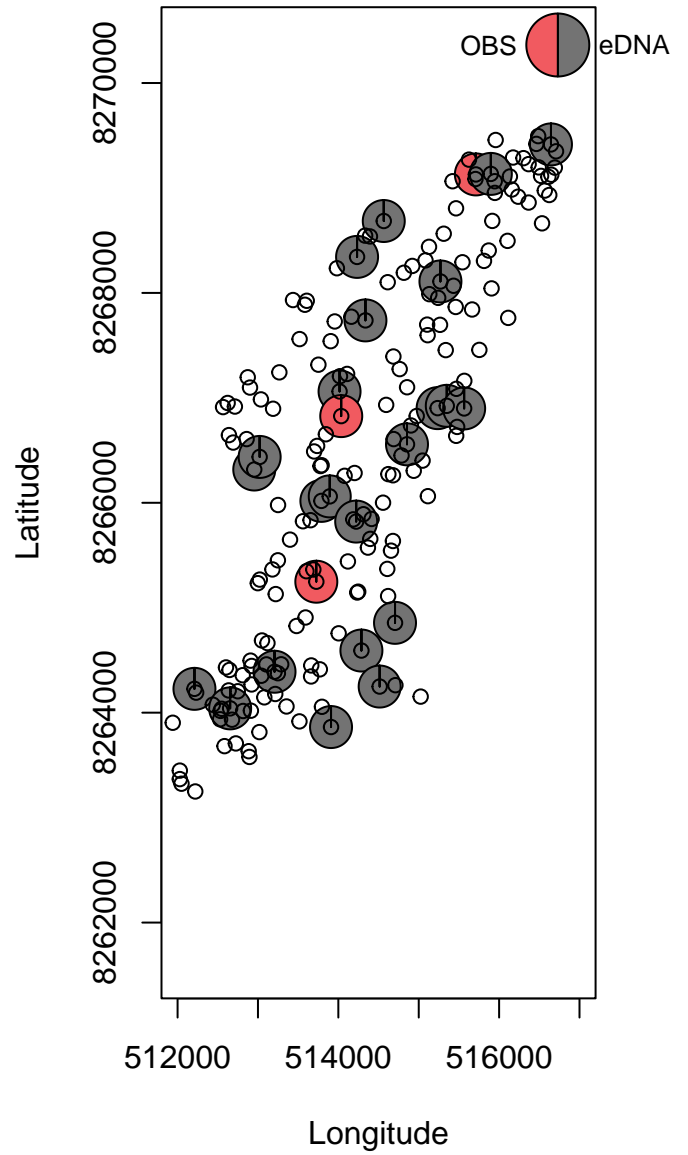
Poa_arctica



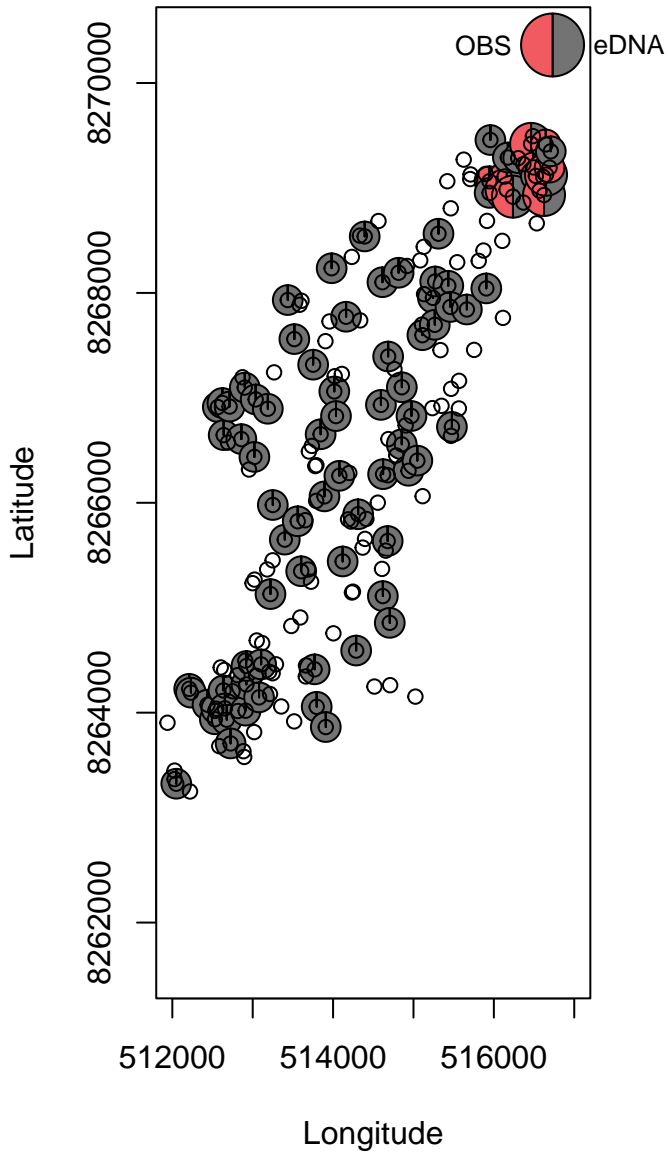
Poa



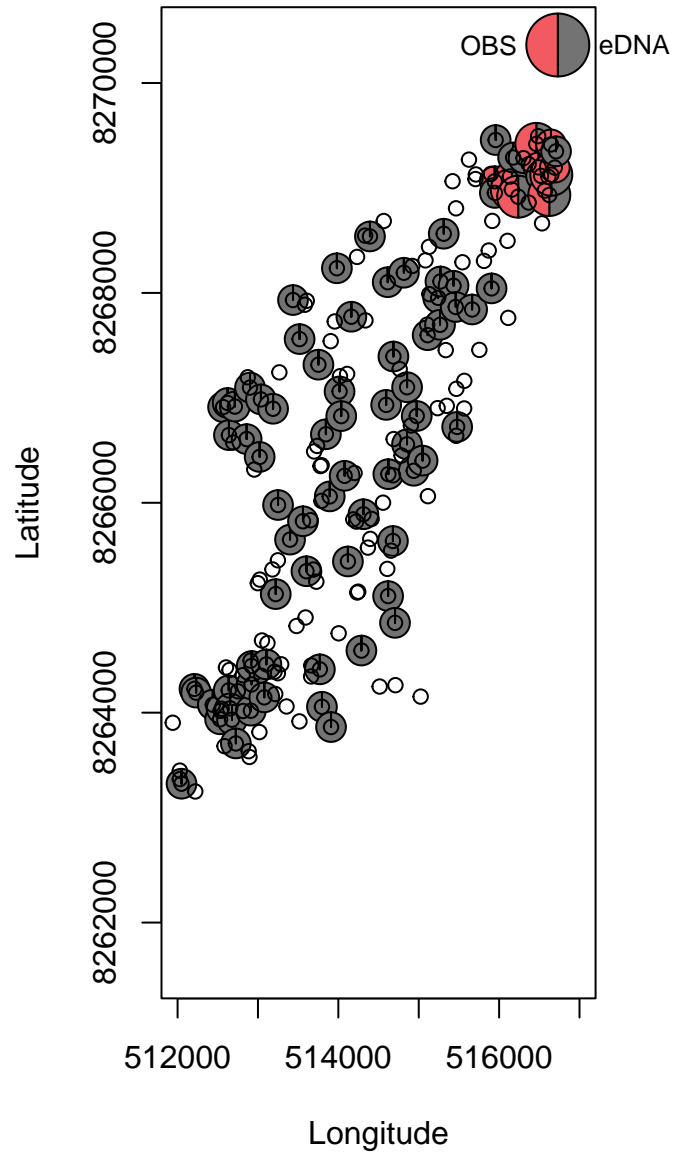
Poa_pratensis



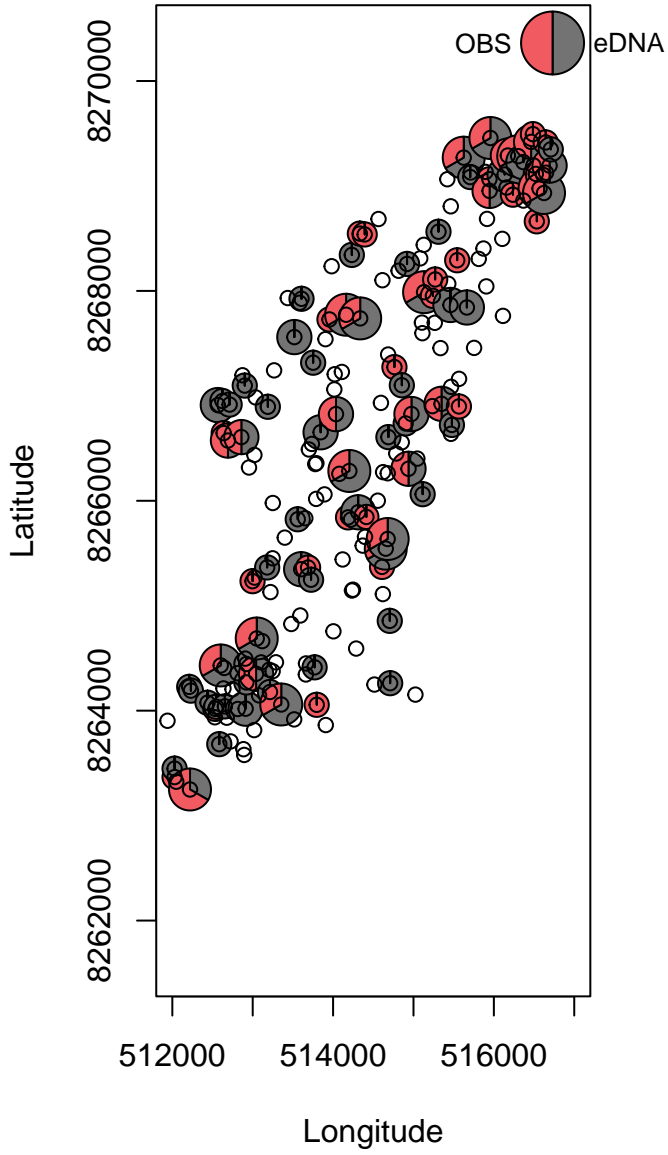
Polemonium



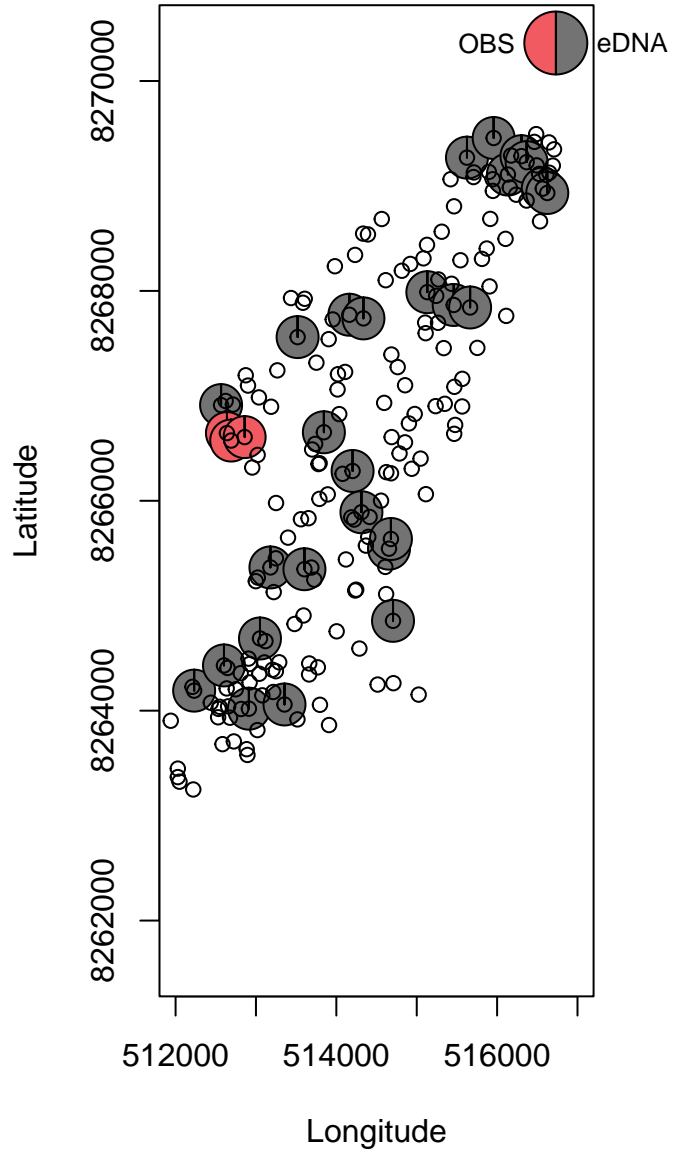
Polemonium_boreale



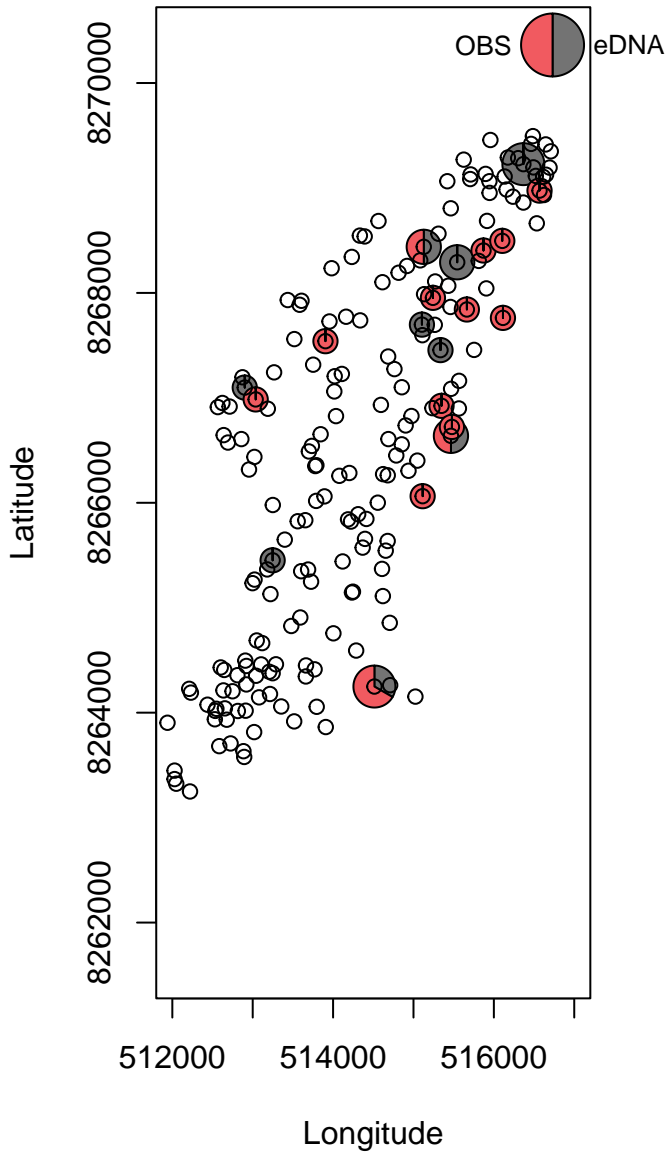
Potentilla



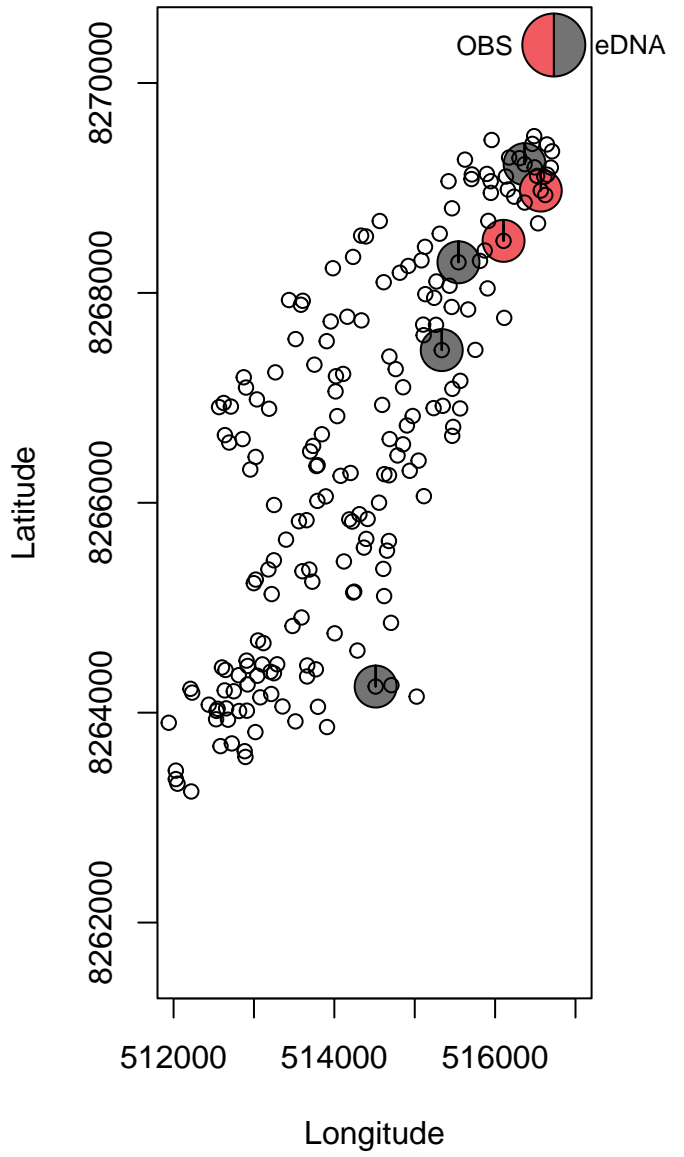
Potentilla_arenosa



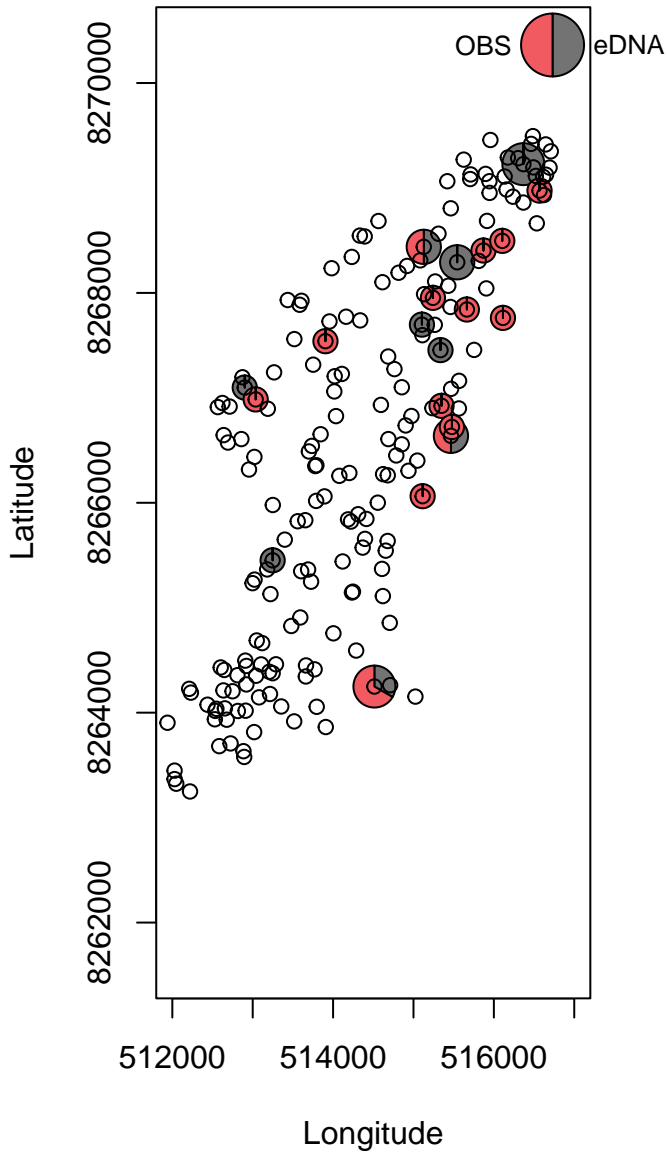
Ranunculus



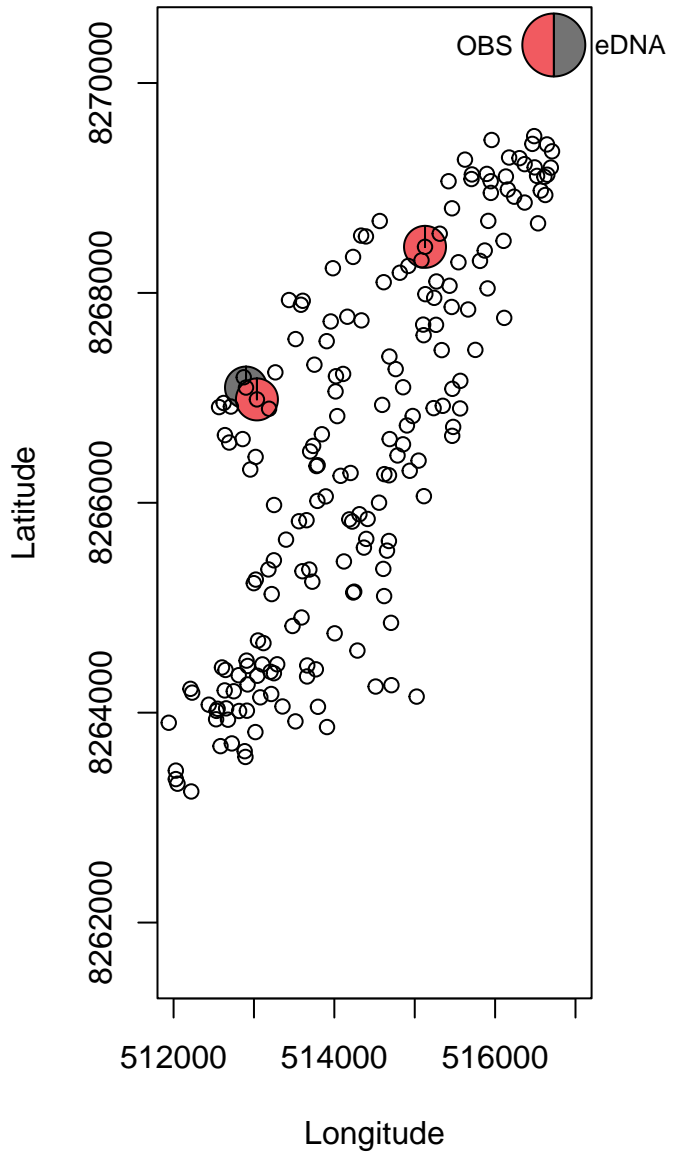
Ranunculus_glacialis



