[*GEB Data papers* - These are short papers (typically 2000 words excluding the abstract, and two figures) that present datasets of broad macroecological interest. The data must be made public at time of publication, by depositing them in a stable online repository. Please use a structured Abstract, not longer than 300 words, with the following headings: Motivation, Main types of variable contained, Spatial location and grain, Time period and grain, Major taxa and level of measurement, Software format.]

Tundra Trait Team: A database of plant traits spanning the tundra biome

Authors: Anne D. Bjorkman, Isla Myers-Smith, Sarah C. Elmendorf, Signe Normand, Haydn Thomas, and the Tundra Trait Team (in order of the number of observations contributed?)

ABSTRACT

Motivation: The Tundra Trait Team (TTT) database includes field-based measurements of key traits related to plant form and function at sites across the tundra biome. This dataset can be used to address theoretical questions about plant strategy and trade-offs, trait-environment relationships and environmental filtering, and trait variation across scales, to validate satellite data, and to inform earth system model parameters.

Main types of variable contained: The database contains 66,308 measurements of 18 plant traits. The most frequently measured traits (>1000 observations each) include plant height, leaf area, specific leaf area, leaf fresh and dry mass, leaf dry matter content, leaf nitrogen content, leaf carbon content, leaf phosphorus content, seed mass, and stem specific density.

Spatial location and grain: Measurements were collected over the entire northern hemisphere tundra biome, including Arctic sites in Alaska, Canada, Greenland, Scandinavia, and Siberia and alpine sites in the European Alps, Colorado Rockies, Caucasus, and Pyrenees. More than 99% of observations are georeferenced.

Time period and grain: All data were collected between 1964 and 2016. A small number of sites have repeated measurements at two or more time periods.

Major taxa and level of measurement: Trait measurements were made on 532 terrestrial plant species growing in tundra habitats. Most observations are on individuals (>95%), while a small number represent plot or site means per species.

Software format: .csv file and GitHub repository with data cleaning scripts in R

BACKGROUND

Plant traits reflect species' ecological strategies and life histories, and underlie differences in the way plants acquire and use resources. Traits related to plant size and the leaf economics spectrum, for example, represent fundamental trade-offs between the capture and conservation of resources (Wright *et al.*, 2004; Díaz *et al.*, 2015). Because plant traits reflect the direct interaction between a plant and its habitat, variation in plant traits is often closely linked to environmental variation (Moles *et al.*, 2006; 2009; Sandel *et al.*, 2010). As such, plant traits can inform predictions of a species' response to environmental change (Soudzilovskaia *et al.*, 2013). Furthermore, many plant functional traits are directly related to

key community and ecosystem processes (Lavorel & Garnier, 2002; Díaz *et al.*, 2009; Reichstein *et al.*, 2014), and are thus considered an Essential Biodiversity Variable (EBV) necessary for assessing biodiversity and ecosystem change globally (Pereira *et al.*, 2013).

Global trait databases (e.g., Kattge *et al.*, 2011) have dramatically increased the accessibility of plant trait data over the past decade, but these databases are heavily geographically biased toward temperate regions. In contrast, the tundra is the most rapidly warming biome on the planet, but until now has been underrepresented in global trait databases. This poor geographic coverage of tundra species is especially pronounced in the most remote (e.g., high Arctic, high alpine) regions. Because intraspecific trait variation is thought to be particularly important in ecosystems like the tundra, where diversity is low and species' ranges are large (Siefert *et al.*, 2015), multi-site trait observations on multiple individuals are needed to capture the full extent of tundra trait variation.

Here, we present the Tundra Trait Team database, which contains more than 66,000 observations of 18 plant traits on 528 tundra species. In addition to filling a major geographical gap, this dataset is unique in its depth and spread. Trait data were collected at 191 unique locations, and include multiple observations on individuals at the same location as well as of the same species at different locations. In addition, 99.8% of the observations in the database are georeferenced, thus allowing trait observations to be linked with environmental data such as gridded climate datasets (e.g., WorldClim, www.worldclim.org, or CHELSA, chelsa-climate.org).

