Leaf gene expression trajectories during the growing season are consistent between sites and years in American beech

U. Uzay Sezen [1*], Jessica E. Shue [1], Samantha J. Worthy[2], Stuart J. Davies[3,4], Sean M. McMahon[1], Nathan G. Swenson[5],

1 Smithsonian Environmental Research Center, 647 Contees Wharf Rd, Edgewater, Maryland 21037 USA.
2 Department of Evolution and Ecology, University of California, Davis, California, 95616 USA.
3 Forest Global Earth Observatory, Smithsonian Tropical Research Institute, Gamboa, Panama.
4 Department of Botany, National Museum of Natural History, Smithsonian Institution, Washington DC 20560 USA.
5 Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556 USA.

***uzay.sezen@gmail.com**

Abstract

Transcriptomics, the quantification of gene expression, provides a versatile tool for ecological monitoring. Here, we show that through genome-guided profiling of transcripts mapping to 33,042 loci, gene expression differences can be discerned among multi-year and seasonal leaf samples collected from American beech trees at two latitudinally separated sites. Despite a bottleneck imposed due to large-scale post-Columbian deforestation, the SNP-based population genetic background analysis has yielded sufficient variation to account for differences between populations and among individuals. Our time series of expression analyses during spring-summer and summer-fall transitions for two consecutive years involved 4197 differentially expressed protein coding genes. A global comparison of 12 seasons has revealed that spring gene expression sets the pace for the rest of the growing season. Using *Populus* orthologs of the differentially expressed genes, we reconstructed a protein-protein interactome as a representation of the leaf physiological states of trees during the seasonal transitions. Gene set enrichment analysis revealed GO terms that highlight molecular functions and biological processes possibly influenced by abiotic forcings such as recovery from drought and response to excess precipitation. Further, based on 324 co-regulated transcripts, we focused on a subset of terms that could be putatively attributed to phenological shifts due to late spring. Our conservative results indicate that extended transcriptome-based monitoring of forests can capture ranges of responses arising from other factors including air quality, chronic disease as well as herbivore outbreaks that require activation and/or downregulation of genes collectively tuning reaction norms needed for the survival of long living trees such as the American beech.

Introduction

Anthropogenic climate change is affecting the phenology of trees, with important implications for the future of forest health [1, 2]. Trees over their long lifespan

experience a multitude of severe conditions, but climate change is shifting essential phenological events and intensifying many of the environmental deviations. These deviations can exceed the boundaries of reaction norms trees have evolved [3–8]. For temperate trees, changes in the arrival of spring, and shifting precipitation regimes can affect the phenological scheduling of vital biological events such as vessel activation [9], the acquisition of soil nutrients [10], flowering [3, 11–13], leaf flushing [14–18], xylogenesis [19], seeding and fruiting [20–22], and senescence induced nutrient re-uptake from leaves [16, 23–26]. Ultimately, the resulting evolutionary dynamics will redefine competition and reshape species ranges [18, 27, 28]. Functional responses to these climatic changes, however, are complex and difficult to demonstrate against the background noise of ecological variation. Because genes fundamentally drive all these mechanisms, gene expression might offer a tool towards identifying patterns to advance our understanding of the mechanisms of organismal response to climate change through phenology.

Leaves are an important component of individual-level performance forming an interface between the phenotype and the environment. Phenotypic variation is an interplay of the genotype and gene expression. Yet, little is known about the relative significance of these two constituents of the phenotype. Fine-scale transcriptomic assays of leaf function of individual trees through space and time are scarce, rarely cover an entire growing season, and/or have not been carried out at more than one site [29, 30]. There is an urgent need for determining the drivers of tree leaf function especially among trees growing in unmanaged natural forest stands. The present study aims to fill this gap through a comprehensive transcriptomic assay of leaf gene expression through space and time.

Transcriptomic analysis captures expressed genes in a tissue at the time of collection and can provide detailed sequence information allowing the quantification, comparison, and accurate interpretation of phenotypes along a phenological sequence. Leaves, in particular, are organs that have evolved adaptations to sense the environment and acclimate in order to maintain homeostasis. Leaves help regulate whole plant performance by balancing the uptake of carbon for photosynthesis and hydraulic flow. Measuring gene expression in leaves, therefore, offers a way to probe into developmental and plastic responses throughout the growing season, linking the influences of climate shifts to the phenotypic response of the trees. Importantly though, no two growing seasons are alike making quantification of gene expression in leaves across space and time vital to further insight into plant function and status under different environmental conditions and phenological stages [3, 31–33].

The timing of seasonal transitions has been shown to differ among years and sites, driven by changes to temperature and precipitation patterns. Unfortunately, these and other climatic variables are expected to be further altered by climate change and become less predictable potentially leading to mismatches and disruptions in pollination services and trophic interactions such as herbivory, predation and parasitizitation [3, 11–13, 34, 35].

American beech (*Fagus grandifolia* Ehrh.) is a common tree, the sole member of the genus naturally found in the northeast United States and a characteristic feature of older growth forests. *Fagus* species are known for their responsiveness to the environment, with an ability to grow slowly when resources are not available, and quickly when conditions reverse such as increased light after gaps form in the canopy or soil humidity is restored [36]. The genus is particularly sensitive to drought as compared to other co-occurring hardwood species [37]. In this study, we investigated patterns of leaf gene expression in two populations of *F. grandifolia*: one in Maryland, USA and the other in central Massachusetts, USA. After collecting two years of expression data, three times each year, we quantified how gene expression differs across sites and seasons. We

leveraged information regarding the up- and down-regulation of genes in each leaf sample collected from the same trees, we asked the following questions: 1) Does leaf gene expression vary with respect to year, site, and season? 2) Can site or annual differences in expression be explained by abiotic factors? 3) Do the identities of the differentially expressed genes suggest mechanisms that control phenology?

Results

Genome alignments and background genetic analysis

Alignment rates of trimmed reads to the draft American beech genome were high. Close to 95 percent of the samples had reads aligned with a rate more than 80% (SI Fig. S3, Table S1). Alignments spanned 33,042 loci corresponding to 27,226 genes in the draft *F. grandifolia* genome (unpublished draft American beech genome sequence). Although the forests of the US east coast are a product of large-scale second-growth regeneration after large-scale deforestation, populations of beech trees at the two ForestGEO FDP sites displayed enough genetic variation to distinguish them based on SNP variants (Fig. S4). A pairwise comparison of SNP profiles of 40 trees generated a kinship pattern where the within population relatedness exceeded that of between population (Fig. S4). Further, a Kendall's rank correlation of association on kinship coefficients had positive values for within population comparisons reflecting local founder effect (Fig. S4, Table S1). In both populations, there are few distantly related individuals that stand out with negative correlations (Fig. S4). On the other hand, correlations were negative using between population comparisons reflecting genetic effect of latitudinal separation by more than 640 km (Fig. S4, Table S1).

Does leaf gene expression vary more with respect to year, site or season?

The first two axes of a Principal Component Analysis (PCA) on the leaf gene expression data showed that the three seasons were transcriptomically distinct (Fig. 1). Site differences were not very pronounced and mostly followed a parallel expression trajectory, except in the spring 2018, which shows the most separation between the research sites (Fig. 1). The yearly dynamics of each season showed a recognizably linear PC1 distribution consistent across both populations. The first two components together represented 73.5 % of the total variation. The spring samples were separated along both PC axes to a greater extent and represent1d the largest magnitude of change during seasonal transitions in both years (Fig. 2). Summer and fall samples together formed tighter clusters, but yearly dynamics were still distinguishable (Fig. 1). Once summer was reached, the expression pattern of genes remained relatively unchanged into the fall. Specifically, summer and fall samples largely overlapped especially in 2017 (Fig. 1). Fall 2017 samples from SERC overlapped with their summer counterparts representing a sustained gene expression pattern along summer-fall transition (Fig. 1 unshaded orange and blue circles). Overall, the 2018 growing season casted a wider gene expression space while 2017 represented a more contracted parallel transcriptional activity in both sites (Fig. 1).

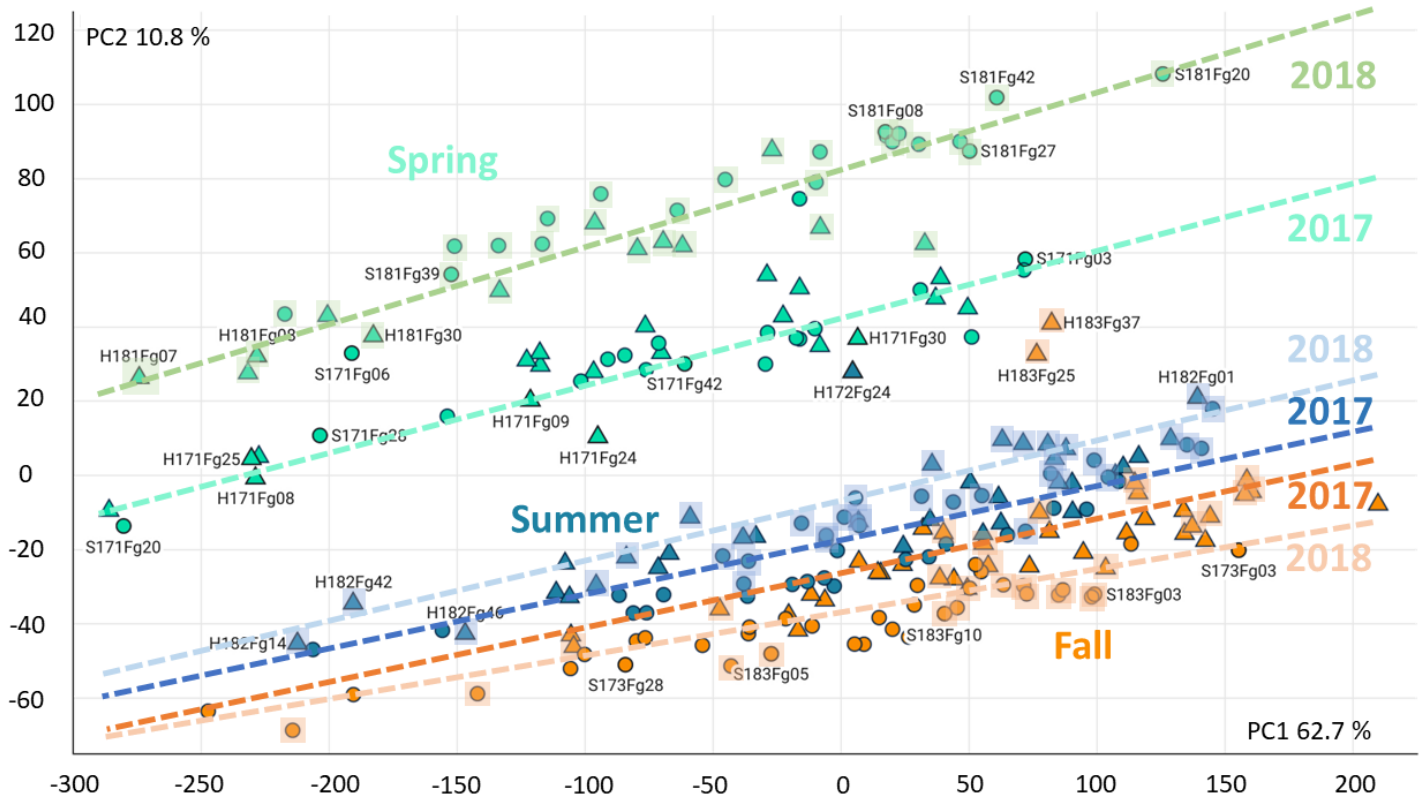


Figure 1. A PCA of differentially expressed genes in 222 *F. grandifolia* samples collected from 40 trees during Spring, Summer and Fall of 2017 and 2018 from HARV (triangles) and SERC (circles). The PCA used singular value decomposition with imputation and includes the top 11.500 most variable loci that were ln(x) transformed and Pareto scaled. Shaded regions represent 2018 samples. The PC1 and PC2 represent 73.5 % of the total variation. Dashed lines were inserted manually to show the approximate trends across years.

