Received: 27 May 2019 | Accepted: 18 September 2019

DOI: 10.1111/1365-2745.13296

RESEARCH ARTICLE

Journal of Ecology | BRITISH ECOLOGICAL SOCIETY

Remote sensing of ploidy level in quaking aspen (*Populus tremuloides* Michx.)

Benjamin Blonder[1,2,3,4] | **Bente J. Graae**[2] | **Burke Greer**[3] | **Marja Haagsma**[5] | **Kenny Helsen**[2,6] | **Rozália E. Kapás**[2,3,7] | **Henry Pai**[8,9] | **Jolanta Rieksta**[2,3] | **Dillon Sapena**[3,10] | **Christopher J. Still**[3,11] | **Richard Strimbeck**[2,3]

[1]Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, UK; [2]Department of Biology, Norwegian University of Science and Technology (NTNU), Trondheim, Norway; [3]Rocky Mountain Biological Laboratory, Crested Butte, CO, USA; [4]School of Life Sciences, Arizona State University, Tempe, AZ, USA; [5]Department of Biological and Ecological Engineering & Water Resources Engineering Graduate Program, Oregon State University, Corvallis, OR, USA; [6]Plant Conservation and Population Biology, Biology Department, University of Leuven, Heverlee, Belgium; [7]Department of Physical Geography, Stockholm University, Stockholm, Sweden; [8]Department of Geological Sciences and Engineering, University of Nevada Reno, Reno, NV, USA; [9]Northwest River Forecast Center, National Oceanic and Atmospheric Administration, Portland, OR, USA; [10]Department of Ecology, Montana State University, Bozeman, MT, USA and [11]Department of Forest Ecosystems and Society, Oregon State University, Corvallis, OR, USA

Correspondence
Benjamin Blonder
Email: bblonder@gmail.com

Present address
Jolanta Rieksta, Terrestrial Ecology Section, Department of Biology, University of Copenhagen, Universitetsparken 15, DK-2100, Copenhagen, Denmark

Funding information
Natural Environment Research Council, Grant/Award Number: NE/M019160/1; National Science Foundation, Grant/Award Number: EAR-1440596 and EAR-1440506; Norwegian Research Council, Grant/Award Number: KLIMAFORSK 250233

Handling Editor: Emily Lines

Abstract

1. Ploidy level in plants may influence ecological functioning, demography and response to climate change. However, measuring ploidy level typically requires intensive cell or molecular methods.
2. We map ploidy level variation in quaking aspen, a dominant North American tree species that can be diploid or triploid and that grows in spatially extensive clones. We identify the predictors and spatial scale of ploidy level variation using a combination of genetic and ground-based and airborne remote sensing methods.
3. We show that ground-based leaf spectra and airborne canopy spectra can both classify aspen by ploidy level with a precision-recall harmonic mean of 0.75–0.95 and Cohen's kappa of *c*. 0.6–0.9. Ground-based bark spectra cannot classify ploidy level better than chance. We also found that diploids are more common on higher elevation and steeper sites in a network of forest plots in Colorado, and that ploidy level distribution varies at subkilometer spatial scales.
4. *Synthesis*. Our proof-of-concept study shows that remote sensing of ploidy level could become feasible in this tree species. Mapping ploidy level across landscapes could provide insights into the genetic basis of species' responses to climate change.

KEYWORDS
adaptation, ploidy level, polyploidy, quaking aspen, reflectance, remote sensing, spectrometry, UAS

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2019 The Authors. *Journal of Ecology* published by John Wiley & Sons Ltd on behalf of British Ecological Society

1 | INTRODUCTION

Species do not respond uniformly to environmental change (Jump & Penuelas, 2005). Genetic variation within populations drives phenotypic variation, creating a mosaic of successful and unsuccessful genotypes under novel conditions (Alberto et al., 2013). Identifying the genetic and phenotypic structure within a population is critical not only for fundamental understanding of evolution by natural selection, but also for conservation and management applications where forecasting or mitigating the effects of environmental change are desired.

For some species, a key component of genotypic variation is ploidy level variation (variation in cytotype or chromosome copy number) among individuals. Polyploidy leads to large effects on organismal phenotype, either directly through changes in genome size, for example by influencing stomatal cell size and thus water-use efficiency in plants (Beaulieu, Leitch, Patel, Pendharkar, & Knight, 2008; Greer, Still, Cullinan, Brooks, & Meinzer, 2017), or indirectly through changes in gene expression (e.g. variation in longevity due to mutational buffering, or variation in growth rate due to increased copy numbers of key genes). As a consequence, individuals with different ploidy levels within a species tend to occupy different environmental and geographical spaces (Otto & Whitton, 2000; Parisod, Holderegger, & Brochmann, 2010), suggesting that knowledge of ploidy levels is important for understanding species response to environmental change. While several hypotheses for the ultimate drivers of polyploidy in plants have been advanced (e.g. related to cold temperatures, reproductive systems, Levin, 1983; Martin & Husband, 2009; Ramsey & Schemske, 1998), few have been tested due to limited available spatial data.

A key example of polyploidy and ploidy level variation occurs in quaking aspen (*Populus tremuloides* [Salicaceae]), the most widely distributed tree species in North America. Quaking aspen occurs over 47° of latitude from central Mexico to northern Alaska, and often forms monodominant stands. The species has high economic and cultural value (Jones & Markstrom, 1973; McCool, 2001), as well as ecological value through provision of habitat and resources to a wide range of organisms (Anderegg, Anderegg, Sherman, & Karp, 2012; Mitton & Grant, 1996). Individual stems (ramets) grow in genetically identical clones (genets), often >1 ha in area. Genets have either two (diploid) or three (triploid) copies of each chromosome (Kemperman & Barnes, 1976; Mock et al., 2012). Phenotypic variation within and across genets is very high, both due to plastic expression of traits and genotypic variation (Barnes, 1969, 1975). Triploids often have different phenology, stem size and bark texture, and compared to diploids have been found to have larger leaves, faster growth rates, higher carbon uptake rate, higher stomatal conductance and higher water-use efficiency (Benson & Einspahr, 1967; Einspahr, Buijtenen, & Peckham, 1963; Every & Wiens, 1971; Greer et al., 2017). Triploids also vary more in their environmental niche (Greer, Still, Howe, Tague, & Roberts, 2016). Bark in both diploids and triploids is photosynthetic (Mitton & Grant, 1996).

The drivers of ploidy level variation in quaking aspen are not completely understood. Diploids are more common in eastern and boreal populations, while triploids are more common in the southwest portion of the range (Callahan et al., 2013; Mock et al., 2012). Nevertheless, co-occurrence of diploids and triploids within sites at <100 m spatial scales is common (Bishop, Furniss, Mock, & Lutz, 2019; Mock, Rowe, Hooten, Dewoody, & Hipkins, 2008). These conclusions are based on a relatively limited set of available data, so that finer-scale spatial patterns of ploidy level variation, or their consequences, remain unknown.

High range-wide mortality of quaking aspen forests has been observed in recent decades, for example 50%–60% in some parts of southwestern Colorado (Worrall et al., 2008). This 'sudden aspen decline' is forecast to become more severe in coming decades (Worrall et al., 2013) and was initiated by a stretch of unusually hot and dry years in the early 2000s. Mortality is thought to be caused by drought weakening trees until they are killed by hydraulic failure or pathogens (Anderegg et al., 2013). These stressors do not equally affect all forests, for unknown reasons (Hogg, Brandt, & Kochtubajda, 2002; Huang & Anderegg, 2012; Michaelian, Hogg, Hall, & Arsenault, 2011). Mortality occurs patchily at small spatial scales, suggesting selection on certain genotypes with potentially large consequences for genetic diversity, range dynamics and phenotypic evolution. Triploids may be at higher risk for drought-induced mortality (Dixon & DeWald, 2015). There is thus a need to better understand fine-scale spatial distribution of ploidy levels, as well as its drivers and consequences, in quaking aspen as well as in other species.

Assessing ploidy level requires intensive laboratory-based work, for example flow cytometry to separate cells with different genome sizes (Greer et al., 2017; Mock et al., 2012), cytotype counts of chromosome number for cells imaged at metaphase (Barnes, 1969) or DNA-based counts of allele (and thus) chromosome number via either microsatellite analysis (Mock et al., 2008) or, more recently, genotyping-by-sequencing on next-gen platforms (Gompert & Mock, 2017). These methodological issues have limited the feasibility of empirical studies of polyploidy in a biogeographical or climate change context.

We propose that remote sensing methods can instead be used to rapidly and inexpensively measure variation in ploidy level in quaking aspen. While our application is focused on this widespread species, the fundamental concepts may be applicable to some other plant species with intraspecific ploidy level variation, for example sagebrush (Pellicer et al., 2010), and many grasses and trees (Keeler, 1998; Wood et al., 2009). The premise of the method is that genetic variation should lead to phenotypic variation in the functional traits of tissues such as bark and leaves (Asner et al., 2017). Furthermore, we assume that this observable chemical (phenotypic) variation is much more pronounced for genotypic variation due to chromosome number across ploidy levels than for allele frequency variation within ploidy levels. This chemical variation should in turn lead to variation in how these tissues absorb or reflect light of different wavelengths (Curran, 1989; Sims & Gamon, 2002). For example, leaf chlorophyll drives absorptance in multiple portions of the visible spectrum,

while leaf nitrogen content drives absorptance in certain portions of the near-infrared spectrum (Yoder & Pettigrew-Crosby, 1995). This spectral variation can be measured easily using optical techniques. Species classification, a similar classification problem to ploidy level discrimination, has been previously addressed using airborne spectral data (Asner et al., 2017; Ustin, Roberts, Gamon, Asner, & Green, 2004). Moreover, prior work has shown that leaf chemical traits (e.g. chlorophyll content) and spectral properties do differ between aspen ploidy levels (Greer et al., 2017), and that quaking aspen genotypes can be discriminated using airborne spectral data (Madritch et al., 2014).

Here we leverage genetic data to determine how topographical variation affects ploidy level in quaking aspen, and to quantify the spatial scales over which ploidy level varies on natural landscapes. We then use these data to ask whether the reflectance spectra of leaf and bark tissue predict ploidy level. We address this question using both ground-based measurements where spectra are obtained from plant tissues, as well as airborne unoccupied aerial system (UAS) measurements.

