bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Strategies of tolerance reflected in two North American maple genomes

Susan L. McEvoy[1], U. Uzay Sezen[2], Alexander Trouern-Trend[1], Sean M. McMahon[2], Paul G. Schaberg[3], Jie Yang[4], Jill L. Wegrzyn[1#], Nathan G. Swenson[5#]

[1] Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut, 06269

[2] Smithsonian Environmental Research Center, Edgewater, Maryland, 21037

[3] Forest Service, U.S. Department of Agriculture, Northern Research Station, Burlington, VT 05405

[4] CAS Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla, 666303, Yunnan, China

[5] Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, 46556

[#]Corresponding authors: Jill Wegryzn, email: jill.wegrzyn@uconn.edu, Nate Swenson, email nswenson@nd.edu

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Maples (the genus *Acer*) represent important and beloved forest, urban, and ornamental trees distributed throughout the Northern hemisphere. They exist in a diverse array of native ranges and distributions, across spectrums of tolerance or decline, and have varying levels of susceptibility to biotic and abiotic stress. Among *Acer* species, several stand out in their importance to economic interest. Here we report the first two chromosome-scale genomes for North American species, *Acer negundo* and *Acer saccharum*. Both assembled genomes contain scaffolds corresponding to 13 chromosomes, with *A. negundo* at a length of 442 Mb, N50 of 32 Mb and 30,491 genes, and *A. saccharum* at 626 Mb, N50 of 46 Mb, and 40,074 genes. No recent whole genome duplications were detected, though *A. saccharum* has local gene duplication and more recent bursts of transposable elements, as well as a large-scale translocation between two chromosomes. Genomic comparison revealed that *A. negundo* has a smaller genome with recent gene family evolution that is predominantly contracted and expansions that are potentially related to invasive tendencies and tolerance to abiotic stress. Examination of expression from RNA-Seq obtained from *A. saccharum* grown in long-term aluminum and calcium soil treatments at the Hubbard Brook Experimental Forest, provided insights into genes involved in aluminum stress response at the systemic level, as well as signs of compromised processes upon calcium deficiency, a condition contributing to maple decline.

Introduction

Acer saccharum (sugar maple) is a long-lived and dominant species in New England forests with a native range representing Eastern Canada, and the Northcentral and Northeastern United States. Widely known for its vibrant autumn hues, high quality timber, and as the preferred species for the production of maple syrup, *A. saccharum* also plays a key role in its native ecosystems, altering soil mineral content (Lucash et al., 2012), moisture levels (Emerman & Dawson, 1996), and mycorrhizae communities (Cong et al., 2015). *A. saccharum* provides food and shelter to many mammals (Godman et al., 1990), resident and migratory birds (Flashpohler and Grosshuesch 1996; Weidensaul et al. 2020), and over 300 species of caterpillar. Although seeds

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

from this predominantly monoecious species are wind dispersed, the early flowers are an important pollen source for bees in late winter (Blitzer et al., 2016).

One of the most phylogenetically (~50 MY) and morphologically distinctive *Acer* species from *A. saccharum* is *Acer negundo* (box elder). *A. negundo* is a short-lived tree that has the largest range of all North American *Acer*. It is native to predominantly lower elevation regions of Canada, the United States, and Mexico. This adaptable pioneer species is often seen in disturbed sites and urban settings (Figure 1). *A. negundo* has soft wood, less concentrated sugars for syrup, grows rapidly, and can tolerate low nutrient soils, moderate salinity, and drought conditions. Its invasive status in large portions of Europe, South America, and Asia are indicative of greater phenotypic plasticity (Lamarque et al., 2013). *A. negundo* has a number of distinctive attributes including dioecy (Renner et al., 2007) and pinnately compound leaves, mostly seen only in close relatives. Within such a significant genus, these two species reflect a breadth of social and ecological diversity and importance that recommends better understanding of their genetic distinctions.

Forest ecosystems of the Northeastern U.S. are facing significant changes in composition driven by climate change (Rogers et al., 2017). “Maple decline” is a term referring to the loss of maple populations, originally referring to *A. saccharum*, but now applicable to *A. platanoides* and *A. rubrum* in the Northeast, and most recently, *A. macrophyllum* in the Northwest. Loss of *A. saccharum* has been documented over the last century, beginning in the late 1950s, leading to the first comprehensive, multidisciplinary study of this condition (Giese & Benjamin, 1964; Horsley et al., 2002). Maple decline is characterized by crown dieback, reduction in overall health and vigor, and a decrease in regeneration (Bishop et al., 2015). Episodic decline has increased in recent decades (Oswald et al., 2018). Decline and crown dieback of dominant *A. saccharum* provides a release for sympatric species such as *Fagus grandifolia* (American beech) which displays a higher level of tolerance to soil conditions and foliar aluminum ratios leading to shifting forest composition (Halman et al., 2015). Studies examining potential factors of maple decline have largely agreed that modified soil conditions, largely due to acid deposition, are the leading cause, compounded by additional climatic, pathogenic, and anthropogenic stressors (Bal

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

et al., 2015). Acidic soils rapidly leach the essential cations calcium, magnesium and potassium, while mobilizing aluminum within the soil and contributing to more phytotoxic forms (Likens et al., 1998; Likens & Lambert, 1998). Competition between aluminum and calcium at the roots further decreases levels of available calcium within tissues, while increasing aluminum damages plasma membranes, cell walls, DNA, and increases the burden of oxidative stress. Such nutrient interactions and their broader consequences on physiology and ecology are studied at the Hubbard Brook Experimental Forest (HBEF), a Long Term Ecological Research (LTER) site. It was here that acid deposition was first discovered in North America (Likens and Bormann 1974) and continues to be studied through the Nutrient Perturbation (NuPert) program (Berger et al., 2001). It provides a replicated high elevation natural ecosystem to examine current, future, and past soil conditions, and has been the site of several key studies on native species, including *A. saccharum*, *A. balsamea*, *F. grandifolia*, and *P. rubens*. At HBEF, no studies to-date on *A. saccharum* have focused at the genomic level, where variation in gene expression or signs of adaptation among gene families may be more immediately informative in these slow-growing organisms. For such analysis, a high-quality, chromosomal-length genome is necessary to more accurately detect these forms of variation.

Genomic resources necessary to guide *Acer* conservation are very limited. Only two genomes exist to-date: *A. yangbiense*, native to the Yuhan Province (J. Yang et al., 2019) and *A. truncatum* (purpleblow maple), widely distributed across East Asia (Ma et al., 2020). Here, were present the first two North American *Acer* genomes, *A. saccharum* and *A. negundo*. With these chromosome-scale references, we describe differences in genomic characteristics that may reflect their alternative tolerance strategies. We conducted a differential expression study with stem tissue from *A. saccharum* individuals from HBEF in order to identify genes that may be involved in aluminum response and calcium deficiency. Identification of key processes in the expression study helped to provide focus to the following analysis of comparative gene family dynamics. Together, these approaches highlighted families associated with various abiotic stress responses, including those that also have significant dynamics or novel isoforms in *A. saccharum* or *A. negundo* relative to other broadleaf tree species. And it allowed investigation of the effects of

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

calcium availability at the molecular level, a significant factor associated with maple decline.