The TTT database can be used to address wide-ranging theoretical and practical ecological questions. Multiple trait observations on individuals and species at numerous sites across the tundra biome enables the quantification of inter- and intraspecific trait variation across scales (e.g., Bjorkman et al. submitted, Thomas et al. in prep). Linking trait observations with environmental data can facilitate our understanding of trait-environment relationships (Bjorkman et al. submitted) and the role of environmental filtering in shaping plant communities (Bernard-Verdier *et al.*, 2012; Asner *et al.*, 2016), and can inform Earth System model parameters in dynamic vegetation models (Wullschleger *et al.*, 2014). We expect that making this dataset publically available will contribute to future research in these and other unforeseen ways.

METHODS

Data acquisition & compilation

Data were submitted directly by the tundra researchers that collected them (see authorship and acknowledgements). These data represent a mix of previously collected data as well as new data collected as part of a multi-site field campaign. In some cases, the submitted trait data have contributed to publications (see supplementary table X for full reference list) but all values in the database are from primary sources (i.e., not extracted from publications) with the exception of some root trait data, as described in Iversen *et al*. (2014). All trait data are collected on plants growing *in situ* under natural conditions (i.e., data from experimental treatments were removed).

Data curation & quality control

All observations were checked to ensure logical latitude and longitude information and converted to standardized units of measurement. We also removed obviously erroneous or impossible values (e.g. leaf dry matter content values greater than 1.0). When possible, suspected errors were checked with the initial data providers and corrected. Species names were standardized to match the accepted names in The Plant List using the R package *Taxonstand v2.0* (Cayuela *et al.*, 2012), but the original names provided by data contributors are also included in the database. The original name may contain additional information about subspecies designations.

For those species with at least 10 observations of the same trait type, we additionally report an "error risk" for each observation (see TRY database protocols for more information on the term "error risk" in this context, https://www.try-db.org/TryWeb/TRY_Data_Release_Notes.pdf). The error risk was calculated as the number of standard deviations that a given value lies from the overall species mean for that trait. We also provide the script used to create the "cleaned" version of the dataset as a GitHub repository (XXXX address), along with both the raw (uncleaned) and cleaned versions of the dataset. The cleaning script can be adapted to vary in its sensitivity to outliers. This script includes code to output histograms that visually identify removed values per species for any traits of interest. It should be noted that this cleaning protocol is primarily useful for species with large numbers of observations of a given trait, and that much of the variation within a species may be due to environmental differences among sites.

Data availability & access

A static version of the cleaned database will be maintained at the Polar Data Catalog and additionally submitted to the TRY plant trait database for inclusion in the next version release. Data retrieved through TRY are fully public but are subject to the usage guidelines outlined in TRY. Trait data collection is ongoing; thus, we will also maintain a periodically updated version of this database at the GitHub repository (XXXX). The complete (uncleaned) dataset is also available at this address.

Data use guidelines

Data are publically available using an Attribution-ShareAlike 4.0 Generic copyright (CC BY-SA 2.5). We ask that data users abide by the 5% rule, that for any future study for which a particular data contributor's data makes up 5% or more of the final dataset, they will be informed about the study and offered to contribute to the publication. If the data contributor declines authorship they should be acknowledged in the acknowledgements. In addition, this publication should be cited.

DESCRIPTION OF DATA

The TTT database contains 66,308 observations on 18 plant traits measured in 191 locations across the tundra biome (Table 1 & Figure 1). A "location" is defined as a unique latitude-longitude combination, when both are rounded to the nearest tenth of a degree. The best represented traits in the database (>1000 observations each) include plant height (both vegetative and reproductive), leaf area, specific leaf area, leaf fresh and dry mass, leaf dry matter content, leaf nitrogen content, leaf carbon content, leaf phosphorus content, seed mass, and stem specific density. In most cases, traits were measured on adult individuals at peak

growing season, but some exceptions exist (e.g. *Rhododendron caucasicum* contains values of LDMC for both young and old leaves).