2 | MATERIALS AND METHODS

2.1 | Site selection

During the summers of 2016 and 2017, we established a network of aspen forest sites spanning a wide range of elevation and

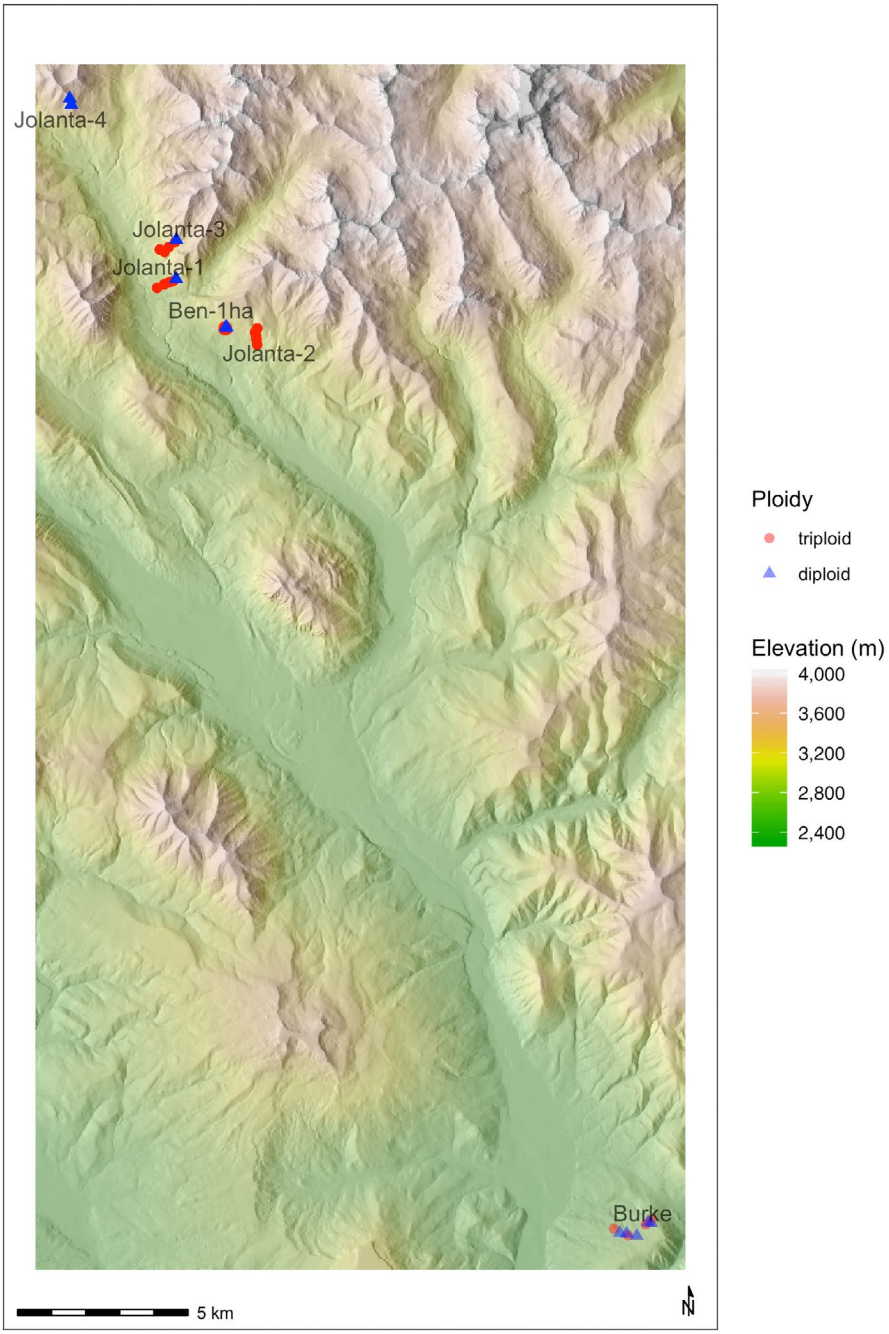

FIGURE 1 Map of sites in southwestern Colorado. Points indicate plots in which ploidy level was determined: triploid, red circles; diploid, blue triangles. Note that many plots overlap each other, for example within the *Ben-1ha* site, at this resolution (see Supporting Information for zoomed versions) [Colour figure can be viewed at wileyonlinelibrary.com]

environmental conditions. Sites were located in southwestern Colorado, over a 37 km maximum distance, near to the towns of Almont, Crested Butte and Gothic, CO (Figure 1). All sites were located within the Gunnison National Forest and were mapped using a handheld GPS unit (Trimble, GeoXT) and/or a laser rangefinder (LaserTech, TruPulse 360R). Sites were chosen to span a locally representative range of environmental conditions for aspen in the region, and spanned an elevation range of 2,730–3,630 m. Forest types ranged from mature stands (>20 m height) with dense understory vegetation to small stunted stands (<0.2 m height) in alpine scree fields. Substrates included a wide range of soil development stages and parent rock materials.

Each site comprised multiple georeferenced plots. Plots were located within a 1 km radius of the site center. A total of 220 total independent measurements of ploidy were made. For cost reasons it was not feasible to obtain an independent ploidy estimate for every stem and leaf in the study, so unique ploidy level measurements were assigned to multiple plots and multiple samples within plots (e.g. for two stems separated by *c*. 1 m distance within an isolated and stunted stand, or for two leaves on the same stem). Thus, the effective sample sizes for each site and type of analysis was variable and often larger than the number of ploidy measurements. We were not concerned about pseudoreplication because the analysis focused on classifier predictive ability rather than statistical significance, and because we controlled for sample size via resampling approaches (see below). Details of the sampling are given in Tables 1 and 2. All data, including the locations of independent ploidy measurements, are available in Files S1 and S2.

TABLE 1 Total number of unique genetic samples per site

Site	Number triploid	Number diploid
Ben-1ha	30	6
Burke	5	5
Jolanta-1	40	10
Jolanta-2	51	0
Jolanta-3	39	10
Jolanta-4	0	22

2.2 | Genotypic analysis

We obtained several healthy and mature canopy leaves using slingshot and rope techniques for tall trees or hand pruners for small trees. Each leaf sample was pressed flat and dried at ambient temperature in silica desiccant for 3–5 days. After drying, samples were analysed for ploidy level. A total of 220 ploidy measurements were obtained. The difference in methodology arose as the study combined data from different co-authors' independent projects. Details are given in Tables 1 and 2.

Ploidy level was determined for 210 samples (at the *Ben-1ha*, *Jolanta-1*, *Jolanta-2*, *Jolanta-3*, and *Jolanta-4* sites) via microsatellite analysis following (Mock et al., 2012, 2008). DNA was extracted from each sample using the E.Z.N.A HP plant DNA mini kit (Omega Bio-tek Inc.). We used 12 unlinked microsatellites, three developed by (Smulders, Schoot, Arens, & Vosman, 2002) (WPMS 014-016), three developed by (Tuskan et al., 2004) (ORPM 028, 059 and 206) and six sourced from http://www.ornl.gov (PMGC 433, 510, 575, 667, 2,571 and 2,658). DNA amplifications were carried out in two multiplexes of six microsatellite markers in 10 μl reactions containing 2.4 μl of one of the multiplexed primer combinations (0.1–0.4 μM primer concentrations), 1 μl template DNA, 5 μl Qiagen Multiplex PCR Master Mix and 1.6 μl RNAse-free water. We used a 'touchdown' PCR protocol adapted from Cole, 2005, with an initial denaturation at 92°C for 5 min, followed by nine cycles of 45 s at 92°C, 45 s at 59°C (dropping by 1°C each cycle to 50°C) and

TABLE 2 Sampling coverage per site and method

Type	Site	Number of genetic samples used	Number of plots	Number of trees	Number of spectra
Bark	*Ben-1ha*	19	5	75	76
Bark	*Jolanta-1*	19	1	3	19
Bark	*Jolanta-2*	50	1	5	50
Bark	*Jolanta-3*	19	1	2	19
Leaf	*Ben-1ha*	36	36	37	110
Leaf	*Burke*	10	10	10	29
Leaf	*Jolanta-1*	50	5	50	400
Leaf	*Jolanta-2*	45	5	45	355
Leaf	*Jolanta-3*	49	5	49	391
Leaf	*Jolanta-4*	22	3	22	179
Canopy	*Ben-1ha*	36	36	36	3,316
Canopy	*Jolanta-1*	50	5	5	6,170
Canopy	*Jolanta-2*	51	5	5	8,463
Canopy	*Jolanta-3*	29	3	3	5,451

60 s at 72°C. This was followed by 21 cycles of 45 s at 92°C, 45 s at 50°C and 60 s at 72°C, with a final extension step of 5 min at 72°C. After PCR, 1 μl of the reaction was added to a solution of 9.35 μl formamide and 0.15 μl of the Applied Biosystems' GeneScan 500 LIZ size standard. Fragments were subsequently sized on a 3,130×L Genetic Analyzer (Applied Biosystems) and scored with GeneMapper Software v4.0 (Applied Biosystems). Markers ORPM 206 and PMGC 2,571 failed to amplify reliably, while marker ORPM 028 was monomorphic across our dataset, resulting in a total of nine informative microsatellite markers. Samples were defined as triploid if three alleles were observed for at least one of the nine markers, or as diploid when a maximum of only two alleles was observed for each marker (Table S1). Note that previous genetic research across the full range of *P. tremuloides* showed high reliability ploidy assessment based on 6–10 microsatellite markers (97% correct classification across 296 individuals) (Mock et al., 2012). This reliability was mainly due to the high genetic diversity (and heterozygosity levels) within and across clones in this species (Mock et al., 2012, 2008), which were also present in our dataset (Table S2). There was also no evidence for linkage disequilibrium (r_d = .008, p = .50 (999 permutations, via the poppr R package), thus indicating that sufficient sexual reproduction does occur among clones (Agapow & Burt, 2001). For these reasons we have confidence in ploidy level inferred from microsatellite data.

Ploidy level was also measured for 10 samples using flow cytometry (at the *Burke* site). One square centimeter sized sections of aspen leaf were combined with equal-sized fresh standard samples of the diploid species, *Hordeum vulgare* (1C genome size is 5.55 pg; Bennett & Leitch, 2005). Nuclei were suspended and stained using the CyStain PI Absolute T kit, Sysmex America, Inc. For nuclei extraction, 150 μl of extraction buffer (with 2% by volume polyvinylpyrrolidone) was added to the chopped leaf material. Then, the suspension was filtered using disposable tube top filters (CellTrics, Sysmex Partec) and 750 μl of stain (CyStain, Sysmex Partec) was added. Filtrates were analysed using a flow cytometer (Accuri C6, BD Biosciences) and excited using a 585 nm laser. Comparisons between samples of the ratios of the median peak fluorescence of each aspen sample relative to the *H. vulgare* peak fluorescence were used to determine the ploidy of each sample. Triploids were classified as those samples with 50% more fluorescence than the diploid standard.