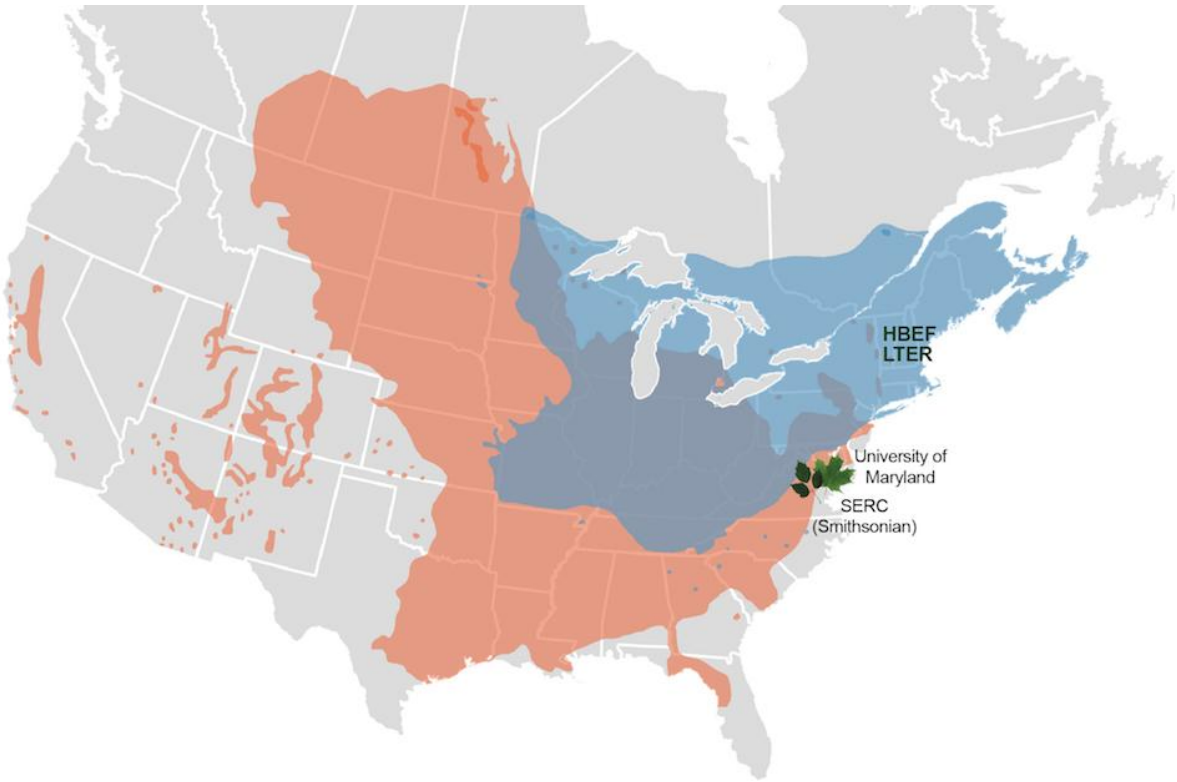


Figure 1. Distribution of *A. saccharum* (blue) and *A. negundo* (orange) in North America. *A. saccharum* range in blue, *A. negundo* in orange. Leaves indicate location of individuals selected for the reference genomes; *A. saccharum* from the University of Maryland campus, and *A. negundo* from the Smithsonian Environmental Research Center. HBEF (Hubbard Brook Experimental Forest) is the location of the 9 individuals used for RNA-seq, three each from calcium, aluminum, and control plots.

Results

Genome size estimation and quality control

A. saccharum DNA sequencing resulted in 63 Gb of PacBio data with an N50 of 21 Kb (max read length 94 Kb) and Illumina PE data totalling 225 Gb. *A. negundo* was similar with 61 Gb of

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

PacBio data, an N50 of 17 Kb (max read length 95 Kb), and 223 Gb of Illumina PE data. Genome size estimation using short reads resulted in smaller than expected estimates (Contreras & Shearer, 2018; Leitch et al., 2019) at 636 Mb and 319 Mb for *A. saccharum* and *A. negundo*, respectively Figure S1. Using the short-read estimations of genome length, DNA sequence read coverage was high, with long reads at 111x and 141x and short reads at 180x and 208x for *A. saccharum* and *A. negundo*, respectively (Table S1). RNA sequencing of the reference individuals resulted in 61 M reads for *A. saccharum* (92% mapped) and 62 M for *A. negundo* (93% mapped). RNA sequencing of samples for the differential expression study resulted in 207 M reads with mapping rates that ranged from 82 to 92%.

Genome assembly

Testing of multiple assembly approaches found the FALCON/Purge Haplotigs assembly to be the most contiguous and closest to *A. saccharum's* estimated size (Figure 2). Statistics for the set of primary contigs from FALCON did not change much between the assembly, unzip, and polishing stages (File S1). The final total length for *A. saccharum* settled at 970 Mb, with an N50 of 691 Kb and 2549 contigs. These last two statistics indicated better contiguity compared to the other assemblers tested. The associated haplotype contigs rose from 109 Mb after assembly, to 320 Mb after unzipping, and dropped to 264 Mb after polishing. Removal of under-collapsed haplotypes reduced the genome size to 668 Mb across 1210 contigs with an N50 of 951 Kb. BUSCO (embryophyte) reported 94.8% complete, but with a somewhat high percentage of duplication (11.7%).

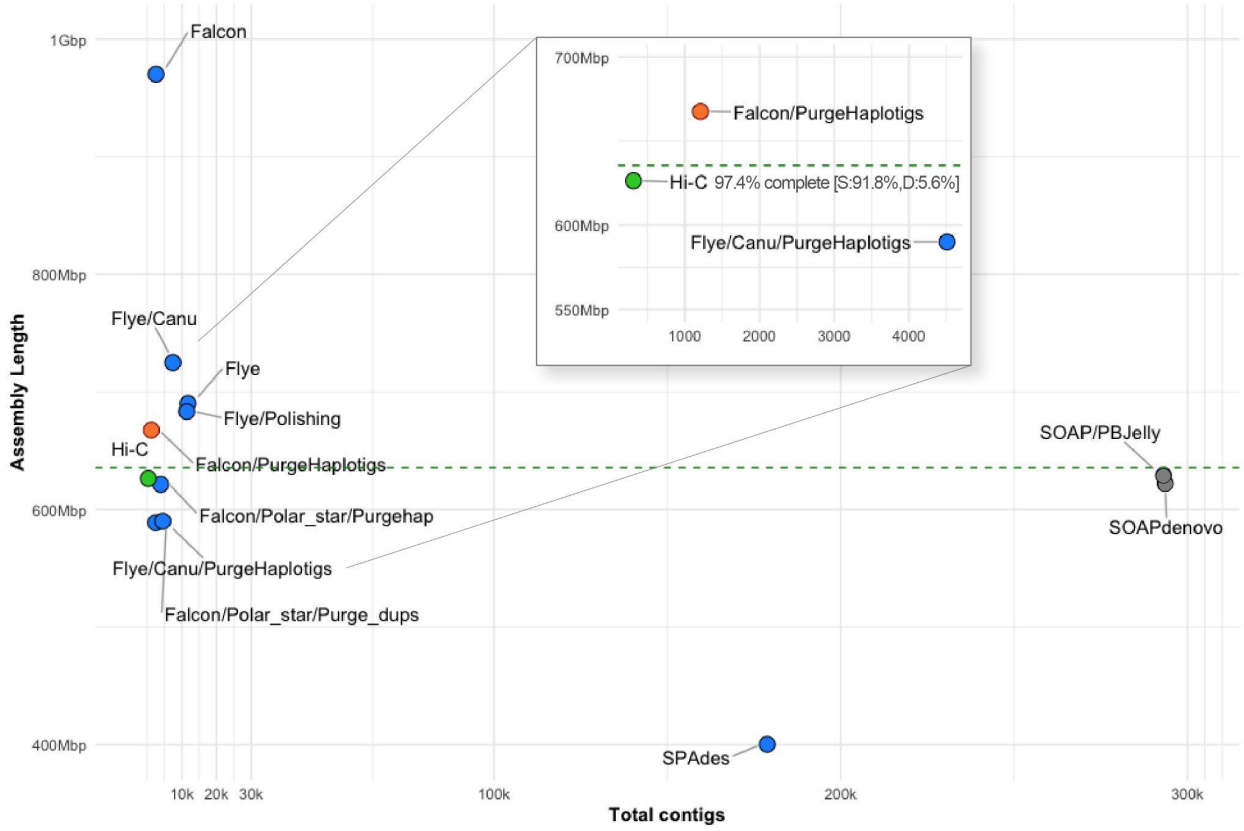
The FALCON primary assembly for *A. negundo* was also consistent in size across pipeline stages, with a total length matching estimates at 481 Mb across 1481 contigs with an N50 of 625 Kb. Removal of haplotype duplication from the primary assembly decreased the overall length to 442 Mb, number of contigs to 1063, and increased the N50 to 700Kb. The BUSCO score was 94.1% with only 5.8% duplication.

Hi-C scaffolding

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The FALCON assembly was selected over Flye due to the substantially more contiguous assembly it produced for *A. saccharum*, though it should be noted that the statistics for both assemblers were comparable for *A. negundo*. Hi-C reads provided 65x coverage of the *A. saccharum* genome. The final assembly was 626.33 Mb in 388 scaffolds with an N50 of 45.72 Mb and GC% of 35.7%. The 13 pseudo-chromosomes represented 97% of the genome length, and BUSCO (embryophyte) scores were 97.7% complete with 3.0% duplicate, 0.7% fragmented, and 1.6% missing.

A. negundo Hi-C reads provided 100x coverage and the FALCON assembly was used for scaffolding due to potential mis-assembly in the Flye version (Figure S2). The final assembly was 442.39 Mb in 108 scaffolds with an N50 of 32.30 Mb and GC% of 34.1%. The thirteen pseudo-chromosomes represented 99.74% of the total length. BUSCO embryophyta scores were 97.4% complete with 5.6% duplicate, 0.9% fragmented, and 1.7% missing. (File S1).



bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2. Results of assembly testing with *A. saccharum*, comparing fragmentation in terms of total contigs versus assembly length. The dashed line represents the estimated genome size. Gray dots are short-read assemblers, shown as highly fragmented. Blue dots are long-read tests of assembly workflows. Canu refers to the use of reads error-corrected by the Canu pipeline. The red dot is the selected draft assembly, and the green dot shows scaffolding results following Hi-C. Detailed assembly statistics are available in File S1.