Mapping differential gene expression on an interactome

The total number of differentially expressed transcripts corresponded to 4197 American beech protein models (SI Dataset S1). A search for these proteins against the black cottonwood (*Populus trichocarpa*), NCBI taxon identifier 3694) matched 2813 orthologs. Of the *Populus* orthologs, a subset of 2421 had a hit in the STRING database generating an interactome network with 2185 nodes and 3231 edges. The largest patch consisted of 1280 nodes and 3185 edges (Fig. 2). The network contained 1329 singleton nodes (Fig. 2, singleton nodes not shown for brevity). Consistent with the PCA results, overall the highest magnitude of change took place during the spring-summer transitions during both 2017 and 2018 (Fig. 2).

Gene expression during spring-summer transitions

The trees at the two sites consistently co-upregulated 31 and co-downregulated 127 transcripts (*Populus* orthologs) in both years during spring-summer transitions (Fig. 2, Venn diagram left). In HARV, 13 transcripts were upregulated exclusively in both years (Fig. 2, green fill with red border). The number of transcripts that were upregulated

solely in SERC for two consecutive years was 74 (Fig 2., yellow fill with red border). In 2017, SERC had 56 and HARV had 41 transcripts upregulated in a site-specific manner. In the following year, the site-specific expression levels were roughly two-fold higher in HARV with 98 transcripts and three-fold in SERC with 176 transcripts (Fig. 2). In parallel, site-specific downregulation in 2018 was 16-fold higher in HARV with 272 transcripts (Fig. 2). SERC had two-fold higher downregulated site-specific expression with 184 transcripts (Fig. 2). Therefore, spring-summer transition in 2018 demonstrated the largest differential gene expression activity in parallel at both sites. There were 127 transcripts co-downregulated in both sites and years (Fig. 2). HARV had 57 transcripts repeatedly downregulated during spring-summer transitions in both years (Fig. 2). SERC had 53 transcripts consistently downregulated in two successive years (Fig. 2).

Gene expression during summer-fall transitions

Gene expression patterns during summer-fall transitions at the two sites did not differ as much as observed in Spring-Summer and therefore it was possible to represent differential expression from both sites in a single panel at a given year (Fig. 2, Venn diagram right). There was a single *Populus* ortholog each for repeatedly upregulated (Acid Phosphatase 1) and downregulated (Nuclear Fusion Defective 4) transcripts found in both years at two sites (Fig. 2). Strikingly, there were no upregulated transcripts in either of the two sites in summer-fall 2017. In 2017, the two sites experienced a relatively homogeneous summer-fall transition with a moderate amount of 10 and 11 downregulated transcripts with only three co-downregulated transcripts (Fig. 2). SERC appears to have a more dynamic summer-fall transition beginning with 28 *Populus* orthologs upregulated as opposed to none in HARV (Fig. 2). There were 94 *Populus* orthologs specifically upregulated in SERC while in HARV there were only four (Fig. 2). No co-expressed orthologs were shared between the two sites in 2018 (Fig. 2). Downregulated transcripts particular to SERC in 2018 corresponded to 98 *Populus* orthologs almost 10-fold higher than that of HARV with 11 (Fig. 2). Co-downregulated *Populus* orthologs in both sites amounted to six (Fig. 2).

Enriched GO terms during spring-summer transitions

Up and downregulated transcripts reflected both latitudinally parallel and divergent patterns in enriched Gene Ontology (GO) terms including biological process (BP), molecular function (MF), cellular compartment (CC) and KEGG (Kyoto Encyclopedia of Genes and Genomes) suggesting a strong environmental influence on response of both forest sites (Fig. S8). In 2017, upregulated transcripts in HARV were grouped under 8 terms belonging to sugar and carbohydrate transport activity (Fig. S8). During the wet growth season of 2018, trees in both sites showed a parallel downregulation of transcripts encompassing 23 cell wall and membrane biosynthetic terms including very long chain fatty acid biosynthesis that control suberin and epicuticular wax quality as a response to abiotic stress (Fig. S8). A list of top 20 differentially expressed transcripts with their draft genome accessions and annotations are provided for all seasonal transitions (Table S1).

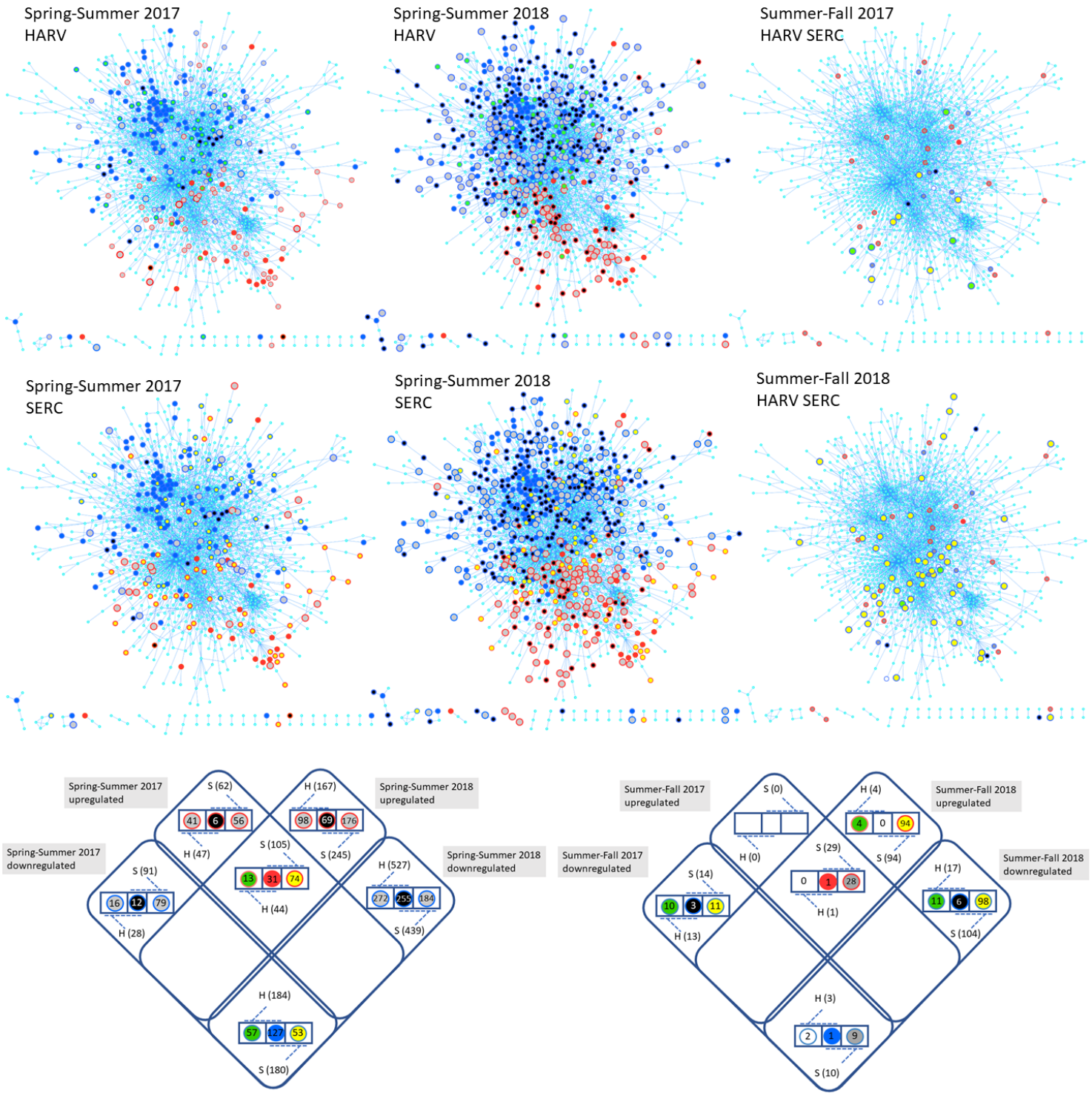


Figure 2. A breakdown of differential and co-expression patterns during spring-summer, summer-fall transitions at two sites during two consecutive years mapped on the *P. trichocarpa* protein-protein interactome. Interactome nodes with their corresponding color combinations are shown inside the respective Venn diagrams where sites are coded as H (HARV) and S (SERC). Due to similarity in differentially expressed genes, summer-fall transitions in two sites are represented in a single interactome panel per year. For brevity, singleton nodes are not shown.

Table 1. STRING output of GO terms from 324 co-regulated transcripts enriched for delayed spring leaf expansion phenology in *F. grandifolia*. Permalink to the analysis: https://version-11-5.string-db.org/cgi/network?networkId=bWUhWsqb5LZt

Biological Process				
GO-term	**description**	**count in network**	**strength**	**false discovery rate**
GO:0098754	Detoxification	10/238	0.73	0.037
GO:0071669	Plant-type cell wall organization or biogenesis	11/297	0.67	0.037
GO:0071555	Cell wall organization	24/716	0.63	1.91E-05
GO:0071554	Cell wall organization or biogenesis	28/894	0.6	8.34E-06
GO:0005976	Polysaccharide metabolic process	17/658	0.52	0.037
GO:0005975	Carbohydrate metabolic process	30/1581	0.38	0.0189
Molecular Function				
GO-term	**description**	**count in network**	**strength**	**false discovery rate**
GO:0004601	Peroxidase activity	8/161	0.8	0.0313
GO:0016747	Transferase activity	16/581	0.55	0.0285
GO:0046906	Tetrapyrrole binding	17/646	0.53	0.0285
GO:0020037	Heme binding	16/607	0.53	0.0285
GO:0003824	Catalytic activity	159/14494	0.15	0.00044
Cellular Component				
GO-term	**description**	**count in network**	**strength**	**false discovery rate**
GO:0009505	Plant-type cell wall	9/195	0.77	0.004
GO:0048046	Apoplast	15/508	0.58	0.0025
GO:0005576	Extracellular region	60/2170	0.55	3.70E-14
GO:0005618	Cell wall	19/766	0.5	0.0025
GO:0005794	Golgi apparatus	26/1479	0.35	0.0149
GO:0031224	Intrinsic component of membrane	96/7034	0.24	9.49E-06
GO:0016021	Integral component of membrane	90/6767	0.23	4.27E-05
GO:0071944	Cell periphery	80/6039	0.23	0.00026
GO:0005886	Plasma membrane	65/5314	0.19	0.0179
GO:0016020	Membrane	130/11031	0.18	3.22E-05
GO:0110165	Cellular anatomical entity	272/29493	0.07	3.14E-05
KEGG pathway				
pathway	**description**	**count in network**	**strength**	**false discovery rate**
pop00062	Fatty acid elongation	5/40	1.2	0.0017
pop00520	Amino sugar and nucleotide sugar metabolism	7/154	0.76	0.0094
pop00940	Phenylpropanoid biosynthesis	8/206	0.7	0.0094
pop01110	Biosynthesis of secondary metabolites	26/1446	0.36	0.0044
pop01100	Metabolic pathways	42/2605	0.31	0.0013

Genes possibly involved in phenology of late-stage leaf expansion affected by delayed spring

A total of 324 transcripts were simultaneously co-regulated (69 upregulated, 255 downregulated) at both sites as a potential response to delayed spring leaf expansion during the spring to summer transition in 2018 (Fig. 2). These transcripts may potentially be controlling the phenology of delayed leaf expansion enriched with 6 GO terms pertaining to biological processes, 5 terms of molecular function and 11 terms involving cellular components. Additionally, these transcripts were enriched for 5

KEGG metabolic pathways (Table 1). Enriched terms related to the membrane (GO:0031224, GO:0016021, GO:0005886, GO:0016020) and metabolic pathway of fatty acid elongation (pop00062) are associated with phenological transitions (Table 1). A full list of STRING identities for these transcripts and their annotations are included in the supplementary file (Table S1).

Discussion

There have been a few transcriptomic time course studies produced to understand multi-year, multi-season dynamics in plants [19, 29, 30, 38, 39]). These works have largely focused on a single site and/or a single growing season. Here, we have produced a detailed study of leaf gene expression in *Fagus grandifolia* in two distinct populations in the eastern United States through the growing season in multiple years. Despite the climatic differences in years and sites, that we describe below, and the potential for a high degree of variation in leaf gene expression profiles, it is notable that several emergent and consistent patterns were found. First, gene expression profiles significantly change between the spring and summer components of the growing season. Second, unlike the spring-summer transition, gene expression profiles are generally consistent from the summer leading into the fall. Third, and perhaps most surprisingly, these results were consistent across years and across genetically distinct populations separated by hundreds of miles and several degrees of latitude. These results indicate that, while leaf gene expression profiles are dynamic, they are largely predictable across seasons, sites and years. In the following, we discuss the results in detail beginning with information regarding the sampling conditions across seasons and years.