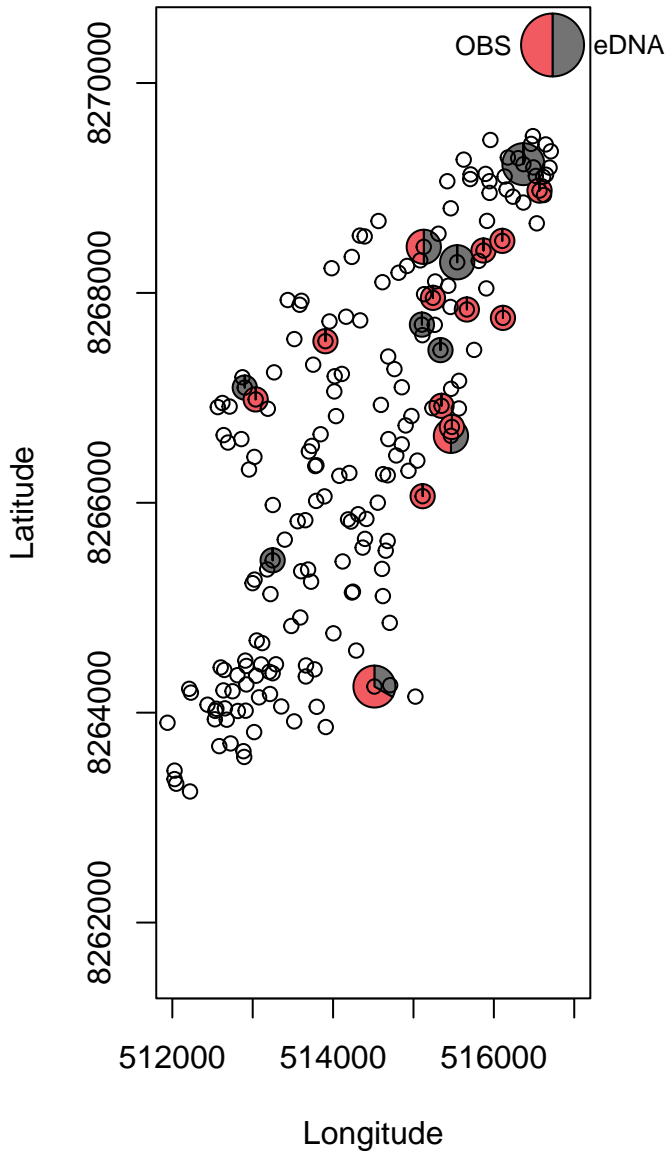
Ranunculus



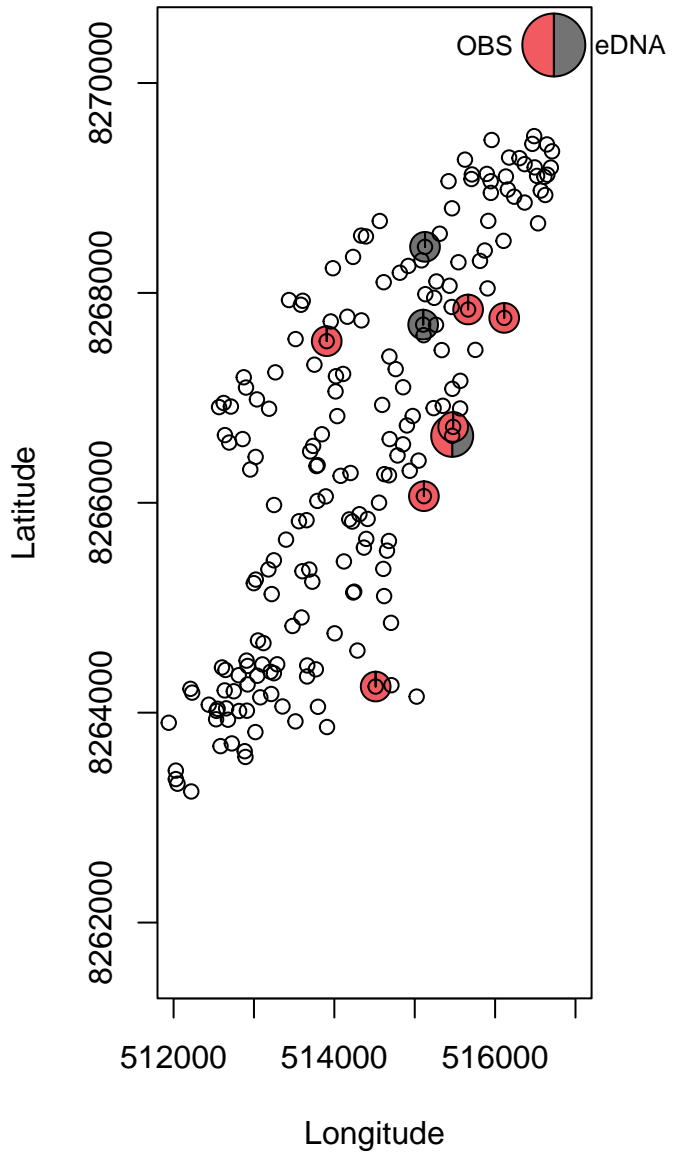
Ranunculus_nivalis



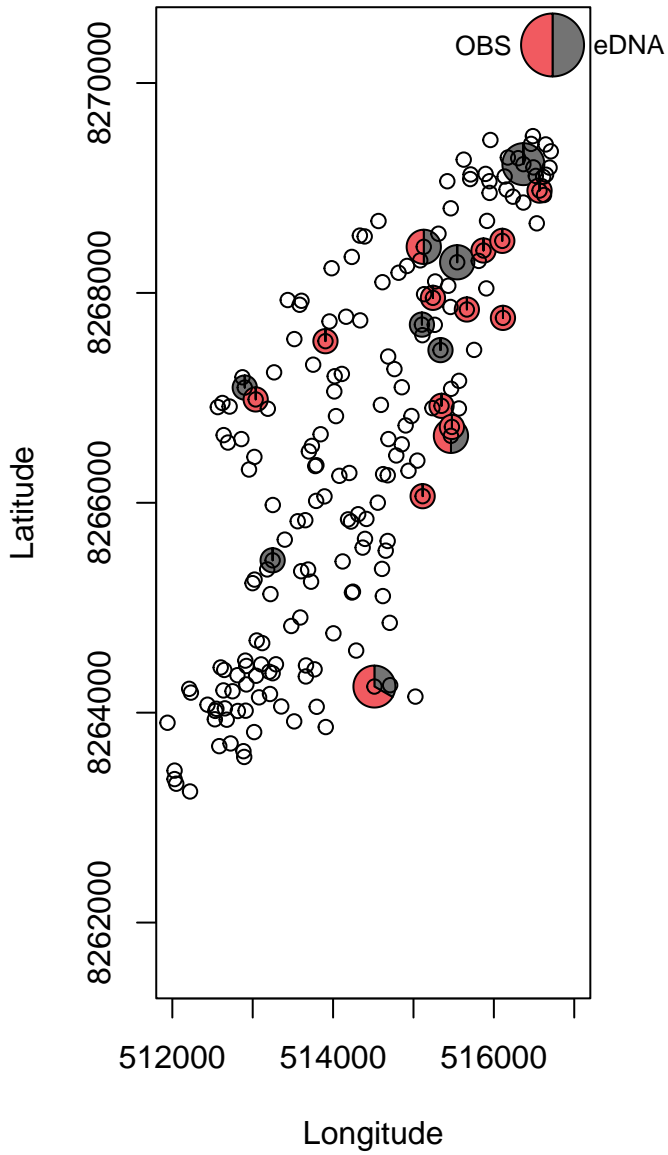
Ranunculus



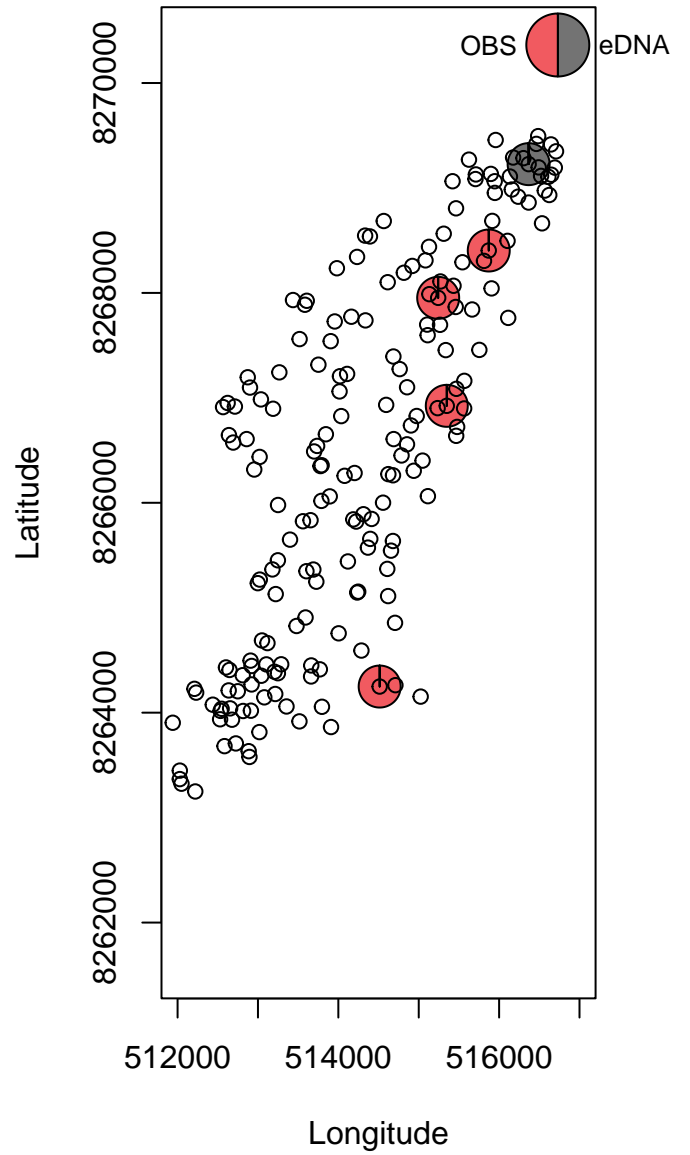
Ranunculus_pygmaeus



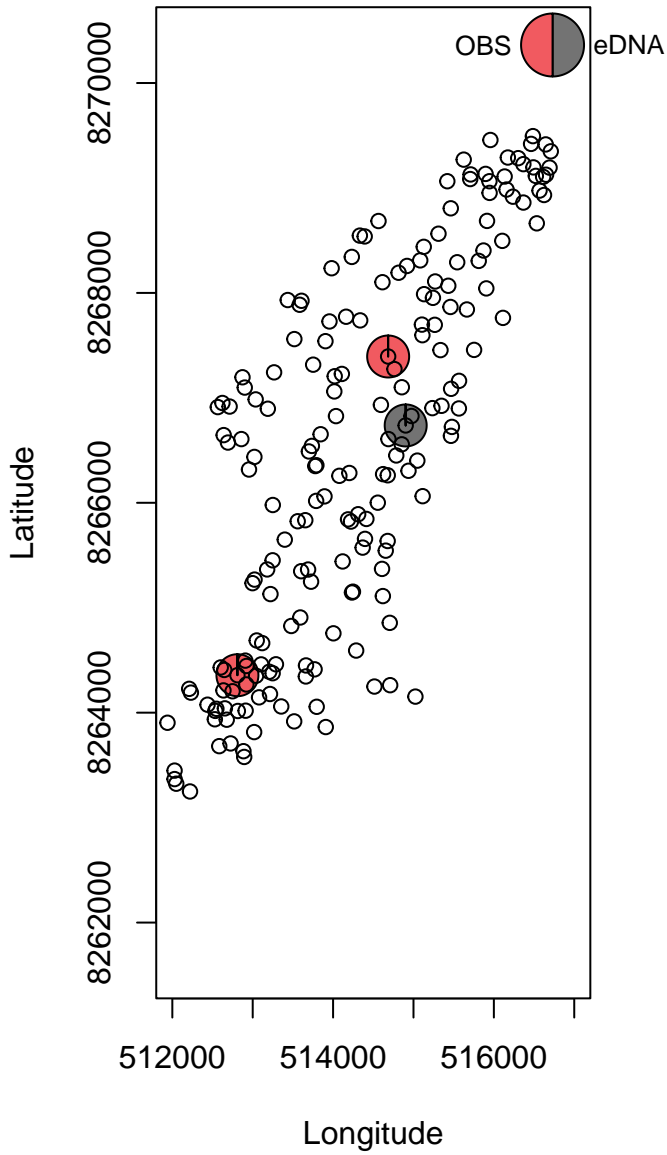
Ranunculus



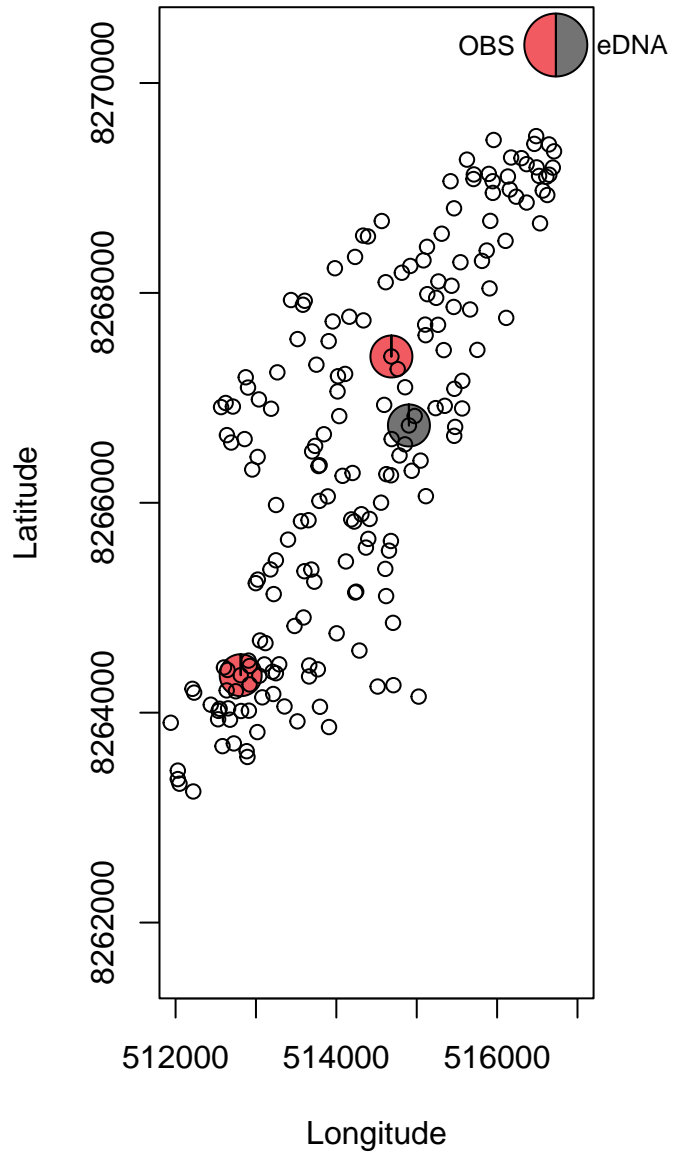
Ranunculus_sulphureus



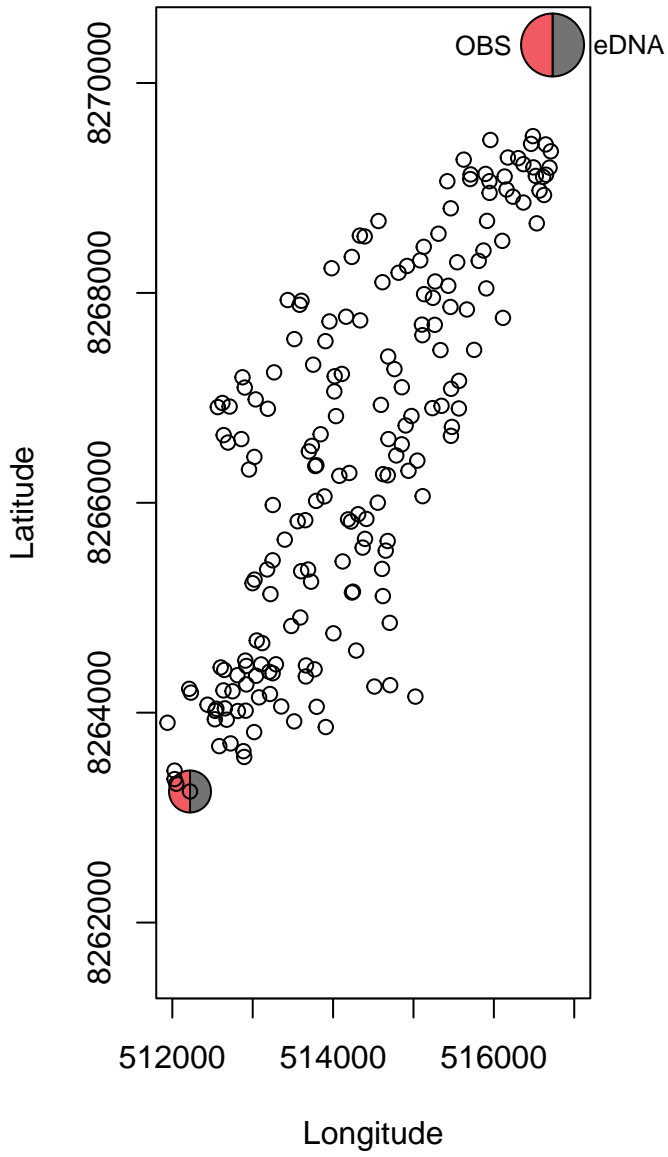
Rhododendron



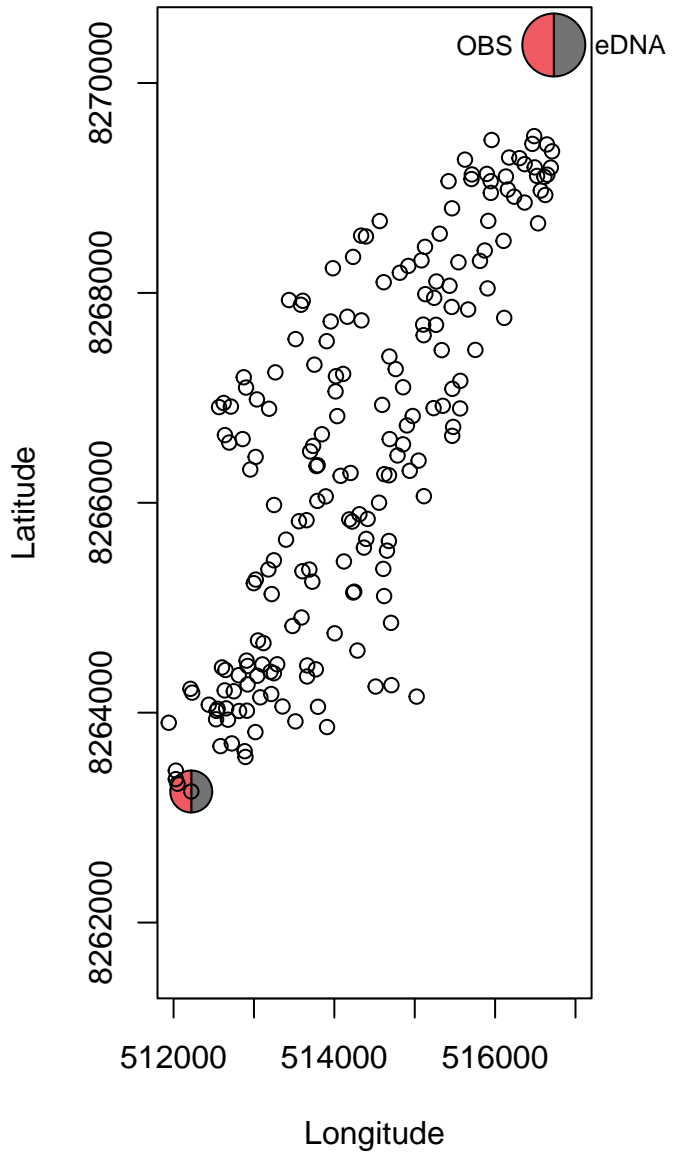
Rhododendron_lapponicum



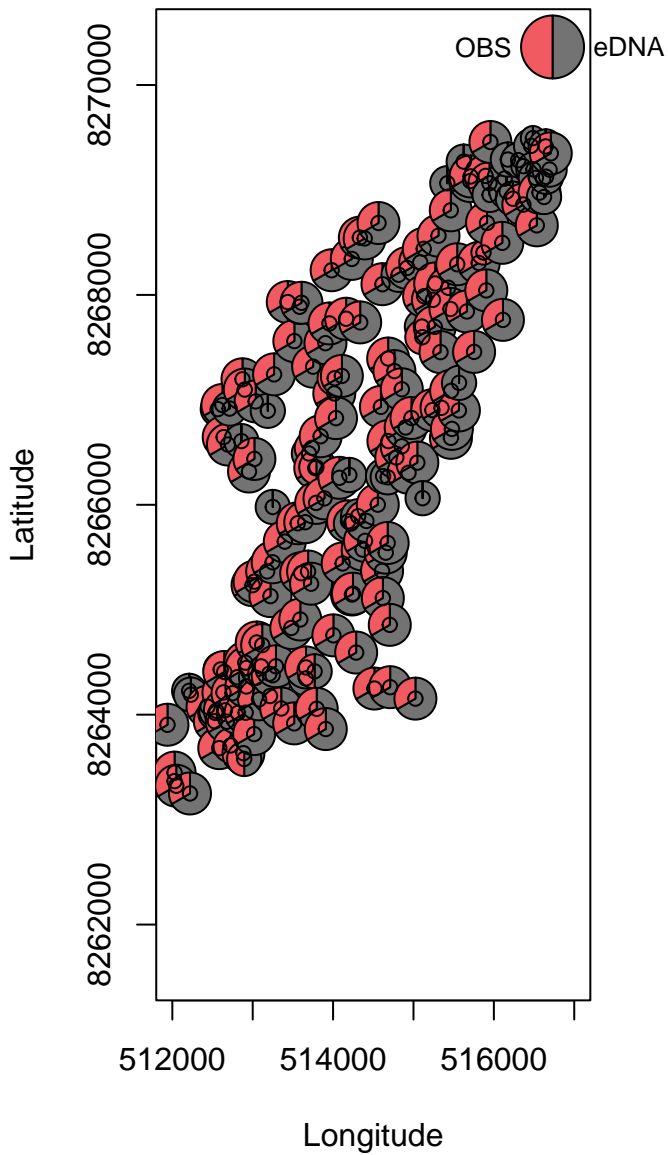
Rumex



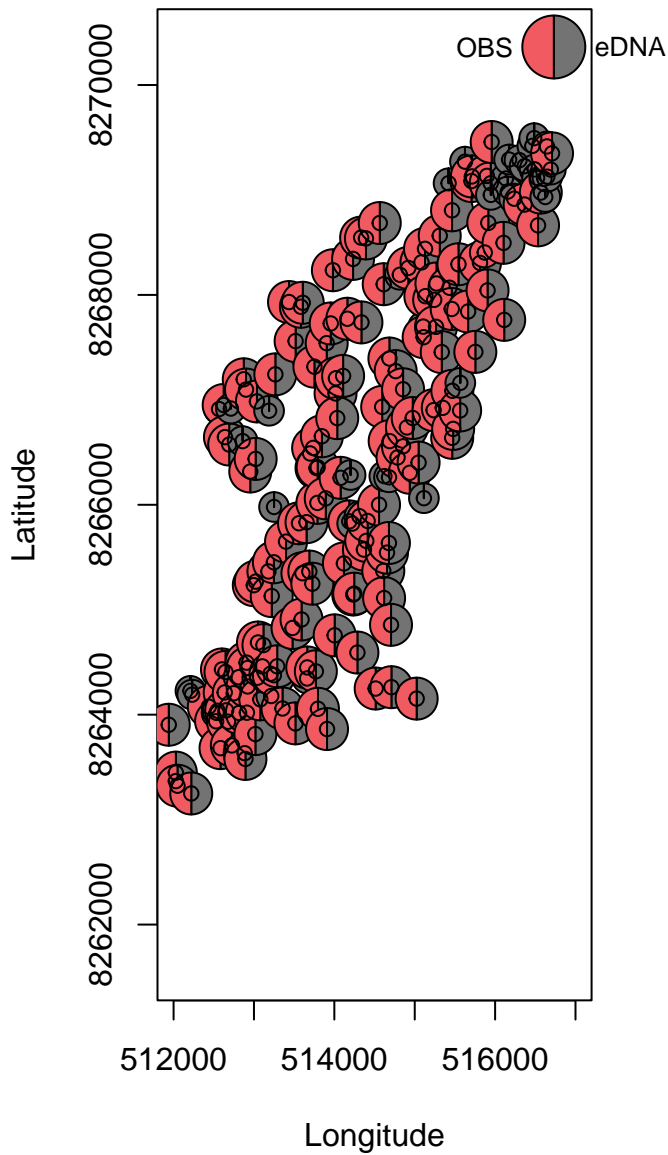