Genome annotation

Annotations for *A. saccharum* resulted in 40,074 gene models of which 8,765 were monoexonics verified by the presence of a protein domain and start and stop codons. Transcriptome comparison based on 15,234 transcript loci supported 13,997 gene models. Functional annotation was applied via similarity search or gene family assignment for 35,304 models. *A. negundo* had 30,491 genes, 5,558 of which were monoexonic, and 16,168 transcript loci supported 14,682 of the *de novo* models. Functional annotations were determined for 27,077 of the models. (File S2). *A. saccharum* repeat content was at 64.4% while *A. negundo* was 58.6%.

Whole genome duplication and Acer synteny

Categorization of putative paralogs revealed a higher percentage of each type in *A. saccharum* relative to *A. negundo*. Plots of Ks distribution for WGD genes in syntenic regions show a single clear peak at a Ks range consistent with the core eudicot WGT reported in other species using the same pipeline (Figure 3a). *A. yangbiense* does not have a recent WGD, and when compared to *A. saccharum* which had an additional small recent peak, further investigation identified small blocks of collinearity, a minimum of five genes in palindromic or tandem arrangements. These blocks are predominantly located on a few scaffolds and are not reflective of the general distribution typical of WGD (File S3). Macrosynteny analysis found that *A. negundo* and *A. yangbiense* are syntenic, while comparisons between each of these and *A. saccharum* revealed a large-scale translocation where two chromosomes from *A. negundo*, including the largest, are split with sections exchanged to form two different chromosomes in *A. saccharum* (Figure 3b, Figure S4).

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

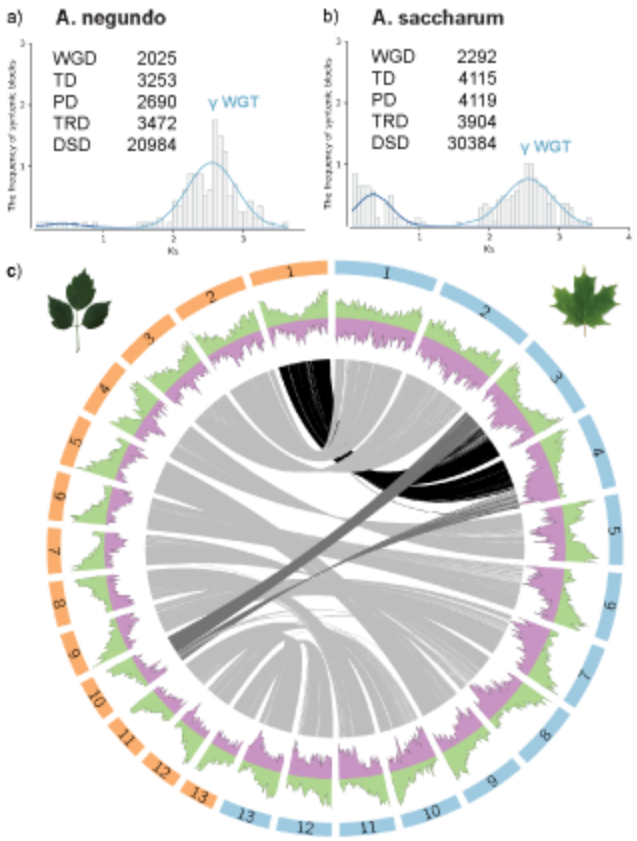


Figure 3. Ks distribution for WGD synteny blocks with a summary of duplication types in (a) *A. negundo* and (b) *A. saccharum*. Abbreviations for categories of duplication: WGD, whole genome duplication; TD, tandem duplication; PD, proximal duplication; TRD, transposed duplication; DSD, dispersed duplication. (c) Circos plot of the thirteen chromosomes ordered largest to smallest for *A. negundo* (orange bars) and *A. saccharum* (blue bars) with distributions of gene density (green) and transposable element frequency (purple). Syntenic regions are linked in gray with darker shades to visually highlight larger recombinations.

Expression analysis of A. saccharum aluminum and calcium treatments

The final annotated genome for *A. saccharum* served as a reference for the expression study. In total, there were 245 unique differentially expressed genes with 181 informatively described by sequence similarity descriptors. Of those with no similarity match, four were completely novel

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with no identifiable protein domain. Initial analysis produced six up and nine downregulated genes comparing the aluminum to calcium treatments, and the other pairwise comparisons had similarly small totals. Clustering of the expression results showed season had a strong effect (Figure S3), so the analysis was repeated for each season individually to remove this variable from treatment comparisons. For brevity, abbreviations are used according to the following definitions: All, across seasons; Fa, fall; Sp, spring; Al, aluminium; Ca, calcium; and Un, unamended. FaAl to FaCa had 26 upregulated and 17 down, FaAl to FaUn had 7 up and 12 down, and FaUn to FaCa had 28 up and 27 down. SpAl to SpCa had 39 up and 19 down, SpAl to SpUn had the greatest number with 33 up and 41 down, and SpUn to SpCa had 28 up and 11 down. (Figure 4, File S4)

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

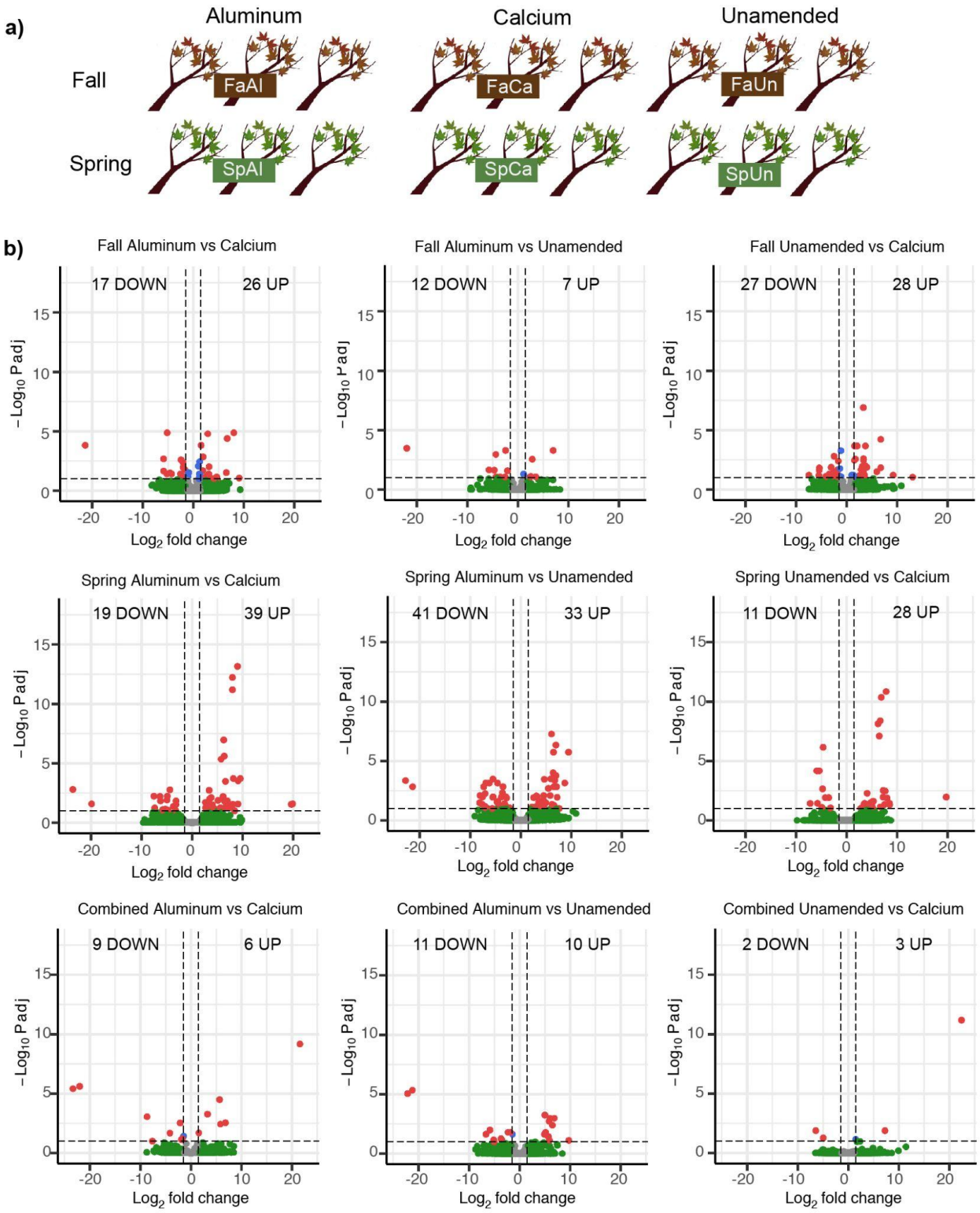


Figure 4. a) Differential expression study design showing number of samples collected in fall and spring from treatment plots at the Hubbard Brook Experimental Forest, Nutrient Perturbation

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

study. b) Differentially expressed genes (up and downregulated) for each treatment and season comparison. Charts display both significance and relative expression denoted as log-fold change. Dotted lines indicate thresholds of significance (0.1 p-adjusted, 1.5 $\log_2$ fold change).