Gene expression is, typically, considered to be highly labile and the quantification of gene expression in the wild may be expected to lead to incredibly noisy signals due to the variable ecological contexts experienced by a tree in space and time. For example, the eastern United States experienced a substantial drought period beginning in 2016, the year prior to our sampling, lasting into the spring of 2017 (SI Fig. S5). Thus, during the 2017 growing season, trees at the HARV site were simultaneously recovering from drought while also dealing with the ongoing impacts of beech bark disease (SI Fig. S5, SI Fig. S7).

In 2018, spring arrived late compared to the previous year (Fig. 3), and many states across the east coast had record breaking rainfall (SI Fig. S6, SI Fig. S5). Specifically, May 2018 to May 2019, the continental United States experienced the most abundant rainfall recorded in 124 years (SI Fig. S6, the US Drought Monitor jointly produced by the National Drought Mitigation Center (USDM) at the University of Nebraska-Lincoln, the United States Department of Agriculture, and the National Oceanic and Atmospheric Administration.)

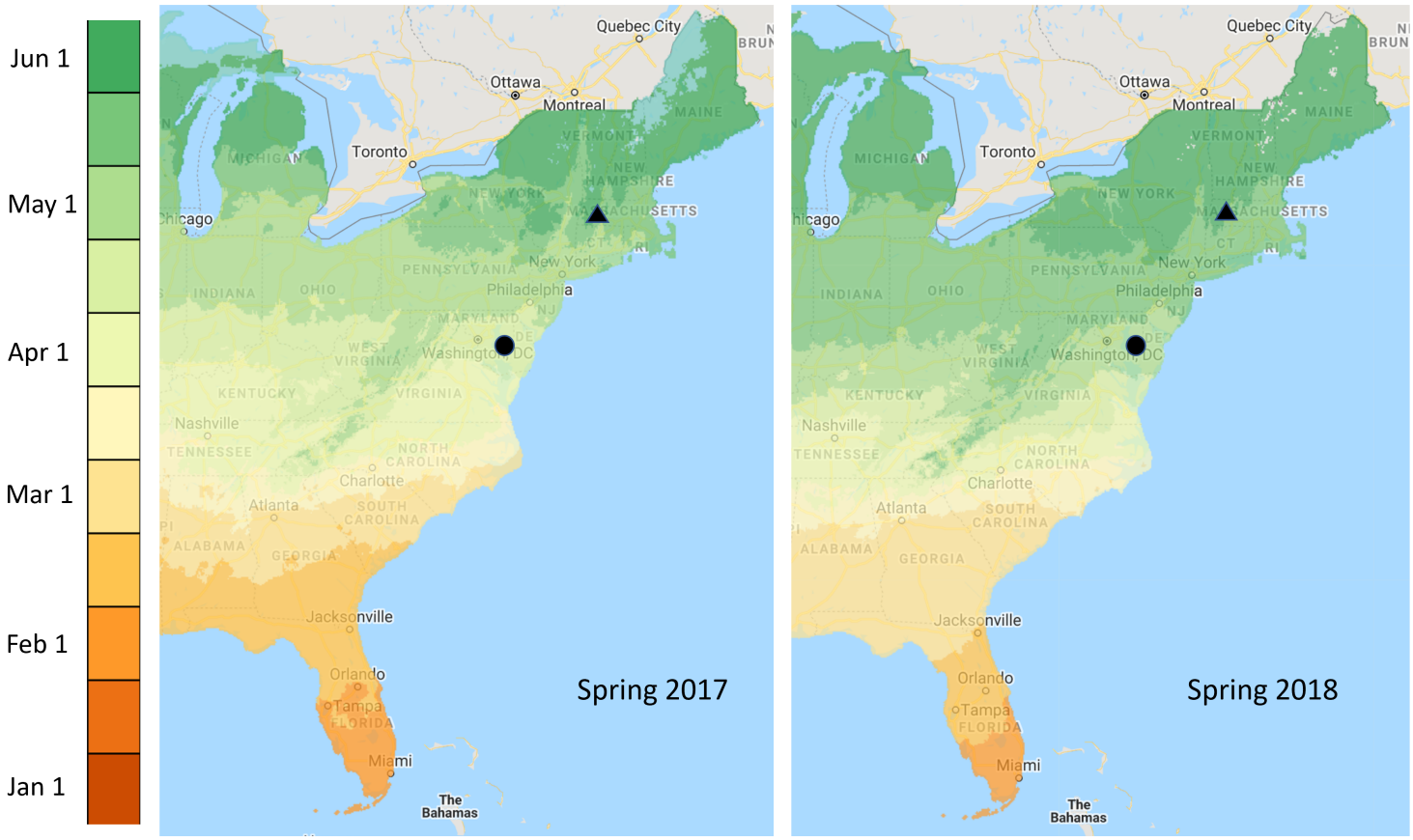


Figure 3. Phenology of Spring arrival comparison based on first bloom between 2017 and 2018. In 2018, the spring arrived late along the entire Northeastern US. The visualization is an annual representation of the days of year that the requirements for the first bloom Spring Index were met, averaged for Red Rothomagensis lilac, Arnold Red honeysuckle and Zabelii honeysuckle; available since 1981, calculated using PRISM Tmin and Tmax data at a particular location. Data from US National Phenology Network (USA-NPN) and NEON.

Taking these yearly and site variations into consideration, one would expect that the beech trees in this study would have highly variable gene expression profiles within and between sites and years. As our results show, the combined effect of abiotic factors such as late spring and precipitation can generate a discernable molecular phenotype between the same seasons of consecutive years. The pattern is visible among Spring 2017 - Spring 2018 samples of both sites separated along the PC2 (Fig. 1). Nearly each spring has a distinct signature with a latitudinally parallel trajectory along both sites (Fig. 1). That is, in each season within a given year, our trees had a transcriptomic pattern following a year-specific trajectory (Fig. 1).

Spring samples casted a wider ordination footprint compared to summer and fall (Fig. 1). This means the largest differential gene expression change is observed during the spring-summer transitions. The vigor observed during the spring-summer transition is evident in the mapping of the differentially expressed transcripts on the *Populus* interactome through their orthologs (Fig. 2). The arrival of spring is a continually moving target and increasingly becoming so with climate change [1, 2, 11–13, 40]. Therefore, understanding spring-summer growth dynamics is important since global warming is expected to induce phenological shifts in flowering and leaf-out timings of most trees and affect mortality [3, 4, 7]. In parallel, we particularly paid attention to a set of 324 co-regulated transcripts in both study sites that could possibly be involved in controlling the phenology of delayed leaf expansion as a consequence of late spring (Fig.

2, Table 1). In the literature, most genetic studies on phenology of trees are focused on bud break [41–44] and phenology of leaf expansion has been limited to early stages [45, 46]. The fact that our putative leaf phenology controlling gene set does not overlap with these studies is probably because we compare fully expanded spring leaves (Fig. 2, Table 1, Table S1).

Summer and fall on the other hand, occur along a more compressed gene expression space, but show a considerable degree of separation along the PC2 (Fig. 1). Slight differences among the transcriptional phenotypes between the two sites could be reflecting chronically diseased states in HARV and/or latitudinal difference (Fig. 1). The unusually wet 2018 growing season appears to assign all trees to an ordination pattern that covers a wider area in the transcriptional phenotypic space (Fig. 1, shaded points). The upper and lower bounds of the ordination are set by 2018 spring and fall samples (Fig. 1).

The growing season of 2017 in HARV started with a large water deficit but followed a steady recovery until April 2018 (Fig. 3, Fig. S5). During this period trees expressed transcripts with GO terms enriched in carbohydrate: proton symporter activity (GO:005351) especially of monosaccharides (GO:0015145) including hexose transmembrane transporter activity (GO:0015145) which include Sugar Transport Protein (STP) subfamily (Fig. S8). Monosaccharide transporters represent an ancestral gene family in land plants and are adapted to overtake diverse physiological roles in vascular plants including carbohydrate partitioning [47, 48]. When overexpressed in the vasculature and mesophyll cells, STPs can increase growth and Nitrogen use in Arabidopsis [49]. In leaves, monosaccharide transporters are highly redundant but essential in non-structural carbon balance [50]. The whole-plant long distance carbon transport is mediated by sucrose. In early spring sucrose travels from roots to the buds for leaf out and is hydrolyzed to monosaccharides via invertases. Once in monosaccharide form, proton-coupled symporters commit sugar uptake [48, 50, 51]. STP-mediated transport of non-structural carbohydrates are key in survival since trees must maintain a certain reserve of mobilizable stored carbon for survival against osmotic stress and pathogen attack [52, 53]. Both stressors were present in HARV during the growing season of 2017. Plant sugar transporters are aggressive enzymes and have been shown to carry three orders of magnitude higher-affinity compared to their pathogenic counterparts and form the backbone of a defense strategy based on apoplastic sugar depletion [53]. Retrieval of carbohydrates leaked into the apoplast traveling to the phloem is a homeostatic activity during regular growth. By upregulating the monosaccharide symporter activity, trees in HARV could be showing a joint response against water deficit and pathogen attack which contrasts with SERC where these stressors didn't exist in 2017 (Fig. S5, Fig. S7, Fig. S8).

Exceedingly wet growing season across the entire East coast of the continental US in 2018 led to simultaneous downregulation of 255 transcripts in SERC and HARV during spring-summer transition (Fig. 2, Fig. S8). These genes span 23 GO terms possibly representing decreased investment on fatty acid biosynthesis including very long chain fatty acids (VLCFAs) which are precursors for cellular membranes, wax and suberin (Fig. S8). This is intriguing since epidermal cell morphology, especially the extension of stomatal guard cell walls in the form of outer cuticular ledge, affects resistance to drought and is dependent on VLCFA deposition [54, 55]. Relaxation from expensive anabolic processes in the leaf tissue allows diversion of resources towards growth in woody tissues or greater allocation to seed storage lipids. VLCFA elongation takes place in the ER membrane through a sequential chain of four enzymes. Our analysis showed that the first enzyme in the chain, the 3-ketoacyl-CoA-synthase activity (GO:0102756) and a set of others including 3-oxo-arachidoyl-CoA synthase activity (GO:0102336) as well as 3-oxo-cerotoyl-CoA synthase activity (GO:0102337), and 3-oxo-lignoceronyl-CoA

synthase activity (GO:0102338) are downregulated (Fig. S8). Chain elongation is an energy demanding anabolic process requiring 7 ATP for building the 16C palmitic acid. The reverse chemistry is also rewarding, the complete beta-oxidation of palmitic acid through Krebs Cycle, electron transport chain and glycolysis can generate 110 ATPs [56]. In less stressful times, such as in 2018, the surplus VLCFA pool could potentially be oxidized in peroxisomes to generate energy to divert growth towards non-leaf tissues. Additionally, VLCFAs are necessary for auxin transport by polar auxin influx carrier PIN1 affecting tissue patterning [57].

Finally, due to recent demographic history, even for unrelated outbred individuals, the knowledge of their population level background relatedness is important for interpreting the observed variation in gene expression [58]. New England has experienced extensive deforestation after the European settlers arrived and lost a fraction of their genetic diversity [59], [60]. Present day forests in the region are largely of second-growth that have expanded from highly bottlenecked founders or few fragmented patches that have maintained old-growth character [61]. The fraction of preserved genetic diversity appears to differ among species with different modes of pollen and seed dispersal [62–66]. The nature of the retained diversity ie. whether it is based on gene flow, emergence of novel variants through mutation and recombination as well as plasticity contribute to this picture [3, 64, 67–77]. Our kinship analysis has shown that SERC and HARV populations were genetically distinct and at the same time harbored some degree of background relatedness characteristic of second-growth forests (Fig. S4). American beech seeds are animal-dispersed. Many vertebrate consumers of its seeds include squirrels, raccoons, porcupines, foxes, black bears, bluejays, pheasants and turkeys [78–80]. Therefore, beech trees are expected to benefit from this adaptation in terms of long-distance gene flow. In fact, sporadic presence of unrelated individuals observed within the two populations could be an indicator of genetic recovery by successful long-distance seed dispersal and seedling establishment (Fig. S4).