Rumex_acetosella



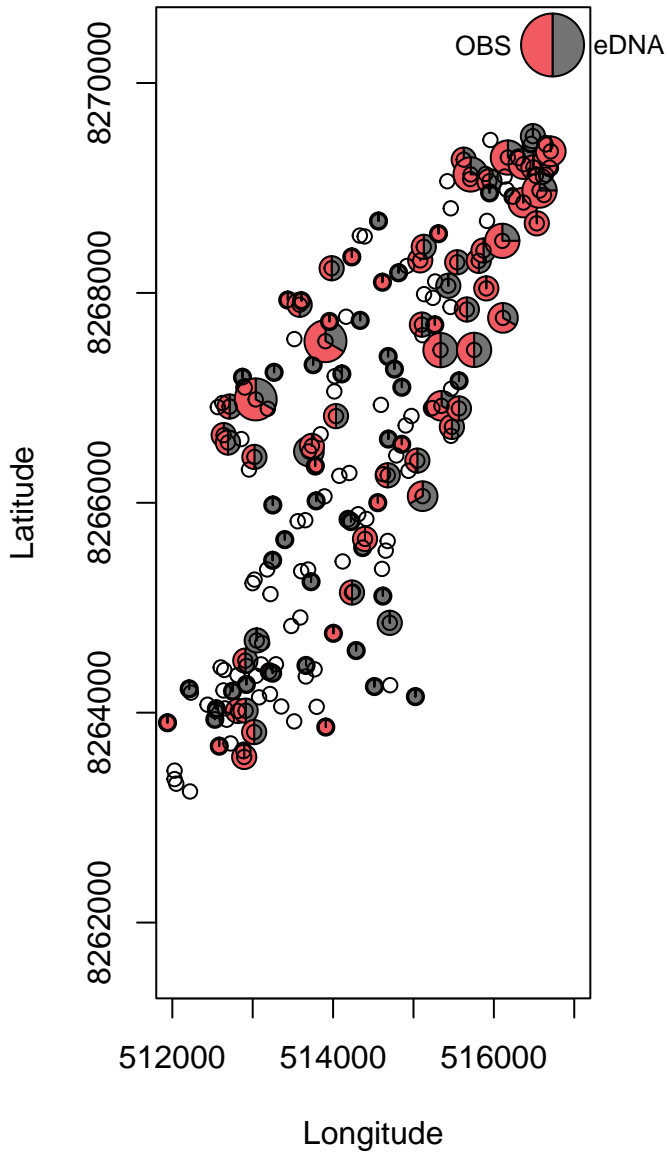
Salix



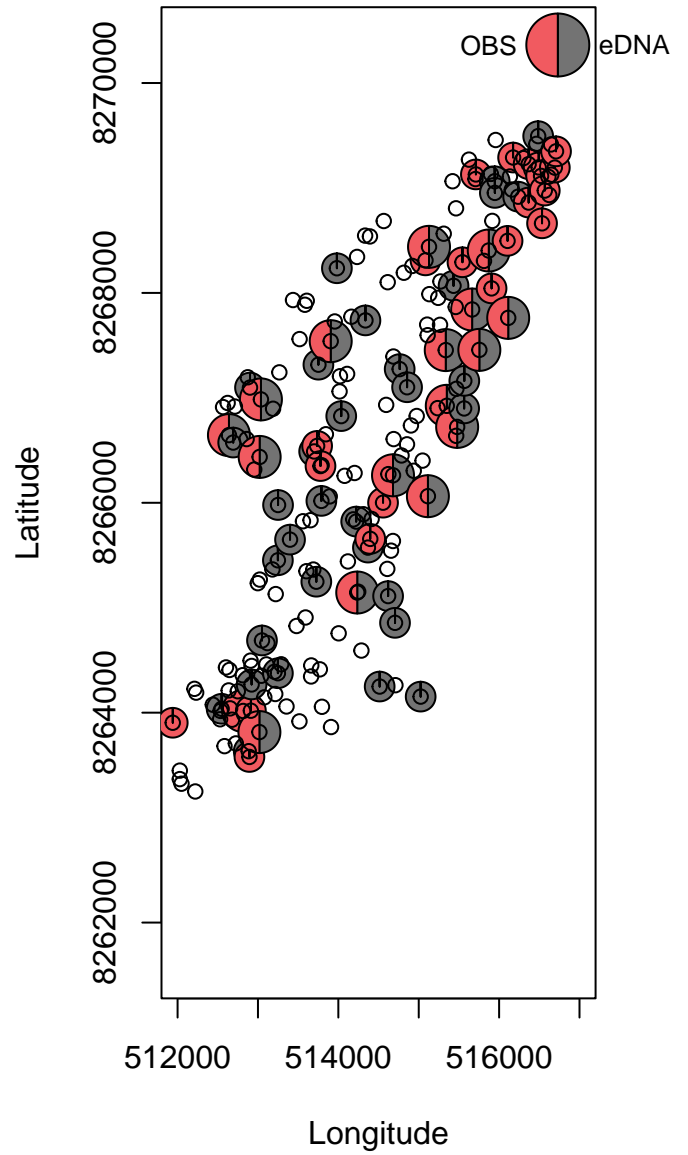
Salix_arctica



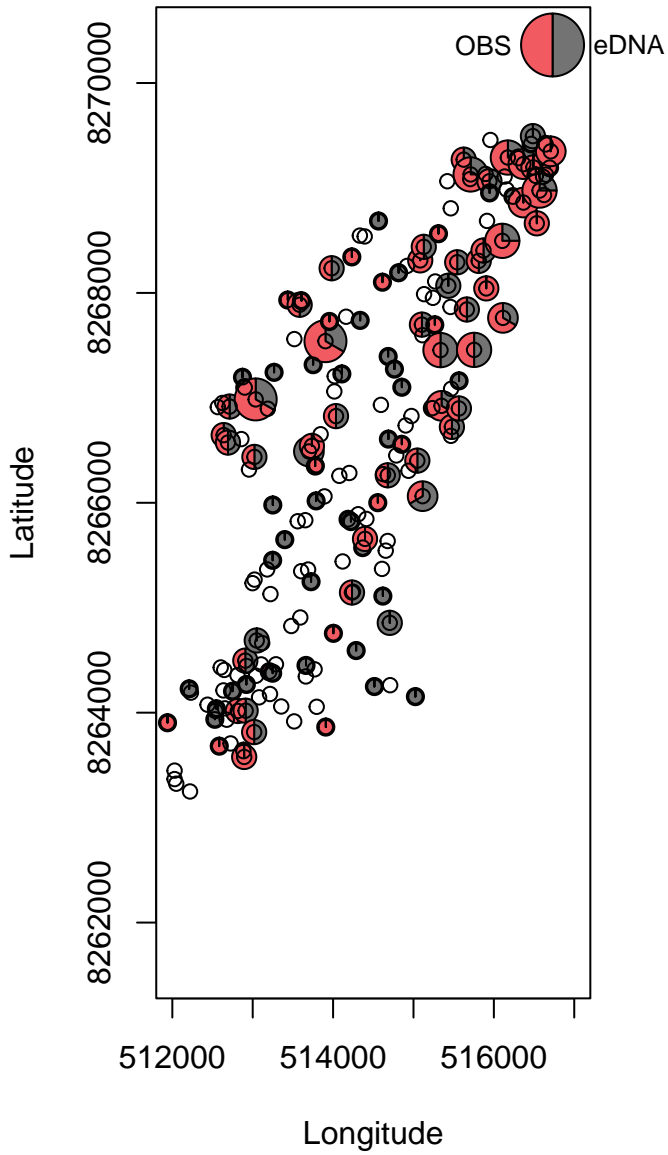
Saxifraga



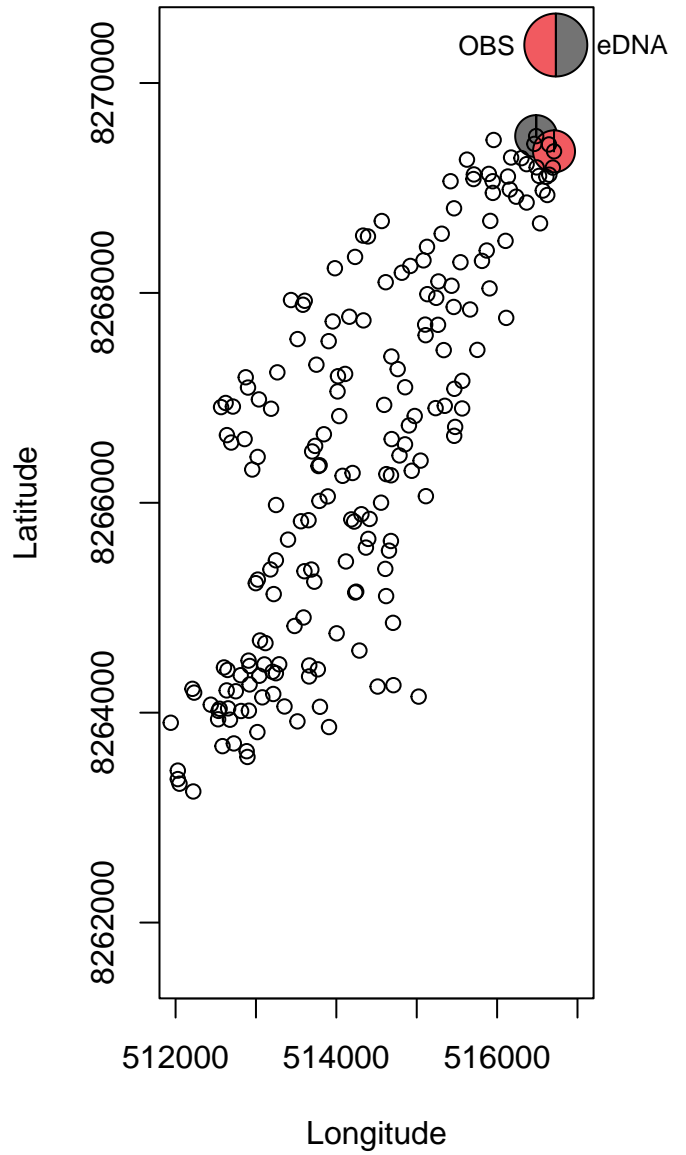
Saxifraga_cernua



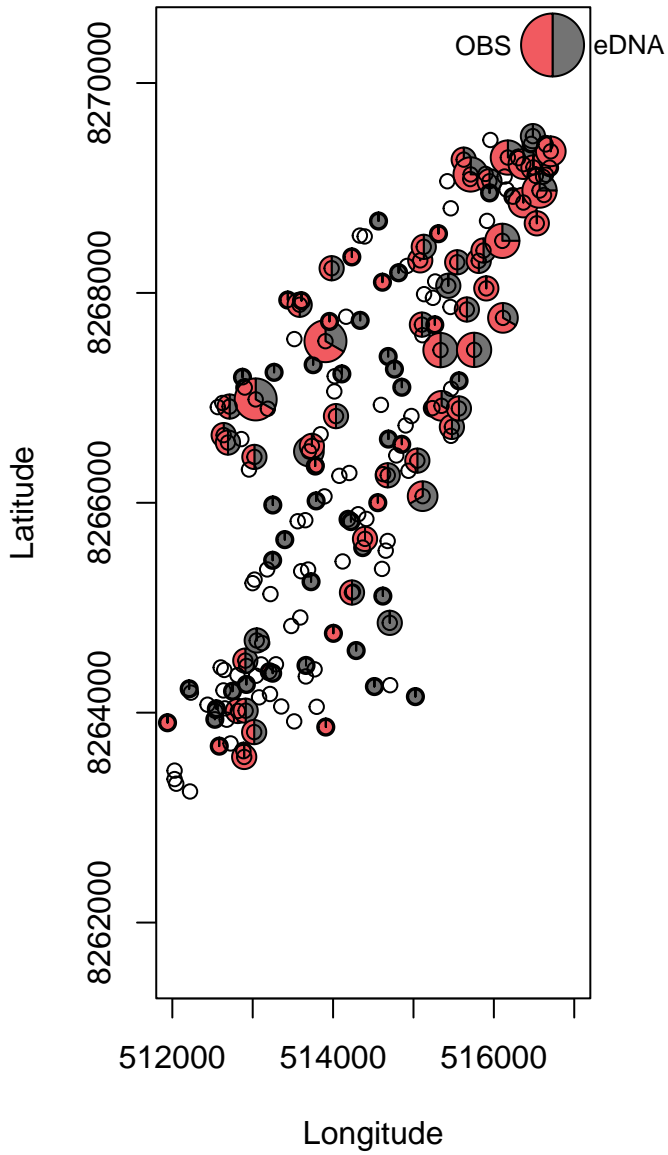
Saxifraga



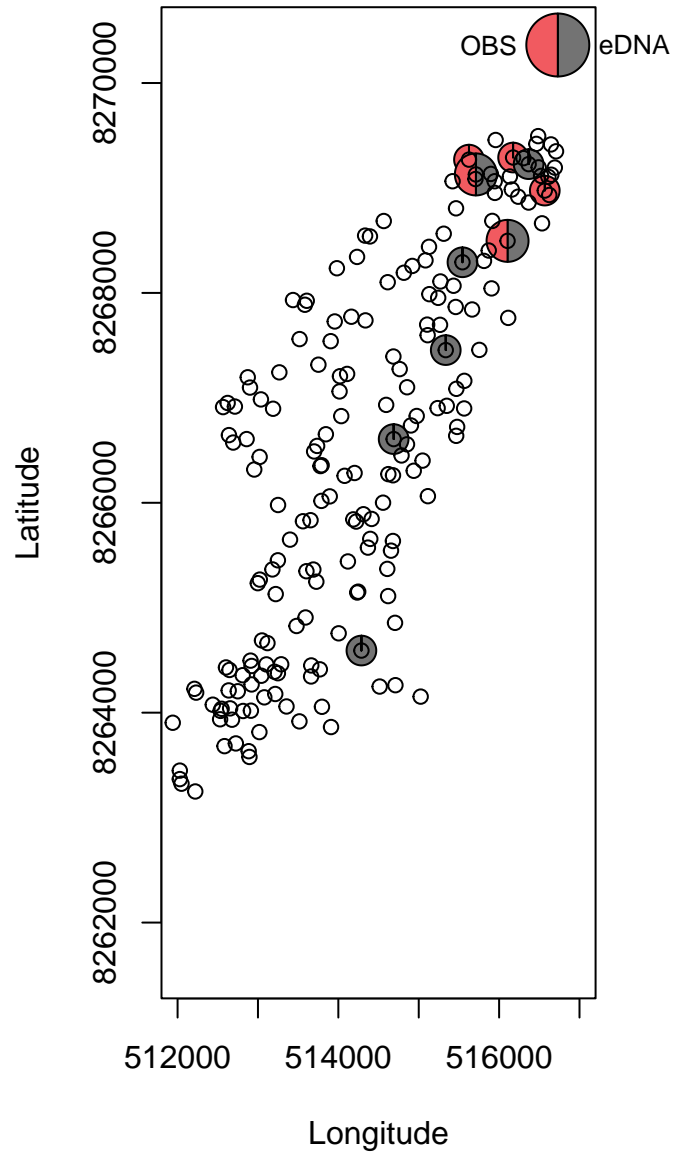
Saxifraga_cespitosa



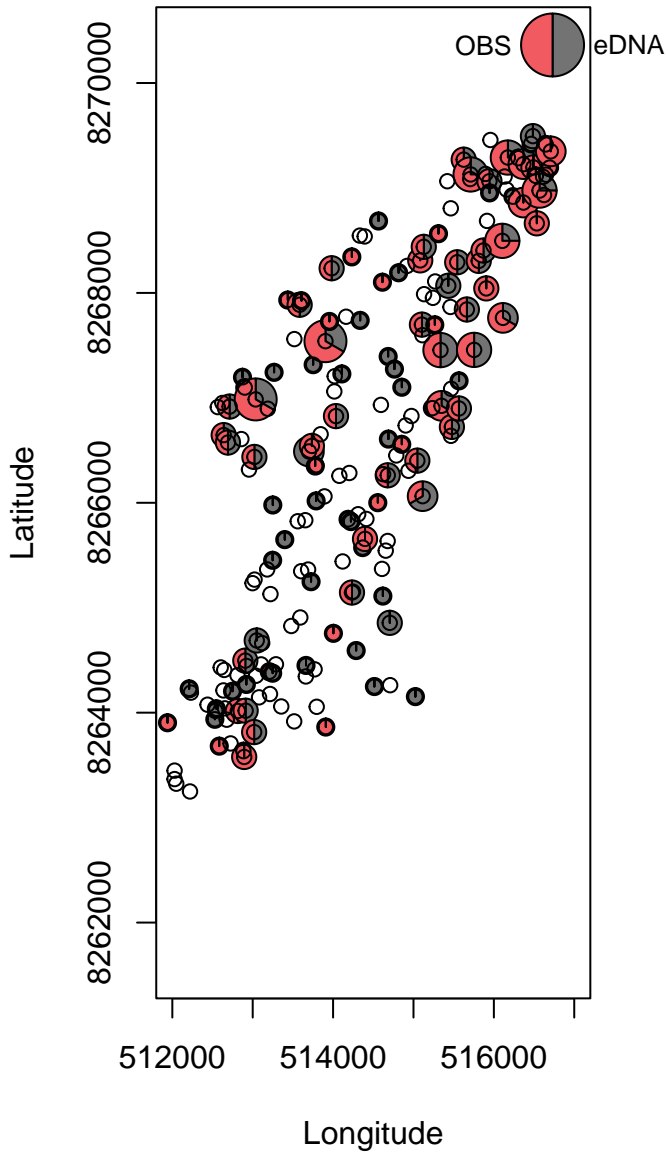
Saxifraga



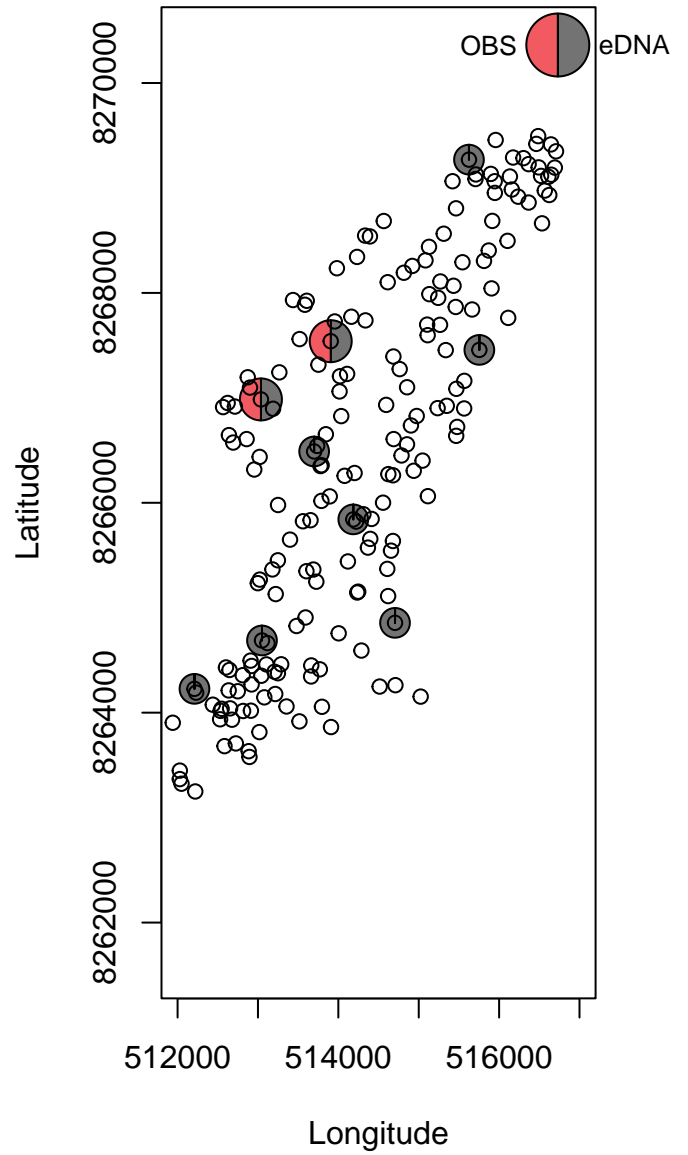
Saxifraga_hirculus



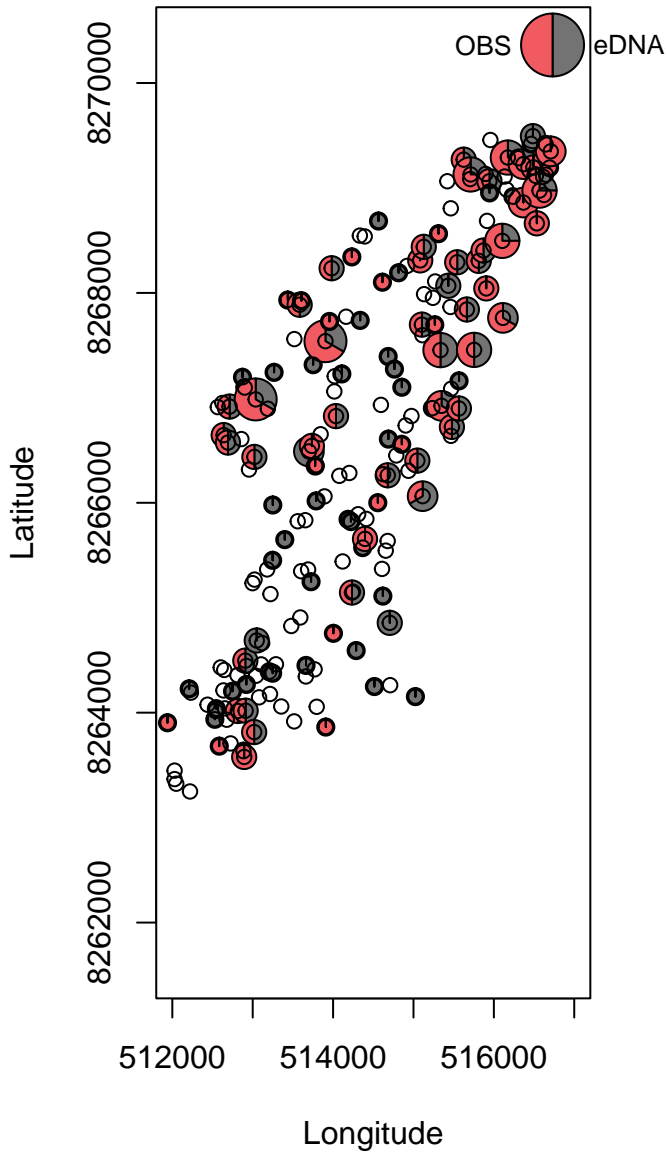
