MOLECULAR ECOLOGY

Molecular Ecology (2012) **21**, 2004–2016 doi: 10.1111/j.1365-294X.2011.05390.x

# Plant species richness belowground: higher richness and new patterns revealed by next-generation sequencing

INGA HIIESALU,* MAARJA ÖPIK,* MADIS METSIS,† LIISA LILJE,† JOHN DAVISON,* MARTTI VASAR,* MARI MOORA,* MARTIN ZOBEL,* SCOTT D. WILSON‡ and MEELIS PÄRTEL*

**Department of Botany, University of Tartu, Lai 40, 51005 Tartu, Estonia, †Centre for Biology of Integrated Systems, Tallinn University of Technology, Akadeemia tee 15A, 12618 Tallinn, Estonia, ‡Department of Biology, University of Regina, Regina, SK S4S 0A2, Canada*

## Abstract

**Variation in plant species richness has been described using only aboveground vegetation. The species richness of roots and rhizomes has never been compared with aboveground richness in natural plant communities. We made direct comparisons of grassland plant richness in identical volumes (0.1 × 0.1 × 0.1 m) above and below the soil surface, using conventional species identification to measure aboveground richness and 454 sequencing of the chloroplast *trn*L(UAA) intron to measure belowground richness. We described above- and belowground richness at multiple spatial scales (from a neighbourhood scale of centimetres to a community scale of hundreds of metres), and related variation in richness to soil fertility. Tests using reference material indicated that 454 sequencing captured patterns of species composition and abundance with acceptable accuracy. At neighbourhood scales, belowground richness was up to two times greater than aboveground richness. The relationship between above- and belowground richness was significantly different from linear: beyond a certain level of belowground richness, aboveground richness did not increase further. Belowground richness also exceeded that of aboveground at the community scale, indicating that some species are temporarily dormant and absent aboveground. Similar to other grassland studies, aboveground richness declined with increasing soil fertility; in contrast, the number of species found only belowground increased significantly with fertility. These results indicate that conventional aboveground studies of plant richness may overlook many coexisting species, and that belowground richness becomes relatively more important in conditions where aboveground richness decreases. Measuring plant belowground richness can considerably alter perceptions of biodiversity and its responses to natural and anthropogenic factors.**

*Keywords*: 454 sequencing, DNA barcoding, plant richness, root identification, species coexistence, *trn*L (UAA)

*Received 1 July 2011; revision received 7 October 2011; accepted 19 October 2011*

## Introduction

Our current understanding of plant coexistence and diversity is based entirely on data from aboveground shoots. In many ecosystems, such as grassland, steppe, desert and tundra, however, the majority of plant growth (i.e. 50–90% of primary production) occurs belowground (Stanton 1988). Further, theories about species coexistence that explain aboveground diversity consistently invoke belowground interactions (Grime 1979; Tilman 1982), emphasizing the need for consideration of the belowground component of plant communities. It is unknown whether the richness patterns described for aboveground vegetation also hold for the large belowground portion.

Here, we focus on temperate grasslands because their aboveground richness has been well studied (Gibson

Correspondence: Inga Hiiesalu, Fax: (+372) 737 6222;
E-mail: inga.hiiesalu@ut.ee

2008) and they are dominated by belowground production (Stanton 1988; Jackson *et al.* 1997), but we expect that the diversity principles will apply to other vegetation types, especially those dominated by belowground productivity. Belowground diversity has received considerable attention for some soil biota (Wardle 2002; Bardgett 2005), but not for plants. We explore variation in richness at the plant neighbourhood and community scales, which are representative of the scales at which herbaceous plants interact and are typically used in empirical studies (Grace 1999). Here, we consider well-known patterns of aboveground plant richness and their relationships to unknown patterns of belowground richness at two spatial scales.

Although no empirical data exist on the relationship between above- and belowground plant species richness, there are several lines of indirect evidence to suggest that belowground richness might exceed that aboveground. First, roots generally persist for a longer time (Eissenstat & Yanai 1997; Wells & Eissenstat 2001; Shefferson 2009; Reintal *et al.* 2010) and occupy more space than shoots (Casper *et al.* 2003), especially in water limited systems (Kummerow *et al.* 1977; Gibbens & Lenz 2001; Schenk & Jackson 2002). Both factors should contribute to higher richness belowground. Second, the diverse nature of the soil environment, including its variety of heterogeneous resources (Hutchings & John 2004) and the abundance of soil microbe–plant interactions (Bever *et al.* 2010) may promote plant species coexistence belowground. Third, the relative symmetry of belowground competition, compared with asymmetric aboveground competition for light (Weiner 1990), may also cause richness to be higher below- than aboveground. It could be assumed that differences in species richness above- and belowground are most pronounced at the plant neighbourhood scale and least so at the community scale. At the larger scale, however, belowground richness may still include some temporarily dormant species which are not detected aboveground.

If belowground plant richness does exceed that aboveground, a question remains whether aboveground richness constitutes a constant or varying proportion of belowground richness. Nonlinear relationships between above- and belowground richness might occur if an increase in belowground richness along an environmental gradient does not occur aboveground. This could occur if the factors noted earlier that cause belowground richness to exceed aboveground richness (greater root dispersion in time and space, soil heterogeneity, symmetrical root competition) vary in their influence along gradients of, for instance, soil fertility. For example, aboveground species richness in temperate grasslands tends to decrease at the highest levels of soil fertility (see Mittelbach *et al.* 2001; Pärtel *et al.* 2007). Currently, the decrease is often explained in terms of asymmetric light competition on fertile soils with high standing crop (Weiner 1990; Hautier *et al.* 2009), but this mechanism might not apply to belowground organs. Alternatively, fertile soils support larger individuals, implying that fewer individuals, and thus, fewer species will be found as fertility increases (Oksanen 1996), but roots, with a great deal of spatial overlap among species (Booth *et al.* 2003), might not show this response. In both cases, in contrast to aboveground richness, belowground richness might not fall as fertility increases. Because belowground plant richness remains unmeasured, it is unknown whether the well-known decrease in aboveground richness with increasing fertility also occurs belowground.

The absence of information about belowground plant richness stems directly from methodological constraints. The roots and rhizomes of different species are mostly morphologically indistinguishable. Thus, previous methods of assigning roots to species have included laborious and time-consuming excavations of root systems in order to trace their linkage to aboveground parts (Wildová 2004). However, this approach does not identify very fine roots, roots attached to dormant meristems, or fragmented roots and rhizomes. More recently, several molecular techniques have been used to identify plant species from single root fragments (Jackson *et al.* 1999; Linder *et al.* 2000; Ridgway *et al.* 2003; Jones *et al.* 2011; Kesanakurti *et al.* 2011), and, more relevantly, for studying natural plant communities from mixed-root samples (Moore & Field 2005; Mommer *et al.* 2008, 2010; Fisk *et al.* 2010; Frank *et al.* 2010). Next-generation sequencing (NGS) technologies, such as 454 sequencing and Illumina, enable DNA-based identification of plants from mixed-species samples with relatively large numbers of sequences per sample (Valentini *et al.* 2009a). NGS has not been used to study plant belowground richness before, although the technique has been applied to identify plants in herbivore guts (Soininen *et al.* 2009; Pegard *et al.* 2009; Valentini *et al.* 2009b) and permafrost samples (Sønstebø *et al.* 2010), and to identify the mycorrhizal fungi inhabiting roots (Öpik *et al.* 2009; Moora *et al.* 2011).