In addition to the trait values themselves, nearly all observations (99.8%) contain information about latitude and longitude of the location where the measurement was taken (Figure 2). Elevation was also provided for some observations. The high degree of georeferencing in the dataset enables the extraction of climate and other environmental data corresponding with each trait measurement.

Table 1. All traits contained in the TTT database, including the number of total observations of each trait, the number of unique locations (rounded to the nearest tenth of a decimal degree) at which each trait was measured, and the total number of species for which that trait was measured. The mean, standard deviation, median, and 95% quantiles for each trait are additionally provided (see also Figure 3).

Trait	*units*	*# obs*	*# locs*	*# spp*	*mean*	*sd*	*q2.5*	*median*	*q97.5*
Height, repro.	m	5910	25	117	0.14	0.11	0.02	0.11	0.43
Height, veg.	m	18037	131	326	0.24	0.40	0.01	0.11	1.55
Leaf dry matter content (LDMC)	ratio	5550	27	279	0.33	0.15	0.10	0.31	0.69
Leaf area	mm2	8496	30	372	726.9	4083.0	5.76	196.3	3872.1
Leaf carbon	mg/g	1743	18	212	459.9	28.0	414.4	455.1	523.5
Leaf C:N ratio	ratio	493	8	68	25.36	14.18	12.63	21.12	67.32
Leaf d13C	ppt	86	2	5	-28.1	1.8	-29.8	-29.0	-24.2
Leaf d15N	ppt	34	1	4	1.08	2.60	-3.51	1.79	5.10
Leaf dry mass	mg	6394	29	275	35.06	82.05	0.41	11.33	222.26
Leaf fresh mass	g	5026	22	272	0.16	0.42	0.00	0.04	1.01
Leaf nitrogen	mg/g	2394	21	238	23.62	9.58	7.88	23.08	45.62
Leaf phosphorus	mg/g	1411	12	206	2.41	1.02	0.85	2.24	4.80
Root:shoot ratio	ratio	269	33	84	3.23	6.59	0.10	1.40	21.27
Rooting depth	cm	77	9	49	0.22	0.22	0.01	0.15	1.00
Seed mass	mg	1281	22	135	1.85	3.75	0.03	0.60	14.90
Specific leaf area (SLA)	mm^2/mg	7518	47	301	14.49	8.15	4.49	12.60	35.41
Stem density (SSD)	mg/mm^3	1181	21	35	0.62	0.14	0.37	0.62	0.91
Stem diameter	m	408	10	13	0.36	0.92	0.01	0.01	3.14