2.3 | Ground-based leaf spectra

Before drying, fresh leaves from the above collection procedure were measured on their adaxial side, avoiding the main vein. Samples were maintained in moist plastic bags in a cooler or refrigerator before measurement. Measurements were obtained using a field spectrometer (ASD Inc., Handheld 2, with leaf clip and internal light source) over the 325–1,075 nm range at 1 nm intervals. The spectrometer was calibrated against a white and a black reference (ASD Inc.) before making leaf measurements. Three replicate spectra were obtained for each leaf sample.

2.4 | Ground-based bark spectra

While bark spectral measurements could be made non-destructively on stems in the field, logistical issues precluded use of our instrument outside of a laboratory. Thus, samples of *c.* 5 cm^2 area and 2 mm depth were cut from trees using a sterilized knife at *c.* 1.3 m above the ground and were used for subsequent measurement in the laboratory. Sampling locations on the stem were chosen to be homogeneous and smooth, avoiding scars, cracks, animal herbivory, and other types of bark damage. Bark samples were kept in a moist paper bag in a cooler or refrigerator prior to measurement. Measurements were obtained over the same wavelength range and with the same instrument as described above, with three spectral replicates per sample.

2.5 | Airborne canopy spectra

In July 2017, we obtained multispectral images covering a total of four sites each of *c.* 500 m length. Sites were overflown by an UAS (Tarot, T560 Sport) flying a raster scan pattern at *c.* 90–120 m above-ground level (*c.* 6–7 cm/pixel ground resolution or *c.* 5–6 cm/pixel treetop resolution). Flights occurred in late-morning conditions during fully sunlight or fully cloudy conditions (i.e. minimizing shadows from partly cloudy skies). Due to weather and permitting issues, flights were not carried out at the *Burke* or *Jolanta-4* sites, nor at the very eastern edge of the *Jolanta-3* site.

Data were collected using a multispectral camera (Micasense, RedEdge) on gimbal mount. The camera obtained co-registered coverage of five spectral bands: blue (475 ± 20 nm), green (560 ± 20 nm), red (668 ± 21 nm), red edge (717 ± 10 nm) and near-infrared (840 ± 40 nm). Immediately prior to data collection, a calibration image was obtained of a ground-based grey reference panel (Micasense, Calibrated Reflectance Panel). UAS flight missions were made using Universal Ground Control Software (SPH Engineering), which automatically accounts for approximate elevation gradients within the flight area by maintaining the elevation above-ground level via an integrated pressure sensor on the flight controller (Pixhawk). The flight missions specified 70%–80% front and side image overlap. Simultaneous to imaging, incident radiation was collected using a downwelling sensor at 1 s intervals. The multispectral images were stitched using Pix4D Mapper software. Pixel values were then converted to reflectance values based on calibration against the downwelling radiation and grey reference data.

2.6 | Spatial analysis of genetic data

We determined whether quaking aspen's realized niche along topographical axes varied with ploidy level. We extracted information for the slope (degrees), cosine aspect (dimensionless) and elevation (m) at the location of each genetic sample, using the USGS National Elevation Dataset. We built linear mixed models using each topographical variable as a response variable, ploidy level as a fixed

effect predictor variable and site as a random intercept predictor variable. We assessed statistical significance of ploidy level using the Satterthwaite approximation.

We also determined the characteristic spatial scale of ploidy level variation, defined as the median distance between points in quaking aspen forest before a change in ploidy level. We calculated the pairwise geographical distance and absolute ploidy level difference between all genetic samples. For each point, we then identified the minimum distance necessary for a change in ploidy level.

2.7 | Spatial analysis of spectral data

We pre-processed the ground-based leaf and bark spectra. We first removed data for 15 spectra (<1%) which included features not characteristic of living vegetation, for example absence of a chlorophyll peak in green bands. These removed spectra were assumed to represent calibration issues or light leaks in the instrument. We then used reflectance data from 400–1,075 nm, cutting off the 325–399 nm bands due to low signal-to-noise ratio. We also averaged the three replicate spectra for each sample into a single composite spectrum and further smoothed this signal with a Savitzky–Golay filter of order 1 and length 21.

We also pre-processed the UAV-based canopy spectra. Images were aggregated to 0.5 m resolution to reduce contrast from small-scale features like shrubs and rocks. Images were then thresholded to include only canopy pixels, by both calculating normalized difference vegetation index (NDVI) and then retaining pixels with NDVI values above a threshold value (0.8), and also retaining pixels with mean reflectance above a lower threshold value (0.05–0.13, dependent on image). Thresholds were manually chosen to best mask unshaded canopy pixels. Spectral information was then extracted from a set of pixel values within a 4 m radius of the focal trees in the masked image. Reflectance values at these pixels were treated as replicates for each plot.

We also spatially interpolated ploidy level data to be able to assign values to samples collected adjacent to the ploidy level measurements. Ploidy measurements were assigned to multiple plots via interpolation over small distances. Interpolation was carried out using a *k*-nearest-neighbour interpolation with $k = 1$ (i.e. where an unknown sample is assigned the same ploidy level as the nearest sample with known ploidy level). Additionally, one or more spectral samples were obtained from each plot, either via ground-based methods (e.g. measurements of multiple leaves on a single stem) or via airborne methods when, for example, multiple adjacent pixels were located around a single location. In the case of spatial interpolation of ploidy, or treatment of adjacent pixels as replicates, we were guided by prior work demonstrating strong spatial homogeneity of clonal identity (and thus also of ploidy level) across landscapes at <50 m spatial scales (Mock et al., 2008), as well as our visual observations of consistent phenotypes at these spatial scales. We never interpolated any data beyond 50 m distance.

We summarized spectral variation using metric multidimensional scaling (Gauch, 1982). Distances between spectra were calculated using a Bray–Curtis metric, and then projected into $k = 2$ dimensions for visualization.

We built random forest models to predict ploidy level based on reflectance spectra predictors for each of the hyperspectral ground-based leaf and ground-based bark, and the multispectral airborne canopy datasets. Random forests are an ensemble learning method for classification that assemble a large set of decision trees based on random subsets of training data, and then make predictions based on votes from the set of decision trees (Breiman, 2001). Random forests were used because they often give good performance on multispectral imagery (Adam, Mutanga, Odindi, & Abdel-Rahman, 2014). The number of bootstrap samples and variables sampled were chosen according to software defaults.

As our datasets were unbalanced, there was a risk that models would be better trained to classify triploids than diploids. We therefore built an ensemble of 10 random forest models, each constructed after resampling triploid data to the same number of observations as the number of diploids in each analysis.

To better compare hyperspectral and multispectral data, we also repeated the above analyses for hyperspectral data after reducing its dimensionality to the first five components via principal components analysis (after scaling and centering data).

To assess the model performance, we report three types of statistics. First, we report overall model performance using the F_1 score (Sammut & Webb, 2010), which is the harmonic mean of precision (true positives divided by sum of true positives and false positives) and recall (true positives divided by true positives and false negatives). Values of F_1 closer to 1 indicate better performance. Second, we report Cohen's kappa (McHugh, 2012), which is the improvement in the model relative to the null expectation of random guessing (In a two-class problem with balanced sampling, 50% of classifications will be correct by chance). Values of kappa above zero (and closer to 1) indicate better performance than random. Third, we report the classification error rate for each of diploids and triploids as the predictive accuracy in out-of-bag predictions, that is on randomly selected data not used in fitting of each tree. This type of model evaluation is similar to a cross-validation and enables unbiased estimate of the testing set error. In practice *c*. one-third of the data is left out of bootstrap samples on each iteration of the model. Error rates closer to zero indicate better performance. These statistics were all summarized across 10 random forest model replicates in order to capture any variation driven by the resampling process.

To determine whether results were driven by site-level rather than ploidy level differences, we also repeated the above analyses via a spatially explicit cross-validation exercise. Models were trained using data for a single site and then tested on data for another single site. This process was repeated for every pairwise combination of sites, and carried out for an ensemble of 10 models, each using a balanced data resample. Model performance was summarized as test classification error rates for diploids and triploids.

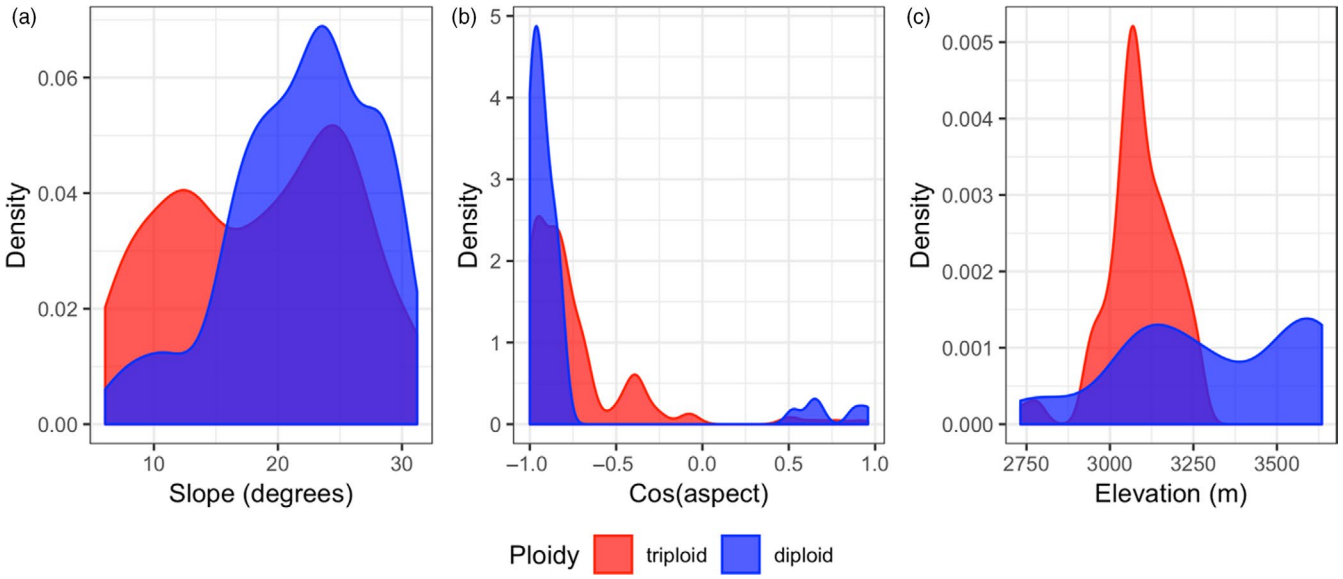

FIGURE 2 Distribution of ploidy levels across topographical gradients of (a) slope, (b) cosine aspect (−1 = south-facing, 1 = north-facing) and (c) elevation [Colour figure can be viewed at wileyonlinelibrary.com]

We also used the original ensemble models to make spatially explicit predictions of ploidy level using the canopy multispectral imagery. We first took each multispectral image and masked out non-aspen pixels, by training an additional random forest model on hand-selected 1 m^2 regions (30 aspen canopy regions, 30 non-aspen canopy regions). Regions were selected to have high spectral diversity. This process yielded good results by visual inspection. We then used spectral values at each non-masked pixel as inputs to the ensemble of 10 random forest models previously trained on data from all sites. Each model made a prediction of ploidy level; the final predicted value at each pixel was chosen as the majority-vote across the ensemble.