Currently applied DNA-based methods work reasonably well to enumerate species richness in samples, but progress remains to be made in the determination of the relative abundances of species, which would in turn allow the calculation of diversity indices (e.g. Mommer *et al.* 2010). NGS is promising in this regard because the number of sequences belonging to a particular species may reflect the relative abundance of the species. However, validation of the accuracy of NGS when

measuring plant abundance in mixed-species samples is still needed (Valentini *et al.* 2009a).

Here, we assess for the first time whether plant belowground species richness differs from aboveground richness in a natural community. We compared neighbourhood- and community-scale richness above- and belowground in natural grassland, a system well known for its relatively high diversity (Gibson 2008). We used conventional species identification to measure aboveground richness and 454 sequencing to measure belowground richness. Our primary hypotheses were the following: (i) belowground richness exceeds aboveground richness; (ii) aboveground richness does not increase proportionally with total belowground richness; and (iii) above- and belowground richness respond differently to soil fertility.

## Material and methods

### *Field site and sampling*

We measured above- and belowground plant species richness in a 2-ha diverse mesophytic grassland in south-eastern Estonia (Põlva County, 58°06′N; 27°04′E). Richness was recorded in 100 volumes ($0.1 \times 0.1 \times 0.1$ m) both above and below the soil surface in mid-June 2007. Samples were arranged contiguously in ten randomly-placed 1-m long transects, with ten samples per transect. Transects were separated by >10 m. These sample volumes were used to capture the scales at which herbaceous plants interact and are representative of the scales used to study aboveground richness (Grace 1999). The soil at the study site is predominantly sandy with a pH (KCl) of 4.6–5.2. Average aboveground biomass is 325 g $m^{-2}$. The most common plant species at the site are *Galium boreale* L., *Geranium pratense* L., *Elymus repens* (L.) Gould, *Festuca rubra* L., *Knautia arvensis* (L.) Coult., and *Veronica chamaedrys* L. The grassland is mowed once per year and the hay removed.

Aboveground species richness was determined by identifying all vascular plant species in each sample. This included species that were rooted in the samples, as well as species that occurred in the sample volumes but were rooted elsewhere.

Belowground plant species richness was measured from root samples by collecting a volume of soil corresponding exactly to that measured for aboveground richness (0.001 $m^3$). The litter layer was removed, roots were sieved from soil and sorted to exclude dead material on the basis of colour and physical appearance (Gregory 2006). Roots were then crushed using liquid nitrogen, and mixed. A root subsample of ca 100 mg dry weight was used for DNA analysis. Soil total nitrogen (N) content (Kjeldahl method) was measured adjacent to each transect. Because we studied a single homogeneous grassland patch, patterns of total N should reflect patterns in available N.

### *Molecular analysis*

*Root DNA extraction and 454 sequencing*. Root subsamples were pulverized with 2.3-mm chrome-steel beads (BioSpec Products, Inc., Bartlesville, OK, USA) in a Mixer Mill 301 (Retsch GmbH, Haan, Germany). DNA was extracted using the High Pure PCR Template Preparation kit (Roche Applied Science, Mannheim, Germany) eluting in a final volume of 200 μL.

Plant sequences were amplified using the chloroplast *trn*L (UAA) gene primers *c* and *d* (Taberlet *et al.* 1991), linked to 454 adapter primers A and B respectively. In order to distinguish 454 sequences coming from different samples, bar-code sequences, 8 bp in length, were inserted between the A primer and *c* primer sequences. Thus, the composite forward primer was: 5′-CCATCTCATCCCTGCGTGTCTCCGACTCAG(NNNNNNNN)*CGAAATCGGTAGACGCTACG*-3′ and the reverse primer was: 5′- CCTATCCCCTGTGTGCCTTGGCAGTCTCAG(NNNNNNNN)*GGGGATAGAGGGACTTGAAC*-3′ where A and B primers are underlined, the bar-code is indicated by *N*-s in parentheses and specific primers *c* and *d* are shown in italics. The same bar-code was incorporated in the reverse primer. Sample preparation for 454 sequencing followed our previously described methods (Öpik *et al.* 2009). The PCRs were performed in a total volume of 40 μL containing 20 μL of HotStarTaq Master Mix (Qiagen Gmbh, Germany), 0.23 μM each of the primers and 2 μL of template DNA. The reactions were run on an Primus 96 Plus Thermal Cycler (MWG-Biotech, Germany) with the following conditions: 95 °C for 15 min; five cycles of 42 °C for 30 s, 72 °C for 60 s, 92 °C for 45 s; 35 cycles of 65 °C for 30 s, 72 °C for 60 s, 92 °C for 45 s, followed by 65 °C for 30 s and 72 °C for 10 min. PCR products were separated by electrophoresis through a 1.5% agarose gel in 0.5× TBE, and the PCR products were purified from the gel using the QIAquick Gel Extraction kit (Qiagen Gmbh). The amount of DNA in the purified PCR products was measured using NanoDrop 1000 (Thermo Scientific, Wilmington, DE, USA). The products were mixed at equimolar concentrations. DNA mix was subjected to sequencing on a Genome Sequencer FLX System, using Titanium Series reagents (Roche Applied Science) at GATC Biotech (Constanz, Germany).

*Reference sequence database.* A custom-made *trn*L(UAA) intron sequence reference database was compiled from three sources: (i) plants sampled at our study site and

sequenced (methods below; sequences deposited in GenBank under the accession numbers HM590228-HM590365); and sequences from species occurring in our study system or closely related taxa that were (ii) available in GenBank; or (iii) generated by the EcoChange Project (EU FP6 Integrated Project EcoChange). Plants collected from our study system and its surroundings were identified and stored as vouchers.

DNA was extracted from leaf samples as described earlier. Samples were further subjected to polymerase chain reaction (PCR) of the chloroplast *trn*L (UAA) intron using the universal primers *c* and *d* (Taberlet *et al.* 1991). PCRs were performed using Ready To Go™ Beads (Amersham Pharmacia Biotech., UK) by adding 1 μL of 10 μM-primers, 1 μL of template DNA and 23 μL of distilled water. Thermocycling conditions were as follows: 95 °C for 5 min, 35 cycles of 95 °C for 30 s, 51 °C for 1 min, 72 °C for 1.5 min, followed by 72 °C for 10 min using a Mastercycler ep Gradient 5341 (Eppendorf AG, Hamburg, Germany). Positive PCR products were directly sequenced in both directions using a 3730XL DNA sequencer (Life Tech. Corp., Carlsbad, CA, USA) at Macrogen Inc. (Seoul, Korea). Raw sequences were trimmed to remove primer sequences and complementary chains assembled using SEQUENCHER 4.8 (GeneCodes Corp., Ann Arbor, MI, USA).