Saxifraga



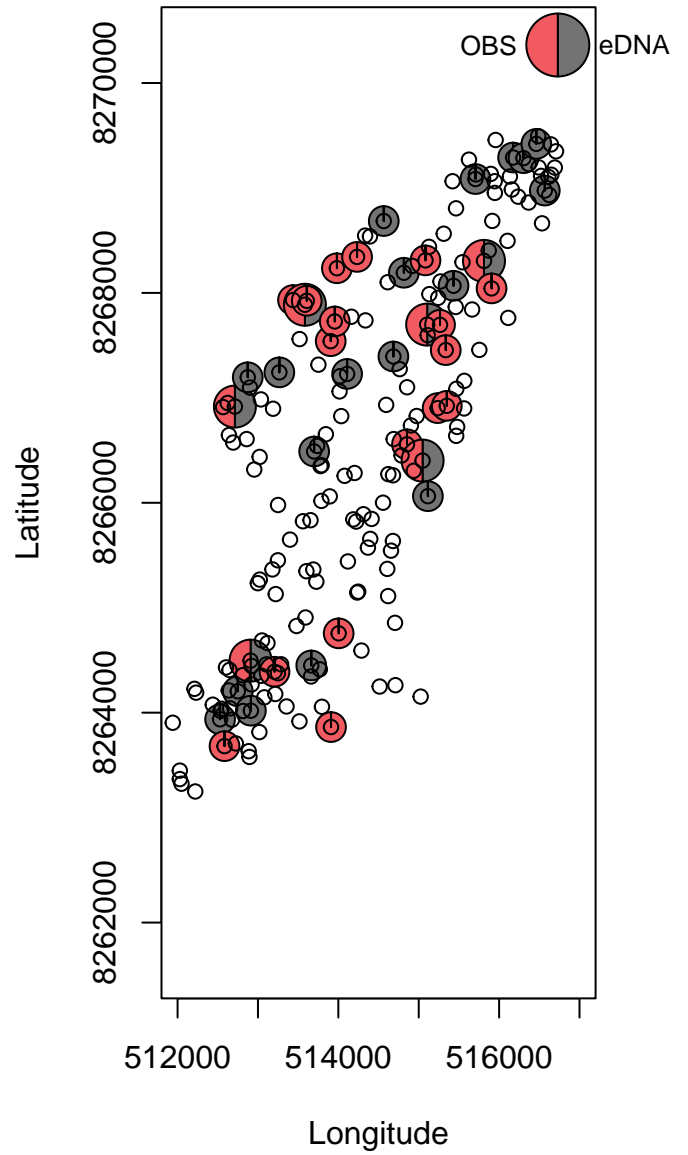
Saxifraga_hyperborea



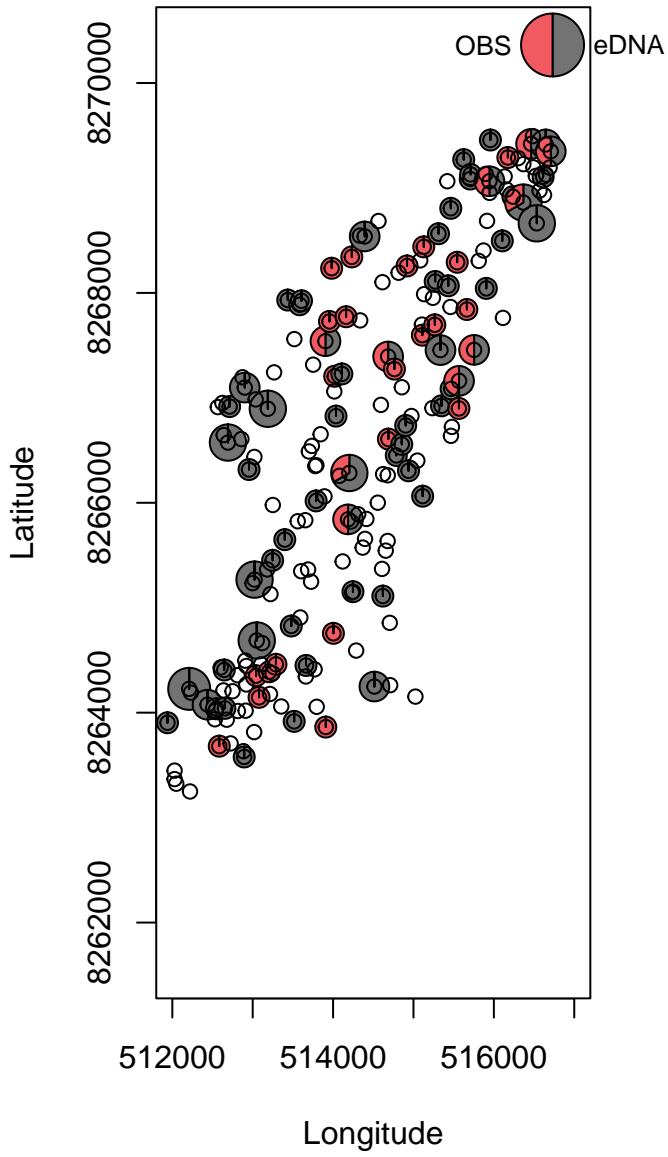
Saxifraga



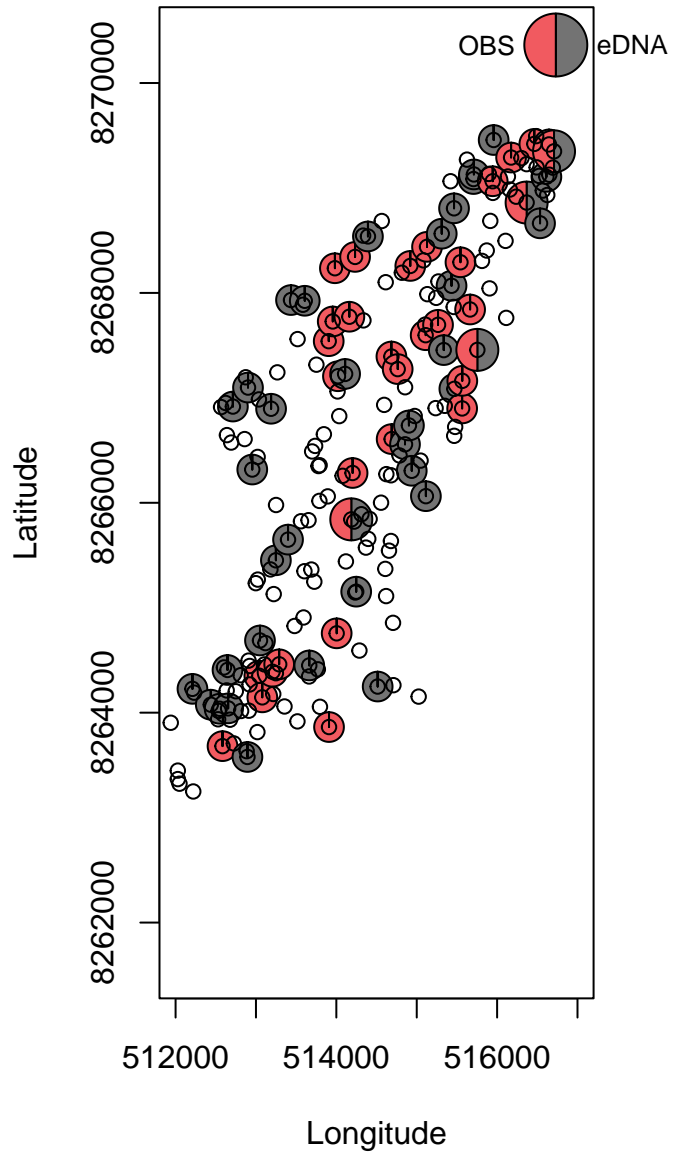
Saxifraga_oppositifolia



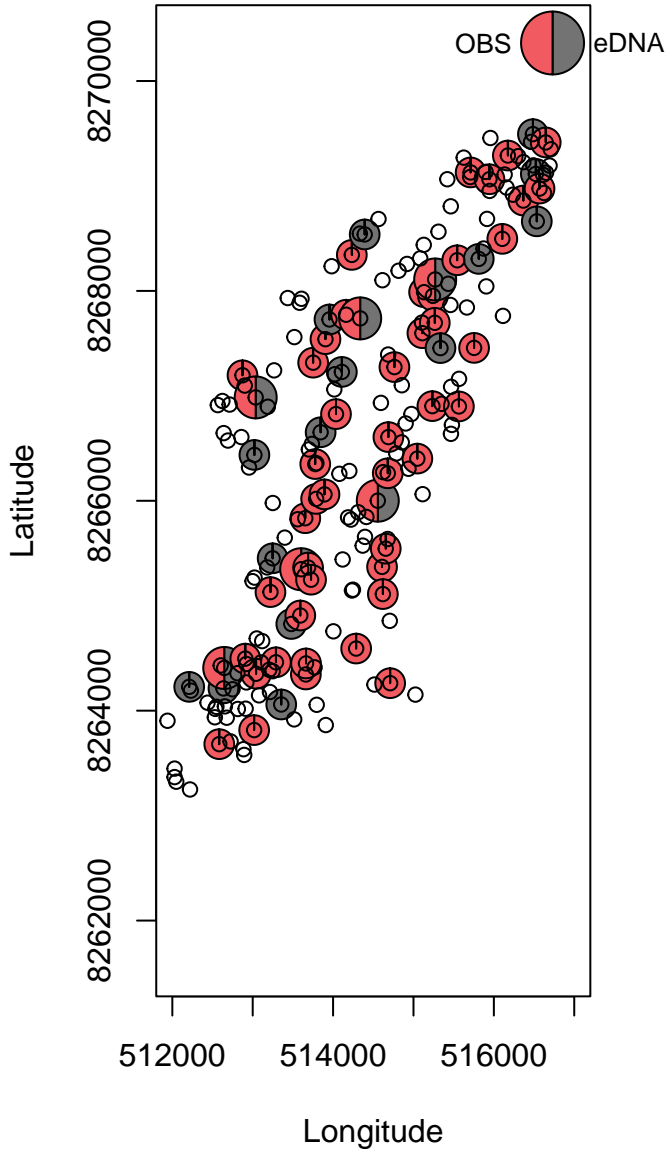
Silene



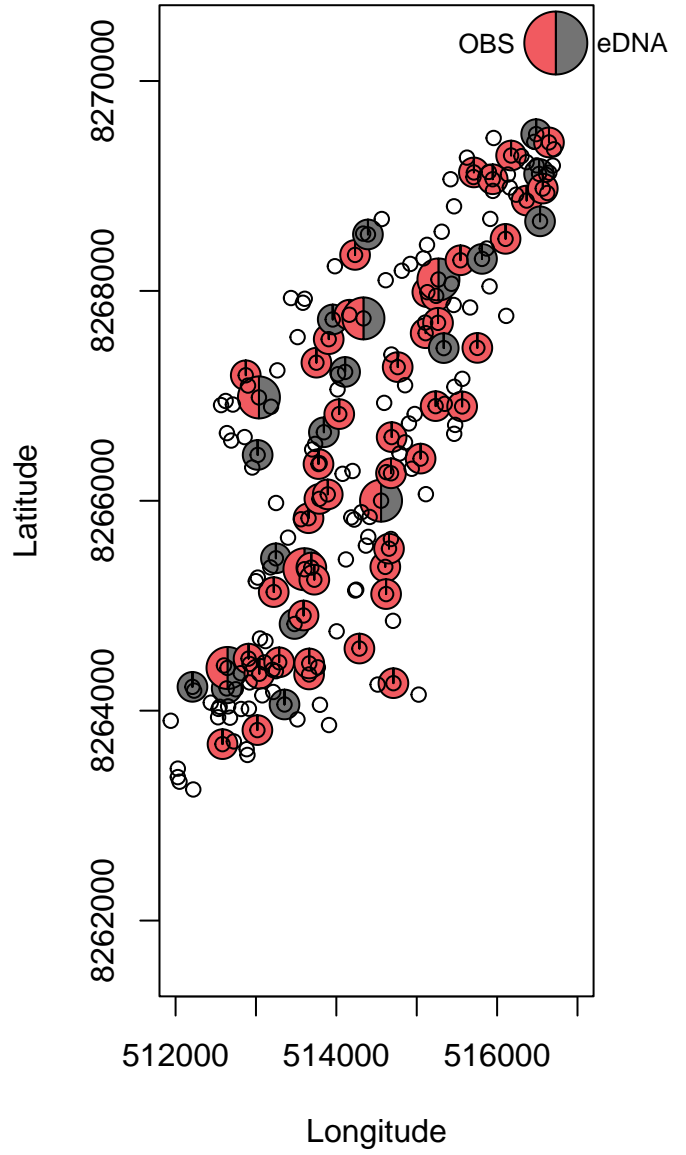
Silene_aucaulis



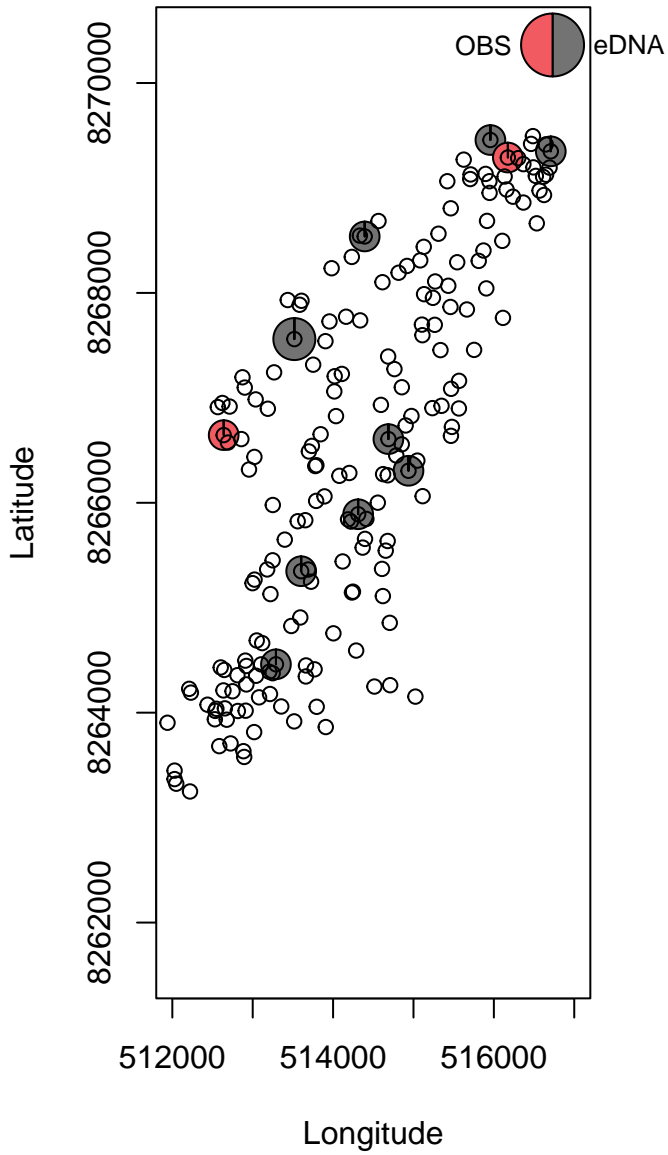
Stellaria



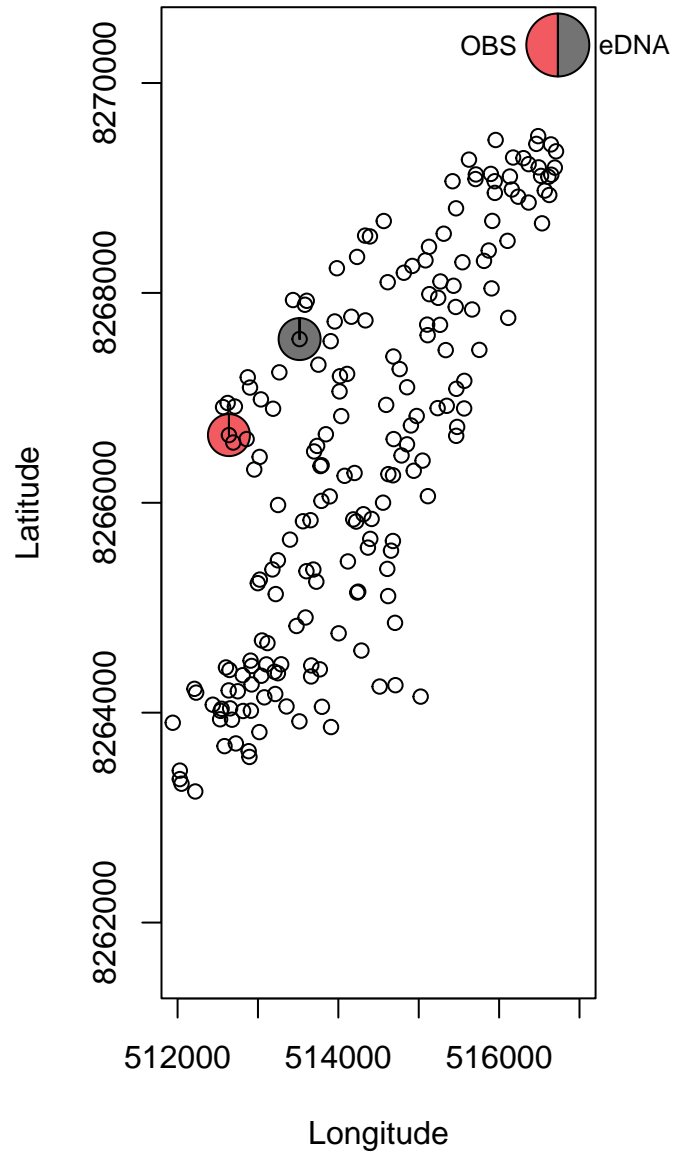
Stellaria_longipes



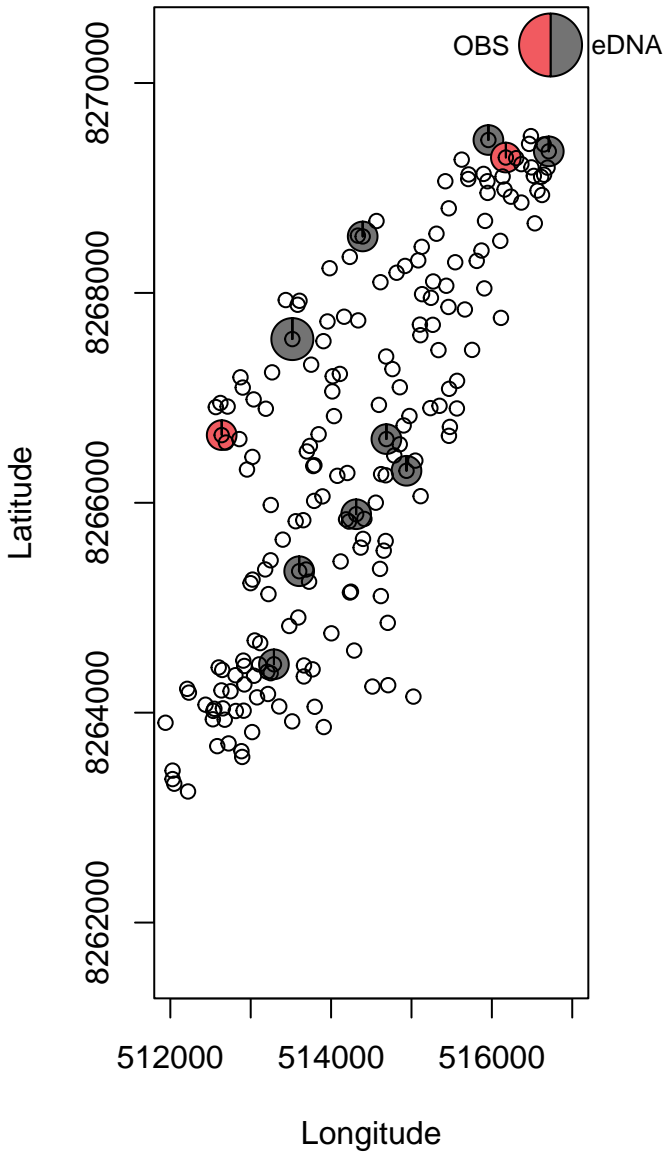
Taraxacum



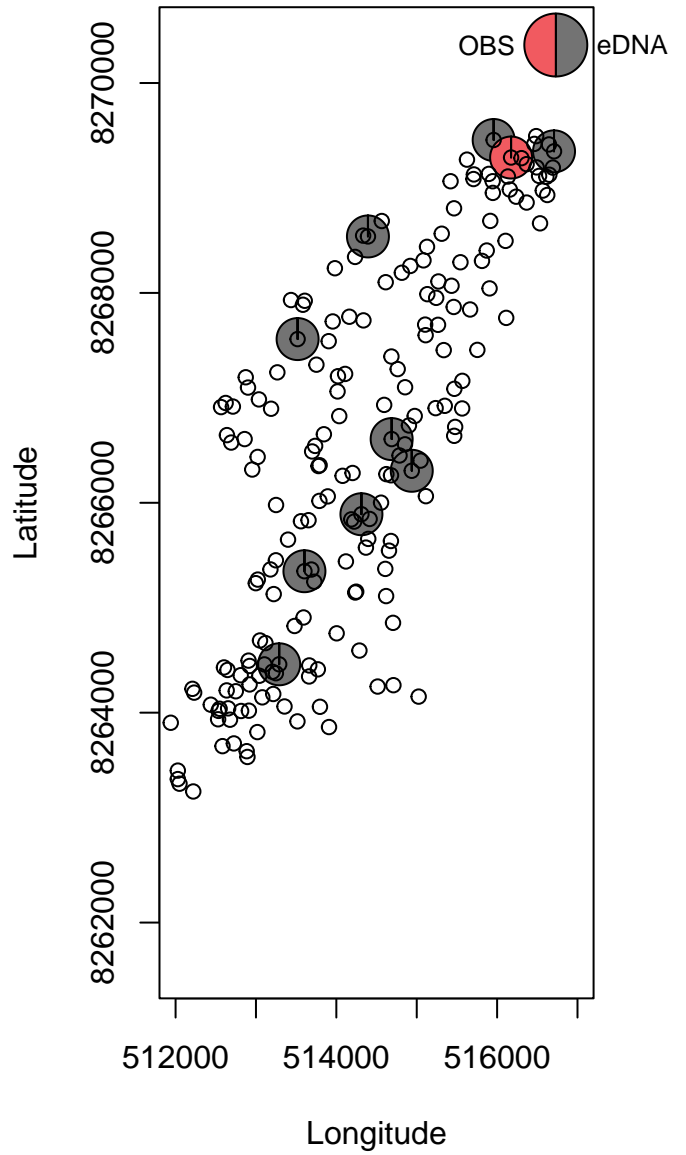
Taraxacum_arcticum