There were only two instances where the same gene was found in both separate seasonal analyses. Transcription factor-like E2FE was upregulated in both unamended and aluminum treatments in fall, spring, and all. It had the most isoforms expressed, and was also the most highly upregulated for fall (8-fold). E2FE represses endoreduplication, reducing this type of growth in response to stress (Hendrix et al., 2018). The second instance was disease resistance protein At4g27190-like, expressed primarily in SpAl, but also FaUn and FaCa. Another At4g27190 isoform was upregulated in SpCa and SpAl, but downregulated in SpUn.

GO enrichment results for fall showed sugar and carbohydrate transmembrane transporter activity upregulated in the FaUn compared to FaCa (Table S3). Several sugar transporters, SWEETs and ERD6-like, were seen in both seasons, though more often in fall. There were two SWEET15 (tandem duplicates), one upregulated in FaUn to FaCa, and SWEET15-like upregulated in SpUn relative to both SpCa and SpAl (7.6-fold). SWEET2a was upregulated in FaAl to FaCa. Three different ERD6 were upregulated in FaUn with a fourth in SpAl. Triterpenoid biosynthetic process was also enriched FaUn to FaCa, supported by two downregulated beta-amyrin synthase genes, one of which was significantly downregulated even further in FaAl to FaUn (5.4-fold). Other strongly differentiated genes included accelerated cell death 6-like (ACD6) upregulated in FaCa (5-fold) relative to FaAl. It is associated with flavin-containing monooxygenase 1 (FMO1), the most highly upregulated DEG in FaUn (22-fold) and FaCa, both relative to FaAl.

In spring, DEGs upregulated in SpUn compared to SpAl were enriched in processes related to biotic and abiotic stress, including gene ontology terms for defense response, and acid, oxygen-containing, and antimicrobial response. There were six disease resistance genes, all largely in SpUn and SpCa compared to two in fall. Also present in large numbers, were

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

serine/threonine kinases, including LRR receptor-like, with seven out of twelve of these in the FaUn. Heat shock proteins were also common, though more equally split between the two seasons and shifted toward unamended or aluminum. There were two copies of ITN1, involved in salicylic acid signaling, one of which was the most upregulated DEG in SpCa to SpAl (23.6-fold). In addition to direct stress response, there was an interesting increase of expression in three Holliday junction resolvases, and a lamin-like gene, which was the second highest DEG in these comparisons (19.7-fold). It can play a role in nuclear membrane integrity, chromatin organization, and gene expression (Hu et al., 2019).

Metal tolerance via ligation, sequestration, and transport

In SpAl, upregulation of both metallothionein-like 3 (6.6-fold), and aluminum-activated malate transporter (ALMT) 10 (4.5-fold) was observed. An EH domain-containing gene (4.6-fold) is involved in endocytosis, vesicle transport, and signal transduction (Naslavsky & Caplan, 2005), and several cytoskeletal related genes were seen such as kinesin in fall (6.8-fold), and myosin-6 in spring (8-fold).

In addition to the sugar transporters, ABC transporters were present as multiple isoforms in FaUn, and one was the most upregulated DEG in FaUn compared to FaCa (13-fold). Two were downregulated in FaAl compared to FaCa, but still elevated in FaUn, and were from family C which contains pumps for glutathione S-conjugates, which have been shown to remove cadmium in *Arabidopsis* (Tommasini et al., 1998). Additional transport included upregulation of ATPases across seasons, cyclic and mechanosensitive ion channels in spring and fall, respectively, and a K+/H antiporter upregulated in FaAl.

Calcium dependent proteins

Three calcium-transporting ATPases were upregulated in either FaCa or FaUn along with an unspecified plasma membrane ATPase. Ca-based external signal relay mechanisms upregulated in SpCa included a cyclic nucleotide-gated ion channel, G-type lectin S-receptor like serine/threonine-protein kinases, and a glutamate receptor (5.5-fold) (Sun et al., 2013; Toyota et

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

al., 2018). A calcium-dependent protein kinase (4.5-fold) was upregulated in FaUn. Calmodulin-binding proteins (60d at 5-fold) were up-regulated in both SpAl and SpCa, and are associated with salicylic acid synthesis and immunity in *Arabidopsis* (Li et al., 2021).

Reactive oxygen species response

Antioxidant and redox genes included glutathione peroxidase, upregulated in FaUn, and glutathione S-transferase, upregulated in both FaCa and SpAl. Two cytosolic sulfotransferases and six thioredoxin genes were upregulated in SpAl. Other redox related genes include a highly expressed epoxide hydrolase (9-fold) and a carotenoid cleavage dioxygenase (4.6-fold) in FaAl; and peroxidase, aldehyde dehydrogenase (6.7-fold), and germin in SpAl. Ascorbate-dependant oxidoreductase SRG1 was upregulated in SpUn, cytochrome P450 71D11 (a monooxygenase) in SpAl, and zinc finger (C2H2 type) in SpUn.

Cell wall and membrane integrity

MDIS1-interacting receptor kinase, upregulated in FaUn relative to FaAl, and expansin, FaUn to FaCa, are associated with the cell wall. Two ADP-ribosylation factor GTPase-activating, one of which is the most highly upregulated gene in spring (19.9 fold-change SpAl to SpCa), assists with cell signaling by recruiting cargo-sorting coat proteins at the membrane and regulating lipid composition in support of development and defense (Donaldson & Jackson, 2011).

Hormone crosstalk

Auxin was indicated by three WAT genes (seen in each season, greatest at 7-fold in SpAl to SpCa) and indole-3-acetic acid (highest of AllAl to AllUn at 9.7-fold). Two ethylene synthesis genes were present: 1-aminocyclopropane-1-carboxylate oxidase, FaAl to FaCa, and methylthioribose kinase, SpAl to SpCa. Ent-kaurenoic acid oxidase is related to brassinosteroid homeostasis and gibberellin biosynthesis (Helliwell et al., 2001) and was present in SpAl to SpUn (7-fold), and obtusifoliol 14α-demethylase, which mediates brassinosteroid synthesis, was upregulated SpUn to SpCa (Xia et al., 2015). Jasmonic acid activity was upregulated in FaCa, negative regulation of cytokinin was up in AllAl to AllUn, and ABA-induced HVA22 which

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

inhibits gibberellin and is possibly involved in vesicular traffic, was in SpAl to SpCa (8-fold) as well as AllAl to AllCa.

Gene family evolution with expression study integration

By leveraging 22 high-quality plant genomes, gene family dynamics between *A. negundo* and *A. saccharum* revealed distinct characteristics. Comparisons among the plant proteomes resulted in 20,234 orthogroups with a mean size of 26.4 genes. Of these, 4,262 were shared by all species, and 79 were single-copy. 88.5% of 603,640 genes were contained in orthogroups, and 0.7% were species-specific. All species had at least 80% of their genes contained in orthogroups with the exception of *Ginkgo biloba, Nymphaea colorata*, and *Oryza sativa* (Table S4).

All three *Acer* shared 11,156 orthogroups. *A. negundo* and *A. saccharum* had the largest overlap, *A. saccharum* and *A. yangbiense* had the second largest overlap, and *A. saccharum* had the most unshared groups. (Figure 5b). Comparing *Acer* against the other woody angiosperms (*B. pendula, C. papaya, C. clementina, C. sinensis, E. grandis, J. hindsii, J. regia, P. vera, P. trichocarpa, P. persica, Q. lobata, Q. robur, T. grandis*), 728 *Acer* orthogroups were expanded, 14 were contracted, 1992 were novel, and *Acer* was estimated to be missing from 202 orthogroups. To clarify, these may not be fully absent, but didn't have representation in orthogroups including those that were more lineage specific. Comparing the Sapindales to the other trees resulted in 340 expanded, 4 contracted, 2788 novel, and 160 missing. Dynamics in common between the two *Acer* included 270 expanded groups, 1 contracted, 0 novel, and 247 missing (File S5).