### *Bioinformatical analysis*

454 sequences were included in subsequent analyses only if they carried a correct bar-code, carried the correct *c* primer sequence, and were ≥170 bp long (including the bar-code and primer *c* sequence). The 454 sequencing reads were shorter than the full amplicon; therefore, we did not consider *d* primer here. In addition, only samples that yielded at least six sequences were included in further analyses. As the *trn*L (UAA) intron sequence between *c* and *d* primers does not distinguish certain closely related species, we defined molecular operational taxonomic units (MOTUs) within our reference database by grouping species that exhibited sequence similarity of ≥97% using the BLASTclust algorithm. 454 sequences were assigned to MOTUs by conducting a BLAST search (soft masking of DUST filter) against the reference database with the following criteria required for a match: sequence similarity ≥97%; an alignment length no more than 10 bp less than the shorter of the query (454 sequence) or subject (reference database sequence) sequence length; and a BLAST e-value < 1e-50. MOTUs containing only a single sequence were omitted following our previously applied procedures (Öpik *et al.* 2009; Moora *et al.* 2011) because these are likely to result from pyrosequencing errors (Huse *et al.* 2007; Tedersoo *et al.* 2010). We investigated those sequences that BLAST did not match against the reference database by conducting a further BLAST search against the GenBank non-redundant database using the same parameters. As the 454 sequences exhibited considerable variation in length, we also examined the influence of sequence length on MOTU identification by artificially shortening a subset of the long 454 sequences (>400 bp) and repeating BLASTs (using the original parameters) against the reference database.

Species recorded aboveground were assigned to the corresponding MOTUs in order to have the same resolution of taxon identification above- and belowground. Consequently, a few groups of closely related species were lumped as single MOTUs (see Results, Table 1). MOTUs are hereafter referred to as species or species groups.

### *Known root mixtures*

We tested the ability of 454 sequencing to detect the presence and abundance of plant species by preparing eight mixtures of roots with 2–5 species from a Canadian grassland (30 km E of Regina, 50° 28′ N, 104° 22′ W). The following species were used in the test mixtures: *Solidago missouriensis* Nutt., *Heterotheca villosa* (Pursh) Nutt. ex DC., *Artemisia frigida* Willd., *Erysimum altum* (Ahti) Tzvelev and *Agropyron cristatum* (L.) Gaertn. We varied the proportion of biomass of added species (range: 10–90%) in order to determine whether sequencing could be used to measure species abundances in mixtures, potentially allowing the calculation of species diversity and evenness in addition to richness. These mixtures were subjected to the same molecular analyses as the samples used to examine our field samples.

### *Statistical analysis*

*Species occurrences above- and belowground.* For each species, we recorded the number of samples where the species occurred above- and belowground. We enumerated three aspects of species richness (Fig. 1): (i) aboveground richness, sampled visually; (ii) total belowground richness (aboveground richness of species rooted in the sample plus additional belowground richness detected by DNA analysis); and (iii) additional belowground richness alone (species detected by DNA analysis of belowground sample but absent from the aboveground sample).

Six species that occurred in aboveground samples and were rooted in our plots were not detected by 454 sequencing: *Anemone nemorosa* L., *Carex* spp., *Elymus*

**Table 1** List of plant molecular operational taxonomic units (MOTUs: referred to as species or species groups in the text) detectable using the *trn*L(UAA) intron and their component taxa

| MOTU | Taxon | NCBI accession no. | Total no. of sequences |
|---|---|---|---|
| 1 | *Solidago virgaurea* | **HM590347** | 5615 |
| | *Solidago missouriensis* | **HM590346** | |
| | *Erigeron acris* | GQ244931 | |
| | *Heterotheca villosa* | **HM590268** | |
| 2 | *Calamagostris deschampsioides* | GQ244659 | 4560 |
| 3 | *Knautia arvensis* | **HM590303** | 4100 |
| 4 | *Veronica chamaedrys* | **HM590360** | 2580 |
| | | AY673632 | |
| 5 | *Achillea millefolium* | **HM590229** | 2380 |
| | | GQ244529 | |
| 6 | *Artemisia frigida* | **HM590243** | 2068 |
| | *Artemisia ludoviciana* | **HM590244** | |
| 7 | *Festuca rubra* | **HM590284** | 1645 |
| | *Festuca ovina* | GQ244979 | |
| 8 | *Aegopodium podagraria* | **HM590230** | 1040 |
| | *Heracleum sibiricum* | **HM590301** | |
| 9 | *Ranunculus polyanthemos* | **HM590338** | 877 |
| 10 | *Geranium pratense* | **HM590294** | 697 |
| | *Geranium palustre* | EU326063 | |
| | | **HM590293** | |
| 11 | *Galium verum* | **HM590289** | 433 |
| 12 | *Anthriscus sylvestris* | **HM590240** | 349 |
| 13 | *Ranunculus acris* | **HM590337** | 344 |
| 14 | *Vicia cracca* | **HM590363** | 270 |
| 15 | *Allium schoenoprasum* | GQ244546 | 94 |
| 16 | *Poa pratensis* | GQ245279 | 86 |
| 17 | *Filipendula ulmaria* | **HM590286** | 70 |
| | | GQ244985 | |
| 18 | *Deschampsia cespitosa* | **HM590272** | 69 |
| | | GQ244835 | |
| 19 | *Centaurea phrygia* | **HM590263** | 63 |
| 20 | *Gaura* sp. | **HM590292** | 48 |
| 21 | *Erysimum altum* | GQ244958 | 29 |
| | *Turritis glabra* | **HM590358** | |
| 22 | *Helictotrichon pubescens* | **HM590300** | 23 |
| 23 | *Dactylis glomerata* | **HM590271** | 12 |
| 24 | *Galium boreale* | **HM590291** | 6 |
| 25 | *Viola tricolor* | **HM590365** | 5 |
| 26 | *Tragopogon dubius* | **HM590356** | 4 |
| 27 | *Equisetum pratense* | **HM590278** | 3 |
| 28 | *Pinus sylvestris* | GQ245260 | 3 |
| 29 | *Myosotis* sp. | **HM590320** | 2 |

Accession nos of reference sequences generated in this study are highlighted in bold. Total 454 sequence counts matching the reference sequences are reported for each MOTU.

*repens* (L.) Gould, *Geum rivale* L., *Rumex acetosa* L., and *Stellaria graminea* L. Two species (*E. repens* and *R. acetosa*) were present in 41 and 63 aboveground plots, respectively, while the remaining species were present in fewer than 10 plots. Overall, an average of 1.5 species per sample went undetected using 454 sequencing. In order to test comparable data sets for above- and belowground richness, we completely omitted these species from the analyses. This avoided a bias of detecting differences between above- and belowground richness based on a methodological inability to detect certain species.

*Known root mixtures.* We explored the ability of 454 sequencing to detect species presence in the known mixtures of roots. We quantified the correspondence between the composition of the known species mixtures and the molecularly detected species in the mixtures.

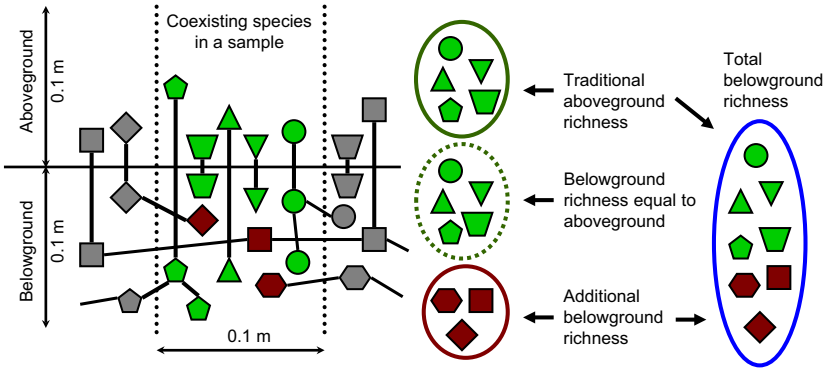


**Fig. 1** The three components of plant richness: (i) aboveground richness, sampled visually; (ii) total belowground richness (aboveground richness of species rooted in the sample plus additional belowground richness detected by DNA analysis); and (iii) additional belowground richness alone.