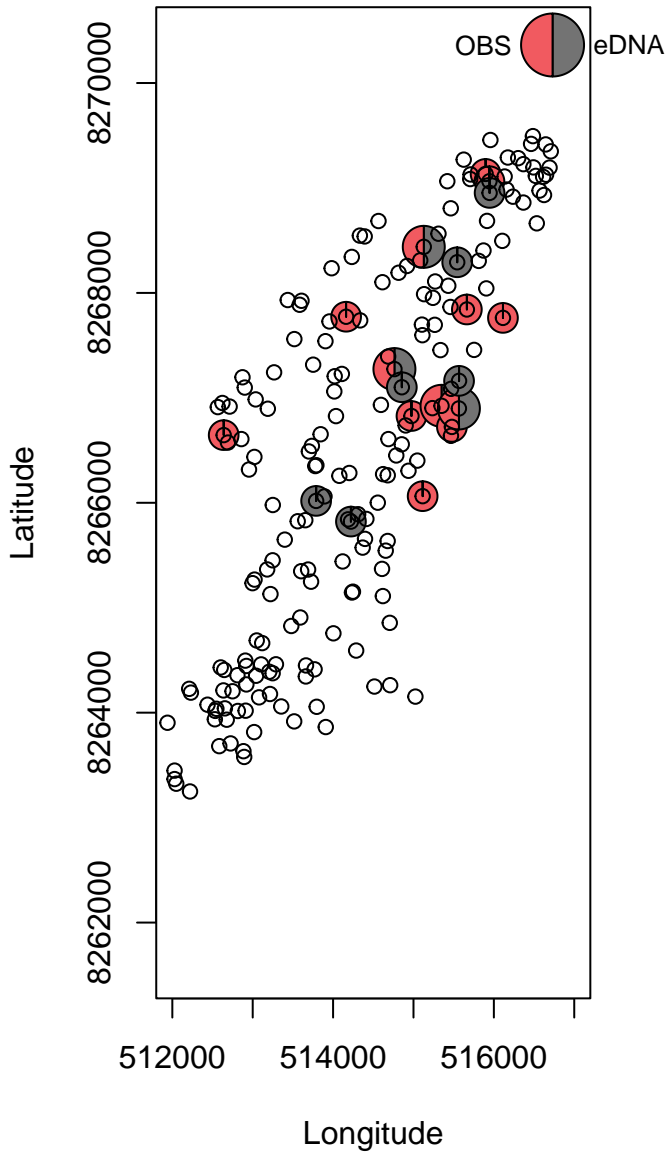
Taraxacum



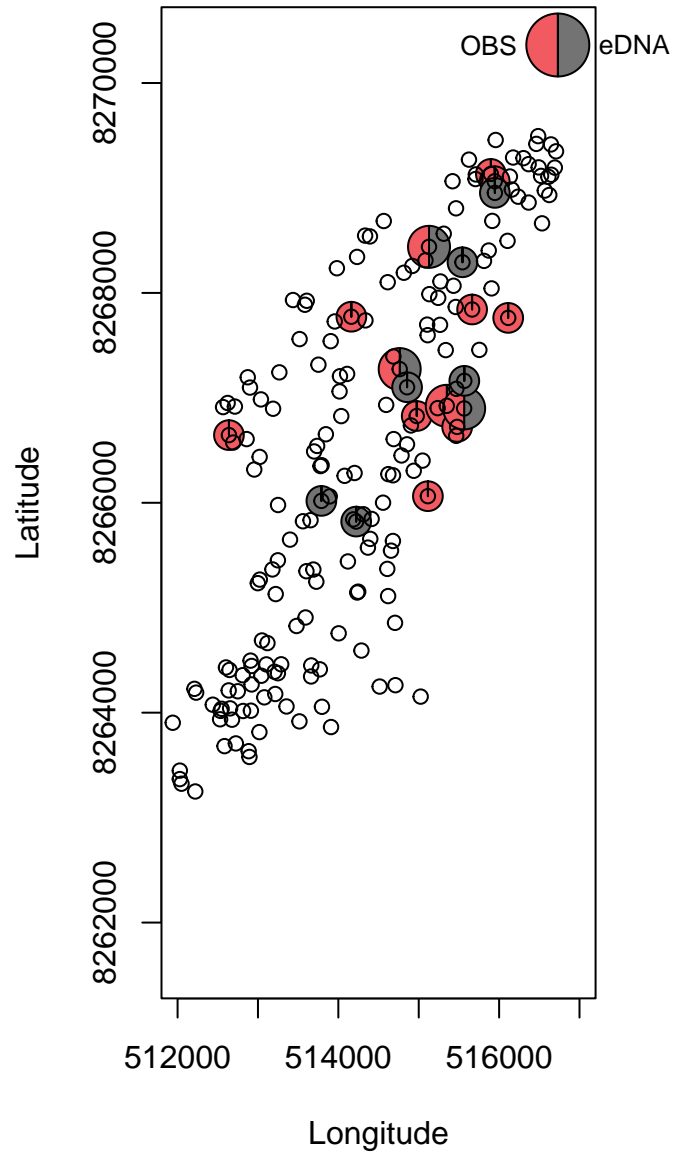
Taraxacum_phymatocarpum



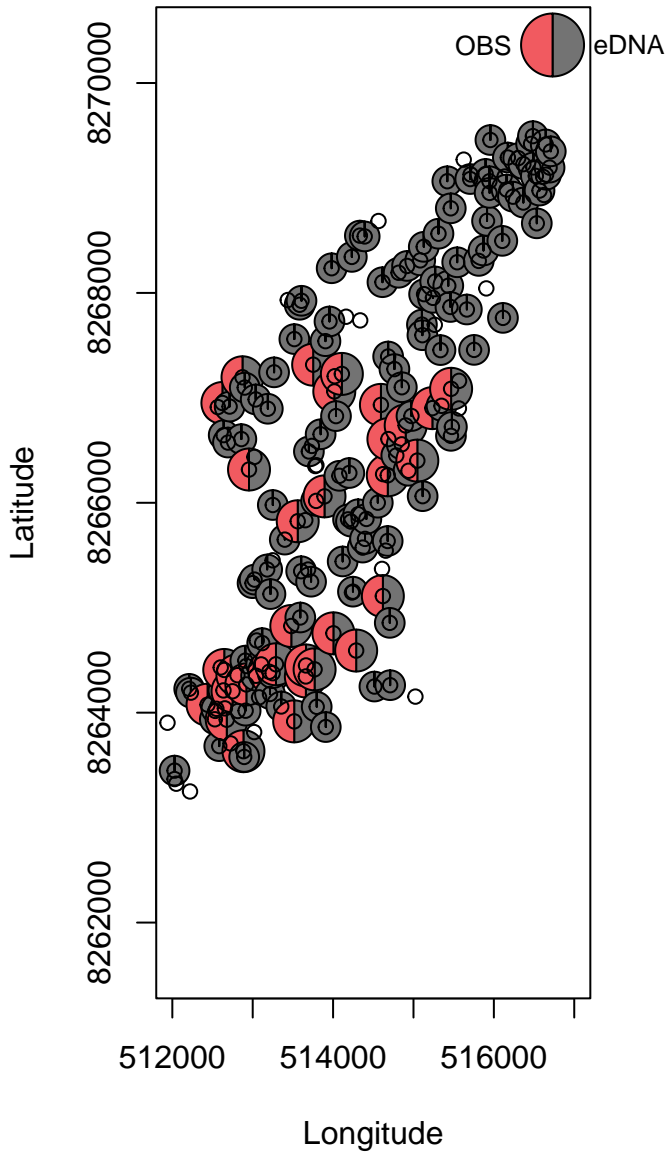
Trisetum



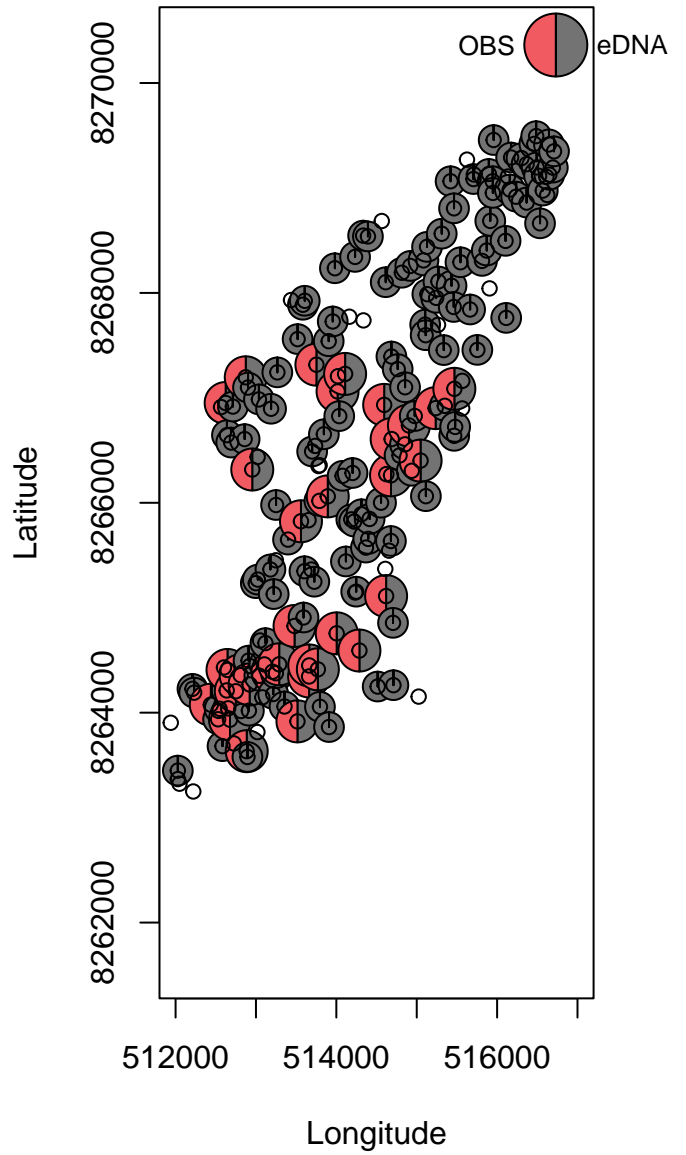
Trisetum_spicatum



Vaccinium



Vaccinium_uliginosum



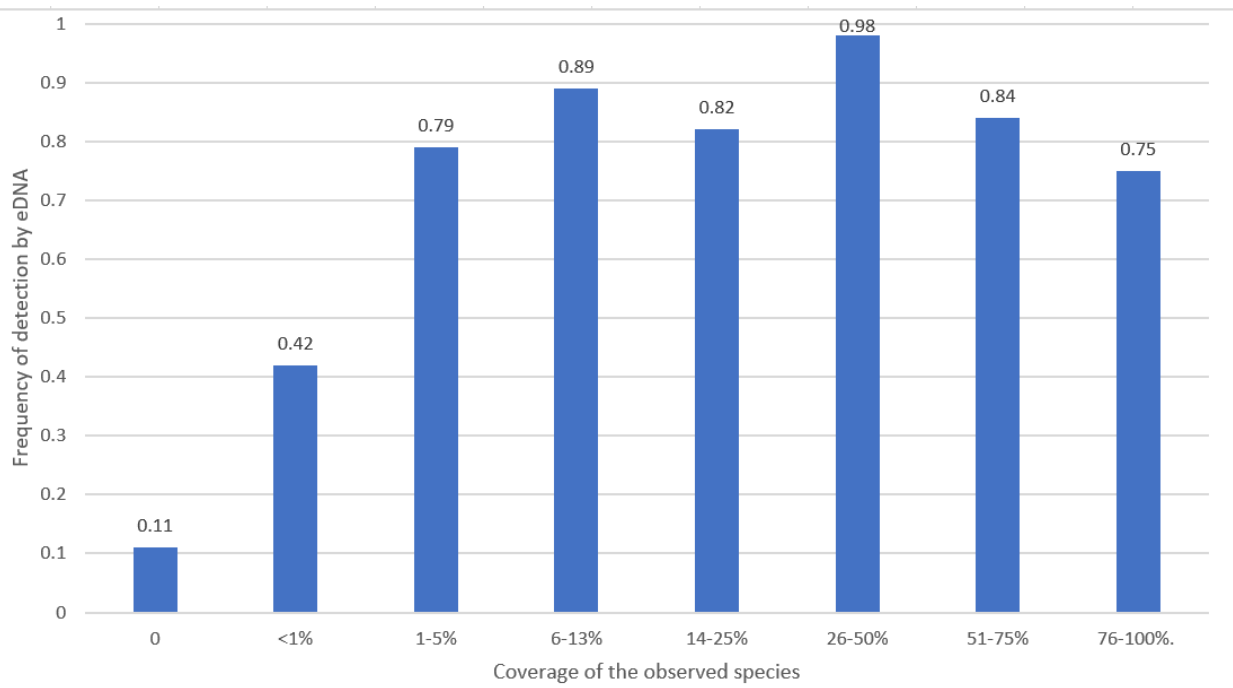


Fig. S4. Summary of the frequency of detection of a species by eDNA (y-axis) as a function of the semiquantitative scoring of the relative coverage of the same species (x-axis) as scored by direct observation.