All statistical, image, and GIS analyses were conducted in R (3.5.1). Image data were processed using the *raster* (2.6-7) and *sp* (1.3-1) packages. Spectra were processed using the *RStoolbox* (0.2.1) and *signal* (0.7-6) packages. Interpolation of ploidy was carried out using the *fnn* (1.1) package. Ordination was carried out using the *vegan* (2.5-3) package. Random forest models were implemented using Breiman's algorithm in the *randomForest* (4.6-14) package, choosing default parameters. Mixed models were built with the *lme4* (1.1-19) and *lmerTest* (3.1-0) packages. Classification statistics were calculated using the *caret* (6.0-80) package.

3 | RESULTS

Genetic analysis indicated that the majority of samples were for triploid (n = 165) rather than diploid (n = 55) individuals (Table 1). Details of numbers of trees and spectra measured within each plot and site are given in Table 2.

Topography had a strong influence on the distribution of ploidy levels (Figure 2). There was a significant effect of ploidy level on elevation ($p < 10^{-12}$, mean shift = 72 m) as well as on slope ($p < .01$, mean shift = 2.8°), with diploids occurring at higher and steeper locations. Diploids also occurred at more south-facing aspects ($p < .001$), though the bimodal distribution of values makes the statistical inference uncertain.

Diploids and triploids co-occurred at most, but not all sites (Figure 1). Ploidy level varied primarily at subkilometer spatial scales (Figure 3). The median distance between ploidy levels was 377 m (interquartile range, 72–886 km).

Each spectral dataset varied in the clarity of separation between diploids and triploids. An example of this separation is shown for the *Jolanta-1* site in Figure 4, with all multispectral sites shown in Figures S1–S4. For bark spectra, overlap between ploidy levels was high (Figure 5a), with diploids appearing to occupy a subset of the triploid spectral space (Figure 5b). For leaf spectra, overlap was lower (Figure 5c), with diploids having shifted and somewhat unique spectra

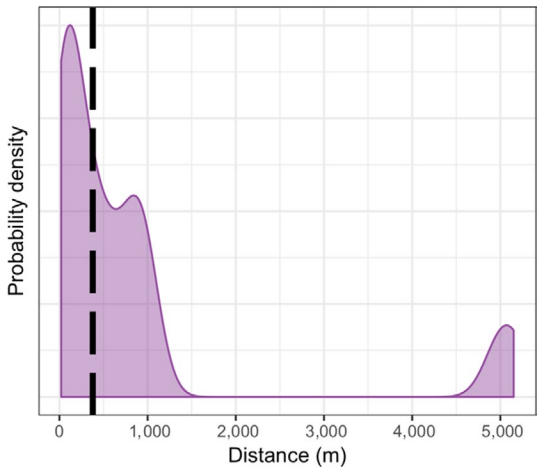

FIGURE 3 Estimates of distances between ploidy levels. Shaded region shows distributions of minimum distances; dashed vertical line indicates median [Colour figure can be viewed at wileyonlinelibrary.com]

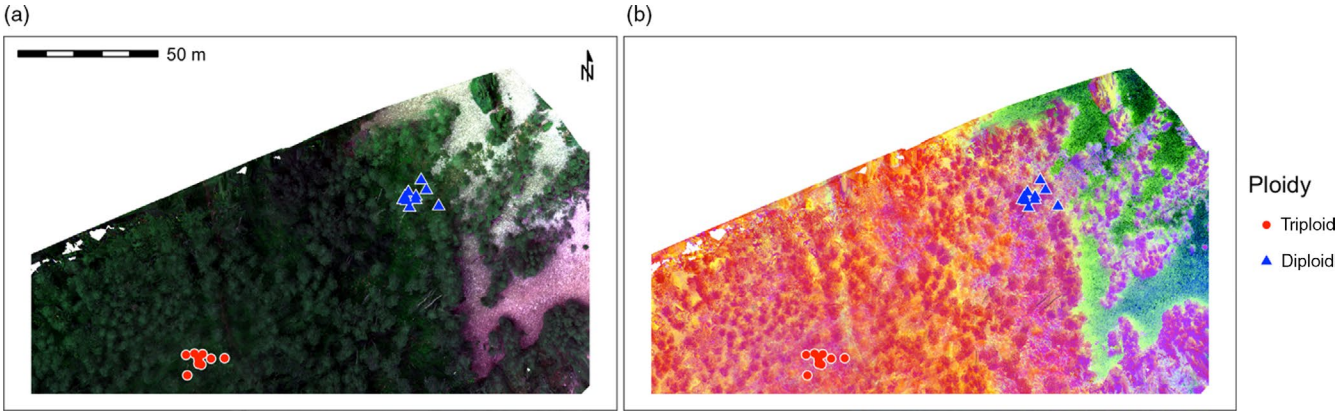

FIGURE 4 Example airborne canopy image subset obtained from the unmanned aerial vehicle and five-band multispectral camera from the upper elevation end of the *Jolanta-1* site. Data are shown using (a) true colour and (b) false colour shaded according to scores from a principal component analysis of all five spectral bands. Georeferenced samples at the left of the image are triploid (red circles), while samples at the right portion of the image are diploid (blue triangles) [Colour figure can be viewed at wileyonlinelibrary.com]

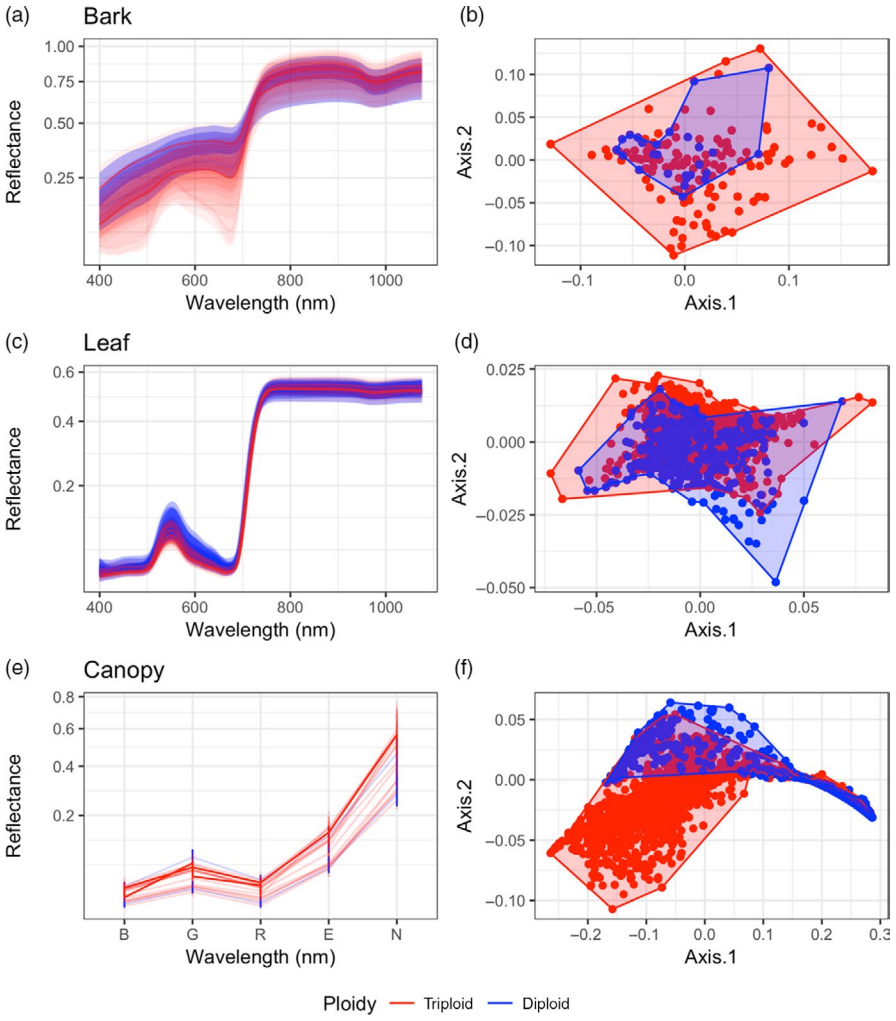

FIGURE 5 Reflectance variation with ploidy level for (a,b) ground-based bark spectra, (c,d) ground-based leaf spectra and (e,f) airborne canopy spectra. Left panels (a,c,e) show reflectance variation across wavelength. Curves are shown for each genetic clone with error bars indicating 95% quantiles of distributions across all stems and leaves within each clone. Right panels (b,d,f) show ordination of the same spectral data (via metric dimensional scaling into $k = 2$ dimensions), with boundaries indicated as alpha hulls. Triploids are shown in red; diploids in blue [Colour figure can be viewed at wileyonlinelibrary.com]

(Figure 5d). In particular, diploids appeared to have higher reflectance in the visible (green) portion of the spectrum. For canopy spectra, overlap between ploidy levels was also low (Figure 5e), with diploids again having shifted and more unique spectra (Figure 5f). Variation among ploidy levels was most apparent in the green and near-infrared spectral bands.