We cast the results for each species as $2 \times 2$ tables (presence or absence in created or detected mixtures), and applied standard meta-analysis for the tables for all species considered together. We used Peto odd ratios which can handle many absences (Borenstein *et al.* 2009).

In order to test whether 454 sequencing can quantify the relative abundance of species in the known mixtures, we compared the log-ratio-transformed proportions of added roots for each species with the numbers of retrieved sequences. Measures of relative abundance allow calculation of belowground plant diversity (Simpson's reciprocal index) and evenness (Simpson's reciprocal index/richness).

*Relationship between aboveground and total belowground richness.* The relationship between aboveground richness and total belowground richness cannot be addressed with traditional statistical tests because theoretically the former is always a subset of the latter (Jackson & Somers 1991). We therefore used a method that addresses a logically similar problem, the relationship between regional and local richness (Szava-Kovats *et al.* 2011). We calculated the log of the ratio of aboveground richness to belowground additional richness and related this to the log of total belowground richness. The absence of a significant slope indicates a linear, proportional relationship between total belowground and aboveground richness, whereas a significant negative slope indicates an asymptotic relationship between these variables (Szava-Kovats *et al.* 2011).

*Above- and belowground richness at multiple spatial scales.* Differences in above- and belowground richness were explored at two spatial scales. The plant neighbourhood scale was investigated using species richness-area (volume) curves obtained by determining the richness in adjacent plots within each transect. A few samples were omitted because of poor amplification of DNA, leaving a range of 0.001–0.008 $m^3$ in volume. Patterns of species occurrence at the community scale were additionally examined by producing species accumulation curves that calculated the cumulative number of species over an increasing number of transects (samples).

*Plant diversity measures related to soil fertility.* We related total belowground richness, aboveground richness and additional belowground richness to soil N content in our transects. By using 454 sequence frequency as a measure of relative abundance, we were able to calculate belowground diversity (Simpson's reciprocal index) and belowground evenness (Simpson's reciprocal index/richness), and relate them to soil N content. Analyses fitted generalized linear models (GLM) using generalized least squares, and accounted for correlated errors (R package nlme, function gls with Gaussian spatial correlation structure, R Development Core Team. 2010).

## Results

We recovered 32 298 sequences of the chloroplast *trn*L (UAA) intron with a length of ≥170 bp (maximum length 596 bp; median length 402 bp, Fig. S1, Supporting information) that carried the correct tag and primer sequence (28 bp in total). Of these, 27 478 sequences were assigned to 29 molecular taxonomic units (MOTUs) that usually comprised one but sometimes more plant taxa (Table 1). Species identification was consistent (>99%) across sequence lengths ranging from 142–400 bp (Table S1, Supporting information). Different species yielded between a few and thousands of sequences, with an average of 950 sequences per species

(Table 1). For the 15% of sequences not assigned to species, a BLAST search of GenBank revealed that 14% were flowering plants not matching our reference database (mostly from the orders *Gentianales*, *Asparagales* and *Brassicales*), <1% were bryophytes, and 85% of the unassigned sequences could not be identified.

Mixtures with known species composition showed a significant correspondence between added and observed species presences (Odds ratio = 23.1, $P < 0.005$). Two species (*Solidago* group and *Artemisia* spp.) were used in most mixtures, and this allowed us to compare relative sequence abundance with relative biomass. We found a good correspondence between proportions of added biomass in the root mixtures and numbers of retrieved sequences ($R^2 = 0.88$, $P < 0.001$, Fig. S2, Supporting information), indicating that 454 sequencing can provide quantitative data on species abundances, allowing the calculation of species diversity and evenness in addition to richness. Two species (*Agropyron cristatum* and *Erysimum altum*) added to the known mixtures were not recovered by 454 sequencing. Further, sequences of two species (*Achillea millefolium* and *Solidago* group) that were not added intentionally were each detected from single prepared samples at very low frequencies—1 and 25 sequences—compared with the hundreds of sequences that were recorded for the intentionally included species.

At the field study site, species frequency (number of samples where species occurred) aboveground was generally related to that of belowground (Fig. 2). The most common species both above- and belowground were *Festuca* spp., *Galium boreale*, *Veronica chamaedrys* and *Galium verum*. Most species (22 out of 29) occurred more frequently below- than aboveground, as shown by their position beneath the 1:1 line on Fig. 2. Nine species or species groups were detected only belowground, the most common being the *Solidago* group, *Artemisia* spp. and the *Turritis* group. Although these species were not found aboveground during this study, they had been detected at our study site later in the season or in previous years (J. Liira, unpublished).

Total belowground richness in a $0.1 \times 0.1 \times 0.1$ m sample had a maximum of 13 species, while aboveground richness had a maximum of nine species (Fig. 3). In a few cases, one or two species projected into aboveground sample volumes from the side, but these additions only increased aboveground richness by 0.5 species on average.

The relationship between aboveground richness and total belowground species richness was significantly nonlinear (Fig. 3) as indicated by a significant negative slope of the log-ratio regression model ($F_{1,71} = 23.5$, $P < 0.001$, Fig. S3, Supporting information). Thus, the increase in total belowground richness was initially associated with an increase in aboveground richness, but average aboveground richness reached an asymptote at about seven species when total belowground richness exceeded 10 species.

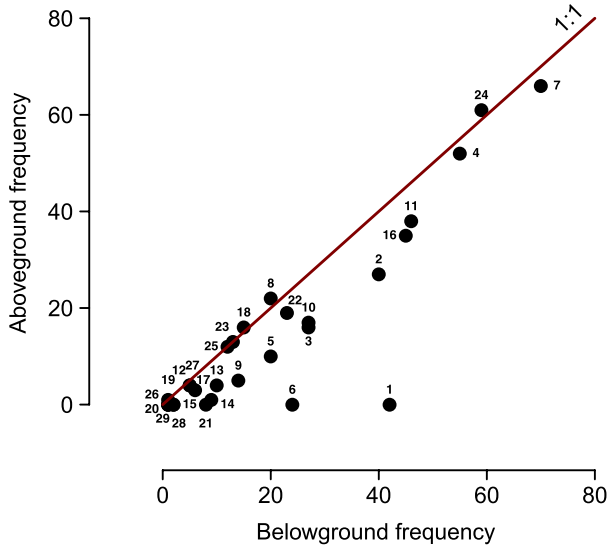


**Fig. 2** Plant species frequencies (number of samples in which species occurred) above- and belowground. Each point corresponds to a species or species group (MOTU); identifying numbers correspond to those in Table 1.