Acer families expanded among the woody angiosperms were enriched for RNA modification, DNA replication strand elongation, and processes of organic cyclic compound, cellular aromatic compound, and heterocycle metabolisms. Novel genes were highly enriched for cell periphery localization and marginally for sesquiterpenoid and triterpenoid biosynthesis (Figure 5a, Table S5). When focusing on absent gene families, none were found to be missing exclusively in Sapindales, which includes the two *Citrus* species and *Pistacia vera*. A total of two families were

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

absent in all *Acer*, phosphatidylcholine transfer protein-like and cellulose synthase interactive 3, but were present in all other species.

Several interesting gene families more novel to *Acer* overlap with the HBEF DEGs. There are twenty orthogroups associated with disease resistance At4g27190, two seen as DEGs, and the specific families containing these DEGs are larger for *A. saccharum* with one novel to the *Acer*. Two additional non-DEG families are rapidly expanding in *A. saccharum*, with one of these also expanding in *A. negundo*, but both contracting in *A. yangbiense*. In fall, another more novel DEG ACD6-like belongs to a family with limited species membership (Sapindales and *V. vinifera*). Compared to *Arabidopsis* ACD6, which has two 3-repeating ankyrin domains, the *A. saccharum* ACD6-like had varying ankyrin positions, as do other members of this family. A third example, seen only in spring, acetyl-coenzyme A synthetase (ACS) is a member of a novel *Acer* family consisting of 1 *A. negundo*, 6 *A. saccharum*, and 18 *A. yangbiense*. Comparison with *A. negundo* and ACS isoforms reveals the *A. saccharum* gene contains a longer ACL domain, with only ~67% query coverage and ~43% percent identity to *A. negundo* which is much more similar to other ACS (87-90% length; ~88% identity).

Within the broader set of species, *A. saccharum* gene families were characterized by more expansion, with 1827 expanded, 18 contracted, 127 novel, and 511 absent, and more rapidly expanding, with 99 compared to 18 contracting. *A. negundo* had 1068 expanded, 23 contracted, 89 novel, and 558 absent orthogroups. Rapidly contracting families were greater in this species with 52 compared to 26 rapidly expanding (Figure 5, File S5, File S6).

A. saccharum gene families

Functional enrichment of *A. saccharum* in the full species comparison revealed that expansions were processes of ncRNA metabolism, RNA modification, organic cyclic compound metabolism, heterocycle metabolism, and intracellular membrane-bound organelle localization (Table S5). Almost half of the *A. saccharum* families had limited annotation information, due to either missing descriptors or uncharacterized protein matches. Relative to *Acer*, *A. saccharum's*

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

expanded families are enriched in a larger list of stress response associated functions that are fairly specific, including water deprivation, hypoxia, salinity, heat, cold, xenobiotic, nematode, karrikin, acid chemical, and hormone (File S13). Other significant processes include regulation of indolebutyric acid stimulus (auxin family), RNA splicing, chloroplast RNA processing, phospholipid translocation, brassinosteroid homeostasis, lignin synthesis, microsporogenesis, phenylpropanoid biosynthesis, cadmium ion transmembrane transport, cyclin-dependent serine/threonine kinase, and calcium-transporting ATPase. Rapidly expanding families are associated with various biotic and abiotic responses, such as fungal, salt stress, and xenobiotic response (File S6). Interesting genes include patatin-like 2, involved in membrane repair via removal of lipids modified by oxidation (Yang et al., 2012) and ALP1 negative regulation of polychrome group chromatin silencing (Liang et al., 2015). Those that overlap with HBEF DEGs include disease resistance At4g27190 and DSC1, FMO1, rapidly expanding SRG1, and rapidly contracting disease resistance At1g50180.

Compared to other *Acer*, contracted families are enriched for pollen wall assembly, extracellular matrix assembly and organization, chlorophyll binding, NADH dehydrogenase (ubiquinone) activity, DNA-directed 5'-3' RNA polymerase activity, programmed cell death, and myb-like transcription factors (File S11, File S12). Genes absent in *A. saccharum* but present in all other species total 28 (File S15), including red chlorophyll catabolite reductase (ACD2) and S-adenosyl-L-homocysteine hydrolase (HOG1), which is necessary to hydrolyze the by-product of the activity of S-adenosyl-L-methionine-dependent methyltransferase, and one of these was also absent. The apparent absence of HOG1 requires further investigation as mutants display a number of problematic phenotypes and variants display association with fiber length in *P. tomentosa* (Du et al., 2014).

A. negundo gene families

In the broad comparison of species, *A. negundo* expanded families were enriched in RNA modification, microgametogenesis, and metabolic processes of nucleobase-containing compound, organic cyclic compound, heterocycle, and cellular aromatic compound (Table S5).

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The small number of rapidly expanding families were by far mostly uncharacterized proteins or missing sequence similarity descriptions with only four out of 26 genes having a description, including glutathione-S-transferase, two disease-resistance proteins, At4g27190-like and At5g66900, a receptor-like 12, and additional functional descriptors such as E3 ubiquitin-protein ligase, LRR receptor-like ser/thr kinase, and more (File S6). Relative to other *Acer*, *A. negundo's* expanded families are enriched in a short list of specific stress response including UV, UV-B, radiation, bacterium, cadmium, metal ion, drug, chemical, and osmotic stress (File S9, File S10). Other processes include proanthocyanidin biosynthesis, lignin synthesis via cinnamyl-alcohol and sinapyl-alcohol dehydrogenase, starch metabolism and glucan catabolism, error-prone translesion synthesis, and other DNA damage response and repair. Processes related to reproduction were present, especially pollen development. For example, decreased size exclusion limit 1 (DSE1, aka aluminum tolerant 2 (ALT2)) is a transcription factor that regulates the size of molecules that can travel through plasmodesmata, as channel aperture is not static, changing in response to stress and decreasing during embryo development (Xu et al., 2012). DSE1 is single copy in all species with expansions only in *A. negundo*, *Q. rober*, and *T. grandis*.

Contractions relative to other *Acer* include transcription by RNA polymerase III, chloroplast RNA processing, and lignin biosynthetic processes (File S7, File S8). Rapidly contracted families were enriched in quercetin 3-O- and 7-O-glucosyltransferase activity (Table S6). Examples of contracted families include 7-ethoxycoumarin O-deethylase, which metabolizes a wide range of xenobiotics (Robineau et al., 1998), and disease resistance-like protein DSC1, both rapidly expanding families in *A. saccharum* (File S6). There are 23 genes absent in *A. negundo* that are present in all other species (File S14) . Several of these are curious, as they appear to be required components of important processes, such as AUGMIN subunits 2 and 7 that help form a complex that plays a role in spindle microtubule generation (Tian & Kong, 2019), and cell division cycle 45-like, which is required for meiosis in *Arabidopsis* (Stevens et al., 2004). The absence of formamidopyrimidine-DNA glycosylase is also interesting as it is involved in base excision repair of DNA damage, a notable area of specific enrichments described below.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