Conclusions

Our investigation shows that based on differential and co-expression patterns, it is possible to reconstruct the leaf physiological states of individual trees to understand forest dynamics with respect to site, year and seasonal differences. This study provides a detailed overview of transcriptomic signatures in the leaf tissue of American beech at a single-gene resolution. We show that it is possible to identify sets of differentially expressed transcripts along a time series, map them on a protein-protein interactome and extract the underpinning biology through enriched GO terms. Our genome-guided transcriptome comparison took place in a setting and at a time interval where overlapping biotic and abiotic factors perturbed the gene expression to interpretable levels and yielded explanations on the collective responses of trees to situations including recovery from drought, late spring induced phenological shift, excessive precipitation and chronic disease states. Individually, many genes may show clearly defined functions seemingly disconnected to others. However collectively, the interactions of their protein products may orchestrate a continuum of shifting phenotypes adapted to carry out life processes in optimum.

As further sequencing data from diverse tree species with more extensive time series accumulate, it will be possible to capture and understand tree responses against rare and extreme incidences together with underlying molecular dynamics shaping the phenotypic space. Here we wanted to set an example for monitoring projects such as the ForestGEO and NEON which can incorporate transcriptomics as an informative cutting-edge tool to generate temporally and spatially resolved genetic data as a legible permanent record of tree responses in forests. We hope our study serves as a template for potential scaled-up transcriptomic studies in natural forest settings.

Materials and Methods

Sites and species studied

The study was carried out at two ForestGEO Forest Dynamics Plots (FDPs): the 30 ha plot at Harvard Forest (HARV), Petersham, MA and the 16 ha plot at the Smithsonian Environmental Research Center (SERC), Edgewater, Maryland. HARV represents a higher latitude inland forest with relatively pristine air quality where almost all *F. grandifolia* trees are under attack from beech bark disease involving scale insects (*Cryptococcus fagisuga* and *Xylococculus betulae*) and fungal pathogen Neonectria (*N. ditissima* and *N. faginata*). HARV also experiences episodic spongy moth (*Lymantria dispar*) outbreaks. SERC represents a lower latitude, coastal forest chronically affected by daily plumes of metropolitan Washington DC area emissions. At SERC, during the time of sampling, all trees were free of beech bark disease, and no outbreak of spongy moth has been observed. The two FDPs are co-located with NEON (National Ecological Observation Network) sites, each harboring a tower equipped with instrumentation for ecological monitoring. At each FDP, 20 trees were selected according to an even size class distribution. All trees had a dendroband installed and growth measurements were taken at weekly intervals.

Sample collection and RNA library construction

Using a modified crossbow with a fishing line and reel attachment, leaves have been felled from the canopy through spring, summer and fall between 2017 and 2018 (SI Fig. S1). For subcanopy trees we used a 12m pruner pole. Sample collections were done in the morning to avoid expression changes throughout the day. For each sample, 10 leaf discs were punched into 2ml extraction vials containing 3 steel balls of 3mm in diameter, and flash frozen in liquid nitrogen in the field (SI Fig. S2). Frozen samples were then transferred to a -80 freezer at the end of the day. RNA extractions were carried out using E.Z.N.A Plant RNA Kit by (Omega Bio-Tek, Norcross, GA, USA). Quality metrics for extracted RNA were determined using the Agilent Bioanalyzer 2100 instrument (Agilent Technologies, Santa Clara, CA, USA) and samples containing genomic DNA contaminants were treated with DNAse I (ThermoFisher Scientific, Waltham, MA, USA). Illumina TruSeq libraries were prepared for each sample RNA (Illumina Inc., San Diego, CA, USA). Libraries were indexed and pooled to be sequenced on the Illumina NovaSeq 6000 platform as 150 nucleotide long paired end reads targeting 14 million total reads per sample (Illumina Inc., San Diego, CA, USA). Sequences are uploaded to the NCBI Sequence Read Archive (SRA) with the bioproject number PRJNA630305. Sample naming convention includes Site (H for HARV, S for SERC), Year (17 for 2017, 18 for 2018), Season (1 for Spring, 2 for Summer, 3 for Fall) Species (Fg for *F. grandifolia*), and tree ID number (between 01-50). The largest tree with dbh of 86.6 cm at the beginning of the 2017 growing season located in SERC (Sample ID SFg10 and ForestGEO FDP treetag ID 43420) served as the high molecular weight DNA source for long-read sequencing and was subsequently used to assemble the UCONN version of the American beech draft genome.

Bioinformatics

Demultiplexed reads were processed with TRIMMOMATIC for adapter removal and quality filtering of the fastq formatted raw data [81]. Trimmed reads were aligned to the *F. grandifolia* draft genome sequence using HISAT2 splice aware aligner [82]. Duplicate identification and read-ID groupings of paired-end reads were carried out with SAMBLASTER, and alignments were sorted by SAMTOOLS [83, 84]. Variants were

called through FREEBAYES to generate individual VCF variant call format files for each season and consolidated into a total of 40 individuals using BCFTOOLS [85, 86]. Population genetic analysis for background genetic relatedness between FDP sites and among individual trees was done by the TASSEL5 suit through calculation of pairwise kinship matrix using centered Identity by State (IBS) genetic distances [87, 88]. The probability of IBS is based on the definition where randomly drawn alleles from two individuals at the same locus are identical. The kinship matrix was used as input to calculate Kendall's correlation for genetic associations within and between the two FDP populations using MORPHEUS matrix visualization and analysis software (Broad Institute, Boston, MA, USA). Alignments from HISAT2 were processed by HTSEQ to generate abundance for transcripts [89]. HTSEQ counts were filtered to include top 11.5k loci with most variance. A PCA was generated from variance filtered loci using CLUSTVIS R package through ln(x) transformation with the default method of singular value decomposition and Pareto scaling [90]. We used the TUXEDO pipeline for differential gene expression analysis [91]. Based on alignments, a merged genome-guided transcriptome assembly and transcript quantification were done by STRINGTIE [92]. Per gene count data for mRNA abundance (FPKM) were calculated using the differential gene expression analysis R package BALLGOWN [93]. Differentially expressed beech transcripts passing the minimum fold change (log2fc >2.32, pval <0.1) were retained. For each year and site, differential expression analyses were carried out in the following comparative frame: Spring (condition 1) vs. Summer (condition 2) and Summer (condition 1) vs Fall (condition 2). Protein sequences of the differentially expressed transcripts were BLASTed against indexed black cottonwood (*Populus trichocarpa*) proteome with the NCBI taxon ID 3694 to serve as orthology template for downstream interactome time course analysis. Top hits were selected through VSEARCH (-ublast -evalue 1e-9 -query_cov 0.9) [94]. BLAST results were used as input into the STRING database for network construction using *P. trichocarpa* interactome [95, 96]. Co-downregulated and co-expressed protein-protein interactome networks were imported into CYTOSCAPE and merged into a single network [97] (Table S1). Gene Ontology (GO) terms were compiled by using g:PROFILER by providing gene lists corresponding to *P. trichocarpa* orthologs [98]. Results generated from g:Profiler are accessible through permalinks in Table S1.

Supporting Information

Supporting information for Sezen et al. 2022, Leaf gene expression trajectories during the growing season are consistent between sites and years in American beech. Dryad accession permalink: https://doi.org/10.5061/dryad.fxpnvx0vp.

Table S1

Spreadsheet. A multi-tab format spreadsheet containing sample list, genome alignments, transcript raw counts, ordination results, list of top 20 differentially expressed genes, g:Profiler output permalinks, kinship matrix, Kendall's correlation values and an annotated list of 324 leaf phenology related gene orthologs.

SI Dataset S1

Sequences. A fasta file of differentially expressed 4197 beech protein sequences.

SI Dataset S2

Variant calling. A consolidated sorted VCF file harboring 40 trees.

Fig. S1

Photographic description of the crossbow sampling set up. Leaves from crowns of the tall trees were sampled using a Barnett International Youth100 draw weight 100 lbs. thrust crossbow with custom-made fishing reel attachment and variable weights on bolt tips.

Fig. S2

Photographic description of leaf sample collection in the field. Leaf discs have been punched out prior to freezing in liquid nitrogen.

Fig. S3

Percentage of samples with alignment rates of trimmed reads to draft American beech genome. Close to 95 percent of the samples had reads aligned with a rate more than 80%.

Fig. S4

A Kinship Analysis of SNP variation and the strength of correlated association among 40 F. grandifolia trees from HARV and SERC. (Top) The pairwise relatedness matrix is calculated based on the centered probability of Identity by State (IBS). Kinship values show a marked decrease in comparisons between the two populations. Within population comparisons show higher kinship reflecting the local founder effect. The diagonal at the bottom shows self-comparisons where kinship coefficient reaches maximum. (Bottom) Kendall's rank correlation of association applied on kinship coefficients for within and between populations. Self-associations in the diagonal assume the value of 1.0. Within population correlation values are dominated by values above zero indicating higher genetic relatedness. Between population correlations drop to negative values indicating higher genetic differentiation. The lowest value (-0.62) was measured between HFg45 and SFg29 indicating the most distantly related trees in our sample. In HARV, three individuals HFg24, HFg25, HFg37 stand out as most distantly related within the population. Similarly, in SERC, SFg10 is the most distantly related tree within the population. The HARV tree HFg37 shows a high genetic relatedness to most trees in SERC highest value (0.74) corresponding to pairwise comparison with SFg07.

Fig. S5

Drought maps from 2016-2018 in the Eastern US compiled from the US Drought Monitor initiative by the National Drought Mitigation Center (NDMC) at the University of Nebraska-Lincoln, the National Oceanic and Atmospheric Administration (NOAA), and the U.S. Department of Agriculture (USDA) and the NASA Gravity Recovery and Climate Experiment (GRACE) satellite that detects changes in the Earth's gravity field caused by water on and beneath the land surface. In 2016, New England states experienced a severe regional drought. 2017 was a recovery year from drought. The visualization switches from NDMC to GRACE in April 2017 (shown by an inset frame). In 2018 the entire US east coast experienced record breaking amounts of precipitation. The map classifications are: abnormally dry (D0), showing areas that

may be going into or are coming out of drought, and four levels of drought: moderate (D1), severe (D2), extreme (D3) and exceptional (D4). The GRACE-Based Root Zone Soil Moisture Drought Indicator colors are based on a wetness percentile relative to the 1948-2012 period as baseline.

Fig. S6

Abundant rain across the Eastern United States in 2018 resulted in the wettest year on record. The shallow shallow groundwater on December 17, 2018, ranks among all the Decembers from 1948 to 2012. Blue areas have more abundant groundwater while orange and red areas have less compared to the mean calculated between 1948-2012 period. Washington, D.C. surpassed the previous annual rainfall record of 61.33 inches set in 1889, accumulating 64.22 inches by the date of this map. Similarly, Wilmington, North Carolina, has received 99.98 inches. Baltimore, Maryland, broke its previous record of 62.66 inches set in 2003 with 68.77 inches. The map is generated by incorporating meteorological data with observations from NASA's Gravity Recovery and Climate Experiment (GRACE) satellites.
https://earthobservatory.nasa.gov/images/144417/soggy-2018-for-the-eastern-us

Fig. S7

A photographic comparison of diseased and healthy state beech trees at the two study sites. Beech bark disease has not been observed on the SERC Forest Dynamics Plot. https://forestgeo.si.edu/sites/north-america/smithsonian-environmental-research-center. The trees in HARV Forest Dynamics Plot however are on the killing front of the disease https://forestgeo.si.edu/sites/north-america/harvard-forest.

Fig. S8

Enriched GO terms of differentially expressed transcripts during spring-summer transitions of 2017 and 2018. Terms enriched among upregulated transcripts are shown in red. Terms related to downregulated transcripts are shown in blue.

Competing Interests

None declared.