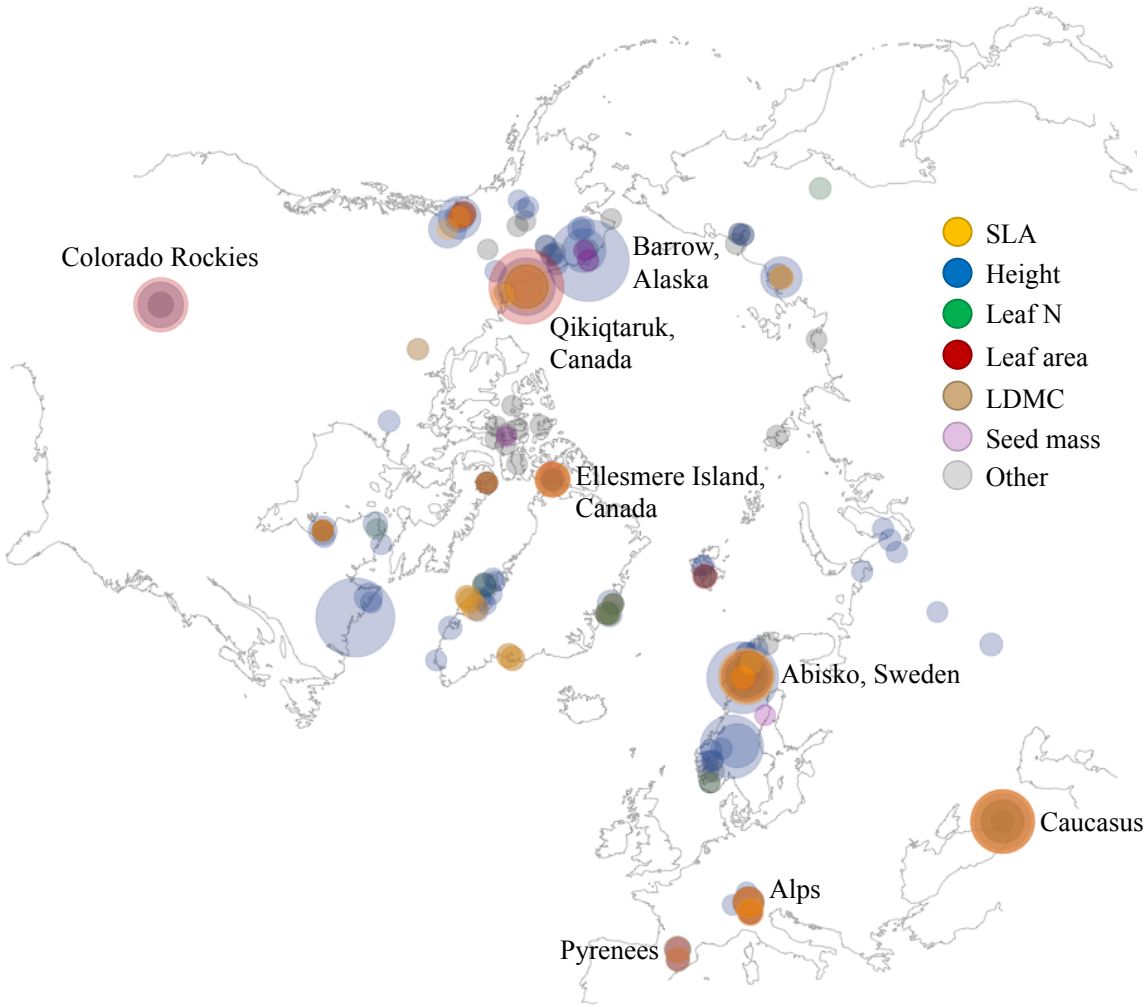

Figure 1. Trait observations span the Arctic and alpine tundra. The size of the circle corresponds to the number of trait observations at a given location (minimum = 1, maximum = 2555), while the color of each circle indicates the measured trait.