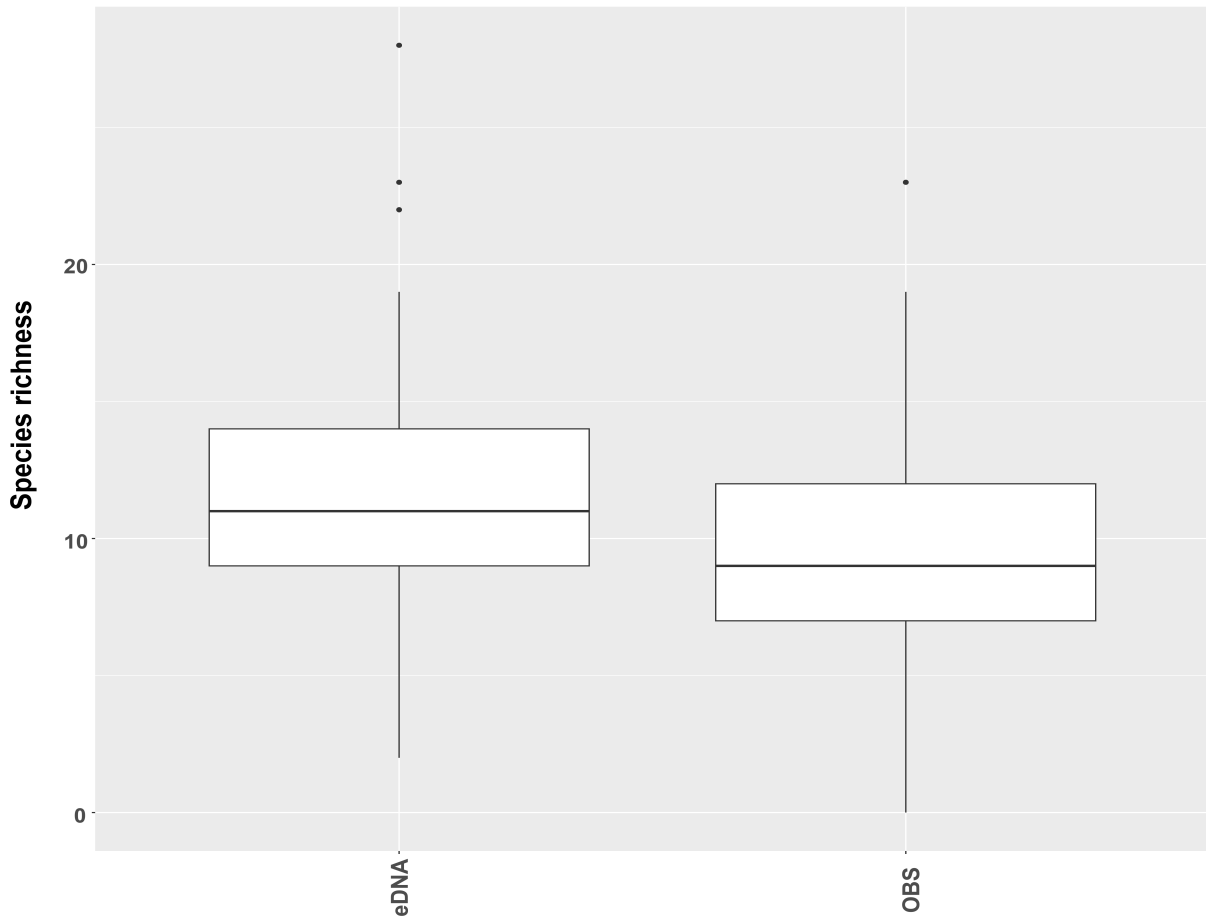
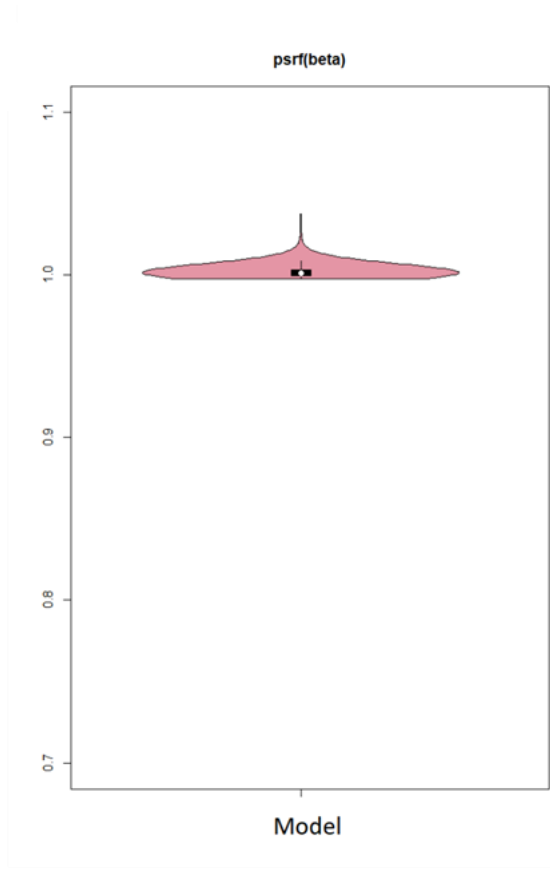


Fig. S5. Boxplot of plant species richness per plot for the two methods of identification. OBS stands for the species richness of plants scored by observation and eDNA for the richness plants by combining ITS2 and rbcLa.

A)

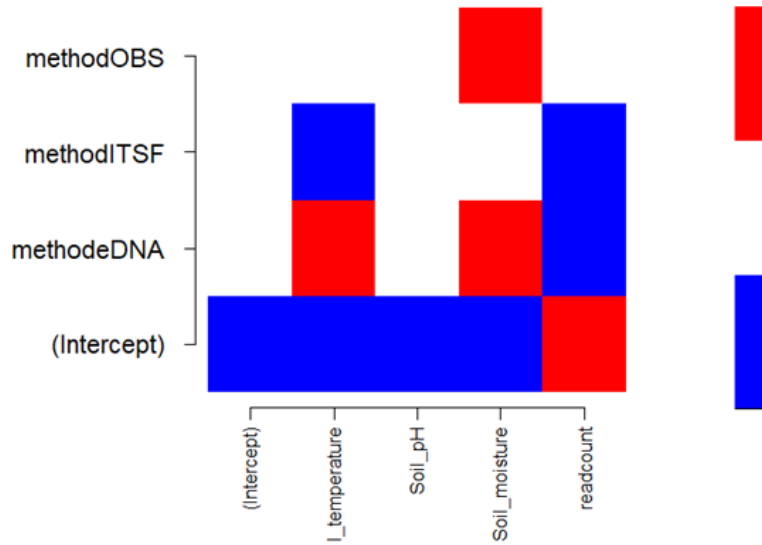


B)

		Model performance indices				Predictors and raw variance explained					
		Tjur R ²	AUC	Cross-validation TjurR ²	Cross-validation AUC	Soil temperature	Soil pH	Soil moisture	Readcount	Random: site	Random: Soil type
Model	Plant-observation	0.20	0.84	0.1	0.72	1.4	1.9	4.6	0.0	8.4	3.8
	Plant-eDNA	0.08	0.75	0.03	0.56	0.6	0.7	1.2	2.3	3.0	0.6
	Bacteria	0.38	0.92	0.09	0.71	0.7	1.0	3.9	6.0	25.5	1.0
	Fungi	0.17	0.85	0.04	0.64	1.6	0.7	3.1	1.8	11.7	1.1
	Total mean model	0.30	0.89	0.08	0.69	0.8	1.1	3.6	4.3	19.3	1.3

Fig. S6. Model convergence and discrimination success achieved for the HMSC model. In (A), the violin plot describes the MCMC convergence of the model. The potential scale reduction factor was close to the theoretical optimum of one. The left-hand part of Table (B) summarizes the discriminatory power of the model, as based on two indices: Tjur R² and AUC (see main text for definitions). Two aspects are evaluated: explanatory power (as reflected by Tjur R² and AUC) and predictive power (as reflected by cross-validation Tjur R² and cross-validation AUC). The right-hand part of the Table shows the average proportion of variation explained by each variable included in the model.

A)



B)

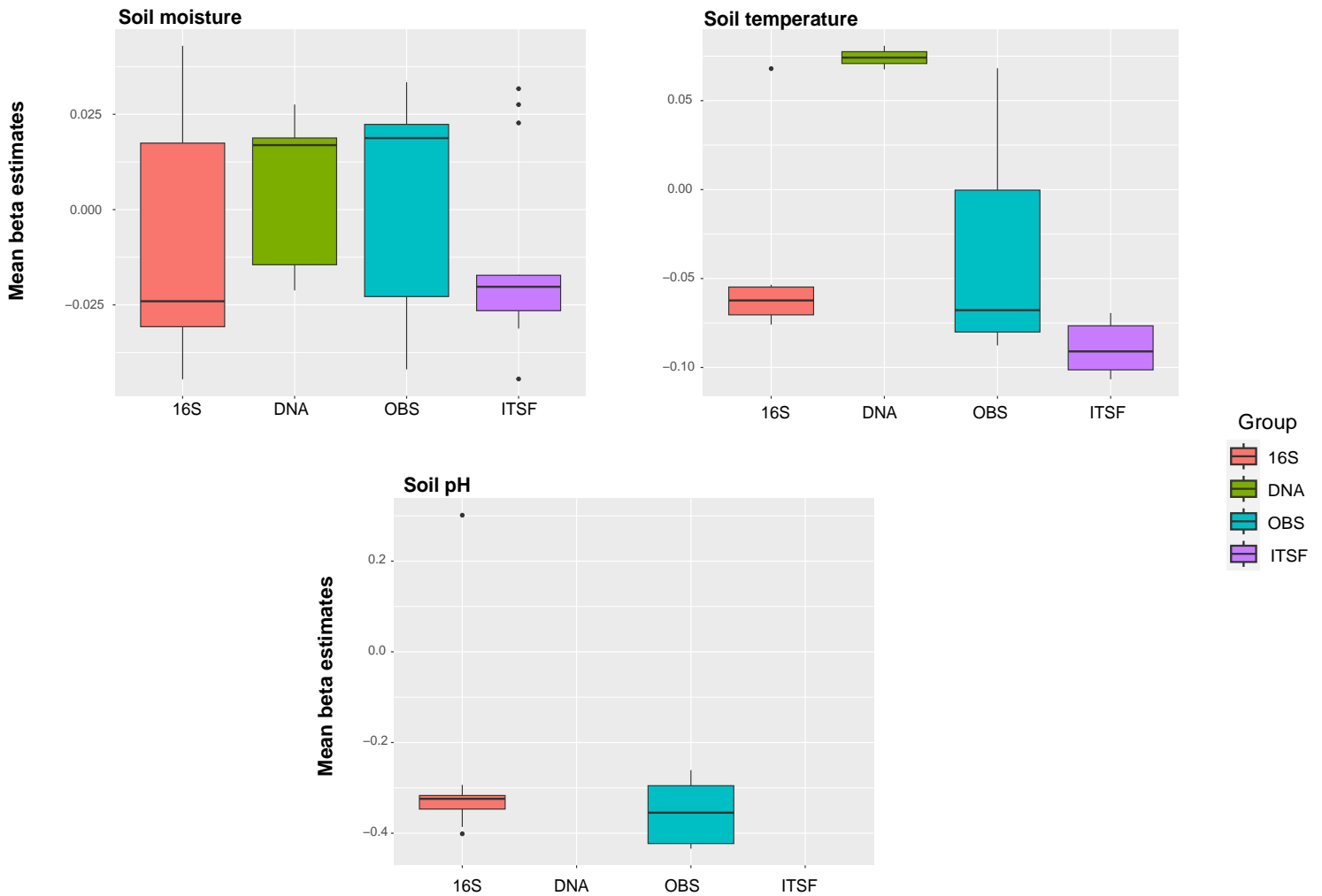
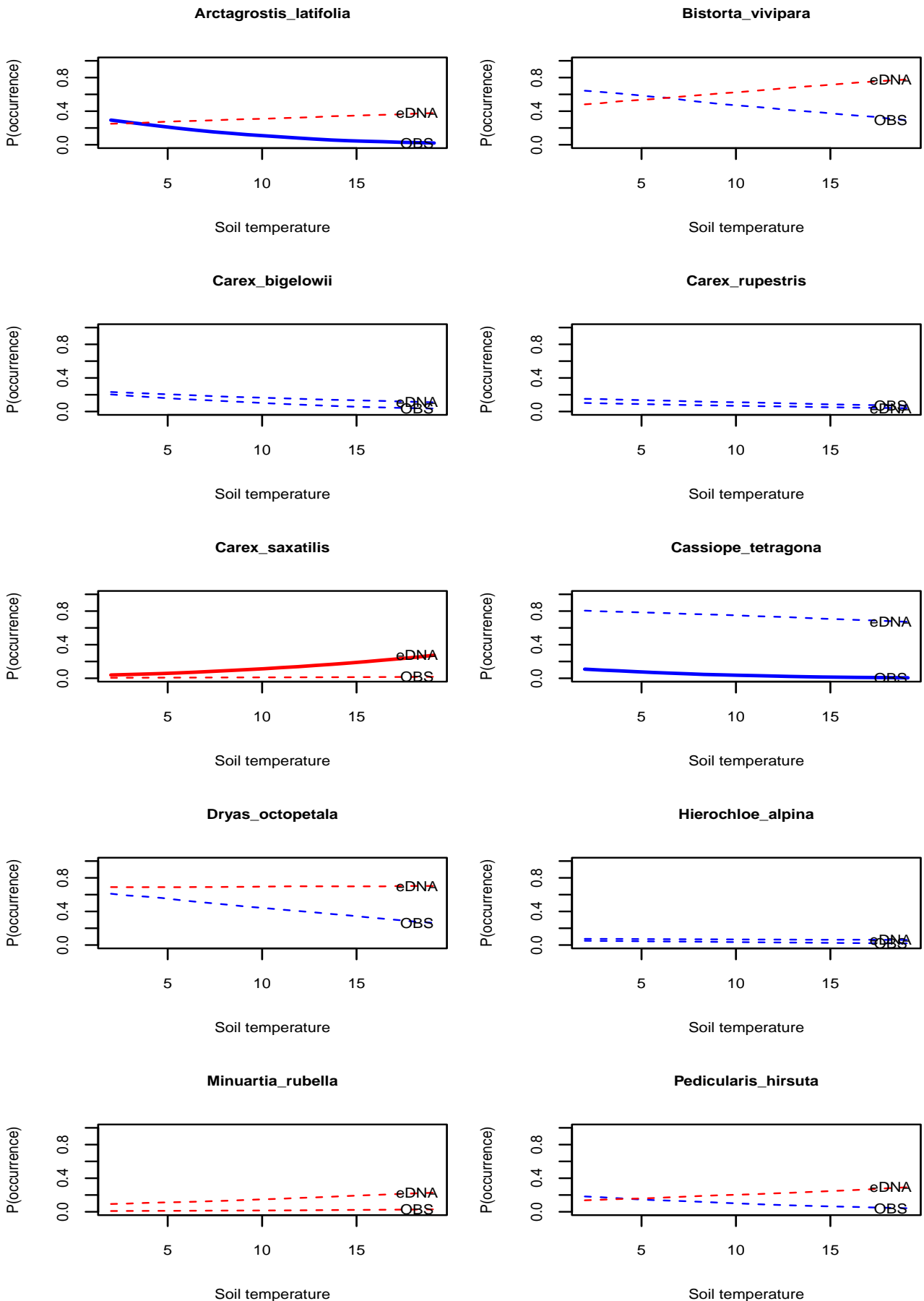


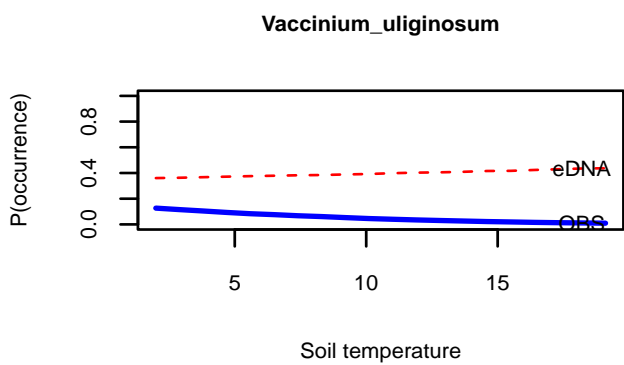
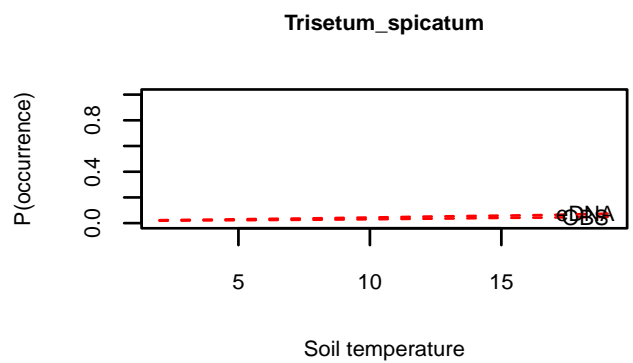
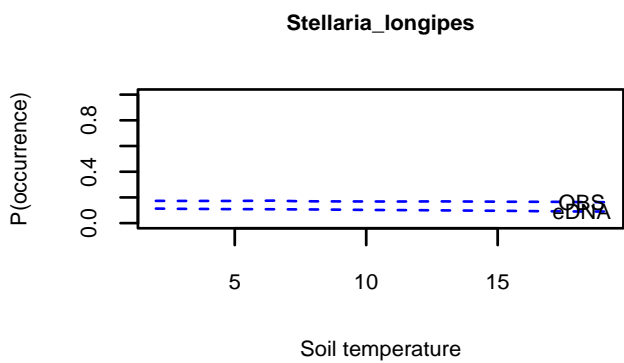
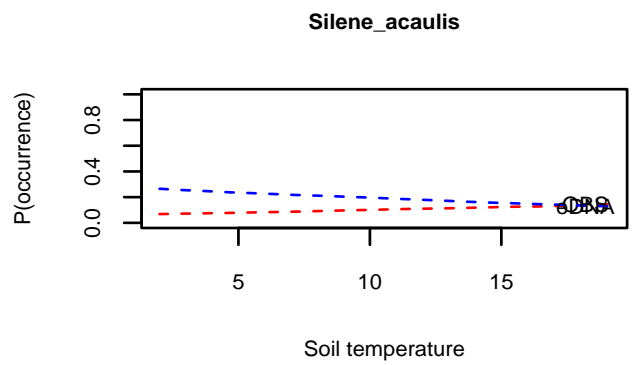
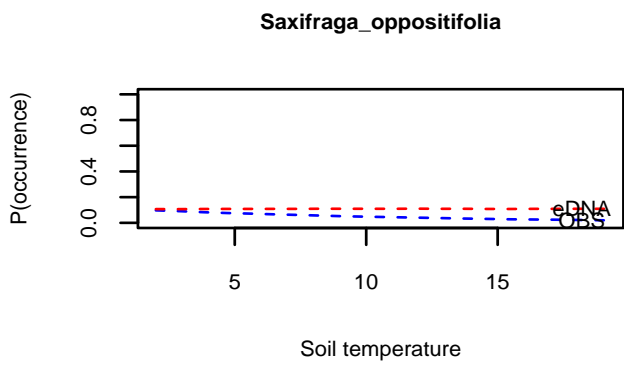
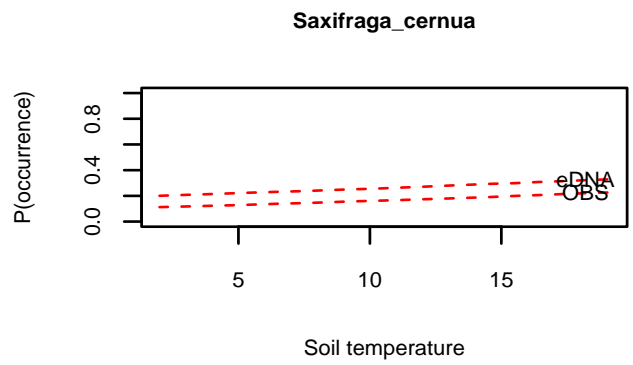
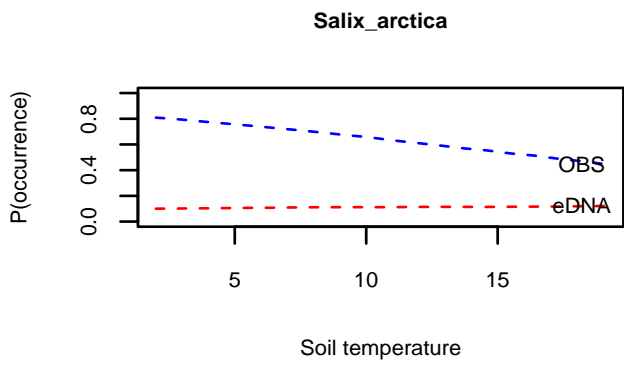
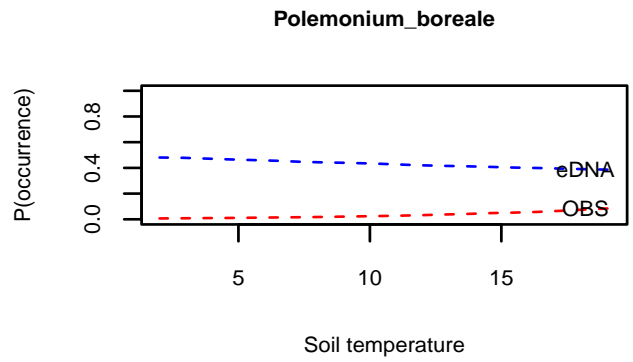
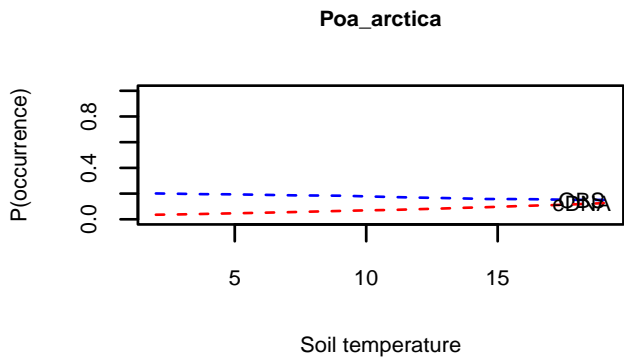
Fig. S7. The top panel (A) shows estimates of the gamma parameter, as reflecting the impact of the trait (i.e. organism group or scoring method) on the estimated response. The intercept corresponds to bacteria detected using locus 16S. Blue or red tiles indicate, when compared to bacteria, that a group of organisms is responding stronger or weaker to a specific covariate (posterior support >0.95). To illustrate the corresponding differences in species-specific responses within the respective groups, we show the distribution of beta parameter (i.e., taxa specific estimates of environmental responses, equivalent to regression coefficients) values as a box plot in panel (B).