These qualitative results were paralleled by the random forest classifier models. The models using the bark dataset had F_1 scores of 0.53 ± 0.07 (*M* ± *SD*) and Cohen's kappa scores of 0.12 ± 0.14. Consistent with these low values, predictive error rates were high: 0.38 ± 0.09 for diploids and 0.50 ± 0.07 for triploids (Figure 6a,b).

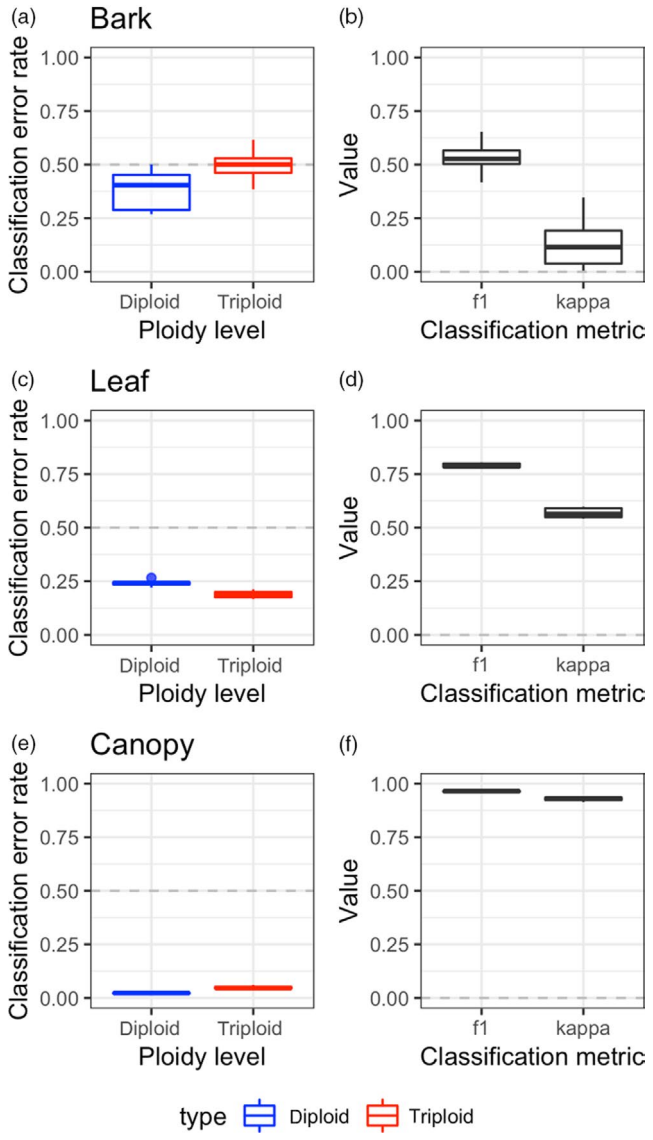

FIGURE 6 Predictive performance of models of ploidy level based on (a,b) bark spectra, (c,d) leaf spectra and (e,f) canopy spectra. Left columns (a,c,e) represent out-of-bag error rates for diploids and triploids. The grey-dashed line indicates the random guessing expectation. Right columns (b,d,f) indicate metrics of overall model performance: The F_1 score (Sørensen–Dice coefficient) and Cohen's kappa. Distributions represent outputs from 10 random forest models each trained on a sample of the full spectral data balanced to include equal number of diploid and triploid data points [Colour figure can be viewed at wileyonlinelibrary.com]

After reducing dimensionality with principal components analysis, F_1 scores increased slightly to 0.64 ± 0.07, and Cohen's kappa increased threefold to 0.30 ± 0.13. In contrast, the models using the leaf dataset had better performance: F_1 scores of 0.79 ± 0.01 and Cohen's kappa scores of 0.57 ± 0.02 (Figure 6c,d). After reducing dimensionality with PCA before analysis, F_1 scores were similar at 0.79 ± 0.01, and Cohen's kappa was also similar at 0.58 ± 0.02. Predictive error rates were lower: 0.24 ± 0.01 for diploids and 0.19 ± 0.02 for triploids. The models using the canopy dataset had the strongest performance: F_1 score of 0.96 ± 0.01 and Cohen's kappa score of 0.93 ± 0.01. Predictive error rates were low: 0.02 ± 0.01 for diploids and 0.05 ± 0.01 for triploids (Figure 6e,f).

The spatially explicit cross-validation showed that many of these results were qualitatively similar when restricted to training data from one site and test data from another site. In general, models performed best when trained and tested at the same site (Figure S5), and with lower error rates for triploids than diploids (Figure S6). However, interpretation of these results should be limited, as carrying out the analyses required a reduced sample size for data within each site. Additionally, some sites were homogenously diploid or triploid, yielding fewer cases in which a site-specific model could be trained or tested.

Spatial predictions of the random forest models trained on data from all sites made reasonable predictions of ploidy level. An example prediction is shown for the *Ben-1ha* site in Figure 7. The analysis at this site indicated that the western side of the site is dominated by triploids, with a diploid patch present on the eastern side. The predictions for the *Jolanta-1* site correctly inferred a diploid patch at the northeastern end of this site and homogenous triploidy across the rest of the site. Similarly, correct predictions of spatially homogeneous triploidy also occurred at the *Jolanta-2* and *Jolanta-3* sites. Some small isolated diploid pixels were also predicted in all images, which we interpret as misclassification based on incomplete masking of shadow or ground pixels. Because these predictions are by construction meant to extrapolate beyond ground-truth data, they could not be explicitly validated. However, the high spatial coherence of predictions was biologically plausible and also not methodologically guaranteed, as random forest algorithms yield pixel-level classifications.

4 | DISCUSSION

Our results clarify the spatial correlates of ploidy level in quaking aspen, and suggest several climate change implications. We found that diploids are more likely to be found at higher elevation sites with steeper slopes (Figure 2). This result builds on prior work at continental spatial scales, which indicated that triploids are more common on drought-stressed sites (Mock et al., 2012), though we found more diploids on steeper slopes, which are also presumably more water-stressed. Previous work has also shown that quaking aspen mortality occurs in lower elevation sites, south aspects, and on flatter slopes (Worrall et al., 2008). Paired with other prior observations that mortality is higher for triploids than diploids (Dixon & DeWald, 2015), our results suggest that ploidy level could provide a key mechanistic link between landscape topography and population performance under changing climates, and that pairing remote sensing of ploidy level with maps of drought stress or topography could yield stronger predictions of mortality risk. Based on our ploidy-distance results (Figure 3), such landscape patterning would be predicted to occur at subkilometer spatial scales, yielding patchy mortality – also consistent with existing airborne surveys of sudden aspen decline (Worrall

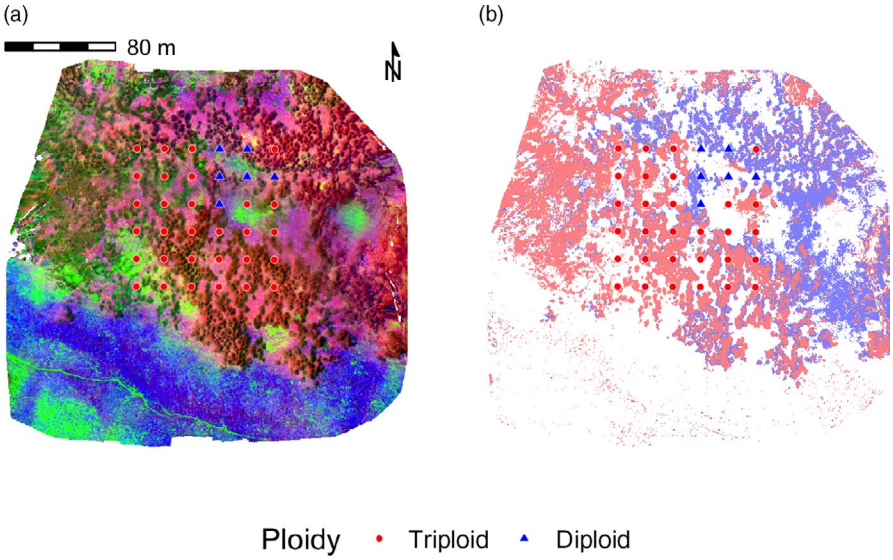

FIGURE 7 Example prediction of ploidy level from multispectral imagery for the *Ben-1ha* site. (a) False colour representation shaded according to scores from a principal component analysis (for visualization only) of all five spectral bands. (b) Spatial predictions of an ensemble of random forest classifiers, with canopy pixels shaded light red (triploids) or light blue (diploids). In both panels, georeferenced genetic samples are shown as red circles (triploids) or blue triangles (diploids) [Colour figure can be viewed at wileyonlinelibrary.com]

et al., 2008). Extant hydraulic and trait-based models (Anderegg et al., 2015; Tai, Mackay, Anderegg, Sperry, & Brooks, 2017) could likely be improved by incorporation of ploidy level information. Similarly, remote sensing analyses of mortality (Huang & Anderegg, 2012) could be complemented by large-scale maps of pre-mortality ploidy levels. In both cases, ploidy level might explain mortality patterns that were previously unexplained or have interactive effects with other variables, for example slope or elevation.

We found that both leaf and canopy spectra can be used to make accurate inferences of ploidy level (Figures 4–6). The spectral variation we observed suggests concomitant variation in pigment concentration (e.g. chlorophyll content in the visible portion of the spectrum), as well as variation in water/dry matter content (in the near-infrared portion of the spectrum). A prior study of trait variation in diploid and triploid aspen leaves matches this perspective (Greer et al., 2017).

Both the canopy data and leaf data were able to classify triploids with low error rates, and with values of F_1 and Cohen's kappa sufficiently high to indicate good predictive ability. The similar results between the hyperspectral and PCA-reduced hyperspectral leaf data indicated that overfitting of features was unlikely to compromise predictive ability, and that a small number of spectral features is sufficient to achieve the classification. This result is promising and suggests that it may become possible to map ploidy level using rapid non-destructive measurements from the ground or from the air. Our approach required only a multispectral sensor, which is far less expensive than the hyperspectral sensors that have been used previously for species discrimination work, or that are analogs for the ground-based spectral data. Further improvements in machine learning methodology are likely to produce more robust and management-ready tools. For example, convolutional neural networks (Brodrick, Davies, & Asner, 2019) may be able to leverage the strong spatial structure in ploidy level (as we identified via our minimum distance analysis) to make cleaner predictions of ploidy level boundaries.