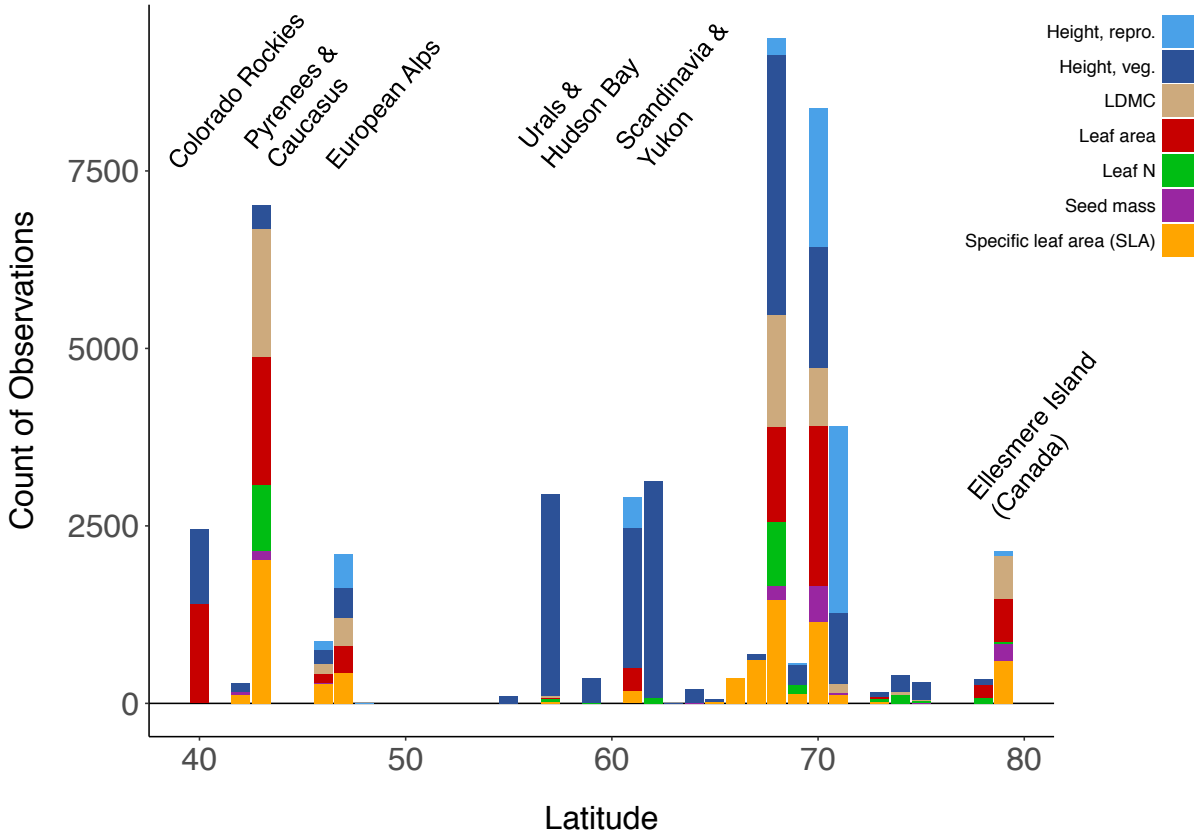

Figure 2. Frequency of the observations across latitudes for the most commonly measured traits. More than 99% of the observations are georeferenced.

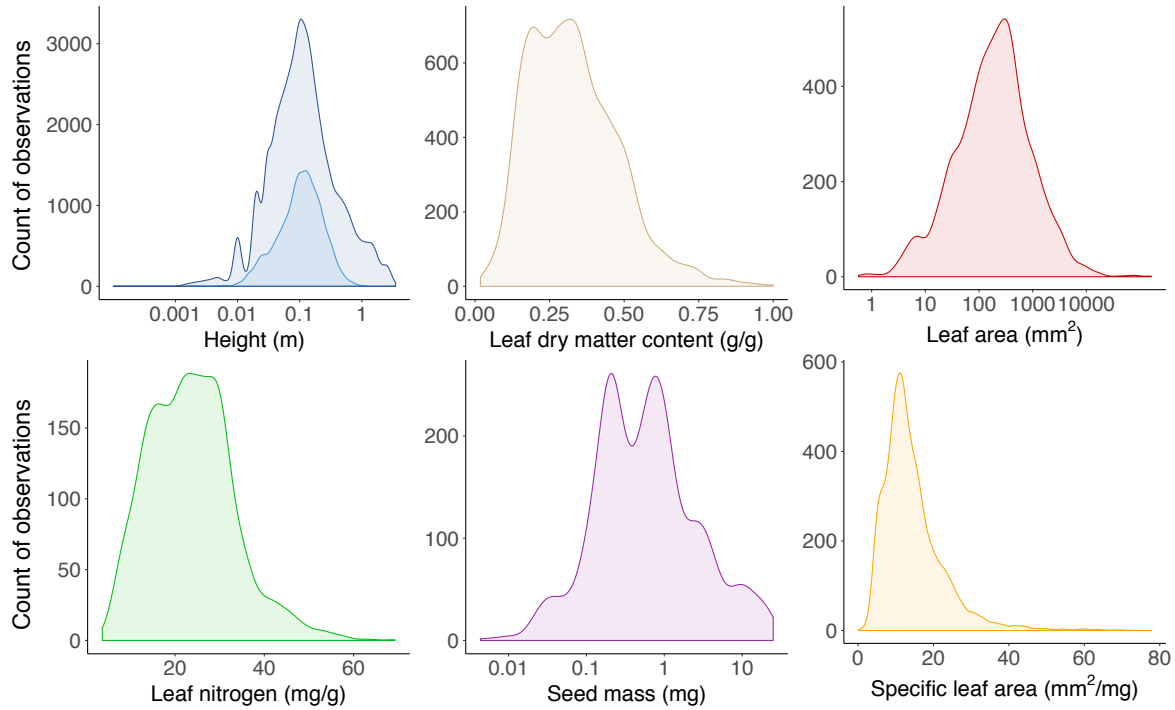

Figure 3. Overall distribution of trait values for the six most commonly measured traits in the TTT database. Note that the x-axis is a log10 scale for certain traits (height, leaf area and seed mass). Plant height is divided into reproductive (light blue) and vegetative (dark blue) height.