Fig. S8. Summary of taxon-specific responses to environmental covariates. Cell entries correspond to the number of taxa for which a statistically supported response was detected. The scoring of taxa is based on a statistically supported beta-parameter (i.e) in the HMSC model, akin to a regression coefficient delectably different from zero. Overall, the analysis includes 44 plant taxa scored by direct observation (Plant_OBS), 37 plant taxa observed by eDNA (Plant_eDNA), 222 bacterial OTUs (Bacteria) and 29 fungal taxa (Fungi).

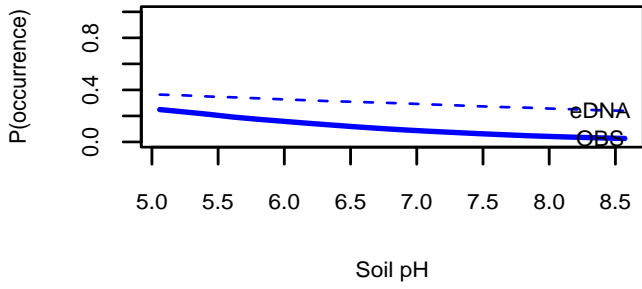
Taxonomic group	Sign of response	Environmental covariables			
		Soil temperature	Soil moisture	Soil pH	Readcount
Plant_OBS	+	5	15	0	0
	0	36	37	33	0
	-	2	10	9	0
Plant_eDNA	+	2	4	0	13
	0	35	36	37	19
	-	0	3	0	5
Bacteria	+	1	29	1	211
	0	217	127	213	12
	-	5	67	9	0
Fungi	+	0	3	0	15
	0	20	16	29	14
	-	9	10	0	0

Fig. S9. Posterior predictions of the mean probabilities of occurrence of 19 plant species based on observational (OBS) vs eDNA data along gradients in soil temperature, pH and moisture. These are marginal predictions, meaning that the values of all fixed effects apart from the focal soil gradient were fixed at their mean value in the dataset.

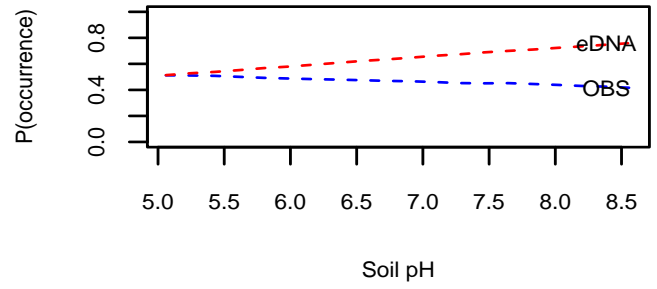




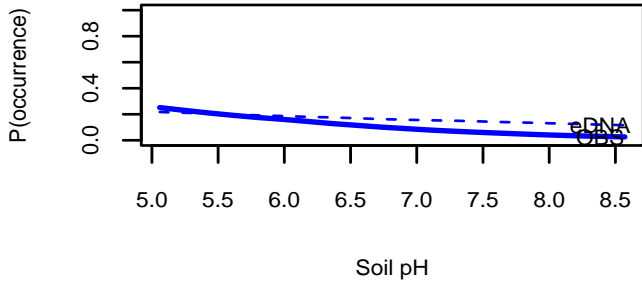
Arctagrostis_latifolia



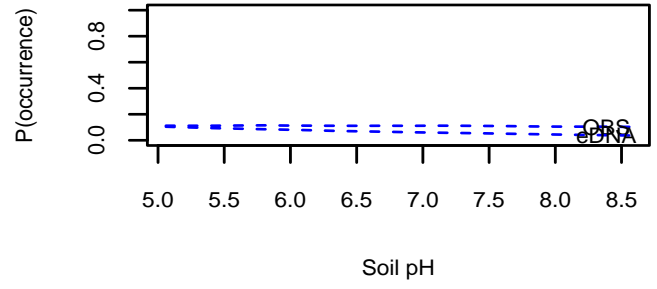
Bistorta_vivipara



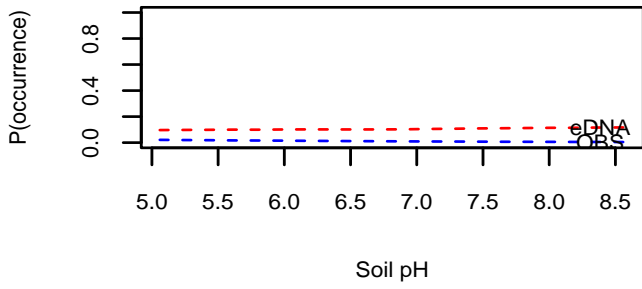
Carex_bigelowii



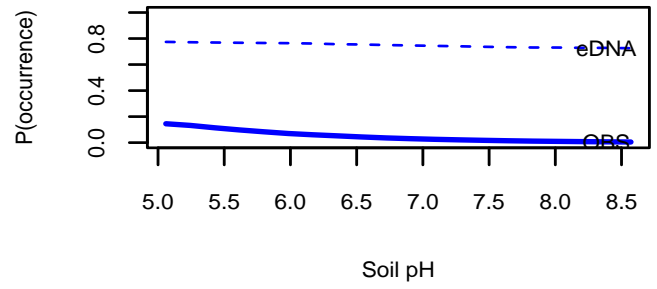
Carex_rupestris



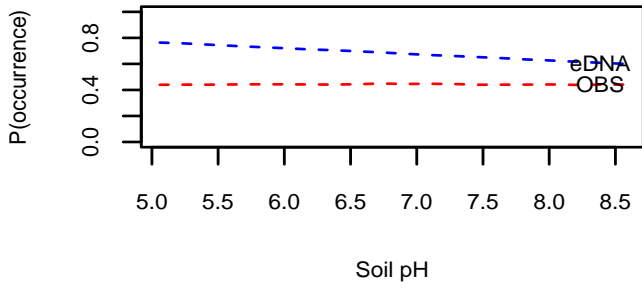
Carex_saxatilis



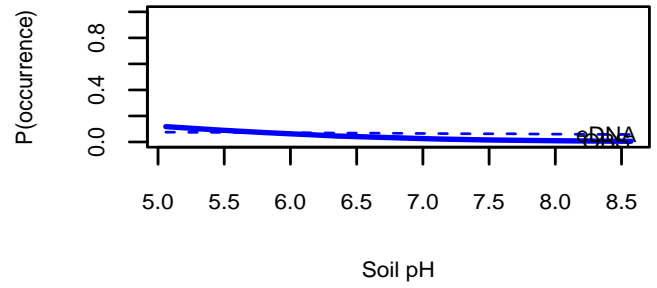
Cassiope_tetragona



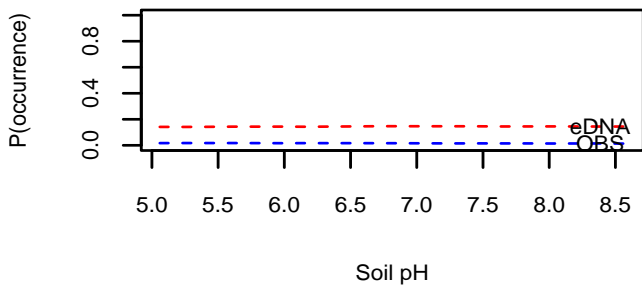
Dryas_octopetala



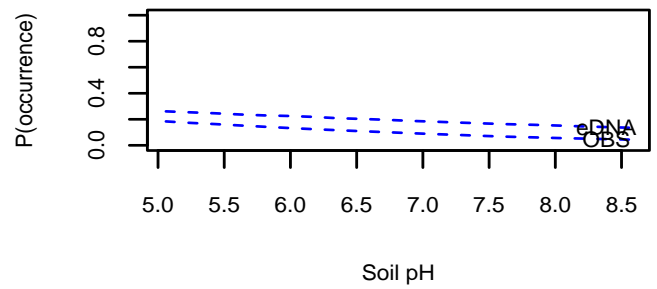
Hierochloe_alpina

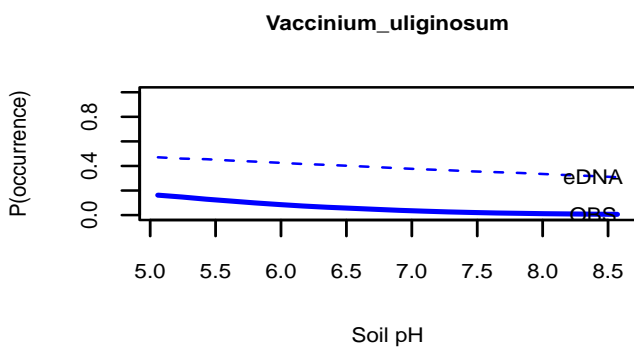
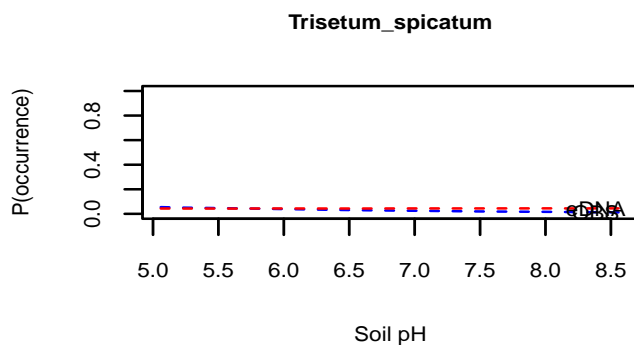
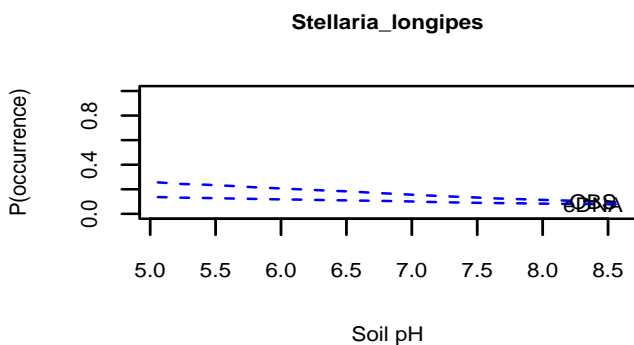
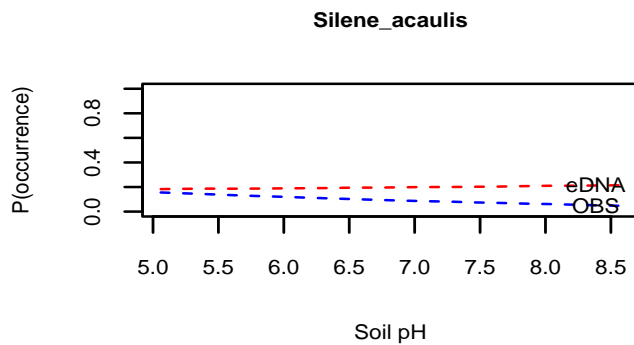
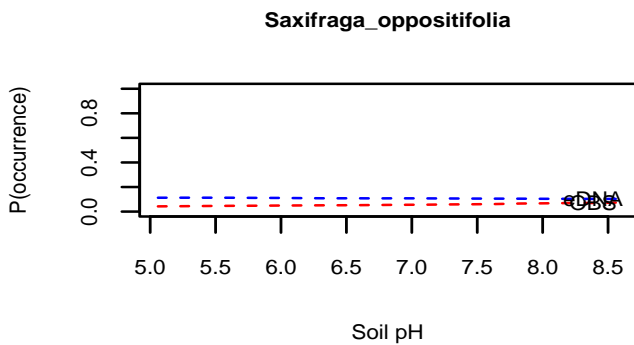
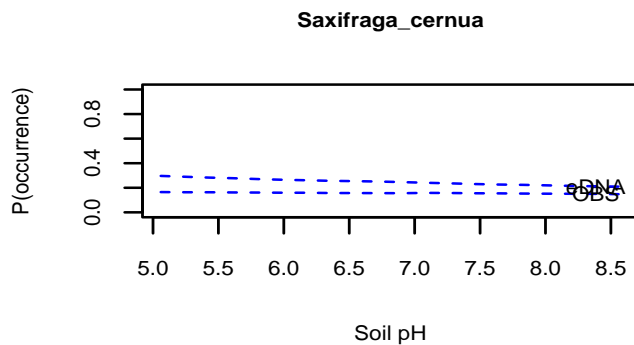
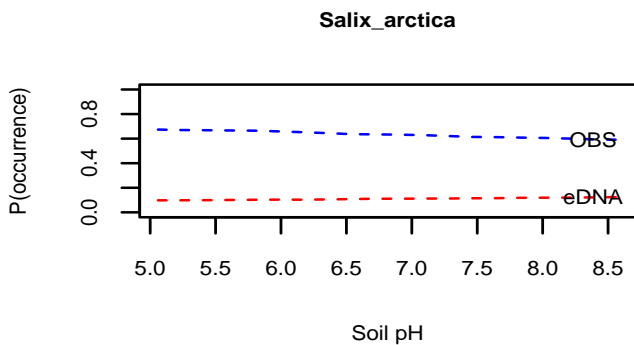
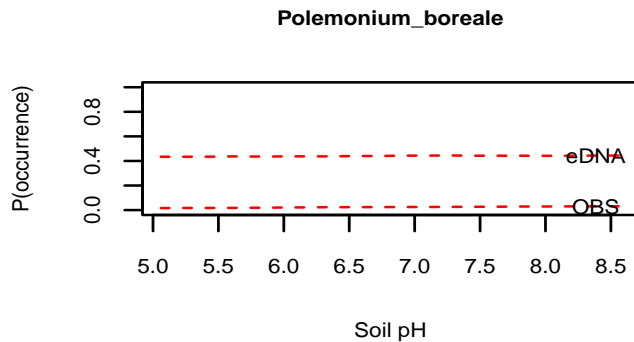
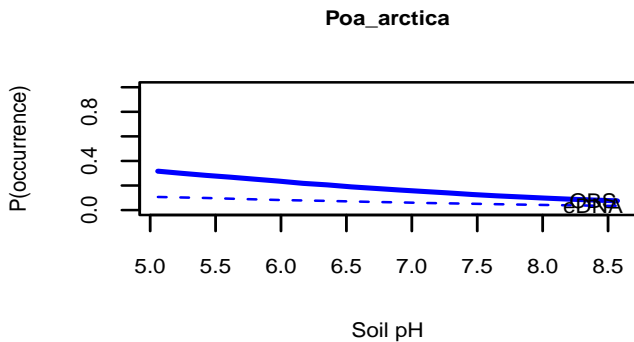


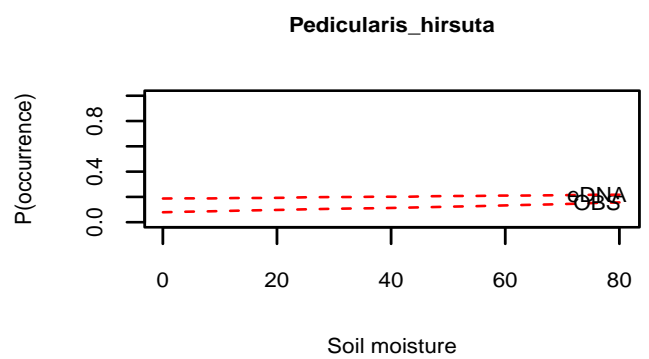
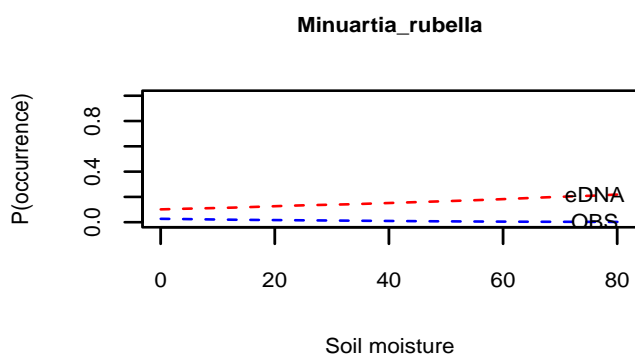
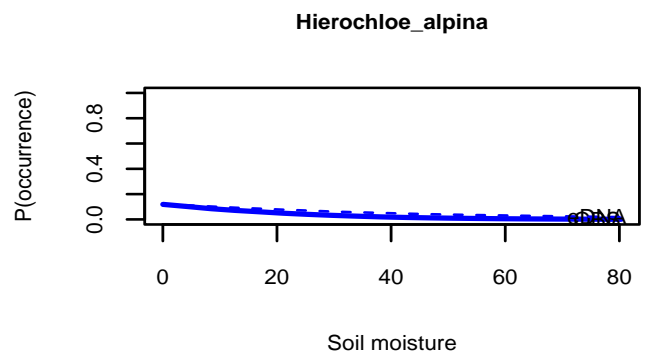
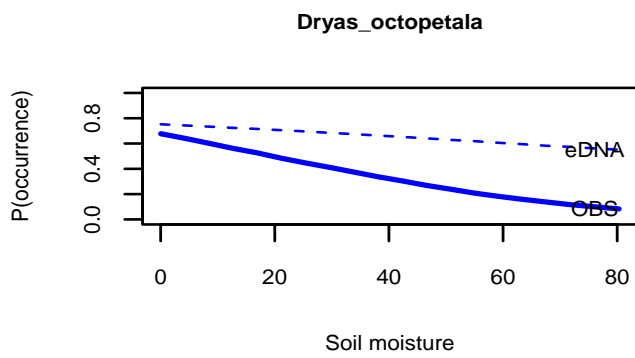
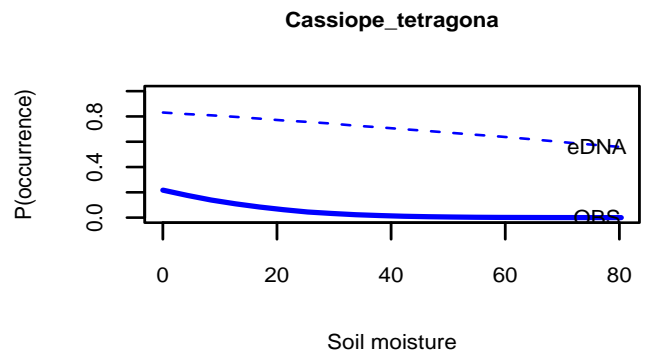
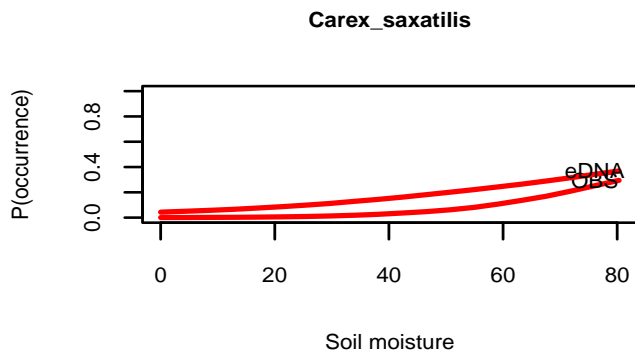
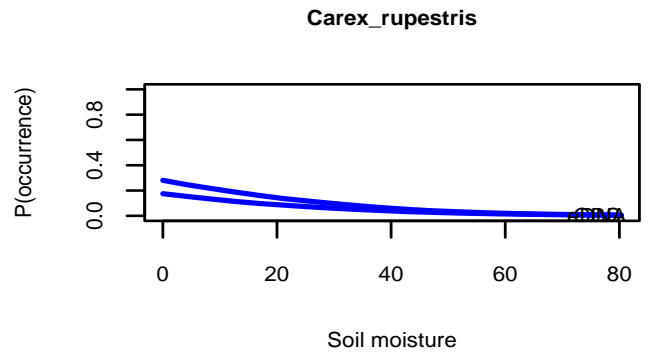
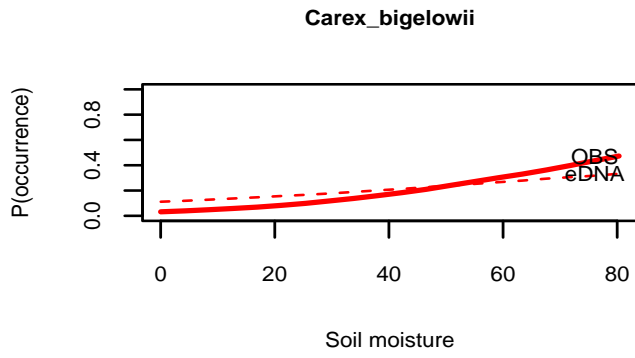
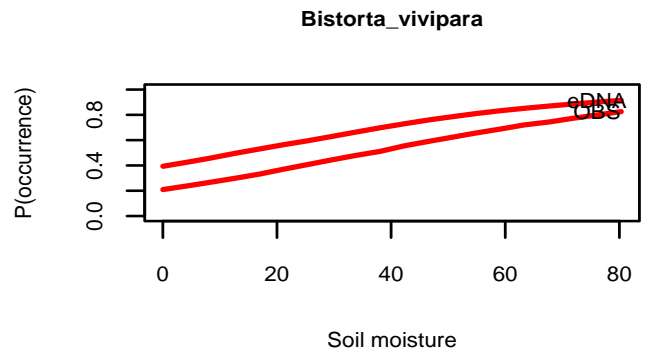
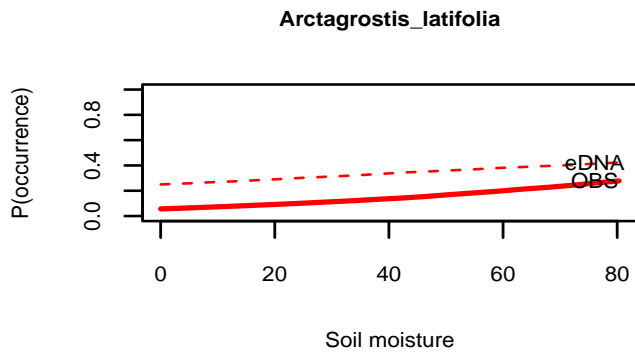
Minuartia_rubella