There were some general limitations to using spectral reflectance that will be relevant to future applications. First, classifying diploids was more difficult than triploids. Diploids appeared visually to occupy a smaller and shifted portion of the multidimensional spectral 'space' occupied by triploids. There are several reasons why triploids are likely to have more spectral diversity than diploids. Tripoids may have higher potential for phenotypic plasticity (Levin, 1983; Parisod et al., 2010). The observation that triploids have a wider range of possible spectral properties than diploids is also consistent with prior evidence that polyploidy leads to increased trait variation in quaking aspen (Einspahr et al., 1963; Greer et al., 2017).

Classification based on leaf spectra was much more successful than classification based on bark spectra. Bark-based error rates were often close to 50%, with F_1 values close to zero – that is no better than random guessing. Performance was increased by reducing the dimensionality of the spectral data with PCA before analysis, suggesting that much of the bark spectral variation was not useful for classification. However, even after dimensionality reduction, classification performance was too low to be useful. Our bark dataset included only a small number of diploids (n = 10), so it is possible that better performance could be achieved with a more comprehensive training dataset. However, bark properties may also vary widely across stems (Barnes, 1975; Einspahr et al., 1963), obscuring ploidy-related signals. Future applications for ploidy level classification are likely to rely on airborne canopy data rather than ground-based leaf or bark spectral data. These limitations are therefore unlikely to be important in practice.

High performance with the airborne spectra is surprising, especially because the canopy data cover only a small subset of the wavelengths quantified by the ground-based data, and also potentially include shading effects from the canopy structure and/or mixing with other non-canopy vegetation features. The first possible explanation is that the spectral variation in triploids is captured by a small number of spectral axes, such that the additional spectral coverage

provides only redundant information. This perspective is supported by the similar performance of leaf spectral data before and after dimensionality reduction. A second possible explanation is that there were more biases in the ground-based data than in the canopy-based data. While the canopy data are limited by shadows, ground pixels and other issues, the canopy data effectively include orders of magnitude more leaves than the ground-based data, and also were obtained for the entire dataset within the span of a week, consistent with all leaves being at a similar phenological stage. Moreover, this canopy structure could potentially be diagnostic of ploidy, for example through variation in canopy openness. In contrast, the ground-based spectra were obtained over two growing seasons for a smaller number of leaves, and mixed both sunlit and shaded leaves. Variation in leaf spectra with leaf age and epiphyll cover has been documented in tropical forests (Chavana-Bryant et al., 2017) as well as with canopy position (Gausman, 1984), and with water stress (Hunt & Rock, 1989). While we focused on healthy mature leaves collected during non-droughted conditions, the ground-based spectra may have included more undesired and unavoidable variation, or could have conflated ploidy level-dependent water stress with ploidy level itself. Another explanation, which we view as also likely, is that the low number of diploids available in the canopy data resulted in some model overfitting. In the absence of further genetic data, it is not possible to test this hypothesis. As such we suggest that our statistical models need further test data before they can be robustly applied at landscape scales.

More strongly, our analyses were not able to determine the mechanistic basis of ploidy level classification. Spectral variation can occur for a range of reasons, some of which may be directly caused by ploidy level variation, and others indirectly (i.e. through interactive effects of environment). In particular, assessing how intra- and inter-annual variation in canopy reflectance influences predictions of ploidy level should be a priority. For example, climate-driven effects on canopy water content between years (Aguilar, Zinnert, Polo, & Young, 2012) or phenological effects within years (Blackburn & Milton, 1995) could obscure or enhance ploidy level signals. While it is likely that spectral variation is linked to ecophysiological (especially hydraulic) traits of these canopies, we did not have data to directly make such linkages. It would be useful to directly measure such functioning, or compare the remotely sensed ploidy level data to other remotely sensed data products, for example microwave data for canopy water content (Konings, Rao, & Steele-Dunne, 2019), or solar-induced fluorescence data for photosynthetic capacity and stress (Magney et al., 2019; Meroni et al., 2009). Such data, if also intra- or inter-annually resolved, could help clarify the direct effects of ploidy level.

Obtaining spectral data for longer wavelengths could further improve classification error rates and potentially better discriminate diploids. Our measurement approaches only extended into the near-infrared portions of the electromagnetic spectrum. Features between 1,100 and 2,500 nm, in the short-wave infrared portion of the spectrum, are known to indicate variation in water content and a range of compounds produced in leaves, and have been necessary for accurately discriminating species and traits in other study systems, as well as for assessing leaf and canopy water content variation based on liquid water absorption features around 1,200 and 1,400 nm (Asner, 1998; Asner et al., 2017; Yoder & Pettigrew-Crosby, 1995). Additionally, variation in canopy structure (e.g. wider tree spacing in mature triploid forests) could also lead to variation in texture and shadowing that would be visible in near-infrared multispectral imagery. Collecting data at a wider range of wavelengths should be a future research priority.

We did not detect any variation in bark coloration linked to ploidy. Bark greenness is known to vary extensively among individuals in this species, sometimes along elevational gradients (Barnes, 1969; Cottam, 1954; Covington, 1975; Mitton & Grant, 1996). Aspen bark is photosynthetic and thus colour variation may correlate with chlorophyll content (Foote & Schaedle, 1976; Pearson & Lawrence, 1958). While the hypothesis that bark spectral properties are linked to ploidy is reasonable, this was not supported by our data. Our bark dataset contained only a very limited set of diploids, so the generalizability of results arising from it is probably low. There may also have been unwanted variation in the data from pooling bark samples from different stem aspects (Pearson & Lawrence, 1958), though the magnitude of these effects is probably small. Regardless, the bark is likely more time-intensive and thus less useful than the leaves or canopies for mapping ploidy at large scales, so we did not pursue this question further.

It is also possible that our results would differ outside Colorado. Previous studies have reported large phenotypic variation between southwestern and boreal populations of this species (Barnes, 1969; Mitton & Grant, 1996), and also between populations growing on marginal talus habitat relative to those on more developed soils (Mueggler, 1985). Because we observed spectral variation in natural populations, we do not know whether the spectral phenotype is more determined by genotype, environment or genotype × environment interactions. The environment case is potentially problematic, as spatial environmental variation could lead to apparent but false effects on ploidy level. We suggest that this scenario is unlikely, as the multispectral imagery (as well as field observations) revealed clear homogeneity of spectral properties of forest patches across 20–100 m spatial scales consistent with boundaries between clones, as well as prior studies demonstrating strong control of aspen leaf traits by genotype (Blonder, Violle, & Enquist, 2013; Kanaga, Ryel, Mock, & Pfrender, 2008) and by ploidy level (Flansburg, 2018; Greer et al., 2017). Regardless, if high model predictive accuracy could be achieved over larger spatial extents, then determining underlying mechanisms is not relevant.

There is also a possibility that diploids and triploids were incorrectly inferred from microsatellite analyses due to low genetic diversity. This is unlikely for several reasons. First, allelic richness was high in our data. Second, the majority of clones observed in our dataset were triploid (69.6%), a pattern that would not be expected in the presence of low genetic diversity and presumable under-identification of triploids due to high homozygosity. Furthermore, we are confident that triploid assignments were reliable and not due

to potential duplicated loci or scoring errors, as triploid clones were identified based, on average, 28.3% markers with three alleles (ranging from 1 to 4 markers).

Our results provide a proof-of-concept analysis that suggests a range of potential applications if predictive ability for ploidy level could be further improved. Using airborne remote sensing, it may soon become possible to map the geographical distribution of ploidy levels at fine spatial grain and large spatial extent (or unmix diploid and triploid prevalence at larger spatial grain). Such work would be fundamental for assessments of the effect of polyploidy on mortality and would assist with clone delineation, as well as physiological studies of drought and heat-related mortality. Furthermore, this approach could contribute to current management and conservation efforts in this ecologically important species.

Many other ecologically important species also show intraspecific ploidy level variation that is associated with ecophysiological variation, and that (based on molecular/cytotyping studies) is associated with strong spatial patterning at landscape scales. For example, many widespread grass species vary in their ploidy level (Keeler, 1998). Examples include *Agrostis stolonifera* (bentgrass) (Björkman, 1984; Kik, Linders, & Bijlsma, 1993), *Bouteloua gracilis* (blue grama) (Fults, 1942), *Deschampsia caespitosa* (tussock grass) (Rothera & Davy, 1986) and *Panicum virgatum* (switchgrass) (McMillan & Weiler, 1959; Nielsen, 1944). Among shrubs, intraspecific variation in polyploidy also occurs, for example in *Artemisia tridentata* (big sagebrush) (Pellicer et al., 2010; Richardson, Page, Bajgain, Sanderson, & Udall, 2012), *Atriplex confertifolia* (salt bush) (Sanderson, 2011) and *Larrea tridentata* (creosote bush) (Laport, Minckley, & Ramsey, 2012). Among ecologically dominant trees, similar variation also occurs, for example in *Betula papyrifera* (paper birch) (Grant & Thompson, 1975), *Ginkgo biloba* (Šmarda et al., 2016), *Inga* spp. (Figueiredo et al., 2014), *Olea europaea* (olive) (Besnard et al., 2007), *Polylepis* spp. (Schmidt-Lebuhn et al., 2010) and *Ulmus americana* (American elm) (Whittemore & Olsen, 2011). We imagine such variation is common in many species than those for which data are presently available. Remote sensing methods like the ones proposed here for quaking aspen may also be relevant to delineation of the genetic structure of populations of other species.

ACKNOWLEDGEMENTS

The project was supported by the Norwegian Research Council (KLIMAFORSK 250233) and the United Kingdom Natural Environment Research Council (NE/M019160/1). Research permits were coordinated by the Rocky Mountain Biological Laboratory and the United States Forest Service. Aircraft, instrumentation and operator support was provided by the National Science Foundation Centers for Transformative Environmental Monitoring Programs under grants EAR-1440596 and EAR-1440506. Ken Williams and his team at the Lawrence Berkeley National Laboratory coordinated UAS permits. Chris Sladek assisted with UAS operations. Yadvinder Malhi provided spectrometer access. Cecilia Chavana-Bryant and Jin Wu assisted with spectrometer protocol development. Marco Castaneda assisted with fieldwork.