Dataset Structure

The cleaned TTT dataset is provided as a .csv file and consists of a single data table. The table structure is as follows:

Column Name	*Description of variable*
AccSpeciesName	Accepted species name as given by The Plant List
OriginalName	Original species name provided by the data contributor
IndividualID	Number associated with each individual measured (as multiple traits were sometimes measured on the same individual)
Latitude	Latitude of the observation location in decimal degrees
Longitude	Longitude of the observation location in decimal degrees
Elevation	Elevation of the observation location in meters
SiteName	Name of the site where the observation was collected (as provided by the data contributor)
SubsiteName	Name of the subsite (nested within the SiteName) where the observation was collected (as provided by the data contributor)
DOY_measured	Day of the year on which the measurement was made
Year_measured	Year in which the measurement was made
DataContributor	Name of the original contributor of the data
ValueKindName	Specificity of the measurement; Single = single observation on an individual, Individual Mean = mean of multiple observations taken on a single individual, Plot mean = mean of multiple observations take on individuals of the same species in a plot, Site specific mean = mean of multiple individuals at the same site, Maximum in plot = maximum of all individuals of that species in a plot
Trait	Name of the trait measured using the TRY trait name convention, or the name reported by the data contributor when a trait is not included in TRY
Value	Value of the trait measured using the reported significant digits
Units	Unit of measurement for each trait (see also Table 1)
ErrorRisk	See description of the error risk variable above, and https://www.try-db.org/TryWeb/TRY_Data_Release_Notes.pdf
Comments	Additional comments provided by the data contributor or collator, usually related to how the measurements were conducted

References

Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E. & Vaughn, N. (2016) Large-scale climatic and geophysical controls on the leaf economics spectrum. *Proceedings of the National Academy of Sciences of the United States of America*, **113**, E4043–51.

Bernard-Verdier, M., Navas, M.-L., Vellend, M., Violle, C., Fayolle, A. & Garnier, E. (2012) Community assembly along a soil depth gradient: contrasting patterns of plant trait convergence and divergence in a Mediterranean rangeland. *Journal of Ecology*, **100**, 1422–1433.

Cayuela, L., Granzow-de la Cerda, Í., Albuquerque, F.S. & Golicher, D.J. (2012) TAXONSTAND: An R package for species names standardisation in vegetation databases. *Methods in Ecology and Evolution*, **3**, 1078–1083.

Díaz, S., Hodgson, J.G., Thompson, K., Cabido, M., Cornelissen, J.H.C., Jalili, A., Montserrat-Martí, G., Grime, J.P., Zarrinkamar, F., Asri, Y., Band, S.R., Basconcelo, S., Castro-Díez, P., Funes, G., Hamzehee, B., Khoshnevi, M., Pérez-Harguindeguy, N., Pérez-Rontomé, M.C., Shirvany, F.A., Vendramini, F., Yazdani, S., Abbas-Azimi, R., Bogaard, A., Boustani, S., Charles, M., Dehghan, M., de Torres-Espuny, L., Falczuk, V., Guerrero-Campo, J., Hynd, A., Jones, G., Kowsary, E., Kazemi-Saeed, F., Maestro-Martínez, M., Romo-Díez, A., Shaw, S., Siavash, B., Villar-Salvador, P. & Zak, M.R. (2009) The plant traits that drive ecosystems: Evidence from three continents. *Journal of Vegetation Science*, **15**, 295–304.