Acknowledgments

We would like to thank Dr. Jill Wegrzyn for being a valuable resource person in bioinformatics and providing access to University of Connecticut's Xanadu High-Performance Computation Cluster. We would like to extend our gratitude for Dr. Susan McEvoy and Dr. Adrian Powell for sharing the early draft beech genome sequence assemblies and for Dr. Mengmeng Lu for providing insight and discussion that has helped shape the final version of this manuscript. We are also grateful to Mark VanScoy for taking dendroband measurements in Harvard Forest. This work was funded by the US National Science Foundation (DEB-1638488).

Author Contributions

UUS analyzed the data, prepared the figures and wrote the manuscript with input from all authors. NGS SMM SJD conceived and articulated the idea and obtained the funding. UUS JES SJW NGS collected the samples. UUS carried out the molecular lab work. NGS coordinated the sequencing, data curation, analytical support. SJW SJD SMM JES NGS contributed to the conceptualization, writing, review and proofreading, JES SJD NGS helped with the project coordination and implementation.

References

1. Katharine L. Gerst, Theresa M. Crimmins, Erin E. Posthumus, Alyssa H. Rosemartin, and Mark D. Schwartz. How well do the spring indices predict phenological activity across plant species? *International Journal of Biometeorology*, 64(5):889–901, May 2020.

2. Catherine J. Chamberlain, Benjamin I. Cook, Iñaki García de Cortázar-Atauri, and Elizabeth M. Wolkovich. Rethinking false spring risk. *Global Change Biology*, 25(7):2209–2220, 2019. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.14642.

3. Steven J. Franks, Jennifer J. Weber, and Sally N. Aitken. Evolutionary and plastic responses to climate change in terrestrial plant populations. *Evolutionary Applications*, 7(1):123–139, 2014. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/eva.12112.

4. Brett R. Scheffers, Luc De Meester, Tom C. L. Bridge, Ary A. Hoffmann, John M. Pandolfi, Richard T. Corlett, Stuart H. M. Butchart, Paul Pearce-Kelly, Kit M. Kovacs, David Dudgeon, Michela Pacifici, Carlo Rondinini, Wendy B. Foden, Tara G. Martin, Camilo Mora, David Bickford, and James E. M. Watson. The broad footprint of climate change from genes to biomes to people. *Science*, 354(6313):aaf7671, November 2016. Publisher: American Association for the Advancement of Science.

5. Enrique Valencia, Francesco de Bello, Thomas Galland, Peter B. Adler, Jan Lepš, Anna E-Vojtkó, Roel van Klink, Carlos P. Carmona, Jiří Danihelka, Jürgen Dengler, David J. Eldridge, Marc Estiarte, Ricardo García-González, Eric Garnier, Daniel Gómez-García, Susan P. Harrison, Tomáš Herben, Ricardo Ibáñez, Anke Jentsch, Norbert Juergens, Miklós Kertész, Katja Klumpp, Frédérique Louault, Rob H. Marrs, Romà Ogaya, Gábor Ónodi, Robin J. Pakeman, Iker Pardo, Meelis Pärtel, Begoña Peco, Josep Peñuelas, Richard F. Pywell, Marta Rueda, Wolfgang Schmidt, Ute Schmiedel, Martin Schuetz, Hana Skálová, Petr Šmilauer, Marie Šmilauerová, Christian Smit, MingHua Song, Martin Stock, James Val, Vigdis Vandvik, David Ward, Karsten Wesche, Susan K. Wiser, Ben A. Woodcock, Truman P. Young, Fei-Hai Yu, Martin Zobel, and Lars Götzenberger. Synchrony matters more than species richness in plant community stability at a global scale. *Proceedings of the National Academy of Sciences*, 117(39):24345–24351, September 2020. Publisher: Proceedings of the National Academy of Sciences.

6. Brunno F. Oliveira, Frances C. Moore, and Xiaoli Dong. Biodiversity mediates ecosystem sensitivity to climate variability. *Communications Biology*, 5(1):1–9, June 2022. Number: 1 Publisher: Nature Publishing Group.

7. Sara J. Germain and James A. Lutz. Climate warming may weaken stabilizing mechanisms in old forests. *Ecological Monographs*, 92(2):e1508, 2022. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/ecm.1508.

8. David Bauman, Claire Fortunel, Guillaume Delhaye, Yadvinder Malhi, Lucas A. Cernusak, Lisa Patrick Bentley, Sami W. Rifai, Jesús Aguirre-Gutiérrez, Imma Oliveras Menor, Oliver L. Phillips, Brandon E. McNellis, Matt Bradford, Susan G. W. Laurance, Michael F. Hutchinson, Raymond Dempsey, Paul E. Santos-Andrade, Hugo R. Ninantay-Rivera, Jimmy R. Chambi Paucar, and Sean M. McMahon. Tropical tree mortality has increased with rising atmospheric water stress. *Nature*, pages 1–6, May 2022. Publisher: Nature Publishing Group.

9. Ute Sass-Klaassen, Clifton R. Sabajo, and Jan den Ouden. Vessel formation in relation to leaf phenology in pedunculate oak and European ash. *Dendrochronologia*, 29(3):171–175, January 2011.

10. Eric A. Nord and Jonathan P. Lynch. Plant phenology: a critical controller of soil resource acquisition. *Journal of Experimental Botany*, 60(7):1927–1937, May 2009.

11. Yasuyuki Aono and Keiko Kazui. Phenological data series of cherry tree flowering in Kyoto, Japan, and its application to reconstruction of springtime temperatures since the 9th century. *International Journal of Climatology*, 28(7):905–914, 2008. _eprint: https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1002/joc.1594.

12. Elizabeth R. Ellwood, Stanley A. Temple, Richard B. Primack, Nina L. Bradley, and Charles C. Davis. Record-Breaking Early Flowering in the Eastern United States. *PLOS ONE*, 8(1):e53788, January 2013. Publisher: Public Library of Science.

13. Ulf Büntgen, Alma Piermattei, Paul J. Krusic, Jan Esper, Tim Sparks, and Alan Crivellaro. Plants in the UK flower a month earlier under recent warming. *Proceedings of the Royal Society B: Biological Sciences*, 289(1968):20212456, February 2022. Publisher: Royal Society.

14. Caroline A. Polgar and Richard B. Primack. Leaf-out phenology of temperate woody plants: from trees to ecosystems. *New Phytologist*, 191(4):926–941, 2011. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-8137.2011.03803.x.

15. Caroline Polgar, Amanda Gallinat, and Richard B. Primack. Drivers of leaf-out phenology and their implications for species invasions: insights from Thoreau's Concord. *New Phytologist*, 202(1):106–115, 2014. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/nph.12647.

16. Yongshuo S. H. Fu, Matteo Campioli, Yann Vitasse, Hans J. De Boeck, Joke Van den Berge, Hamada AbdElgawad, Han Asard, Shilong Piao, Gaby Deckmyn, and Ivan A. Janssens. Variation in leaf flushing date influences autumnal senescence and next year's flushing date in two temperate tree species. *Proceedings of the National Academy of Sciences*, 111(20):7355–7360, May 2014. Publisher: Proceedings of the National Academy of Sciences.

17. Constantin M Zohner, Lidong Mo, and Susanne S Renner. Global warming reduces leaf-out and flowering synchrony among individuals. *eLife*, 7:e40214, November 2018. Publisher: eLife Sciences Publications, Ltd.

18. Rebecca A. Montgomery, Karen E. Rice, Artur Stefanski, Roy L. Rich, and Peter B. Reich. Phenological responses of temperate and boreal trees to warming depend on ambient spring temperatures, leaf habit, and geographic range. *Proceedings of the National Academy of Sciences*, 117(19):10397–10405, May 2020. Publisher: Proceedings of the National Academy of Sciences.

19. Soile Jokipii-Lukkari, Nicolas Delhomme, Bastian Schiffthaler, Chanaka Mannapperuma, Jakob Prestele, Ove Nilsson, Nathaniel R. Street, and Hannele Tuominen. Transcriptional Roadmap to Seasonal Variation in Wood Formation of Norway Spruce. *Plant Physiology*, 176(4):2851–2870, April 2018.

20. Thomas Caignard, Antoine Kremer, Cyril Firmat, Manuel Nicolas, Samuel Venner, and Sylvain Delzon. Increasing spring temperatures favor oak seed production in temperate areas. *Scientific Reports*, 7(1):8555, August 2017. Number: 1 Publisher: Nature Publishing Group.

21. Janet S. Prevéy, Lauren E. Parker, and Constance A. Harrington. Projected impacts of climate change on the range and phenology of three culturally-important shrub species. *PLOS ONE*, 15(5):e0232537, May 2020. Publisher: Public Library of Science.

22. Manette E. Sandor, Clare E. Aslan, Liba Pejchar, and Judith L. Bronstein. A Mechanistic Framework for Understanding the Effects of Climate Change on the Link Between Flowering and Fruiting Phenology. *Frontiers in Ecology and Evolution*, 9, 2021.

23. Yann Vitasse, Annabel Josée Porté, Antoine Kremer, Richard Michalet, and Sylvain Delzon. Responses of canopy duration to temperature changes in four temperate tree species: relative contributions of spring and autumn leaf phenology. *Oecologia*, 161(1):187–198, August 2009.

24. Marc Estiarte and Josep Peñuelas. Alteration of the phenology of leaf senescence and fall in winter deciduous species by climate change: effects on nutrient proficiency. *Global Change Biology*, 21(3):1005–1017, 2015. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.12804.

25. Allison L. Gill, Amanda S. Gallinat, Rebecca Sanders-DeMott, Angela J. Rigden, Daniel J. Short Gianotti, Joshua A. Mantooth, and Pamela H. Templer. Changes in autumn senescence in northern hemisphere deciduous trees: a meta-analysis of autumn phenology studies. *Annals of Botany*, 116(6):875–888, November 2015.

26. Jos H.M. Schippers, Romy Schmidt, Carol Wagstaff, and Hai-Chun Jing. Living to Die and Dying to Live: The Survival Strategy behind Leaf Senescence1. *Plant Physiology*, 169(2):914–930, October 2015.

27. Isabelle Chuine. Why does phenology drive species distribution? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1555):3149–3160, October 2010. Publisher: Royal Society.

28. Fernando Valladares, Cristina C. Bastias, Oscar Godoy, Elena Granda, and Adrián Escudero. Species coexistence in a changing world. *Frontiers in Plant Science*, 6, 2015.

29. Haiwei Lu, Michael I. Gordon, Vindhya Amarasinghe, and Steven H. Strauss. Extensive transcriptome changes during seasonal leaf senescence in field-grown black cottonwood (Populus trichocarpa Nisqually-1). *Scientific Reports*, 10(1):6581, April 2020. Number: 1 Publisher: Nature Publishing Group.

30. Richard Cronn, Peter C. Dolan, Sanjuro Jogdeo, Jill L. Wegrzyn, David B. Neale, J. Bradley St. Clair, and Dee R. Denver. Transcription through the eye of a needle: daily and annual cyclic gene expression variation in Douglas-fir needles. *BMC Genomics*, 18, July 2017.

31. J. Núñez-Farfán and C. D. Schlichting. Evolution in Changing Environments: The "Synthetic" Work of Clausen, Keck, and Hiesey. *The Quarterly Review of Biology*, 76(4):433–457, December 2001. Publisher: The University of Chicago Press.

32. Juha Merilä and Andrew P. Hendry. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evolutionary Applications*, 7(1):1–14, 2014. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/eva.12137.

33. Alexandre Marchal, Carl D Schlichting, Rémy Gobin, Philippe Balandier, Frédéric Millier, Facundo Muñoz, Luc E Pâques, and Leopoldo Sánchez. Deciphering Hybrid Larch Reaction Norms Using Random Regression. *G3 Genes|Genomes|Genetics*, 9(1):21–32, January 2019.

34. Emily K. Meineke, Charles C. Davis, and T. Jonathan Davies. Phenological sensitivity to temperature mediates herbivory. *Global Change Biology*, 27(11):2315–2327, 2021. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gcb.15600.

35. Marcel E Visser and Christiaan Both. Shifts in phenology due to global climate change: the need for a yardstick. *Proceedings of the Royal Society B: Biological Sciences*, 272(1581):2561–2569, December 2005. Publisher: Royal Society.