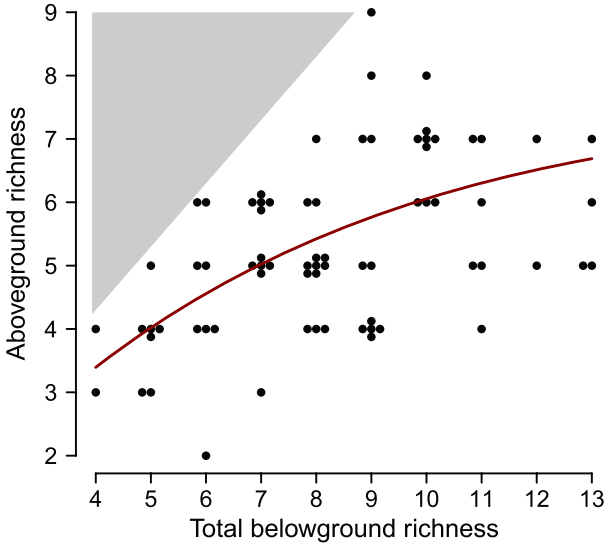


**Fig. 3** Aboveground plant richness rooted in a sample compared with total belowground richness (both in 0.001 $m^3$ sample volumes). The shaded area represents a region outside the operational space (aboveground richness cannot exceed total belowground richness). The trend line demonstrates the relationship between these two variables (inverse transformation from the log-ratio function). Overlapping points are slightly shifted for better visibility.

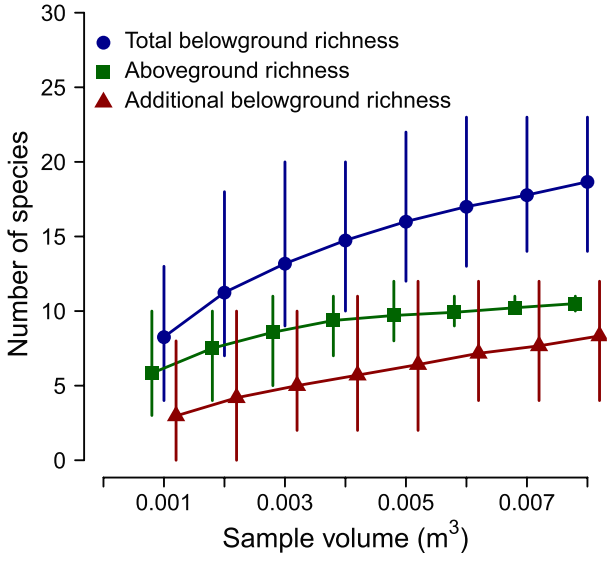


**Fig. 4** Contrasting patterns of total belowground, aboveground and additional belowground plant richness in relation to increasing nested sample size. Richness is presented as means with minimum and maximum.

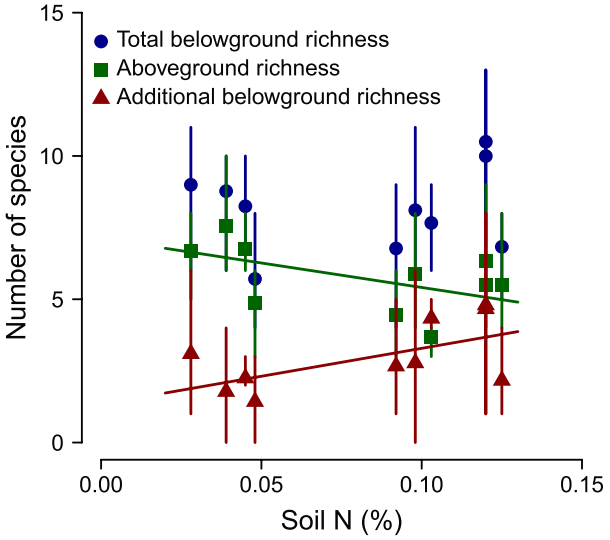


**Fig. 6** Contrasting patterns of total belowground, aboveground and additional belowground plant richness in relation to increasing soil fertility (N content). Number of species (pooled means over transects with minimum and maximum values) from 0.001 $m^3$ sample volume are shown.

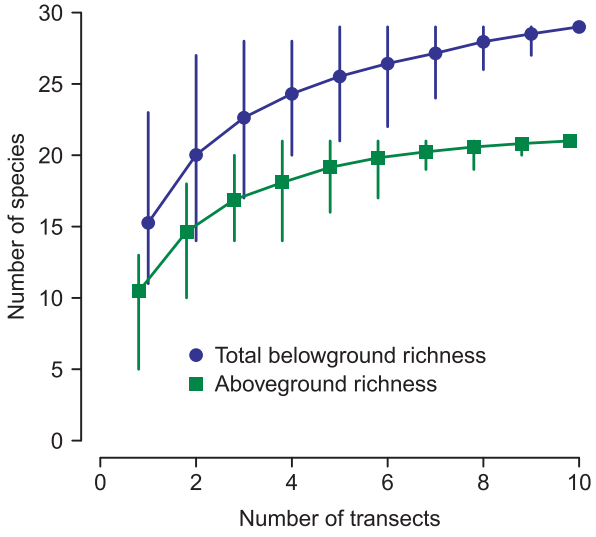


**Fig. 5** Cumulative number of plant species above- and belowground over ten sample groups located across the 2 ha study site. Each transect constitutes a volume of $0.1 \times 0.1 \times 1$ m above and below the soil surface. Cumulative number of species is presented as means with minimum and maximum.

Total belowground richness exceeded aboveground richness at all scales investigated. At the plant neighbourhood scale (0.001–0.008 $m^3$), total belowground richness was, on average, 1.8 times higher than aboveground richness (Fig. 4). At the smallest sampled scale (0.001 $m^3$), differences in above- and total belowground richness were least pronounced (total belowground richness was on average 1.4 times higher than aboveground). Total belowground richness also exceeded aboveground richness at the community scale, as indicated by the pattern of species accumulation from 1 to 10 transects (Fig. 5).

Aboveground richness decreased significantly with increasing soil total N (Fig. 6, GLM analysis $F_{1,71} = 7.5$, $P = 0.03$). In contrast, additional belowground richness increased with soil N ($F_{1,71} = 7.3$, $P = 0.02$), while total belowground richness did not change along a gradient of soil N ($F_{1,71} = 0.1$, $P = 0.59$). Belowground diversity increased with soil N behaving similarly to additional belowground richness ($F_{1,71} = 9.3$, $P < 0.01$), but belowground evenness was not related to soil N ($F_{1,71} = 0.01$, $P = 0.90$).

## Discussion

Our results indicate that conventional aboveground measures of plant richness may omit many coexisting species at both small (plant neighbourhood) and large (community) scales. Furthermore, variation in belowground richness may not be mirrored aboveground. The number of species found only belowground increased significantly along the soil N gradient, whereas aboveground richness decreased, suggesting that above- and belowground richness can differ in how they vary along environmental gradients, and that belowground richness might be relatively more important in fertile soil patches where aboveground richness decreases.

### *Root identification with 454 sequencing*

In total, we obtained >32 000 chloroplast *trn*L (UAA) intron sequences, of which 85% could be assigned to 29 species (or species groups) known to be present in the local species pool of our study site. The majority of the remaining unidentified sequences were likely to be of poor quality (as suggested by no matches against GenBank). Sequencing errors have been reported using a similar 454 metagenomic approach to study ectomycorrhizal fungi (Tedersoo *et al.* 2010). Almost all of the 15% of unidentified sequences that BLAST matched against GenBank were assigned to flowering plants not matching our reference database, but including related taxa. We did not include these matches in our analyses as it was not possible to determine whether they represented sequencing errors or a level of natural variability within reference database species that exceeded the similarity level used in our analysis. Indeed, the 97% similarity threshold is not applicable to all plant species using the *trn*L region sequence between primers *c* and *d*; in a few previous works, this marker has not been able to distinguish between certain related species and genera, especially within *Poaceae*, *Cyperaceae* and *Asteraceae* (Ridgway *et al.* 2003; Frank *et al.* 2010), and this was also the case in our analysis. Taxonomic resolution can be improved by increasing marker sequence variation among species, for example by using a more variable DNA region of the same length or a longer DNA fragment. In addition, the imperfect resolution of the *trn*L (UAA) intron can be compensated for by standardizing the method, wherein the number of possible plant species in a study system is restricted (Taberlet *et al.* 2007).