Díaz, S., Kattge, J., Cornelissen, J.H.C., Wright, I.J., Lavorel, S., Dray, S., Reu, B., Kleyer, M., Wirth, C., Prentice, I.C., Garnier, E., Bönisch, G., Westoby, M., Poorter, H., Reich, P.B., Moles, A.T., Dickie, J., Gillison, A.N., Zanne, A.E., Chave, J., Wright, S.J., Sheremet'ev, S.N., Jactel, H., Baraloto, C., Cerabolini, B., Pierce, S., Shipley, B., Kirkup, D., Casanoves, F., Joswig, J.S., Günther, A., Falczuk, V., Rüger, N., Mahecha, M.D. & Gorné, L.D. (2015) The global spectrum of plant form and function. *Nature*.

Iversen, C.M., Sloan, V.L., Sullivan, P.F., Euskirchen, E.S., McGuire, A.D., Norby, R.J., Walker, A.P., Warren, J.M. & Wullschleger, S.D. (2014) The unseen iceberg: plant roots in arctic tundra. *New Phytologist*, **205**, 34–58.

Kattge, J., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönisch, G., Garnier, E., Westoby, M., Reich, P.B., Wright, I.J., Cornelissen, J.H.C., Violle, C., Harrison, S.P., van Bodegom, P.M., Reichstein, M., Enquist, B.J., Soudzilovskaia, N.A., Ackerly, D.D., ANAND, M., ATKIN, O., Bahn, M., BAKER, T.R., BALDOCCHI, D., BEKKER, R., BLANCO, C.C., Blonder, B., BOND, W.J., BRADSTOCK, R., BUNKER, D.E., CASANOVES, F., CAVENDER-BARES, J., CHAMBERS, J.Q., Chapin, F.S., III, CHAVE, J., COOMES, D., Cornwell, W.K., CRAINE, J.M., DOBRIN, B.H., DUARTE, L., DURKA, W., ELSER, J., ESSER, G., ESTIARTE, M., FAGAN, W.F., FANG, J., FERNÁNDEZ-MÉNDEZ, F., FIDELIS, A., FINEGAN, B., FLORES, O., FORD, H., FRANK, D., Freschet, G.T., FYLLAS, N.M., GALLAGHER, R.V., GREEN, W.A., GUTIERREZ, A.G., HICKLER, T., HIGGINS, S.I., Hodgson, J.G., Jalili, A., JANSEN, S., JOLY, C.A., KERKHOFF, A.J., KIRKUP, D., KITAJIMA, K., KLEYER, M., KLOTZ, S., Knops, J.M.H., KRAMER, K., KÜHN, I., KUROKAWA, H., LAUGHLIN, D., LEE, T.D., Leishman, M., LENS, F., LENZ, T., LEWIS, S.L., LLOYD, J., LLUSIÀ, J., LOUAULT, F., MA, S., Mahecha, M.D., MANNING, P., MASSAD, T., MEDLYN, B.E., Messier, J., MOLES, A.T., MÜLLER, S.C., NADROWSKI, K., Naeem, S., NIINEMETS, Ü., NÖLLERT, S., NÜSKE, A., OGAYA, R., Oleksyn, J., Onipchenko, V.G., ONODA, Y., ORDOÑEZ, J., OVERBECK, G., OZINGA, W.A., PATIÑO, S., PAULA, S., PAUSAS, J.G., Peñuelas, J., PHILLIPS, O.L., PILLAR, V., POORTER, L., Poorter, H., POSCHLOD, P., PRINZING, A., PROULX, R., RAMMIG, A., REINSCH, S., Reu, B., SACK, L., Shipley, B., SIEFERT, A., SOSINSKI, E., SOUSSANA, J.F., SWAINE, E., SWENSON, N., Thompson, K., THORNTON, P., WALDRAM, M., Weiher, E., WHITE, M., WHITE, S., WRIGHT, S.J., YGUEL, B., ZAEHLE, S., ZANNE, A.E. & Wirth, C. (2011) TRY–a global database of plant traits. *Global Change Biology*, **17**, 2905–2935.

Lavorel, S. & Garnier, E. (2002) Predicting changes in community composition and ecosystem functioning from plant traits: revisiting the Holy Grail. *Functional Ecology*, **16**, 545–556.