AUTHORS' CONTRIBUTIONS

Authors were ordered by last name after the first author. R.E.K., J.R. and D.S. contributed ground-based spectra measurements. M.H. and H.P. contributed airborne spectra measurements. K.H. and B.J.G. contributed the genetic analyses. B.J.G., C.J.S. and R.S. contributed resources. B.B. analysed the data and wrote the manuscript. All the authors contributed to fieldwork and to editing the manuscript.

DATA AVAILABILITY STATEMENT

Data are archived at the UK NERC Environmental Information Data Centre at: https://catalogue.ceh.ac.uk/documents/d663aeb9-1e3e-40f7-ab9e-494e7646faeb

ORCID

Benjamin Blonder https://orcid.org/0000-0002-5061-2385
Kenny Helsen https://orcid.org/0000-0001-6856-7095

REFERENCES

Adam, E., Mutanga, O., Odindi, J., & Abdel-Rahman, E. M. (2014). Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: Evaluating the performance of random forest and support vector machines classifiers. *International Journal of Remote Sensing*, *35*, 3440–3458. https://doi.org/10.1080/01431161.2014.903435

Agapow, P.-M., & Burt, A. (2001). Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes*, *1*, 101–102. https://doi.org/10.1046/j.1471-8278.2000.00014.x

Aguilar, C., Zinnert, J. C., Polo, M. J., & Young, D. R. (2012). NDVI as an indicator for changes in water availability to woody vegetation. *Ecological Indicators*, *23*, 290–300. https://doi.org/10.1016/j.ecolind.2012.04.008

Alberto, F. J., Aitken, S. N., Alía, R., González-Martínez, S. C., Hänninen, H., Kremer, A., … Savolainen, O. (2013). Potential for evolutionary responses to climate change – evidence from tree populations. *Global Change Biology*, *19*, 1645–1661. https://doi.org/10.1111/gcb.12181

Anderegg, W. R. L., Anderegg, L. D. L., Sherman, C., & Karp, D. S. (2012). Effects of widespread drought-induced aspen mortality on understory plants. *Conservation Biology*, *26*, 1082–1090. https://doi.org/10.1111/j.1523-1739.2012.01913.x

Anderegg, W. R., Flint, A., Huang, C.-Y., Flint, L., Berry, J. A., Davis, F. W., … Field, C. B. (2015). Tree mortality predicted from drought-induced vascular damage. *Nature Geoscience*, *8*, 367. https://doi.org/10.1038/ngeo2400

Anderegg, W. R., Plavcová, L., Anderegg, L. D., Hacke, U. G., Berry, J. A., & Field, C. B. (2013). Drought's legacy: Multiyear hydraulic deterioration underlies widespread aspen forest die-off and portends increased future risk. *Global Change Biology*, *19*, 1188–1196. https://doi.org/10.1111/gcb.12100

Asner, G. P. (1998). Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment*, *64*, 234–253. https://doi.org/10.1016/S0034-4257(98)00014-5

Asner, G. P., Martin, R. E., Knapp, D. E., Tupayachi, R., Anderson, C. B., Sinca, F., … Llactayo, W. (2017). Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science*, *355*, 385–389. https://doi.org/10.1126/science.aaj1987

Barnes, B. V. (1969). Natural variation and delineation of clones of *Populus tremuloides* and *P. grandidentata* in northern Lower Michigan. *Silvae Genetica*, *18*, 130.

Barnes, B. V. (1975). Phenotypic variation of trembling aspen in western North America. *Forest Science*, *21*, 319–328.

Beaulieu, J. M., Leitch, I. J., Patel, S., Pendharkar, A., & Knight, C. A. (2008). Genome size is a strong predictor of cell size and stomatal density in angiosperms. *New Phytologist*, *179*, 975–986. https://doi.org/10.1111/j.1469-8137.2008.02528.x

Bennett, M., & Leitch, I. (2005). *Plant DNA C-values database*. Kew, UK: Royal Botanic Gardens.

Benson, M. K., & Einspahr, D. W. (1967). Early growth of diploid, triploid and triploid hybrid aspen. *Forest Science*, *13*, 150–155.

Besnard, G., Garcia-Verdugo, C., Rubio De Casas, R., Treier, U. A., Galland, N., & Vargas, P. (2007). Polyploidy in the olive complex (*Olea europaea*): Evidence from flow cytometry and nuclear microsatellite analyses. *Annals of Botany*, *101*, 25–30. https://doi.org/10.1093/aob/mcm275

Bishop, M., Furniss, T. J., Mock, K. E., & Lutz, J. A. (2019). Genetic and spatial structuring of *Populus tremuloides* in a mixed-species forest of southwestern Utah, USA. *Western North American Naturalist*, *79*, 63–71. https://doi.org/10.3398/064.079.0107

Björkman, S. O. (1984). Chromosome studies in Agrostis. II. *Hereditas*, *40*, 254–268.

Blackburn, G., & Milton, E. (1995). Seasonal variations in the spectral reflectance of deciduous tree canopies. *International Journal of Remote Sensing*, *16*, 709–720. https://doi.org/10.1080/01431169508954435

Blonder, B., Violle, C., & Enquist, B. J. (2013). Assessing the causes and scales of the leaf economics spectrum using venation networks in *Populus tremuloides*. *Journal of Ecology*, *101*, 981–989.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32.

Brodrick, P. G., Davies, A. B., & Asner, G. P. (2019). Uncovering ecological patterns with convolutional neural networks. *Trends in Ecology & Evolution*, *34*, 734–745. https://doi.org/10.1016/j.tree.2019.03.006

Callahan, C. M., Rowe, C. A., Ryel, R. J., Shaw, J. D., Madritch, M. D., & Mock, K. E. (2013). Continental-scale assessment of genetic diversity and population structure in quaking aspen (*Populus tremuloides*). *Journal of Biogeography*, *40*, 1780–1791.

Chavana-Bryant, C., Malhi, Y., Wu, J., Asner, G. P., Anastasiou, A., Enquist, B. J., … Gerard, F. F. (2017). Leaf aging of Amazonian canopy trees as revealed by spectral and physiochemical measurements. *New Phytologist*, *214*, 1049–1063. https://doi.org/10.1111/nph.13853

Cole, C. T. (2005). Allelic and population variation of microsatellite loci in aspen (*Populus tremuloides*). *New Phytologist*, *167*, 155–164. https://doi.org/10.1111/j.1469-8137.2005.01423.x

Cottam, W. P. (1954). Prevernal leafing of aspen in Utah mountains. *Journal of the Arnold Arboretum*, *35*, 239–250.

Covington, W. W. (1975). Altitudinal variation of chlorophyll concentration and reflectance of the bark of *Populus tremuloides*. *Ecology*, *56*, 715–720. https://doi.org/10.2307/1935507

Curran, P. J. (1989). Remote sensing of foliar chemistry. *Remote Sensing of Environment*, *30*, 271–278. https://doi.org/10.1016/0034-4257(89)90069-2

Dixon, G. B., & DeWald, L. E. (2015). Microsatellite survey reveals possible link between triploidy and mortality of quaking aspen in Kaibab National Forest, Arizona. *Canadian Journal of Forest Research*, *45*, 1369–1375. https://doi.org/10.1139/cjfr-2014-0566

Einspahr, D. W., Van Buijtenen, J., & Peckham, J. (1963). Natural variation and heritability in triploid aspen. *Silvae Genetica*, *12*, 51–58.

Every, A. D., & Wiens, D. (1971). Triploidy in Utah Aspen. *Madroño*, *21*, 138–147.

Figueiredo, M. F., Bruno, R., Barros e Silva, A. E., Nascimento, S., Oliveira, I. G., & Felix, L. P. (2014). Intraspecific and interspecific polyploidy of Brazilian species of the genus Inga (Leguminosae: Mimosoideae). *Genetics and Molecular Research*, *13*, 3395–3403. https://doi.org/10.4238/2014.April.29.18

Flansburg, A. L. (2018). *Trembling aspen triploid advantage has a physiological basis with geographic caveats*. Madison, WI: MS (Forestry), University of Wisconsin.

Foote, K. C., & Schaedle, M. (1976). Diurnal and seasonal patterns of photosynthesis and respiration by stems of *Populus tremuloides* Michx. *Plant Physiology*, *58*, 651–655.

Fults, J. L. (1942). Somatic chromosome complements in bouteloua. *American Journal of Botany*, *29*, 45–55. https://doi.org/10.1002/j.1537-2197.1942.tb13969.x

Gauch, H. G. (1982). *Multivariate analysis in community ecology*. Cambridge, UK: Cambridge University Press.

Gausman, H. W. (1984). Evaluation of factors causing reflectance differences between sun and shade leaves. *Remote Sensing of Environment*, *15*, 177–181. https://doi.org/10.1016/0034-4257(84)90045-2

Gompert, Z., & Mock, K. E. (2017). Detection of individual ploidy levels with genotyping-by-sequencing (GBS) analysis. *Molecular Ecology Resources*, *17*, 1156–1167. https://doi.org/10.1111/1755-0998.12657

Grant, W., & Thompson, B. (1975). Observations on Canadian birches, *Betula cordifolia*, *B. neoalaskana*, *B. populifolia*, *B. papyrifera*, and *B.× caerulea*. *Canadian Journal of Botany*, *53*, 1478–1490.

Greer, B. T., Still, C., Cullinan, G. L., Brooks, J. R., & Meinzer, F. C. (2017). Polyploidy influences plant–environment interactions in quaking aspen (*Populus tremuloides* Michx.). *Tree Physiology*, *38*, 630–640.

Greer, B. T., Still, C., Howe, G. T., Tague, C., & Roberts, D. A. (2016). Populations of aspen (*Populus tremuloides* Michx.) with different evolutionary histories differ in their climate occupancy. *Ecology and Evolution*, *6*, 3032–3039.

Hogg, E., Brandt, J. P., & Kochtubajda, B. (2002). Growth and dieback of aspen forests in northwestern Alberta, Canada, in relation to climate and insects. *Canadian Journal of Forest Research*, *32*, 823–832. https://doi.org/10.1139/x01-152

Huang, C. Y., & Anderegg, W. R. (2012). Large drought-induced aboveground live biomass losses in southern Rocky Mountain aspen forests. *Global Change Biology*, *18*, 1016–1027. https://doi.org/10.1111/j.1365-2486.2011.02592.x

Hunt, Jr., E. R., & Rock, B. N. (1989). Detection of changes in leaf water content using near-and middle-infrared reflectances. *Remote Sensing of Environment*, *30*, 43–54. https://doi.org/10.1016/0034-4257(89)90046-1

Jones, J., & Markstrom, D. (1973). *Aspen: An American wood*. Madison, WI: United States Department of Agriculture, Forest Service.