Sequence identification accuracy can be affected by sequence length (Huse *et al.* 2008). We found that the assignment of artificially shortened *trn*L sequences (142–350 bp) very closely matched (99–100%) assignments based on long sequences (400 bp). This result suggests that taxon identification using the *trn*L amplicon is robust to sequence length variation within the range we considered.

Overall, there was good correspondence in the species composition of the added and observed species in the prepared root mixtures. However, two species added to these samples were not detected by 454 sequencing, even though they both amplified well in single species samples (I. Hiiesalu, personal communication). Non-recovery may have been because of competition during mixed-species PCR, whereby the DNA of some taxa is amplified more efficiently than that of others (Schlichter & Bertioli 1996). This mechanism may also have been responsible for the fact that six species in the field that were observed to be rooted in the sampled plots were not detected using 454 sequencing. In known root mixtures, we did not detect *Agropyron cristatum*, which is closely related to *Elymus repens* another species which could not be detected in the field samples regardless of aboveground dominance of this species. In a similar manner, substantial proportions of graminoids morphologically observed in gut samples were not evident by sequences of P6 loop of *trn*L intron (Soininen *et al.* 2009), a short portion of the *c* and *d* primer amplicon used in our study. We can only speculate that secondary structure and/or length of the amplicon might have impeded with the detection of these species here.

Measuring evenness and diversity, which requires measurement of relative abundances, may be possible using 454 sequencing (Valentini *et al.* 2009b; Amend *et al.* 2010). Evidence for this is that the proportion of sequences of the two most common species in the known root mixtures (*Solidago* group and *Artemisia* spp.) was strongly related to the initial proportions that were added to the mixtures. However, further research is needed to confirm this with more taxa.

We also detected sequences of two species in known mixtures that were not added intentionally. These unexpected sequences were only present in low numbers and were probably derived from small fragments of live root or dead plant material that were attached to the roots of the intentionally added species (species added intentionally to the mixtures were collected from the field, and unexpectedly found species were present in that community).

Overall, these results indicate that 454 sequencing has potential for measuring belowground plant species richness from environmental samples with a possibility for quantitative analysis, as long as certain limits are recognized. Difficulties with taxon recovery and species resolution (ca 20% of our molecular taxonomic units grouped two or more closely related species) remain as constraints of the chloroplast *trn*L(UAA) intron marker as used in this study. Alternative markers have been sought for plant bar-coding, but as yet there are no known barcode regions that would consistently separate plants at the species level (Fazekas *et al.* 2009; Hollingsworth *et al.* 2009). However, it is now possible to acquire 454 sequence read length of up to 1000 bp that would improve the species resolution of *trn*L(UAA) intron considerably by increasing the usable length of the amplicon and thus the number of variable sites per sequence.

### *Plant species richness above- and belowground*

Molecular belowground sampling can reveal species which would otherwise only be recorded by repeated aboveground surveys during different seasons and

vegetation periods (Fridley *et al.* 2006). Within the entire study site, most species occurred more frequently below- than aboveground, and nine species were only detected belowground. These included ephemerals (e.g. *Turritis glabra*), species that produce aboveground shoots and flower later in the vegetative period (e.g. *Solidago virgaurea*), or clonal plants with widespread root systems (e.g. *Vicia cracca*). All species that were found belowground, but not detected aboveground, previously have been found at our study site (J. Liira, unpublished).

Measuring belowground richness should detect species that are present as belowground meristems but absent as shoots in the aboveground community in a particular year of sampling (the 'Carousel Model', van der Maarel & Sykes 1993; Wilson & Tilman 2002). Thus, an additional advantage of belowground measurements is that a relatively short-term single-year study can replace many years of sampling that would be required to detect all members of a community that appear aboveground. On one hand, it could be argued that multiple-year sampling is a cheaper method of detecting dormant species. On the other hand, multiple-year sampling can be confounded by successional trends and does not address other causes of greater belowground diversity.

At the plant neighbourhood scale, a nonlinear pattern between above- and belowground richness indicates saturation of the aboveground community (Cornell & Lawton 1992; Srivastava 1999), even as belowground richness continued to increase. Further, higher total belowground richness compared with aboveground richness, suggests that traditional measures of aboveground plant richness greatly underestimates the number of coexisting species at small scales. Greater belowground richness was also apparent at the community scale: we did not detect any convergence of cumulative species richness with increasing scale over the 10 transects in our 2-ha study site.

Several biological mechanisms might contribute to a higher richness of plants belowground. Virtually all grassland plant species found aboveground have roots or rhizomes in nearby soil, but the converse is not necessarily true: aboveground shoots might not be present at every location where there are roots or rhizomes belowground. A number of processes may lead to an absence of aboveground shoots. For example, clonal plants can become temporarily 'invisible' to the aboveground observer while persisting as rhizome networks with few aboveground shoots (Wildová *et al.* 2007). Thus, high clonal mobility might enhance coexistence belowground (Zobel *et al.* 2010). Also, roots and rhizomes are generally more persistent than shoots, and can survive during unfavourable periods (e.g. winter, heavy grazing), while some species can be dormant for several years (Klimešová & Klimeš 2007). The fact that most species occurred more frequently below- than aboveground confirms our suggestion that spatial and temporal dispersion is greater for roots than shoots. Further, the soil environment contains heterogeneous resources and is rich in micro-organisms that interact with plant roots and influence plant species coexistence (Bever *et al.* 2010).

### *Above- and belowground plant richness related to soil fertility*

Because of the relatively high range of soil nitrogen values at our site, we observed the negative (right-hand side) slope of the unimodal relationship between aboveground richness and soil fertility. The same pattern has been recorded in many earlier experiments and observations (see Mittelbach *et al.* 2001 for a review). In marked contrast, additional belowground richness as well as belowground diversity increased significantly with soil fertility. These results support the idea that aboveground exclusion of species in fertile soils is probably caused by asymmetric light competition where tall plants gain a disproportionate advantage over small ones (Zobel 1992). Roots, however, preferentially grow into fertile patches (Hodge 2004), which may result in symmetric root competition, as all plants are relatively equal in their ability to acquire soil resources (Weiner 1990; Cahill & Casper 2000). Many perennial species can stay dormant belowground until environmental conditions are favourable again (Shefferson *et al.* 2005). In this way, plant species may be buffered against local extinction in fertile soils. This scenario appears plausible in our study site where reduced aboveground biomass in very dry years (Cahill 2003) allows all plant species to capture enough light.

Total belowground richness did not change along the soil fertility gradient because aboveground richness decreased but additional belowground richness increased, causing the overall relationship to remain neutral. Belowground evenness was not related to soil fertility, showing that relative variation in species' abundances belowground was not as important as variation in the number of species. However, this negative result must be treated with caution until the suitability of 454 technology for estimating species abundance is clearer (Amend *et al.* 2010).