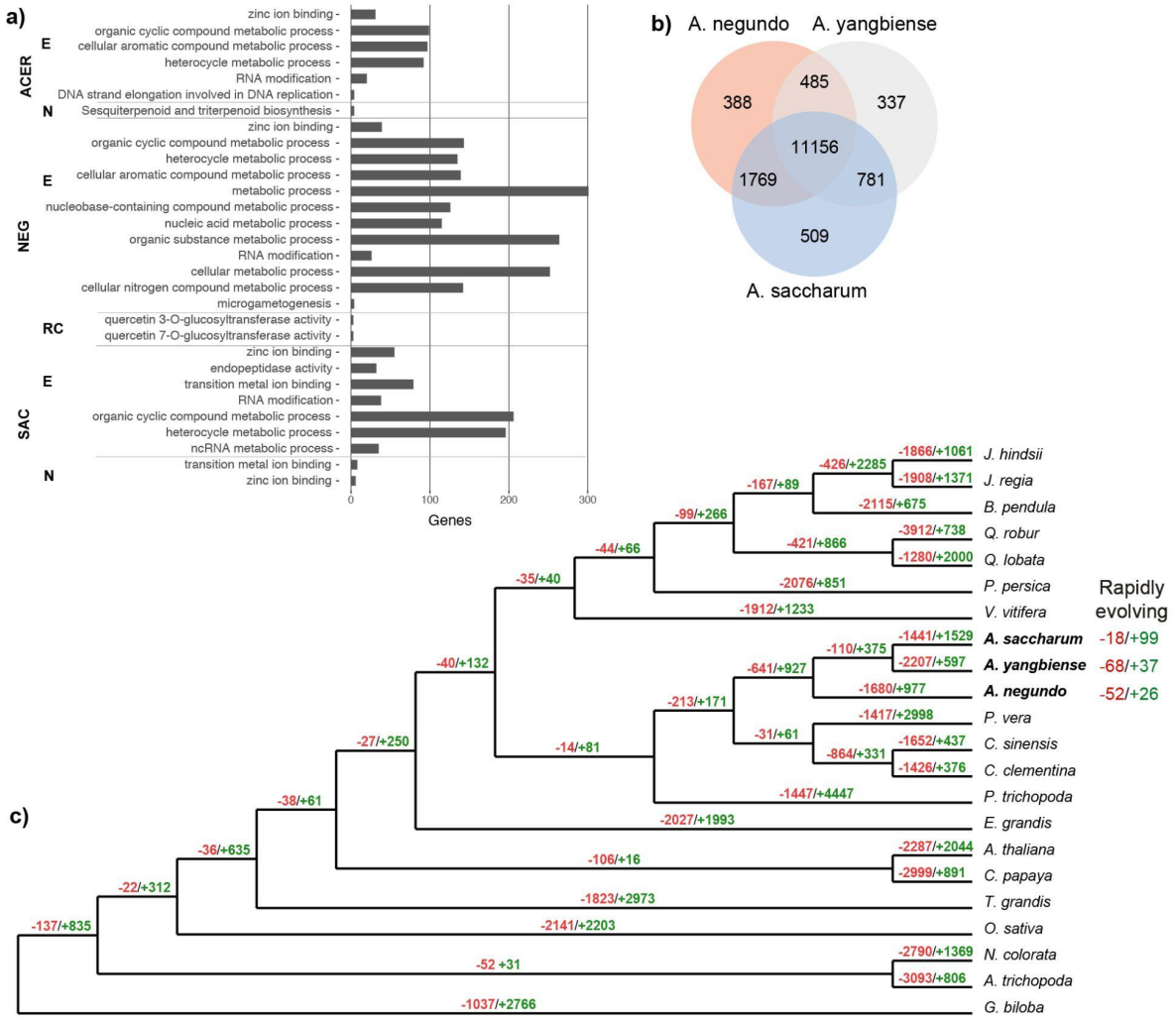


Figure 5. a) Gene ontology enrichments for *Acer* (all three species combined), *A. negundo*, and *A. saccharum*. Abbreviations for gene family dynamics: E, expanded; N, novel; RC, rapidly contracting. b) Total gene families, shared and unique, among the *Acer*. c) Reconstructed gene tree showing contracted gene families in red and expanded in green.

Discussion

Completion of the first chromosome-scale genomes for two North American maples brings Sapindales to a total of 15 available reference genomes. Sequenced members in this group contain citrus, mangos, pistachio, and poison ivy. Sapindaceae is morphologically diverse and is known for having opposite leaves, colorful fall foliage, and samaras. The maple genus *Acer* represents over 150 species, most native to Eastern Asia and a small number in Eastern North America (10)

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and Europe (22) (J. Li et al., 2019). Today, one out of five species are endangered in their native range (Crowley et al., 2020) and will continue to face challenges from both abiotic and biotic threats resulting from a rapidly changing climate.

At the trait level, both *A. saccharum* and *A. negundo* differ in distribution and tolerance to abiotic stressors. *A. negundo* grows quickly, is able to reproduce after only five years, and has a shorter lifespan of 60 years. It is moderately tolerant of a range of conditions and is widespread throughout North America. In contrast, *A. saccharum* has slow growth until release of canopy coverage, doesn't achieve reproductive maturity until 40 years, and is able to live 300 to 400 years. It requires high nutrient soils, prefers mesic environments, and is tolerant of cold, but not salinity. Its range crosses a more narrow latitudinal, but stronger elevational, gradient, whereas *A. negundo* tends to be more limited by elevation. *A. negundo* is considered an aggressive invasive in Europe, South Africa, and parts of Asia and North America, and rapidly colonizes and dominates disturbed habitats leading to loss of native species (CABI, 2021; Lamarque et al., 2015). The sequenced individual in this study is from the native range where *A. negundo* is highly plastic in growth, leaf unfurling, leaf mass area, maximum assimilation rate, as well as nitrogen content and photosynthetic efficiency (Lamarque et al., 2015). *A. negundo* also exhibits sexual dimorphism in photosynthetic rates, leaf size and allocation, and growth form, where males are more successful in dry environments due to enhanced stomatal sensitivity and females are found in more mesic environments (Dawson & Ehleringer, 1993). Invasive populations maximize growth in high light and nutrients, with reduced performance in deficient conditions. Even in optimal light, photosynthetic capacity and leaf nitrogen content remain low in *A. negundo*. Such disadvantageous traits were surprising, and studies attributed its plasticity in growth to morphology as increased leaf area allocation is a minimal investment allowing for adjustment to changing conditions (Lamarque et al., 2013, 2015; Porté et al., 2011). A reciprocal common garden study, examining invasive and native populations, found that in addition to plasticity, post-invasion genetic differentiation was a factor in later stages of invasion success (Lamarque et al., 2015). At least six varieties of *A. negundo* exist across the native range as well, based on morphological characteristics (Rosario, 1988).

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The maple genomes reveal support for their contrasting life histories. While both are small in size and diploid, *A. saccharum* is 42% larger, containing 38% more gene duplications, many very recent, and twice as many transposable elements. Gene families tend to be larger, more diverged, and undergoing rapid expansion in *A. saccharum* compared to *A. yangbiense* or *A. negundo*, which is characterized by contracting families, particularly among those rapidly evolving.
The *A. negundo* reference genome is a small diploid with high heterozygosity and lower repeat (LTR) content. Synteny with *A. yangbiense* indicates there isn't much large-scale structural variation and supports its reduced character (Figure S4). Invasive plant species are often associated with smaller genomes. Traits such as fast growth rate, germination time, stomatal responsiveness, and dispersal ability are cell size or division rate dependent (Pyšek et al., 2018). The greater surface area to volume ratio of small cells, derived from small genomes, reduces the metabolic and signaling requirements, but does not preclude additional growth or activation, thus extending the range of capacity, or plasticity, for a wider set of traits (Roddy et al., 2019; Suda et al., 2015). The adaptive potential conferred by polyploidism can also be leveraged for invasion, and while polyploidism has not been documented in the native species, we cannot rule this out as a factor in *A. negundo's* invasion success. The *A. negundo* genome also has a lower GC% relative to *A. saccharum*, and within phylogenetically close relatives, lower GC% is typically associated with smaller genomes. GC content indicates DNA base composition in terms of guanine and cytosine, which have different biochemical properties in base pair and higher order structure, nucleotide synthesis requirements, and methylation and mutation rates (Šmarda et al., 2014). Earlier work based on low coverage sequencing postulated they were at higher percentages, *A. negundo* in particular (Contreras & Shearer, 2018; Staton et al., 2015), but whole genome sequencing supports their place at the lower range among angiosperm plants (Trávníček et al., 2019). Higher relative GC% is associated with increased size, often as a result of increased LTR content and adaptation to colder climates or greater annual temperature fluctuations (Veleba et al., 2017). Extreme cold and/or desert environments are also conditions where lower metabolic rates might be selected for, accompanied by a larger genome (Roddy et al., 2019). While *A. saccharum* does not occupy extreme regions, it is found at higher elevations, and the genome is

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

enriched in cold and water-related response relative to *A. negundo*. In spite of *A. saccharum's* larger size, GC%, and functional enrichments, it remains challenged by abiotic stress in its current range. It is similar in size and GC% to *A. yangbiense*, an endangered species found at high elevations in a very limited region of Yunnan Province (J. Yang et al., 2019). If larger genome size presupposes increased metabolic, transport, and nutrient demands (Pellicer et al., 2018) it is possible *A. saccharum's* susceptibility to nutritional deficiencies, and calcium in particular, may be due to the extra burden of a larger genome and the various mechanisms within. The decrease in soil nutrient availability in its native range over the past decades is at odds with resources necessary to tolerate stressors brought on by a changing environment. From the foundational differences between *A. negundo* and *A. saccharum* in morphological, physiological, and genomic characteristics, we extend to the integration of gene family dynamics and expression data to further illuminate the contrasting strategies of competition versus resistance seen in these species.