Jump, A. S., & Penuelas, J. (2005). Running to stand still: Adaptation and the response of plants to rapid climate change. *Ecology Letters*, *8*, 1010–1020. https://doi.org/10.1111/j.1461-0248.2005.00796.x

Kanaga, M. K., Ryel, R. J., Mock, K. E., & Pfrender, M. E. (2008). Quantitative-genetic variation in morphological and physiological traits within a quaking aspen (*Populus tremuloides*) population. *Canadian Journal of Forest Research*, *38*, 1690–1694.

Keeler, K. H. (1998). Population biology of intraspecific polyploidy in grasses. In G. P. Cheplick (Ed.), *Population biology of grasses* (pp. 183–206). Cambridge, UK: Cambridge University Press.

Kemperman, J. A., & Barnes, B. V. (1976). Clone size in American aspens. *Canadian Journal of Botany*, *54*, 2603–2607. https://doi.org/10.1139/b76-280

Kik, C., Linders, T. E., & Bijlsma, R. (1993). Ploidy level and somatic chromosome number variation in Agrostis stolonifera. *Acta Botanica Neerlandica*, *42*, 73–80.

Konings, A. G., Rao, K., & Steele-Dunne, S. C. (2019). Macro to micro: Microwave remote sensing of plant water content for physiology and ecology. *New Phytologist*, *223*, 1166–1172. https://doi.org/10.1111/nph.15808

Laport, R. G., Minckley, R. L., & Ramsey, J. (2012). Phylogeny and cytogeography of the North American creosote bush (*Larrea tridentata*, Zygophyllaceae). *Systematic Botany*, *37*, 153–164.

Levin, D. A. (1983). Polyploidy and novelty in flowering plants. *The American Naturalist*, *122*, 1–25. https://doi.org/10.1086/284115

Madritch, M. D., Kingdon, C. C., Singh, A., Mock, K. E., Lindroth, R. L., & Townsend, P. A. (2014). Imaging spectroscopy links aspen genotype with below-ground processes at landscape scales. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*, 20130194. https://doi.org/10.1098/rstb.2013.0194

Magney, T. S., Bowling, D. R., Logan, B. A., Grossmann, K., Stutz, J., Blanken, P. D., … Frankenberg, C. (2019). Mechanistic evidence for tracking the seasonality of photosynthesis with solar-induced fluorescence. *Proceedings of the National Academy of Sciences of the United States of America*, *116*, 11640–11645. https://doi.org/10.1073/pnas.1900278116

Martin, S. L., & Husband, B. C. (2009). Influence of phylogeny and ploidy on species ranges of North American angiosperms. *Journal of Ecology*, *97*, 913–922. https://doi.org/10.1111/j.1365-2745.2009.01543.x

McCool, S. F. (2001). *Quaking aspen and the human experience: Dimensions, issues, and challenges*. Fort Collins, CO: USDA Forest Service.

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*, 276–282. https://doi.org/10.11613/BM.2012.031

McMillan, C., & Weiler, J. (1959). Cytogeography of *Panicum virgatum* in central North America. *American Journal of Botany*, *46*, 590–593. https://doi.org/10.1002/j.1537-2197.1959.tb07057.x

Meroni, M., Rossini, M., Guanter, L., Alonso, L., Rascher, U., Colombo, R., & Moreno, J. (2009). Remote sensing of solar-induced chlorophyll fluorescence: Review of methods and applications. *Remote Sensing of Environment*, *113*, 2037–2051. https://doi.org/10.1016/j.rse.2009.05.003

Michaelian, M., Hogg, E. H., Hall, R. J., & Arsenault, E. (2011). Massive mortality of aspen following severe drought along the southern edge of the Canadian boreal forest. *Global Change Biology*, *17*, 2084–2094. https://doi.org/10.1111/j.1365-2486.2010.02357.x

Mitton, J. B., & Grant, M. C. (1996). Genetic Variation and the Natural History of Quaking Aspen: The ways in which aspen reproduces underlie its great geographic range, high levels of genetic variability, and persistence. *BioScience*, *46*, 25–31. https://doi.org/10.2307/1312652

Mock, K. E., Callahan, C. M., Islam-Faridi, M. N., Shaw, J. D., Rai, H. S., Sanderson, S. C., … Wolf, P. G. (2012). Widespread triploidy in western North American aspen (*Populus tremuloides*). *PLoS ONE*, *7*, e48406. https://doi.org/10.1371/journal.pone.0048406

Mock, K. E., Rowe, C., Hooten, M. B., Dewoody, J., & Hipkins, V. (2008). Clonal dynamics in western North American aspen (*Populus tremuloides*). *Molecular Ecology*, *17*, 4827–4844.

Mueggler, W. F. (1985) Vegetation associations. In N. V. DeByle & R. P. Winokur (Eds.), *Aspen: Ecology and management in the Western United States* (pp. 45–55). US Forest Service General Technical Report RM-119, Fort Collins, CO.

Nielsen, E. L. (1944). Analysis of variation in *Panicum virgatum*. *Journal of Agricultural Research*, *69*, 327–353.

Otto, S. P., & Whitton, J. (2000). Polyploid incidence and evolution. *Annual Review of Genetics*, *34*, 401–437. https://doi.org/10.1146/annurev.genet.34.1.401

Parisod, C., Holderegger, R., & Brochmann, C. (2010). Evolutionary consequences of autopolyploidy. *New Phytologist*, *186*, 5–17. https://doi.org/10.1111/j.1469-8137.2009.03142.x

Pearson, L., & Lawrence, D. (1958). Photosynthesis in aspen bark. *American Journal of Botany*, *45*, 383–387. https://doi.org/10.1002/j.1537-2197.1958.tb13141.x

Pellicer, J., Garcia, S., Canela, M., Garnatje, T., Korobkov, A. A., Twibell, J., & Vallès, J. (2010). Genome size dynamics in *Artemisia* L. (Asteraceae): Following the track of polyploidy. *Plant Biology*, *12*, 820–830. https://doi.org/10.1111/j.1438-8677.2009.00268.x

Ramsey, J., & Schemske, D. W. (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annual Review of Ecology and Systematics*, *29*, 467–501. https://doi.org/10.1146/annurev.ecolsys.29.1.467

Richardson, B. A., Page, J. T., Bajgain, P., Sanderson, S. C., & Udall, J. A. (2012). Deep sequencing of amplicons reveals widespread intraspecific hybridization and multiple origins of polyploidy in big sagebrush (*Artemisia tridentata*; Asteraceae). *American Journal of Botany*, *99*, 1962–1975.

Rothera, S. L., & Davy, A. J. (1986). Polyploidy and habitat differentiation in *Deschampsia cespitosa*. *New Phytologist*, *102*, 449–467. https://doi.org/10.1111/j.1469-8137.1986.tb00822.x

Sammut, C., & Webb, G. I. (2010). *Encyclopedia of machine learning*. Boston, MA: Springer.

Sanderson, S. C. (2011). The ploidy races of *Atriplex confertifolia* (Chenopodiaceae). *Western North American Naturalist*, *71*, 67–78.

Schmidt-Lebuhn, A. N., Fuchs, J., Hertel, D., Hirsch, H., Toivonen, J., & Kessler, M. (2010). An Andean radiation: Polyploidy in the tree genus *Polylepis* (Rosaceae, Sanguisorbeae). *Plant Biology*, *12*, 917–926. https://doi.org/10.1111/j.1438-8677.2009.00297.x

Sims, D. A., & Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, *81*, 337–354. https://doi.org/10.1016/S0034-4257(02)00010-X

Šmarda, P., Veselý, P., Šmerda, J., Bureš, P., Knápek, O., & Chytrá, M. (2016). Polyploidy in a 'living fossil' *Ginkgo biloba*. *New Phytologist*, *212*, 11–14.

Smulders, M. J. M., Van Der Schoot, J., Arens, P., & Vosman, B. (2002). Trinucleotide repeat microsatellite markers for black poplar (*Populus nigra* L.). *Molecular Ecology Notes*, *1*, 188–190. https://doi.org/10.1046/j.1471-8278.2001.00071.x

Tai, X., Mackay, D. S., Anderegg, W. R. L., Sperry, J. S., & Brooks, P. D. (2017). Plant hydraulics improves and topography mediates prediction of aspen mortality in southwestern USA. *New Phytologist*, *213*, 113–127. https://doi.org/10.1111/nph.14098

Tuskan, G. A., Gunter, L. E., Yang, Z. K., Yin, T., Sewell, M. M., & DiFazio, S. P. (2004). Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. *Canadian Journal of Forest Research*, *34*, 85–93.

Ustin, S. L., Roberts, D. A., Gamon, J. A., Asner, G. P., & Green, R. O. (2004). Using imaging spectroscopy to study ecosystem processes and properties. *AIBS Bulletin*, *54*, 523–534.

Whittemore, A. T., & Olsen, R. T. (2011). *Ulmus americana* (Ulmaceae) is a polyploid complex. *American Journal of Botany*, *98*, 754–760.

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., & Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 13875. https://doi.org/10.1073/pnas.0811575106

Worrall, J. J., Egeland, L., Eager, T., Mask, R. A., Johnson, E. W., Kemp, P. A., & Shepperd, W. D. (2008). Rapid mortality of *Populus tremuloides* in southwestern Colorado, USA. *Forest Ecology and Management*, *255*, 686–696. https://doi.org/10.1016/j.foreco.2007.09.071

Worrall, J. J., Rehfeldt, G. E., Hamann, A., Hogg, E. H., Marchetti, S. B., Michaelian, M., & Gray, L. K. (2013). Recent declines of *Populus tremuloides* in North America linked to climate. *Forest Ecology and Management*, *299*, 35–51. https://doi.org/10.1016/j.foreco.2012.12.033

Yoder, B. J., & Pettigrew-Crosby, R. E. (1995). Predicting nitrogen and chlorophyll content and concentrations from reflectance spectra (400–2500 nm) at leaf and canopy scales. *Remote Sensing of Environment*, *53*, 199–211. https://doi.org/10.1016/0034-4257(95)00135-N

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Blonder B, Graae BJ, Greer B, et al. Remote sensing of ploidy level in quaking aspen (*Populus tremuloides* Michx.). *J Ecol*. 2020;108:175–188. https://doi.org/10.1111/1365-2745.13296