Moles, A.T., Ackerly, D.D., Tweddle, J.C., Dickie, J.B., Smith, R., Leishman, M.R., Mayfield, M.M., Pitman, A., Wood, J.T. & Westoby, M. (2006) Global patterns in seed size. *Global Ecology and*

Biogeography, **16**, 109–116.
Moles, A.T., Warton, D.I., Warman, L., Swenson, N.G., Laffan, S.W., Zanne, A.E., Pitman, A., Hemmings, F.A. & Leishman, M.R. (2009) Global patterns in plant height. *Journal of Ecology*, **97**, 923–932.
Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., Coops, N.C., Dulloo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Hoft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., ONODA, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M. & Wegmann, M. (2013) Essential Biodiversity Variables. *Science*, **339**, 277–278.
Reichstein, M., Bahn, M., Mahecha, M.D., Kattge, J. & Baldocchi, D.D. (2014) Linking plant and ecosystem functional biogeography. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 13697–13702.
Sandel, B., Goldstein, L.J., Kraft, N.J.B., Okie, J.G., Shuldman, M.I., Ackerly, D.D., Cleland, E.E. & Suding, K.N. (2010) Contrasting trait responses in plant communities to experimental and geographic variation in precipitation. *New Phytologist*, **188**, 565–575.
Siefert, A., Violle, C., Chalmandrier, L., Albert, C.H., Taudiere, A., Fajardo, A., Aarssen, L.W., Baraloto, C., Carlucci, M.B., Cianciaruso, M.V., de L Dantas, V., de Bello, F., Duarte, L.D.S., Fonseca, C.R., Freschet, G.T., Gaucherand, S., Gross, N., Hikosaka, K., Jackson, B., Jung, V., Kamiyama, C., Katabuchi, M., Kembel, S.W., Kichenin, E., Kraft, N.J.B., Lagerström, A., Bagousse-Pinguet, Y.L., Li, Y., Mason, N., Messier, J., Nakashizuka, T., Overton, J.M., Peltzer, D.A., Pérez-Ramos, I.M., Pillar, V.D., Prentice, H.C., Richardson, S., Sasaki, T., Schamp, B.S., Schöb, C., Shipley, B., Sundqvist, M., Sykes, M.T., Vandewalle, M. & Wardle, D.A. (2015) A global meta-analysis of the relative extent of intraspecific trait variation in plant communities. *Ecology Letters*, **18**, 1406–1419.
Soudzilovskaia, N.A., Elumeeva, T.G., Onipchenko, V.G., Shidakov, I.I., Salpagarova, F.S., Khubiev, A.B., Tekeev, D.K. & Cornelissen, J.H. (2013) Functional traits predict relationship between plant abundance dynamic and long-term climate warming. *Proceedings of the National Academy of Sciences of the United States of America*, **110**, 18180–18184.
Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, F.S., III, Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., Navas, M.-L., Niinemets, 2.L., Oleksyn, J., Osada, N., Poorter, H., Poot, P., Prior, L., Pyankov, V.I., Roumet, C., Thomas, S.C., Tjoelker, M.G., Veneklaas, E.J. & Villar, R. (2004) The worldwide leaf economics spectrum. *Nature*, **428**, 821–827.
Wullschleger, S.D., Epstein, H.E., Box, E.O., Euskirchen, E.S., Goswami, S., Iversen, C.M., Kattge, J., Norby, R.J., van Bodegom, P.M. & Xu, X. (2014) Plant functional types in Earth system models: past experiences and future directions for application of dynamic vegetation models in high-latitude ecosystems. *Annals of Botany*, **114**, 1–16.

TTT REFERENCES:

Milbau A, Graae BJ, Shevtsova A & Nijs I (2009) Effects of a warmer climate on seed germination in the subarctic. Annals of Botany, 104, 287-296.
Buchwal, A., Rachlewicz, G., Fonti, P., Cherubini, P., Gärtner, H., 2013. Temperture modulates intra-plant growth of *Salix polaris* from a high Arctic site (Svalbard). Polar Biology 36(9), 1305-1318.