Anthropogenic eutrophication is currently a major threat to natural vegetation (Sala *et al.* 2000). Our results suggest that aboveground plant richness might decline relatively rapidly with increasing soil fertility, whereas belowground richness might remain high for a longer period. There must evidently be a temporal limit

to this phenomenon: after some time, a decrease in belowground richness should follow the aboveground decrease. Nonetheless, this might provide a buffer period during which it is possible to restore eutrophic sites by decreasing productivity and easing strong asymmetric light competition.

Future opportunities for exploring belowground diversity will benefit from method refinement. Primary among these will be an ability to detect all species belowground that are known from their aboveground presence. Similarly, we need to identify all species belowground that are not known aboveground. A second major opportunity is to be able to distinguish species currently detectable only as species groups. Here, we have dealt with these concerns by excluding problematic species in order to maintain objectivity, but a complete enumeration is far more desirable. There is evidence that measuring abundance, and consequently being able to measure diversity and evenness, is possible, but this needs to be tested on a much larger data set. All of these challenges are likely to be addressed as techniques continue to improve. Other opportunities for refinement lie in practical matters of obtaining root samples for sequencing that comprise all species, ensuring that only recently living roots are included, and obtaining equal sample volumes of roots and shoots in systems characterized by deep tap roots or tall shoots.

## Conclusions

Application of next-generation sequencing to roots from the natural communities can shed new light on plant biodiversity. Many more plant species coexist within a limited area than are detected using conventional aboveground methods, and plant richness is not proportionally related above- and belowground. Moreover, we observed a decline in aboveground richness in fertile conditions, but this pattern was offset by an increasing number of species occurring only belowground. Incorporating belowground plant diversity into future studies is likely to uncover new patterns that can refine predictions of vegetation responses to biodiversity threats and may stimulate a reassessment of ecological theory.

## Acknowledgements

This study was supported by the European Union through the European Regional Development Fund (Center of Excellence FIBIR), EU 6th framework project ECOCHANGE (FP6-036866), Marie Curie European Reintegration Grant (GLOBAM, PERG03-GA-2008-231034), the Natural Sciences and Engineering Research Council of Canada and the Estonian Science Foundation (grants 7738, 8323).

## References

Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Molecular Ecology*, **19**, 5555–5565.

Bardgett RD (2005) *The Biology of Soil: A Community and Ecosystem Approach*. Oxford University Press, Oxford.

Bever JD, Dickie IA, Facelli E *et al.* (2010) Rooting theories of plant community ecology in microbial interactions. *Trends in Ecology and Evolution*, **25**, 468–478.

Booth MS, Stark JM, Hart SC (2003) Overlapping resource use in three Great Basin species: implications for community invasibility and vegetation dynamics. *Journal of Ecology*, **91**, 36–48.

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR (2009) *Introduction to Meta-Analyses*. John Wiley & Sons Ldt, Chichester, UK.

Cahill JF (2003) Neighbourhood-scale diversity, composition and root crowding do not alter competition during drought in a native grassland. *Ecology Letters*, **6**, 599–603.

Cahill JF, Casper BB (2000) Investigating the relationship between neighbor root biomass and belowground competition: field evidence for symmetric competition belowground. *Oikos*, **90**, 311–320.

Casper BB, Schenk HJ, Jackson RB (2003) Defining a plant's belowground zone of influence. *Ecology*, **84**, 2313–2321.

Cornell HV, Lawton JH (1992) Species interactions, local and regional processes, and limits to the richness of ecological communities: a theoretical perspective. *Journal of Animal Ecology*, **61**, 1–12.

Eissenstat DM, Yanai RD (1997) The ecology of root lifespan. *Advances in Ecological Research*, **27**, 1–60.

Fazekas AJ, Kesanakurti PR, Burgess KS *et al.* (2009) Are plant species inherently harder to discriminate than animal species using DNA barcoding markers? *Molecular Ecology Resources*, **9**, 130–139.

Fisk MC, Yanai RD, Fierer N (2010) A molecular approach to quantify root community composition in a northern hardwood forest – testing effects of root species, relative abundance, and diameter. *Canadian Journal of Forest Research*, **40**, 836–841.

Frank DA, Pontes AW, Maine EM *et al.* (2010) Grassland root communities: species distributions and how they are linked to aboveground abundance. *Ecology*, **91**, 3201–3209.

Fridley JD, Peet RK, Van der Maarel E, Willems JH (2006) Integration of local and regional species-area relationships from space-time species accumulation. *American Naturalist*, **168**, 133–143.

Gibbens RP, Lenz JM (2001) Root systems of some Chihuahuan desert plants. *Journal of Arid Environments*, **49**, 221–263.

Gibson DJ (2008) *Grasses and Grassland Ecology*. Oxford University Press, New York, NY.

Grace JB (1999) The factors controlling species density in herbaceous plant communities: an assessment. *Perspectives in Plant Ecology, Evolution and Systematics*, **2**, 1–28.

Gregory JP (2006) *Plant Roots: Growth, Activity, and Interaction with Soils*. Blackwell Publishing, Oxford.

Grime JP (1979) *Plant Strategies and Vegetation Processes*. John Wiley & Sons, London.

Hautier Y, Niklaus PA, Hector A (2009) Competition for light causes plant biodiversity loss after eutrophication. *Science*, **324**, 636–638.

Hodge A (2004) The plastic plant: root responses to heterogeneous supplies of nutrients. *New Phytologist*, **162**, 9–24.

Hollingsworth PM, Forrest LL, Spouge JL *et al.* (2009) A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 12794–12797.

Huse SM, Huber JA, Morrison HG *et al.* (2007) Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biology*, **8**, R143.

Huse SM, Dethlefsen L, Huber JA, Welch DM, Relman DA, Sogin ML (2008) Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing. *PLoS Genetics*, **4**, e1000255.

Hutchings MJ, John EA (2004) The effects of environmental heterogeneity on root growth and root/shoot partitioning. *Annals of Botany*, **94**, 1–8.

Jackson DA, Somers KM (1991) The spectre of 'spurious' correlations. *Oecologia*, **86**, 147–151.

Jackson RB, Mooney HA, Schulze ED (1997) A global budget for fine root biomass, surface area, and nutrient contents. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 7362–7366.

Jackson RB, Moore LA, Hoffmann WA, Pockman WT, Linder CR (1999) Ecosystem rooting depth determined with caves and DNA. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 11387–11392.

Jones FA, Erickson DL, Bernal MA *et al.* (2011) The roots of diversity: below ground species richness and rooting distributions in a tropical forest revealed by DNA barcodes and inverse modeling. *PLoS ONE*, **6**, e24506.

Kesanakurti PR, Fazekas AJ, Burgess KS *et al.* (2011) Spatial patterns of plant diversity below-ground as revealed by DNA barcoding. *Molecular Ecology*, **20**, 1289–1302.

Klimešová J, Klimeš L (2007) Bud banks and their role in vegetative regeneration – a literature review and proposal for simple classification and assessment. *Perspectives in Plant Ecology, Evolution and Systematics*, **8**, 115–129.

Kummerow J, Krause D, Jow W (1977) Root systems of chaparral shrubs. *Oecologia*, **29**, 163–177.