Herbivory, Reproduction, Light, DNA Damage

Compared to the other *Acer*, *A. negundo's* smaller set of response related functional enrichment is mainly limited to UV, bacteria, metal, cadmium, chemical, and osmotic response rather than the more extensive set seen in *A. saccharum*. Expression of these more specific response-related gene families could protect *A. negundo* from pests and pollutants, providing benefits conducive to life in an urban environment. A high portion of proanthocyanidin synthesis related genes were observed, which are precursors to condensed tannins that protect against herbivory, bacteria and fungal pathogens, and encroaching of neighboring plants (He et al., 2008). All four families were relatively novel with absence in all but one or two other species. Proanthocyanidins also have antioxidant and radical scavenging functions, so it would be interesting to see if these were differentially expressed in abiotic stress conditions similar to other flavonoids in *A. saccharum*. They have been compared to lignins in terms of pathogen defense mechanisms (Stafford, 1988), and some enrichment of expanded lignin precursor monolignol genes does exist, also assigned to relatively novel gene families.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Successful reproductive strategies are characteristic of invasive plants. Of the 62 reproduction related gene families expanded within *Acer*, over a third were also significantly expanded across the full set of species, with one third as pollen development, two thirds as fruit development, and a few related to embryonic development and seed germination. Pollen development related gene families included an ortholog of transcription factor DUO1, a key step in male germline specification that has been conserved yet heavily diverged in the evolution of male gametes (Higo et al., 2018). This expanded family is novel to the three *Acer* and contains multiple copies for the two dioecious species, with only one for *A. saccharum*. Within fruit development, LEUNIG, APETALA2-like (AP2), and another AP2 domain-containing family were significantly expanded, with yet a third AP2 domain-containing expanded within *Acer* for *A. negundo* but missing in *A. saccharum*. AP2 and LEUNIG are corepressors of homeobox gene Agamous, and in the absence of this repression, sepals and petals become stamens and carpels, as studied in *Arabidopsis* (Conner & Liu, 2000). The AP2 family is large, containing some members associated with germination, growth, and stress response (Krizek, 2015; Shu et al., 2018).

Both species are enriched in gene families with DNA damage recognition and repair functionality, but *A. negundo* has a greater number, largely categorized as either error-prone translesion repair or meiosis-related. Within these are additional enrichments in regulation of leaf, seed, flower development, and post-embryonic structures, response to abiotic and biotic stress, and growth such as cell division and endoreduplication (Figure 6), all of which can be altered in response to DNA damage that can be UV, genotoxic, or oxidative in nature. UV damage can activate translesion synthesis, an error-prone repair mechanism designed to quickly eliminate lesions that might otherwise stall replication and lead to double-stranded breaks (Sakamoto, 2019). The mutation rate from this type of repair is much higher than others (Kunkel, 2000) and depends on additional repair for correction. *A. negundo* has three families in this category expanded relative to *Acer,* two of which are significantly expanded among the full comparison of species. Metal toxicity, including Al^{3+}, also causes double stranded breaks resulting in inhibition of cell cycle progression and cessation of growth, notably in the root

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Zhang et al., 2018). The ATR gene family partially responsible for that type of cell cycle inhibition is actually larger in *A. saccharum.* This could contribute to decreased aluminum tolerance if expression is likewise increased. Too much DNA damage can initiate either programmed cell death or continuing growth without replication via endoreduplication, which is a way for plants to enlarge size via increased nuclear content without the usual cell division step, thus preventing the spread of heavily damaged DNA to new cells (Nisa et al., 2019).

Figure 6. *A. negundo* gene families with ontology related to DNA damage and repair, and secondary enrichments categorized by color. Circles with multiple colors indicate multiple ontology assignments. Lines indicated known or predicted interactions, or other association via text-mining, co-expression, or protein homology.

In addition to UV-related DNA repair, there is also potential adaptation for photooxidative stress. The early light induced protein 1 (ELIP) gene family stood out as significantly expanded in *A. negundo* relative to all other species, with twice as many genes compared to other *Acer*. Research has focused on ELIP's protection against high light (Huang et al., 2019), but it also

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

regulates seed germination in response to environmental factors (Rizza et al., 2011). ELIPs are expressed in roots, and in response to some abiotic stressors including nitrogen in *Populus* (Luo et al., 2015), indicating additional protective functions. Capacity for regulation of photosynthesis and DNA repair mechanisms, within a largely streamlined genome, could represent investments with large-scale effects, altering growth according to resources or stress as described for *A. negundo* in various nutrient and light conditions.

Integration of A. saccharum expression studies

In contrast to *A. negundo's* plasticity and invasion success, many *A. saccharum* populations are facing maple decline due to limited tolerance of abiotic stress. Integration of gene expression sampling with gene family dynamics highlighted the genetic factors that may influence maple decline and adaptation. HBEF NuPert saplings grown in native ecological and environmental conditions within replicated long-term soil treatments provided an opportunity to examine differences in gene expression related to soil acidity, calcium deficiency, and increased aluminum availability. Previous experiments involving aluminum are typically short-term time series measurements of more immediate responses in very young seedlings maintained in controlled environments (Cardoso et al., 2019; Liu et al., 2009). Examination of these trees on the landscape allowed us to study the effects of long-term nutrient stress in conjunction with other life-history processes across seasons.

Calcium is key for signaling

A. saccharum requires nutrient-rich soils, and calcium is the common limiting element underlying maple decline. Increased acid deposition in regions of low base cation concentration limits calcium availability (Long et al., 2019; Schaberg et al., 2001, 2006; Sullivan et al., 2013), but trees improve long-term with the addition of calcium (Moore & Ouimet, 2021). Additional nutrient imbalances, such as magnesium, phosphorus, and potassium and nitrogen deficiency, as well as high aluminum and manganese concentrations, climatic effects, and biotic stressors such as defoliation, are compounding factors, culminating in widespread decline that has been studied throughout most of *A. saccharum's* native range (Bal et al., 2015). The effects of calcium can be

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

studied using the HBEF NuPert plots, where calcium amendment was designed to recreate previous levels of soil availability. Annual amendments of $CaCl_2$ from 1995-1997 followed by applications of wollastonite in 1999 and 2015 (Table S7) resulted in a 50% increase of calcium in foliar tissues of maples, while also decreasing aluminum concentrations non-significantly. Trees in calcium plots devoted more carbon to growth than storage, were better able to flush after a late spring frost, produced more flowers, and increased seed germination (Halman et al., 2013). Calcium-dependent signaling mechanisms and abiotic stress are well studied in model species with recent work on homologous pathways in poplar and a wide variety of gymnosperm and angiosperm trees (Estravis-Barcala et al., 2020). Calcium transport and binding can activate and regulate primary components of stress response systems, including internalization of external signals, modulation of cytosolic levels for transient signaling, and lead to initiation and perpetuation of ROS responses, enzyme activation, and pathways of hormone modulation and secondary metabolism. Calcium and salicylic acid in combination can improve aluminum tolerance by increasing exudates, preventing root accumulation, decreasing root growth inhibition, and stimulating antioxidants as seen in soybean (Lan et al., 2016). The wide variety of processes, including genes such as Ca-transporting ATPases, calcium-dependent protein kinases, and calmodulin-binding proteins, were all important in the differential expression response and also observed in the gene family dynamics.

Calcium is necessary for primary and secondary metabolic processes

Seasonal differences between the differentially expressed genes were striking, revealing the effects of calcium treatments on the larger activities at work during the time of sampling. Additional calcium facilitates natural life cycle processes, while unamended treatments show an increase in stress response perhaps reflective of the additional burden of calcium deficiency combined with external stressors. In fall calcium treatments, genes involved in leaf senescence and remobilization of nutrients were highly expressed. The lack of other differential expression in this season tied to clear processes made it all the more interesting. ACD6 is active in both natural and stress response senescence and is calcium signaling dependent (Jasinski et al., 2021; Zhu et al., 2021). Given the context of where these genes were expressed, considering treatments

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and other neighboring DEGs such as calcium-dependent protein kinase, the likelihood seems to be the occurrence of natural leaf senescence in the presence of adequate calcium. The ACD6 isoform expressed here is quite varied from the *Arabidopsis* form, and its gene family membership is limited to Sapindales. The FMO1 highly expressed in the same comparative context is associated with ACD6 and senescence in other studies, and is also a new variant very expanded in *A. saccharum*. FMO1 is a flavonoid antioxidant that appears in other aluminum tolerance studies and is associated with auxin regulation in addition to cell death, reflecting the multipurpose nature of these types of redox genes (Schlaich, 2007). Another senescence gene is absent in *A. saccharum*, red chlorophyll catabolite reductase (RCCR aka ACD2), a non-green chlorophyll degradation gene (Chen et al., 2019). The exact role of RCCR is unclear (Jockusch & Kräutler, 2020), but it appears possible that it provides additional fitness advantages beyond chlorophyll degradation, such as increased tolerance to infection (Mach et al., 2001). This, combined with the ACD6 susceptibility to calcium deficiency, potentially creates complications for nutrient reabsorption or other senescence related activities in fall that could affect spring growth, development, and ability to respond to stress.