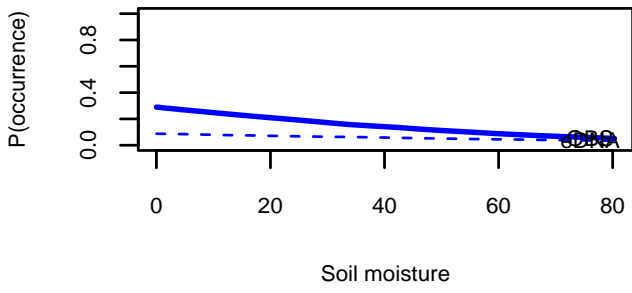
Pedicularis_hirsuta



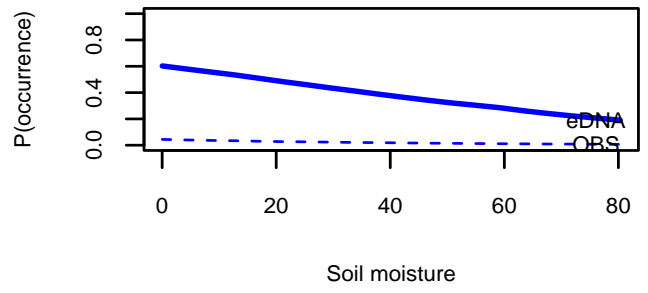




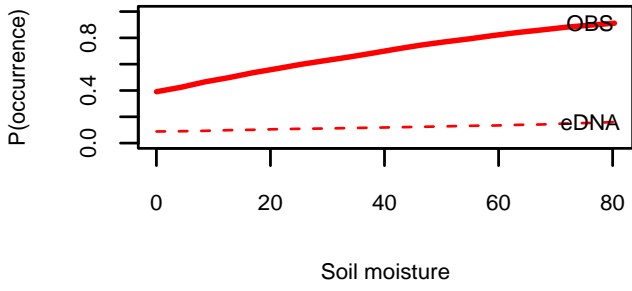
Poa_arctica



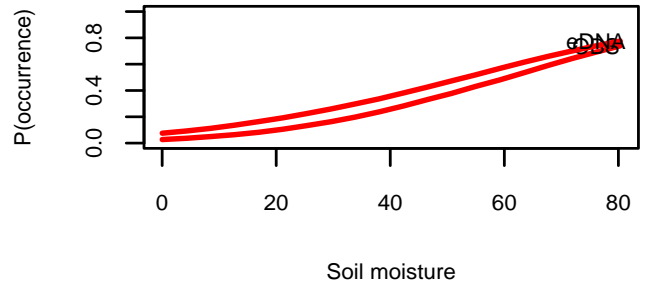
Polemonium_boreale



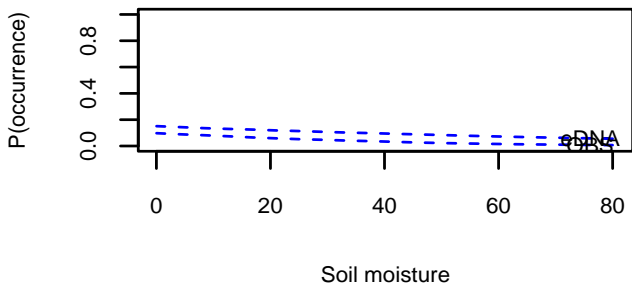
Salix_arctica



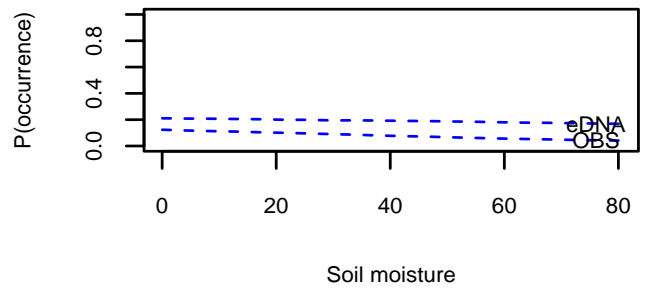
Saxifraga_cernua



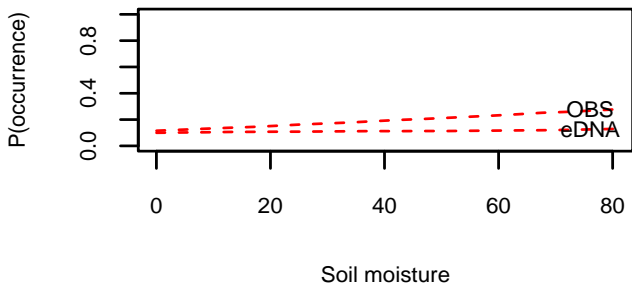
Saxifraga_oppositifolia



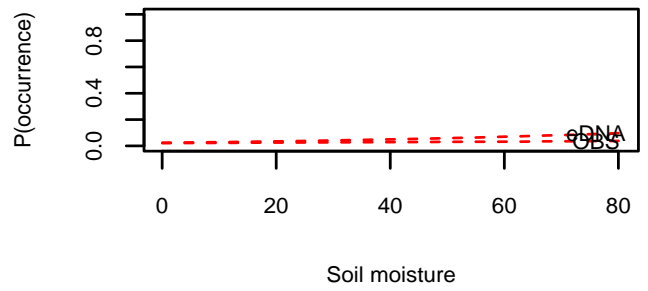
Silene_aucaulis



Stellaria_longipes



Trisetum_spicatum



Vaccinium_uliginosum

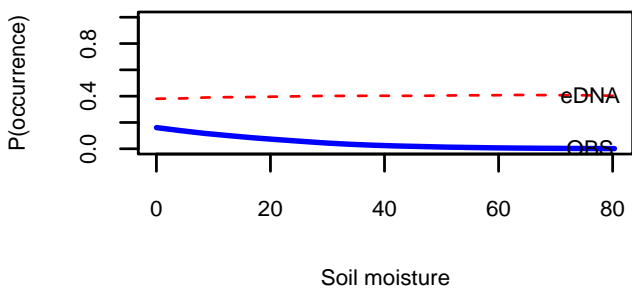


Fig. S10. Numerical summary of associations detected between taxa. The top panel (A) summarizes all possible associations between taxa included in our model. Proportions are calculated from the numbers in brackets, as based on the number of statistically supported associations out of all possible taxon-pairs between each group. The bottom panel (B) summarizes the number of associations detected between taxa, as visually represented in Fig. 5 of the main text and in Fig S11-A&B (below). Proportions are calculated from the numbers in brackets, as based on the number of statistically supported associations out of all possible taxon-pairs for each type of association.

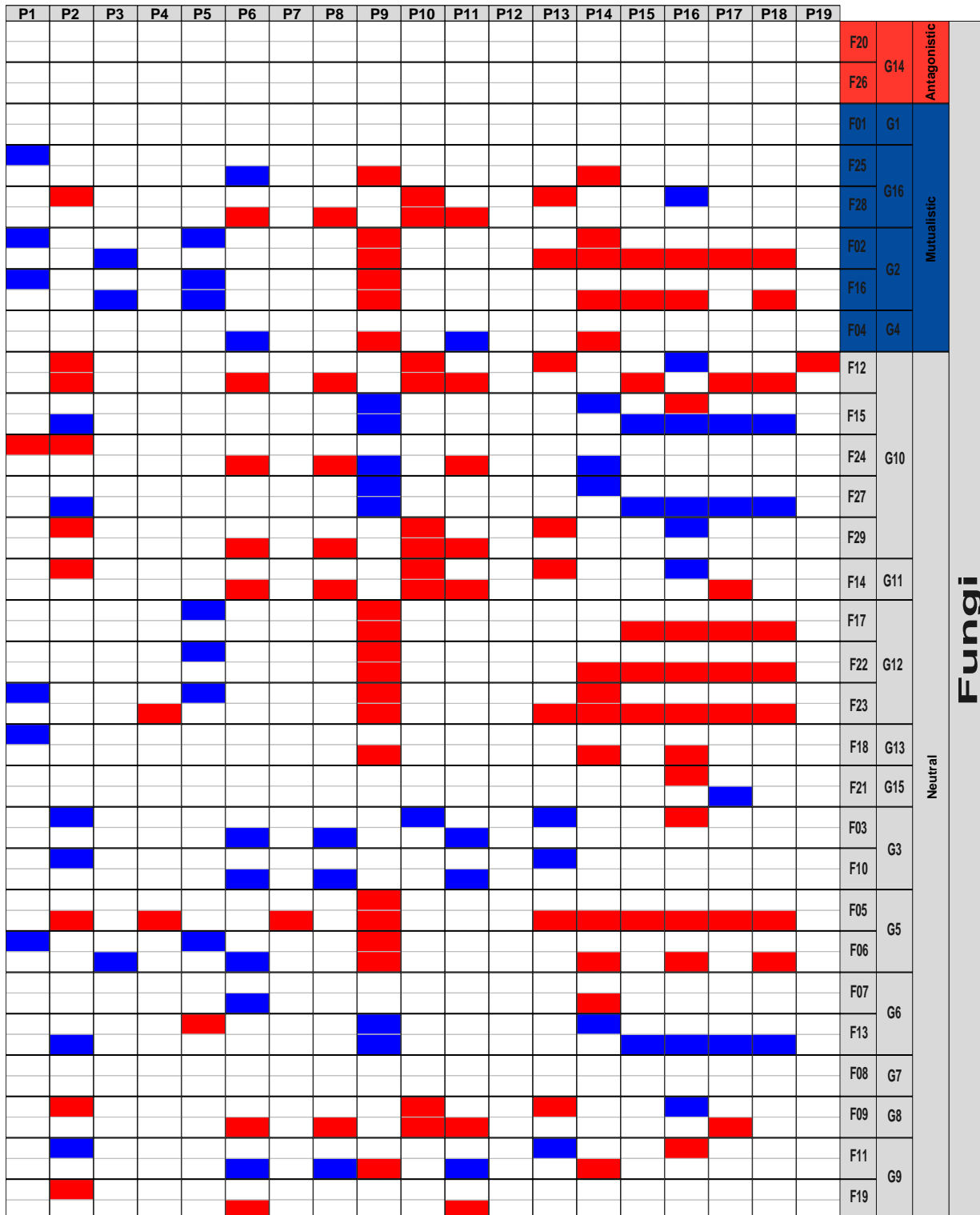
A)

		Plant_OBS (n=44)	Plant_eDNA (n=37)	Bacteria (n=222)	Fungi (n=29)
Plant_OBS (n=44)	Total	21% (406/1936)	6% (103/1628)	33% (3236/9812)	27% (343/1276)
	Negative	5% (104/1936)	2% (32/1628)	14% (1369/9812)	9% (122/1276)
	Positive	16% (302/1936)	4% (71/1628)	18% (1867/9812)	17% (221/1276)
Plant_eDNA (n=37)	Total		6% (87/1369)	15% (1276/8251)	13% (137/1073)
	Negative		1% (16/1369)	7% (575/8251)	5% (55/1073)
	Positive		5% (71/1369)	8% (701/8251)	8% (82/1073)
Bacteria (n=222)	Total			59% (29203/49729)	44% (2866/6467)
	Negative			28% (13912/49729)	20% (1286/6467)
	Positive			31% (15291/49729)	24% (1580/6467)
Fungi (n=29)	Total				38% (317/841)
	Negative				16% (132/841)
	Positive				22% (185/841)

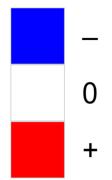
B)

		Proportion of supported associations between the subset of 19 plant taxa and other organism groups					
		Bacteria (n=222)			Fungi (n=29)		
		Antagonistic	Mixed	Mutualistic	Antagonistic	Mixed	Mutualistic
Plant_OBS (n=19)	Total	29.7% (192/646)	24.7% (913/3705)	27.3% (166/608)	0% (0/38)	24.8% (99/399)	22.8% (26/114)
	Negative	9.9% (64/646)	9.5% (351/3705)	11.2% (68/608)	0% (0/38)	8.3% (33/399)	5.3% (6/114)
	Positive	19.8% (128/646)	15.2% (562/3705)	16.1% (98/608)	0% (0/38)	16.5% (66/399)	17.5% (20/114)
Plant_eDNA (n=19)	Total	14.2% (92/646)	13% (481/3705)	14.7% (89/608)	0% (0/38)	12.8% (51/399)	10.6% (12/114)
	Negative	4.6% (30/646)	6.1% (227/3705)	6.6% (40/608)	0% (0/38)	6% (24/399)	5.3% (6/114)
	Positive	9.6% (62/646)	6.9% (254/3705)	8.1% (49/608)	0% (0/38)	6.8% (27/399)	5.3% (6/114)

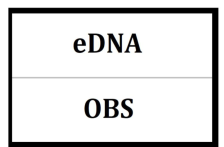
Fig. S11A. Estimated pairwise residual associations among plants and different functional groups of fungi. Here, each plant species is shown as a column, including the 19 plant species identified by both direct observation and eDNA and thus allowing direct comparisons between methods. Rows correspond to individual microbial genera, as sorted by functional groups. Red fields indicate presumptively antagonistic relationships, as based on the functional classification of taxa, blue fields presumptively mutualistic associations, and grey fields indicate neutral or mixed interactions (i.e. the same genus being associated with several different functions). For visual comparison, each cell is divided in two, with the upper part describing the association estimated when plant occurrence was detected by eDNA and the bottom part describing the association estimated when plant occurrence was detected by Observation. G corresponds to the functional group, F to the Fungal taxon, and P to the plants taxon. For the identity of individual taxa, see key in Supplementary Information 3. For a numerical summary of figure contents, see Fig S10-B



Sign of association

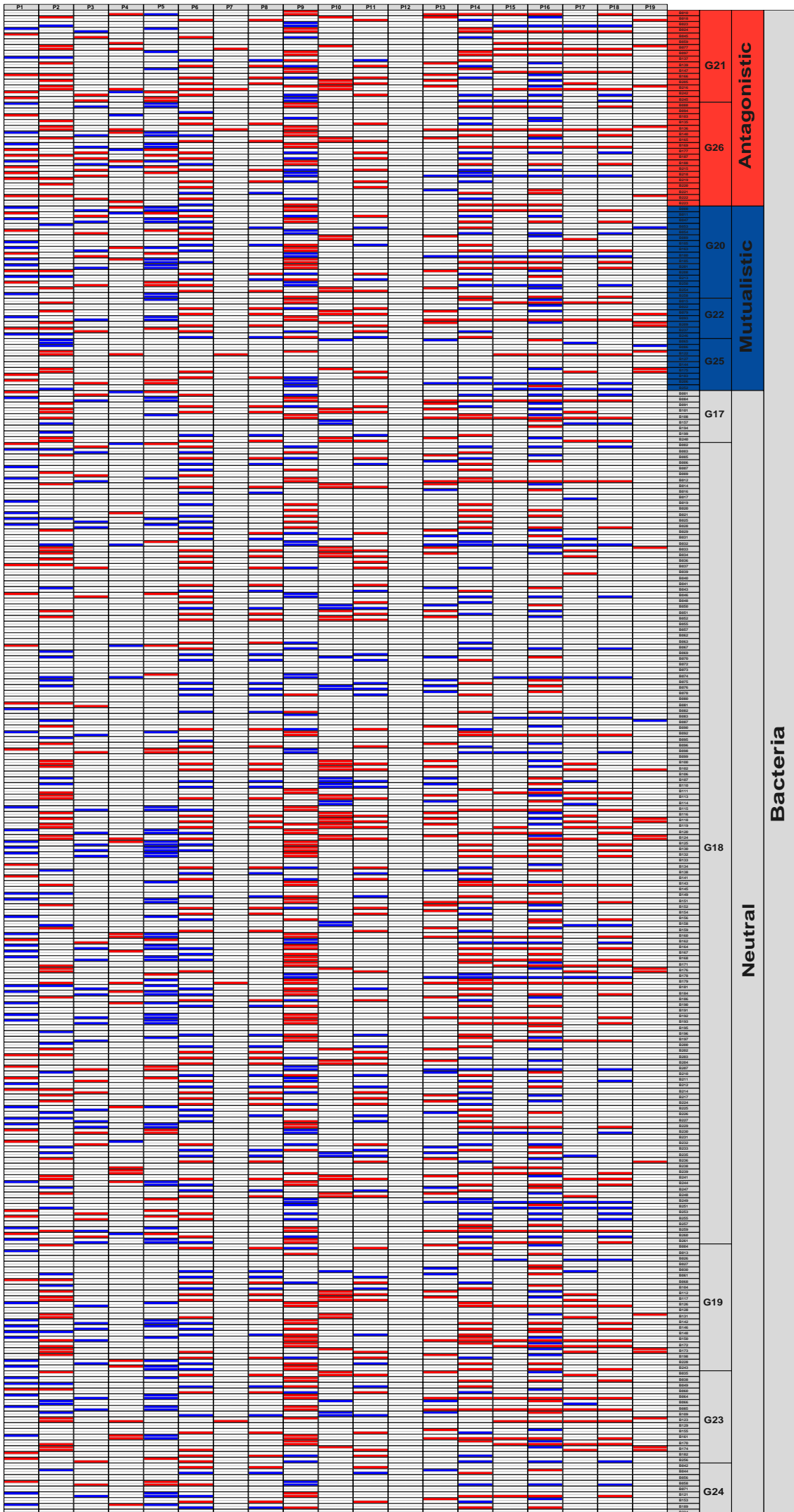


Method

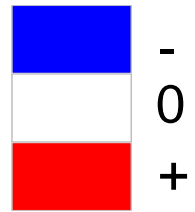


Fungi

Neutral



Sign of association



Method