36. Sven Wagner, Catherine Collet, Palle Madsen, Tohru Nakashizuka, Ralph D. Nyland, and Khosro Sagheb-Talebi. Beech regeneration research: From ecological to silvicultural aspects. *Forest Ecology and Management*, 259(11):2172–2182, May 2010.

37. Marieke van der Maaten-Theunissen, Hanna Bümmerstede, Janette Iwanowski, Tobias Scharnweber, Martin Wilmking, and Ernst van der Maaten. Drought sensitivity of beech on a shallow chalk soil in northeastern Germany – a comparative study. *Forest Ecosystems*, 3(1):24, September 2016.

38. Atsushi J. Nagano, Tetsuhiro Kawagoe, Jiro Sugisaka, Mie N. Honjo, Koji Iwayama, and Hiroshi Kudoh. Annual transcriptome dynamics in natural environments reveals plant seasonal adaptation. *Nature Plants*, 5(1):74, January 2019.

39. Susan L. McEvoy, U. Uzay Sezen, Alexander Trouern-Trend, Sean M. McMahon, Paul G. Schaberg, Jie Yang, Jill L. Wegrzyn, and Nathan G. Swenson. Strategies of tolerance reflected in two North American maple genomes. *The Plant Journal*, 109(6):1591–1613, 2022. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/tpj.15657.

40. Quansheng Ge, Huanjiong Wang, Jingyun Zheng, Rutishauser This, and Junhu Dai. A 170 year spring phenology index of plants in eastern China. *Journal of Geophysical Research: Biogeosciences*, 119(3):301–311, 2014. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2013JG002565.

41. Jihua Ding and Ove Nilsson. Molecular regulation of phenology in trees—because the seasons they are a-changin'. *Current Opinion in Plant Biology*, 29:73–79, February 2016.

42. Annarita Marrano, Gina M. Sideli, Charles A. Leslie, Hao Cheng, and David B. Neale. Deciphering of the Genetic Control of Phenology, Yield, and Pellicle Color in Persian Walnut (Juglans regia L.). *Frontiers in Plant Science*, 10, 2019.

43. Jihua Ding, Bo Zhang, Yue Li, Domenique André, and Ove Nilsson. Phytochrome B and PHYTOCHROME INTERACTING FACTOR8 modulate seasonal growth in trees. *New Phytologist*, 232(6):2339–2352, 2021. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/nph.17350.

44. Rami-Petteri Apuli, Thomas Richards, Martha Rendón-Anaya, Almir Karacic, Ann-Christin Rönnberg-Wästljung, and Pär K. Ingvarsson. The genetic basis of adaptation in phenology in an introduced population of Black Cottonwood (Populus trichocarpa, Torr. & Gray). *BMC Plant Biology*, 21(1):317, July 2021.

45. Joanna Meger, Bartosz Ulaszewski, and Jaroslaw Burczyk. Genomic signatures of natural selection at phenology-related genes in a widely distributed tree species Fagus sylvatica L. *BMC Genomics*, 22(1):583, July 2021.

46. Xiaoyan Sheng, Chuan-Yu Hsu, Cathleen Ma, and Amy M. Brunner. Functional Diversification of Populus FLOWERING LOCUS D-LIKE3 Transcription Factor and Two Paralogs in Shoot Ontogeny, Flowering, and Vegetative Phenology. *Frontiers in Plant Science*, 13, 2022.

47. Deborah A. Johnson, Jeffrey P. Hill, and Michael A. Thomas. The monosaccharide transporter gene family in land plants is ancient and shows differential subfamily expression and expansion across lineages. *BMC Evolutionary Biology*, 6(1):64, August 2006.

48. Thomas L. Slewinski. Diverse Functional Roles of Monosaccharide Transporters and their Homologs in Vascular Plants: A Physiological Perspective. *Molecular Plant*, 4(4):641–662, July 2011.

49. R. A. Schofield, Y.-M. Bi, S. Kant, and S. J. Rothstein. Over-expression of STP13, a hexose transporter, improves plant growth and nitrogen use in Arabidopsis thaliana seedlings. *Plant, Cell & Environment*, 32(3):271–285, 2009. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-3040.2008.01919.x.

50. Benjamin T. Julius, Kristen A. Leach, Thu M. Tran, Rachel A. Mertz, and David M. Braun. Sugar Transporters in Plants: New Insights and Discoveries. *Plant and Cell Physiology*, 58(9):1442–1460, September 2017.

51. Peter Aasted Paulsen, Tânia F. Custódio, and Bjørn Panyella Pedersen. Crystal structure of the plant symporter STP10 illuminates sugar uptake mechanism in monosaccharide transporter superfamily. *Nature Communications*, 10(1):407, January 2019. Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 1 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Monosaccharides;Plant development;Plant molecular biology;X-ray crystallography Subject_term_id: monosaccharides;plant-development;plant-molecular-biology;x-ray-crystallography.

52. Audrey Barker Plotkin, Meghan Blumstein, Danelle Laflower, Valerie J. Pasquarella, Jennifer L. Chandler, Joseph S. Elkinton, and Jonathan R. Thompson. Defoliated trees die below a critical threshold of stored carbon. *Functional Ecology*, 35(10):2156–2167, 2021. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1365-2435.13891.

53. Hong Gil Lee and Pil Joon Seo. Transcriptional activation of SUGAR TRANSPORT PROTEIN 13 mediates biotic and abiotic stress signaling. *Plant Signaling & Behavior*, 16(8):1920759, August 2021. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/15592324.2021.1920759.

54. Lee Hunt and Julie E. Gray. How the stomate got his pore: very long chain fatty acids and a structural cell wall protein sculpt the guard cell outer cuticular ledge. *New Phytologist*, 228(6):1698–1700, 2020. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/nph.16843.

55. Antoine De Bigault Du Granrut and Jean-Luc Cacas. How Very-Long-Chain Fatty Acids Could Signal Stressful Conditions in Plants? *Frontiers in Plant Science*, 7:1490, 2016.

56. Paras Jain, Sneha Singh, and Aditya Arya. A student centric method for calculation of fatty acid energetics: Integrated formula and web tool. *Biochemistry and Molecular Biology Education*, 49(3):492–499, 2021. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/bmb.21486.

57. François Roudier, Lionel Gissot, Frédéric Beaudoin, Richard Haslam, Louise Michaelson, Jessica Marion, Diana Molino, Amparo Lima, Liên Bach, Halima Morin, Frédérique Tellier, Jean-Christophe Palauqui, Yannick Bellec, Charlotte Renne, Martine Miquel, Marco DaCosta, Julien Vignard, Christine Rochat, Jonathan E. Markham, Patrick Moreau, Johnathan Napier, and Jean-Denis Faure. Very-Long-Chain Fatty Acids Are Involved in Polar Auxin Transport and Developmental Patterning in Arabidopsis. *The Plant Cell*, 22(2):364–375, February 2010.

58. Bruce S. Weir, Amy D. Anderson, and Amanda B. Hepler. Genetic relatedness analysis: modern data and new challenges. *Nature Reviews Genetics*, 7(10):771–780, October 2006. Number: 10 Publisher: Nature Publishing Group.

59. Alexander S. P. Pfaff. From Deforestation to Reforestation in New England, United States. In Matti Palo and Heidi Vanhanen, editors, *World Forests from Deforestation to Transition?*, World Forests, pages 67–82. Springer Netherlands, Dordrecht, 2000.

60. Paula E. Marquardt, Craig S. Echt, Bryan K. Epperson, and Dan M. Pubanz. Genetic structure, diversity, and inbreeding of eastern white pine under different management conditions. *Canadian Journal of Forest Research*, 37(12):2652–2662, December 2007. Publisher: NRC Research Press.

61. Bruce Kershner, Robert T. Leverett, and 1941. Sierra Club guide to the ancient forests of the Northeast. Technical report, Sierra Club Books, 2004. ISBN: 9781578050666.

62. Om P. Rajora and John W. R. Zinck. Genetic Diversity, Structure and Effective Population Size of Old-Growth vs. Second-Growth Populations of Keystone and Long-Lived Conifer, Eastern White Pine (Pinus strobus): Conservation Value and Climate Adaptation Potential. *Frontiers in Genetics*, 12:650299, August 2021.

63. J. B. Gerwein and R. V. Kesseli. Genetic diversity and population structure of Quercus rubra (Fagaceae) in old-growth and secondary forests in southern New England. *Rhodora*, 108(933):1–18, March 2006. Publisher: New England Botanical Club.

64. F. A. Jones, J. L. Hamrick, C. J. Peterson, and E. R. Squiers. Inferring colonization history from analyses of spatial genetic structure within populations of Pinus strobus and Quercus rubra. *Molecular Ecology*, 15(3):851–861, 2006. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-294X.2005.02830.x.

65. Emily V. Moran and James S. Clark. Between-Site Differences in the Scale of Dispersal and Gene Flow in Red Oak. *PLOS ONE*, 7(5):e36492, May 2012. Publisher: Public Library of Science.

66. Juan J. Robledo-Arnuncio, Etienne K. Klein, Helene C. Muller-Landau, and Luis Santamaría. Space, time and complexity in plant dispersal ecology. *Movement Ecology*, 2(1):16, August 2014.

67. Martin F. Breed, Kym M. Ottewell, Michael G. Gardner, and Andrew J. Lowe. Clarifying climate change adaptation responses for scattered trees in modified landscapes. *Journal of Applied Ecology*, 48(3):637–641, 2011. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2664.2011.01969.x.

68. B. D. Dow and M. V. Ashley. Microsatellite analysis of seed dispersal and parentage of saplings in bur oak, Quercus macrocarpa. *Molecular Ecology*, 5(5):615–627, 1996. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-294X.1996.tb00357.x.

69. John D. Nason, E. Allen Herre, and J. L. Hamrick. The breeding structure of a tropical keystone plant resource. *Nature*, 391(6668):685–687, February 1998. Number: 6668 Publisher: Nature Publishing Group.

70. Preston R. Aldrich and J. L. Hamrick. Reproductive Dominance of Pasture Trees in a Fragmented Tropical Forest Mosaic. *Science*, 281(5373):103–105, July 1998. Publisher: American Association for the Advancement of Science.

71. Yuji Isagi, Tatsuo Kanazashi, Wajirou Suzuki, Hiroshi Tanaka, and Tetsuto Abe. Microsatellite analysis of the regeneration process of Magnolia obovata Thunb. *Heredity*, 84(2):143–151, February 2000. Number: 2 Publisher: Nature Publishing Group.

72. Uzay U. Sezen, Robin L. Chazdon, and Kent E. Holsinger. Genetic Consequences of Tropical Second-Growth Forest Regeneration. *Science*, 307(5711):891–891, February 2005. Publisher: American Association for the Advancement of Science.

73. Britta Denise Hardesty, Stephen P. Hubbell, and Eldredge Bermingham. Genetic evidence of frequent long-distance recruitment in a vertebrate-dispersed tree. *Ecology Letters*, 9(5):516–525, 2006. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1461-0248.2006.00897.x.

74. Christopher W. Dick, Olivier J. Hardy, F. Andrew Jones, and Rémy J. Petit. Spatial Scales of Pollen and Seed-Mediated Gene Flow in Tropical Rain Forest Trees. *Tropical Plant Biology*, 1(1):20–33, March 2008.

75. Céline Born, Olivier J. Hardy, Marie-Hélène Chevallier, Simon Ossari, Christiane Attéké, E. Jean Wickings, and Martine Hossaert-Mckey. Small-scale spatial genetic structure in the Central African rainforest tree species Aucoumea klaineana: a stepwise approach to infer the impact of limited gene dispersal, population history and habitat fragmentation. *Molecular Ecology*, 17(8):2041–2050, 2008. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-294X.2007.03685.x.

76. Uzay U. Sezen, Robin L. Chazdon, and Kent E. Holsinger. Proximity is not a proxy for parentage in an animal-dispersed Neotropical canopy palm. *Proceedings of the Royal Society B: Biological Sciences*, 276(1664):2037–2044, June 2009. Publisher: Royal Society.

77. Luke Browne, Kym Ottewell, Victoria L. Sork, and Jordan Karubian. The relative contributions of seed and pollen dispersal to gene flow and genetic diversity in seedlings of a tropical palm. *Molecular Ecology*, 27(15):3159–3173, 2018. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/mec.14768.