Linder CR, Moore LA, Jackson RB (2000) A universal molecular method for identifying underground plant parts to species. *Molecular Ecology*, **9**, 1549–1559.

van der Maarel E, Sykes MT (1993) Small-scale plant species turnover in a limestone grassland: the carousel model and some comments on the niche concept. *Journal of Vegetation Science*, **4**, 179–188.

Mittelbach GG, Steiner CF, Scheiner SM *et al.* (2001) What is the observed relationship between species richness and productivity? *Ecology*, **82**, 2381–2396.

Mommer L, Wagemaker CAM, De Kroon H, Ouborg NJ (2008) Unravelling below-ground plant distributions: a real-time polymerase chain reaction method for quantifying species proportions in mixed root samples. *Molecular Ecology Resources*, **8**, 947–953.

Mommer L, van Ruijven J, de Caluwe H *et al.* (2010) Unveiling below-ground species abundance in a biodiversity experiment: a test of vertical niche differentiation among grassland species. *Journal of Ecology*, **98**, 1117–1127.

Moora M, Berger S, Davison J *et al.* (2011) Alien plants associate with widespread generalist arbuscular mycorrhizal fungal taxa: evidence from a continental-scale study using massively parallel 454 sequencing. *Journal of Biogeography*, **38**, 1305–1317.

Moore LA, Field CB (2005) A technique for identifying the roots of different species in mixed samples using nuclear ribosomal DNA. *Journal of Vegetation Science*, **16**, 131–134.

Oksanen J (1996) Is the humped relationship between species richness and biomass an artefact due to plot size? *Journal of Ecology*, **84**, 293–295.

Öpik M, Metsis M, Daniell TJ, Zobel M, Moora M (2009) Large-scale parallel 454 sequencing reveals host ecological group specificity of arbuscular mycorrhizal fungi in a boreonemoral forest. *New Phytologist*, **184**, 424–437.

Pärtel M, Laanisto L, Zobel M (2007) Contrasting plant productivity-diversity relationships across latitude: the role of evolutionary history. *Ecology*, **88**, 1091–1097.

Pegard A, Miquel C, Valentini A, *et al.* (2009) Universal DNA-based methods for assessing the diet of grazing livestock and wildlife from feces. *Journal of Agricultural and Food Chemistry*, **57**, 5700–5706.

R Development Core Team. (2010) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.

Reintal M, Tali K, Haldna M, Kull T (2010) Habitat preferences as related to the prolonged dormancy of perennial herbs and ferns. *Plant Ecology*, **210**, 111–123.

Ridgway KP, Duck JM, Young JPW (2003) Identification of roots from grass swards using PCR-RFLP and FFLP of the plastid trnL (UAA) intron. *BMC Ecology*, **3**, e (e8)8.

Sala OE, Chapin FS, Armesto JJ *et al.* (2000) Biodiversity – global biodiversity scenarios for the year 2100. *Science*, **287**, 1770–1774.

Schenk HJ, Jackson RB (2002) Rooting depths, lateral root spreads and below-ground/above-ground allometries of plants in water-limited ecosystems. *Journal of Ecology*, **90**, 480–494.

Schlichter UHA, Bertioli D (1996) Abundance and competition in PCR. *Nature Biotechnology*, **14**, 1202.

Shefferson RP (2009) The evolutionary ecology of vegetative dormancy in mature herbaceous perennial plants. *Journal of Ecology*, **97**, 1000–1009.

Shefferson RP, Kull T, Tali K (2005) Adult whole-plant dormancy induced by stress in long-lived orchids. *Ecology*, **86**, 3099–3104.

Soininen EM, Valentini A, Coissac E *et al.* (2009) Analysing diet of small herbivores: the efficiency of DNA barcoding coupled with high-throughput pyrosequencing for deciphering the composition of complex plant mixtures. *Frontiers in Zoology*, **6**, 16.

Sønstebø JH, Gielly L, Brysting AK *et al.* (2010) Using next-generation sequencing for molecular reconstruction of past Arctic vegetation and climate. *Molecular Ecology Resources*, **10**, 1009–1018.

Srivastava DS (1999) Using local-regional richness plots to test for species saturation: pitfalls and potentials. *Journal of Animal Ecology*, **68**, 1–16.

Stanton NL (1988) The underground in grasslands. *Annual Review of Ecology, Evolution, and Systematics*, **19**, 573–589.

Szava-Kovats R, Zobel M, Pärtel M (2011) The local-regional species richness relationship: new perspectives on the null-hypothesis. *Oikos*, doi: 10.1111/j.1600-0706.2011.19603.x.

Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*, **17**, 1105–1109.

Taberlet P, Coissac E, Pompanon F *et al.* (2007) Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*, **35**, e14.

Tedersoo L, Nilsson RH, Abarenkov K *et al.* (2010) 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytologist*, **188**, 291–301.

Tilman D (1982) *Resource Competition and Community Structure*. Monographs in Population Biology, Princeton University Press, Princeton, NJ.

Valentini A, Pompanon F, Taberlet P (2009a) DNA barcoding for ecologists. *Trends in Ecology and Evolution*, **24**, 110–117.

Valentini A, Miquel C, Nawaz MA *et al.* (2009b) New perspectives in diet analysis based on DNA barcoding and parallel pyrosequencing: the trnL approach. *Molecular Ecology Resources*, **9**, 51–60.

Wardle DA (2002) *Communities and Ecosystems: Linking the Aboveground and Belowground Components*. Princeton University Press, Princeton.

Weiner J (1990) Asymmetric competition in plant populations. *Trends in Ecology and Evolution*, **5**, 360–364.

Wells CE, Eissenstat DM (2001) Marked difference in survivorship among apple roots of different diameters. *Ecology*, **82**, 882–892.

Wildová R (2004) Below-ground spatial pattern of rhizomes in a grassland community and its relevance to above-ground spatial pattern. *Plant Ecology*, **174**, 319–336.

Wildová R, Wild J, Herben T (2007) Fine-scale dynamics of rhizomes in a grassland community. *Ecography*, **30**, 264–276.

Wilson SD, Tilman D (2002) Quadratic variation in old-field species richness along gradients of disturbance and nitrogen. *Ecology*, **83**, 492–504.

Zobel M (1992) Plant species coexistence: the role of historical, evolutionary and ecological factors. *Oikos*, **65**, 314–320.

Zobel M, Moora M, Herben T (2010) Clonal mobility and its implications for spatio-temporal patterns of plant communities: what do we need to know next? *Oikos*, **119**, 802–806.

---

This research is perfomed in cooperation of different research groups joining together ecologists and molecular biologists. Our research interests include biodiversity patterns and processes, functioning and conservation of grassland ecosystems, molecular ecology of plants and fungi, and metagenomic environmental research.

---

## Data accessibility

Reference database sequences: GenBank accession nos HM590228–HM590365.

454 Sequences and species richness data: Dryad doi: 10.5061/dryad.cg8q67q5.

## Supporting information

Additional supporting information may be found in the online version of this article.

**Fig. S1** Frequency distribution of 454 sequence lengths.

**Fig. S2** Scatterplot of logratio-transformed proportions of added biomass and detected number of sequences for two most common species (*Solidago* group and *Artemisia* spp.) in known root mixtures.

**Fig. S3** Scatterplot of logratio-transformed aboveground richness and total belowground richness and total belowground species richness (mean from 0.001 $m^3$ sample volume).

**Table S1** The effect of sequence length on MOTU (molecular operational taxonomic unit) identification.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.