78. W. Carter Johnson and Curtis S. Adkisson. Dispersal of Beech Nuts by Blue Jays in Fragmented Landscapes. *The American Midland Naturalist*, 113(2):319–324, 1985. Publisher: University of Notre Dame.

79. C. S. (Clifford Scharff) Schopmeyer, United States, and Forest Service. Seeds of woody plants in the United States. pages viii, 883 , [8] leaves of plates : ill. (some col.) ; 27 cm.–USDA, 1974.

80. Stephen B. Vander Wall. The evolutionary ecology of nut dispersal. *The Botanical Review*, 67(1):74–117, January 2001.

81. Anthony M. Bolger, Marc Lohse, and Bjoern Usadel. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120, August 2014.

82. Daehwan Kim, Joseph M. Paggi, Chanhee Park, Christopher Bennett, and Steven L. Salzberg. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nature Biotechnology*, 37(8):907–915, August 2019. Number: 8 Publisher: Nature Publishing Group.

83. Gregory G. Faust and Ira M. Hall. SAMBLASTER: fast duplicate marking and structural variant read extraction. *Bioinformatics*, 30(17):2503–2505, September 2014.

84. Petr Danecek, James K Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O Pollard, Andrew Whitwham, Thomas Keane, Shane A McCarthy, Robert M Davies, and Heng Li. Twelve years of SAMtools and BCFtools. *GigaScience*, 10(2):giab008, February 2021.

85. Erik Garrison and Gabor Marth. Haplotype-based variant detection from short-read sequencing. Technical Report arXiv:1207.3907, arXiv, July 2012. arXiv:1207.3907 [q-bio] type: article.

86. Heng Li. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics (Oxford, England)*, 27(21):2987–2993, November 2011.

87. Peter J. Bradbury, Zhiwu Zhang, Dallas E. Kroon, Terry M. Casstevens, Yogesh Ramdoss, and Edward S. Buckler. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics*, 23(19):2633–2635, October 2007.

88. Jeffrey B Endelman and Jean-Luc Jannink. Shrinkage Estimation of the Realized Relationship Matrix. *G3 Genes|Genomes|Genetics*, 2(11):1405–1413, November 2012.

89. Simon Anders, Paul Theodor Pyl, and Wolfgang Huber. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*, 31(2):166–169, January 2015.

90. Tauno Metsalu and Jaak Vilo. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Research*, 43(W1):W566–W570, July 2015.

91. Mihaela Pertea, Daehwan Kim, Geo M. Pertea, Jeffrey T. Leek, and Steven L. Salzberg. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nature Protocols*, 11(9):1650–1667, September 2016. Number: 9 Publisher: Nature Publishing Group.

92. Mihaela Pertea, Geo M. Pertea, Corina M. Antonescu, Tsung-Cheng Chang, Joshua T. Mendell, and Steven L. Salzberg. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nature Biotechnology*, 33(3):290–295, March 2015. Number: 3 Publisher: Nature Publishing Group.

93. Alyssa C. Frazee, Geo Pertea, Andrew E. Jaffe, Ben Langmead, Steven L. Salzberg, and Jeffrey T. Leek. Ballgown bridges the gap between transcriptome assembly and expression analysis. *Nature Biotechnology*, 33(3):243–246, March 2015. Number: 3 Publisher: Nature Publishing Group.

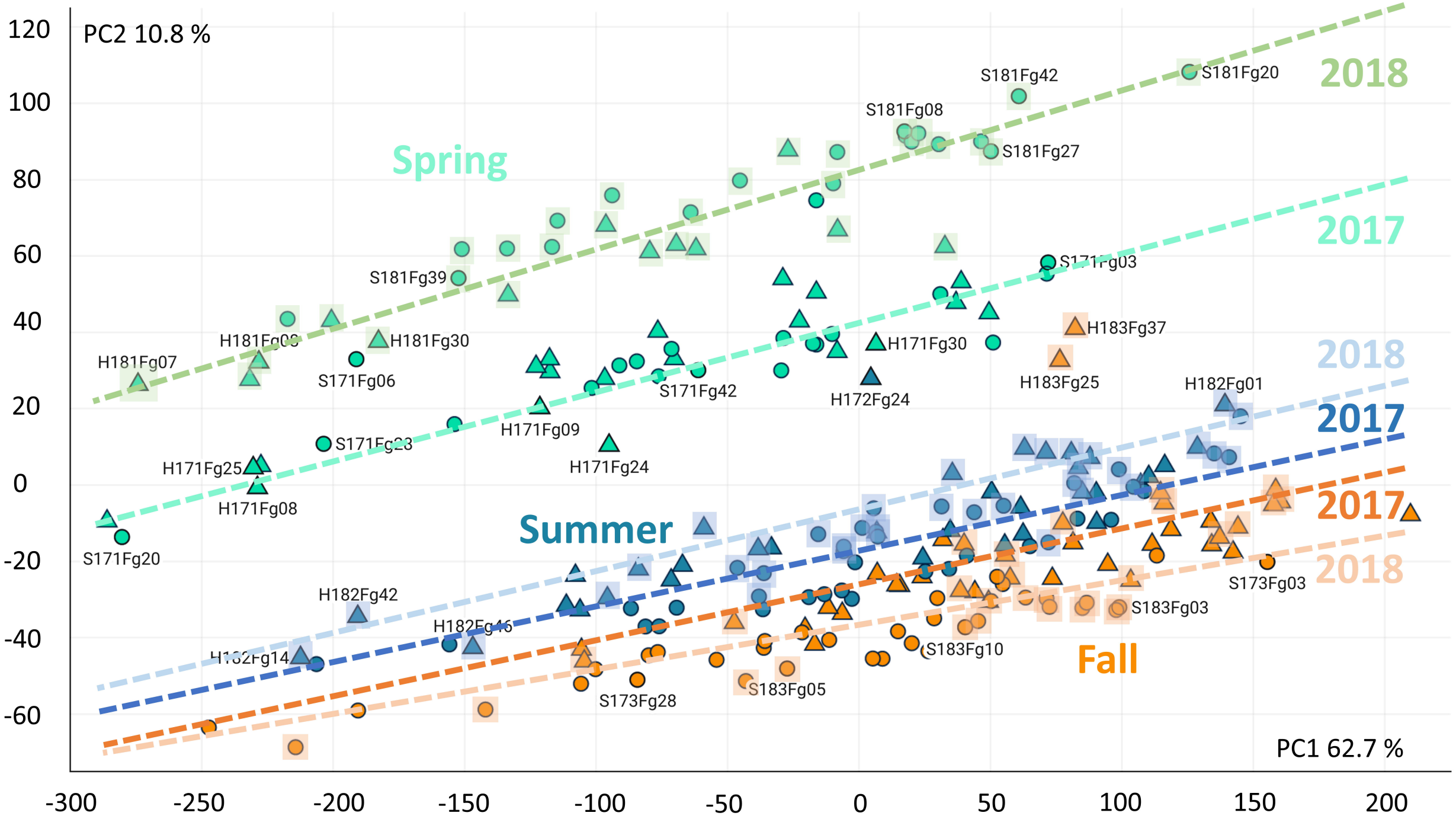
94. Torbjørn Rognes, Tomáš Flouri, Ben Nichols, Christopher Quince, and Frédéric Mahé. VSEARCH: a versatile open source tool for metagenomics. *PeerJ*, 4, October 2016.

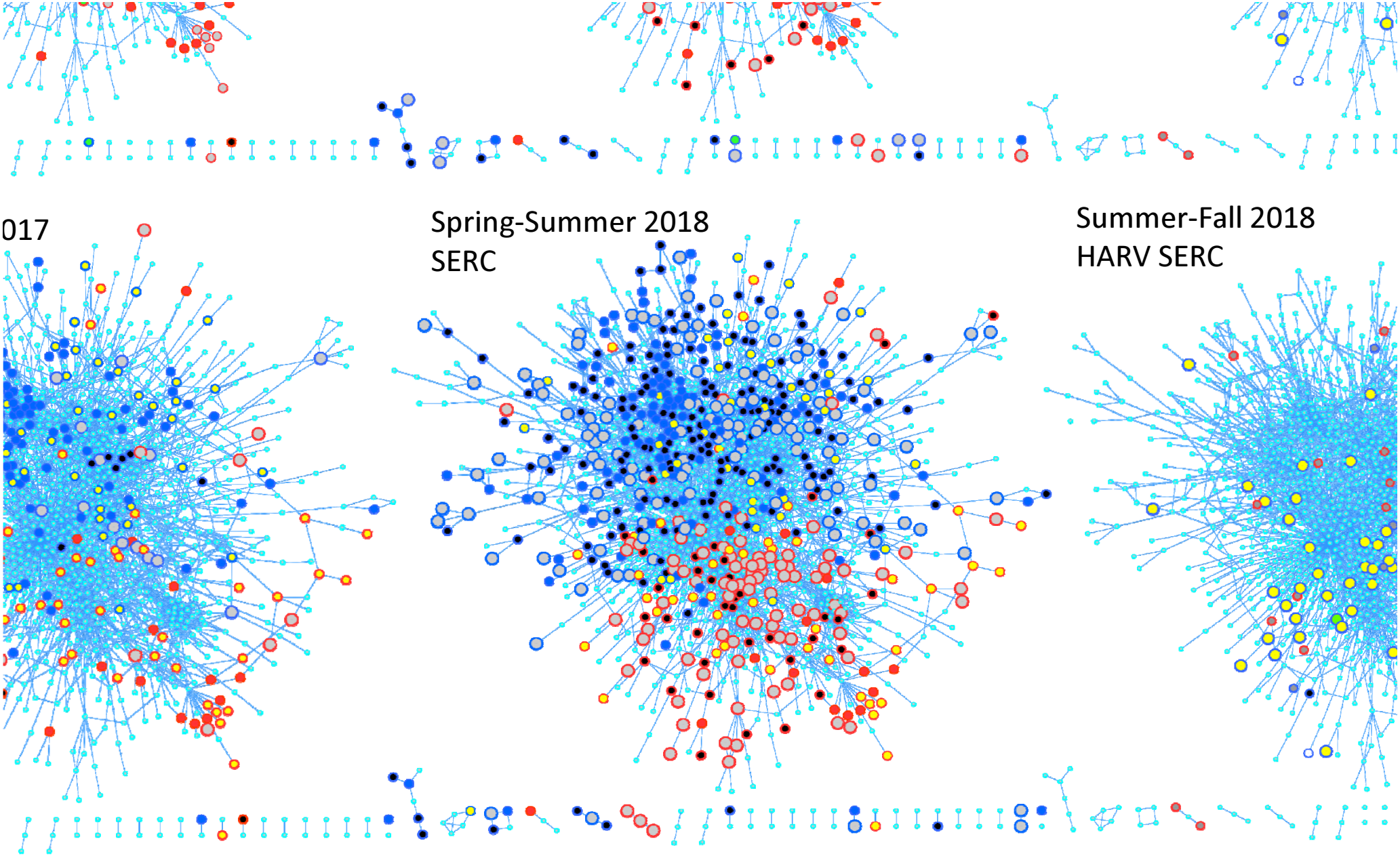
95. Christian von Mering, Lars J. Jensen, Berend Snel, Sean D. Hooper, Markus Krupp, Mathilde Foglierini, Nelly Jouffre, Martijn A. Huynen, and Peer Bork. STRING: known and predicted protein–protein associations, integrated and transferred across organisms. *Nucleic Acids Research*, 33(suppl_1):D433–D437, January 2005.

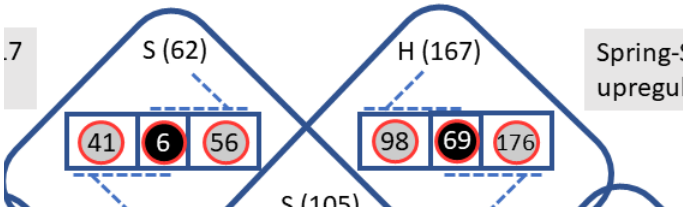
96. Andrea Franceschini, Damian Szklarczyk, Sune Frankild, Michael Kuhn, Milan Simonovic, Alexander Roth, Jianyi Lin, Pablo Minguez, Peer Bork, Christian von Mering, and Lars J. Jensen. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Research*, 41(D1):D808–D815, January 2013.

97. Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S. Baliga, Jonathan T. Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, 13(11):2498–2504, November 2003.

98. Uku Raudvere, Liis Kolberg, Ivan Kuzmin, Tambet Arak, Priit Adler, Hedi Peterson, and Jaak Vilo. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, 47(W1):W191–W198, July 2019.

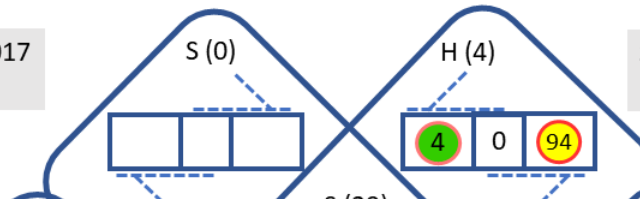

PC2 10.8 %
PC1 62.7 %
120
100
80
60
40
20
0
-20
-40
-60
-300
-250
-200
-150
-100
-50
0
50
100
150
200
Spring
Summer
Fall
2018
2017
2018
2017
2017
2018
S181Fg42
S181Fg20
S181Fg08
S181Fg27
S181Fg39
S171Fg03
H181Fg30
H181Fg07
H171Fg30
H183Fg37
S171Fg06
S171Fg42
H172Fg24
H183Fg25
H182Fg01
H171Fg09
S171Fg23
H171Fg24
H171Fg25
H171Fg08
S171Fg20
H182Fg42
H182Fg46
H182Fg14
S173Fg03
S183Fg03
S183Fg10
S173Fg28
S183Fg05

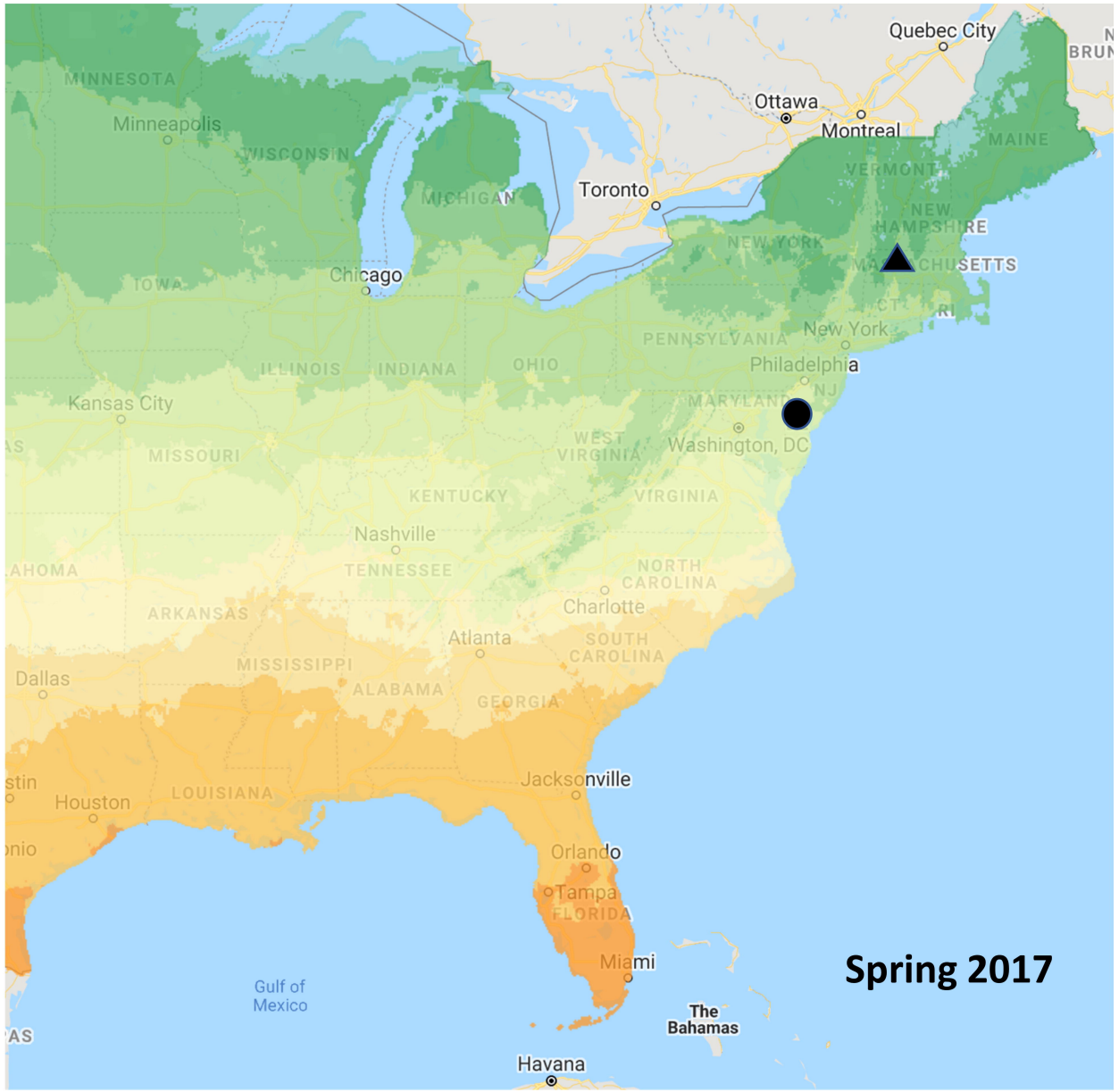


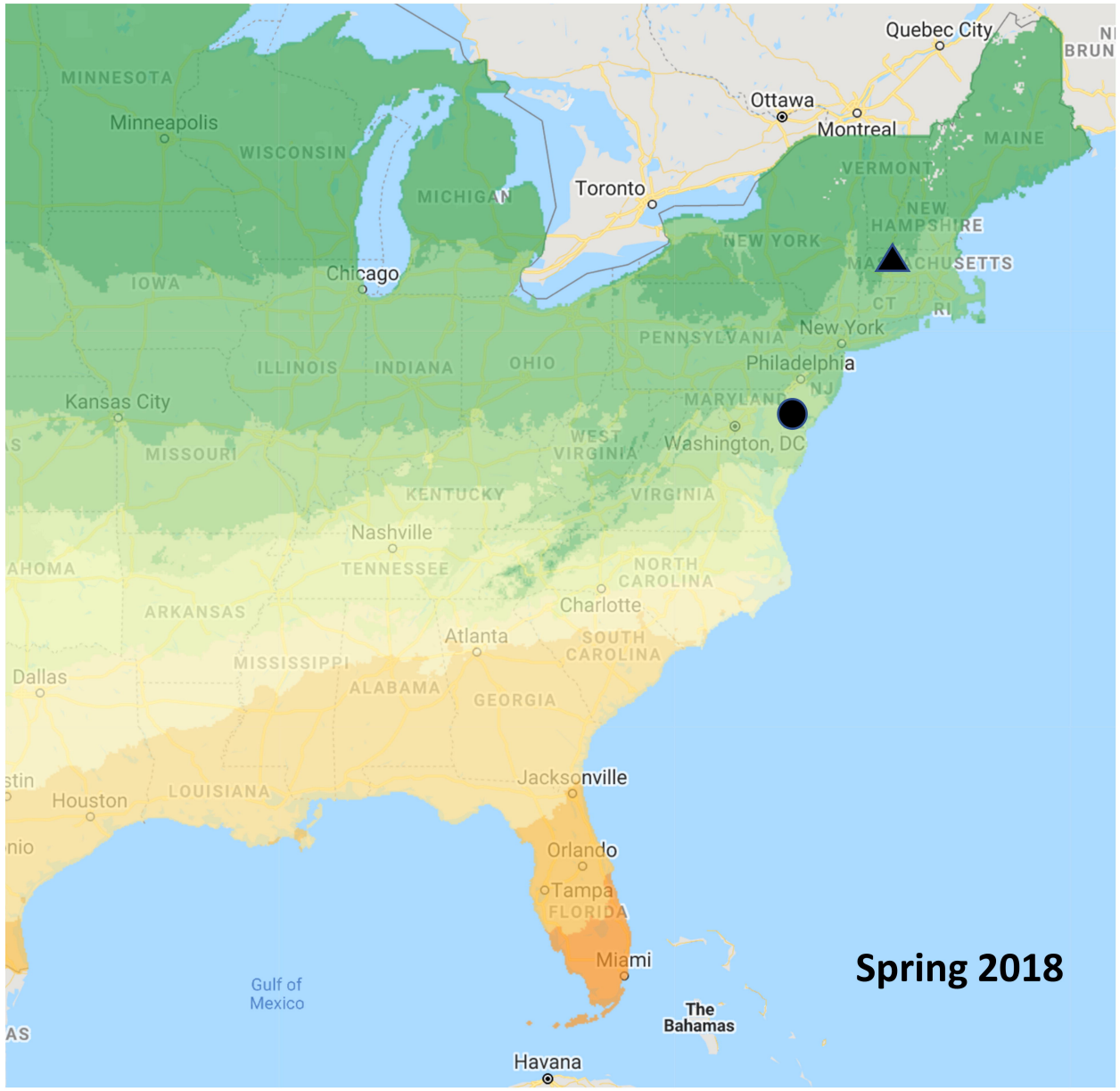


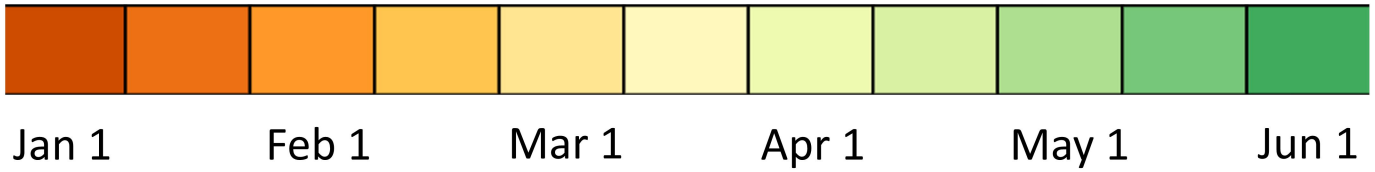
Spring-Summer 2018
upregulated

Summer-Fall 2017
upregulated




Quebec City
Ottawa
Montreal
Toronto
Minneapolis
Chicago
New York
Philadelphia
Washington, DC
Kansas City
Nashville
Charlotte
Atlanta
Dallas
Houston
Jacksonville
Orlando
Tampa
Miami
Havana
MINNESOTA
WISCONSIN
MICHIGAN
IOWA
ILLINOIS
INDIANA
OHIO
PENNSYLVANIA
NEW YORK
VERMONT
NEW HAMPSHIRE
MAINE
MASSACHUSETTS
CT
RI
NJ
MARYLAND
WEST VIRGINIA
VIRGINIA
KENTUCKY
TENNESSEE
NORTH CAROLINA
SOUTH CAROLINA
GEORGIA
ALABAMA
MISSISSIPPI
ARKANSAS
MISSOURI
LOUISIANA
FLORIDA
Gulf of Mexico
The Bahamas
Spring 2017


Quebec City
Ottawa
Montreal
Toronto
Minneapolis
Chicago
New York
Philadelphia
Washington, DC
Kansas City
Nashville
Charlotte
Atlanta
Dallas
Houston
Jacksonville
Orlando
Tampa
Miami
Havana
MINNESOTA
WISCONSIN
MICHIGAN
IOWA
ILLINOIS
INDIANA
OHIO
PENNSYLVANIA
NEW YORK
VERMONT
NEW HAMPSHIRE
MAINE
MASSACHUSETTS
CT
RI
NJ
MARYLAND
WEST VIRGINIA
VIRGINIA
KENTUCKY
TENNESSEE
NORTH CAROLINA
SOUTH CAROLINA
GEORGIA
ALABAMA
MISSISSIPPI
ARKANSAS
MISSOURI
LOUISIANA
FLORIDA
Gulf of Mexico
The Bahamas
Spring 2018


Jan 1
Feb 1
Mar 1
Apr 1
May 1
Jun 1