Characterization of aluminum response in A. saccharum stem tissue

Aluminum amendment in NuPert was designed to create the effect of future acidification. Al and Ca compete for uptake by the root, so Al amendment and low Ca/Al molar ratios essentially increase Ca deficient conditions. This is exacerbated by acid deposition which contributes to leaching of base cations and increases the availability of more toxic forms of Al. The focus of current literature on *A. saccharum* mortality, health, and regeneration is calcium depletion, with improvements seen upon calcium amendment (Cleavitt et al., 2017; Huggett et al., 2007). Previous studies examining physiological effects of aluminum versus calcium with a multi-tissue study on mature *A. saccharum* found foliar levels of Al did not vary much by treatment and all were well below thresholds of toxicity. Within this tissue, concentrations of Al dropped slightly in Ca treatments, while Al treatments caused Ca to drop more strongly in dominant than non-dominant trees (Halman et al., 2015). The main effects noted in aluminum treated trees were moderate root damage and foliar antioxidant activity (Halman et al., 2013). Some studies from

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

other locations have also linked foliar Al and Ca levels (Schaberg et al., 2006) and stem Al levels to branch dieback as seen in maple decline (Mohamed et al., 1997). Within the differential expression analysis we can detect levels of stress response between the unamended and aluminum treatments, as well as signs related to improved functionality with calcium amendment.

Aluminum targets the cell wall, plasma membranes, DNA, RNA, and proteins. Accumulation occurs primarily in the root, binding to pectin in the cell wall, increasing its rigidity and decreasing permeability. Downregulation of cellulose and upregulation of callose, oxidative stress, hormones, and DNA damage signaling all result in root growth inhibition (Sade et al., 2016). The plasma membrane is another frequent target, where Al disrupts membrane potential and membrane-bound solute transporters, affecting symplastic and apoplastic concentrations (Kar et al., 2021). Damage within proteins and DNA reduces enzymatic activity and can trigger cell cycle checkpoints. Plants can be susceptible, resistant, or tolerant of aluminum, and have various abilities in avoidance or accumulation related to the different targets.

Genes specifically associated with aluminum resistance or tolerance were not present in the stem tissue as they are typically expressed in the root, but upregulated genes related to metal transport and sequestration are likely involved in aluminum remediation. Metallothionein 3 was highly expressed in AllAl, though unevenly distributed throughout the replicates. Similar to other metallothioneins, it works via ROS scavenging and metal ion homeostasis via chelation, while potential transport and vacuolization capabilities are unknown (Hasan et al., 2017). In salinity-tolerant *Oryza sativa*, this gene responded to cadmium, salinity, and oxidative stress (Mekawy et al., 2018). Expression of growth inhibitors was another interesting finding. ACC oxidase and IAA-amido synthetase are known to be interacting root growth inhibitors, and if expression of these is elevated in the root as well, this would result in decreased tolerance in *A. saccharum* regardless of other response mechanisms. Genes functionally adjacent to aluminum resistance were present, such as CAMTA4. Its role in regards to aluminum is unknown, though it is responsive to cold and stress-related hormones (Kidokoro et al., 2017), and CAMTA2 is a

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

positive regulator of ALMT1 (Tokizawa et al., 2015). ALMTs are anion channels involved in a wide variety of processes, and although ALMT1 is known to be expressed in roots in response to aluminum (Hoekenga et al., 2006), most ALMT are not thought to be involved in aluminum tolerance (Liu & Zhou, 2018). Recent characterization of ALMT10, upregulated in SpAl in a pattern similar to metallothionein, proposes involvement in homeostasis of Cl^- efflux and $NO3^-$ assimilation, induced by water deficit (Racero & J, 2020). ALMT9 is significantly expanded in *A. saccharum*, and though not a stem tissue DEG, it is thought to be a vacuolar malate channel involved in guard cell regulation (De Angeli et al., 2013).

Most of *A. saccharum's* aluminum resistance families are modestly sized, with the exception of PEP carboxylase and a large family of mixed MATEs, with members likely similar to homologs studied in *P. trichocarpa* and *Arabidopsis* (N. Li et al., 2017). Overexpression of either of these increases efflux of organic acids at the root (Begum et al., 2009). The metallothionein, CAMTA4, and growth inhibitors ACC oxidase and IAA-amido synthetase were also expanded. This is in line with broader gene family comparisons which revealed fewer expanded aluminum resistance and tolerance families in some, but not all, species reported as low-tolerance phenotypes, including *C. papaya*, *P. persica*, and *V. vinifera* (Figure 7, File S16; Jaillon et al., 2007; Ming et al., 2008; Verde et al., 2017). Aluminum accumulator *A. trichopoda* was also contracted in these families, but accumulators specialize in Al sequestering and it is likely that other genes are involved (Jansen et al., 2002). Expanded families were seen more frequently in high-tolerance species such as *P. vera*, *P. trichocarpa*, *T. grandis*, *Q. lobata*, and *E. grandis*, though the mix of families varies (Figure 7; Q. Li et al., 2015; Sork et al., 2016; Tuskan et al., 2006; Zeng et al., 2019; Zhao et al., 2019). *A. negundo* had expansions related to tolerance that likely contribute to its ability to manage multiple agents of toxicity. Gene family size is only a partial view, for example, modification to promoter or intronic regions of Al resistance genes have been shown to increase or decrease the expression of organic acid efflux in barley, wheat, rice, and sorghum (Pereira & Ryan, 2019). Antioxidants and redox genes are also important factors of tolerance due to the increase in oxidative stress caused by internalized aluminum. It is

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

interesting that many of these DEGs are also from expanded families, in particular the significantly expanded thioredoxins H3 and YLS8 and SRG1.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 7. Orthogroup sizes for aluminum tolerance gene families are presented by species. Families were selected for inclusion based on documented aluminum tolerance and/or presence in the HBEF RNA-Seq differential expression results. Color represents the proportion of gene membership per species, with darker purple equating to more contracted families relative to the median, and dark green indicating expansion. (H) Family contains HBEF differentially expressed gene; (E) Expanded in *A. saccharum*; (C) Contracting; (M) Missing; (N) Novel; (*) Rapidly expanding; Categorization of tolerance is according to literature describing aluminum stress phenotypes. The undetermined category contains species where tolerance to aluminum or acidic soils has not been reported. [1]*B. pendula* is undetermined due to high variability in tolerance by genotype.

Contribution of flavonoids and specialist metabolites as antioxidants

Aluminum derived oxidative stress is a factor in inhibition of cell growth, and an early signal of aluminum toxicity (Yamamoto et al., 2003). Reactive oxygen species (ROS) are generated as a result of normal cellular processes, and reactive oxygen can take different forms, with varying toxicities. ROS is reduced by a variety of potential "scavengers" to prevent accumulation of levels that are damaging to lipids, proteins, DNA, and RNA, leading to cell death. ROS production and redox activity vary according to the primary processes of each cellular compartment (i.e., metabolism in mitochondria), and also in response to different stressors or stress combinations (i.e., drought and heat), so it is speculated that cells develop complex ROS signatures or hotspots that influence signal transduction and metabolic regulation in nuanced ways (Castro et al., 2021; Choudhury et al., 2017).

Transcriptomic studies of cadmium accumulators *Salix integra* and *Populus x canadensis* 'Neva' both expressed superoxide dismutase (SOD), with glutathione pathway genes and peroxidases, respectively (X. Li et al., 2021; Shi et al., 2016). Nickel stress in resistant versus susceptible genotypes of *Betula papyrifera* found GST and TRX in resistant trees (Theriault et al., 2016), and *A. rubrum*, a nickel avoider, had root-level expression of SOD, but downregulated GST (Nkongolo et al., 2018). In comparisons of aluminum treated *Citrus* root expression, peroxidases

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and germin-like proteins were upregulated. The HBEF NuPert trees in aluminum plots were under oxidative stress that was at relatively low levels in controls. Halman et al. (2013) found elevated glutathione reductase in aluminum and ascorbate peroxidase higher in control and significantly so in aluminum. Similar to the other species, this transcriptomic study found antioxidants (TRX, two peroxidases, and a germin-like) all upregulated in response to the aluminum treatment, implying significant stress response activation. Aluminum increases peroxidation of lipids in membranes and produces H2O2 which can participate in retrograde signalling, regulating expression of additional genes (Castro et al., 2021). The expression in stem tissue implies that aluminum has translocated from the root to other tissues where it is causing peroxidation of membranes. Glutathione system members, which can also act as chelators, are more broadly seen in both unamended and aluminum, with one actually downregulated in Al similar to *A. rubrum*.

In calcium treatments, there are alternative forms of oxidative reduction, perhaps associated with different processes when sufficient levels of calcium are maintained. NADPH-dependent 2-alkenal reductase, one of two DEGs seen in *A. saccharum's* oxidative stress gene family enrichment relative to *Acer*, is associated with mediating photooxidative injury and improved photosynthesis, nutrient use efficiency, and biomass (Mano et al., 2005; Wang et al., 2021). It was very highly upregulated in AllUn and AllCa. There is also a variant zinc finger cysteine-2/histidine-2-type transcription factor. These are general transcription factors, but in *P. euphratica* stem tissue, it promoted the expression of an ascorbate peroxidase to scavenge ROS, which resulted in greater freezing tolerance while maintaining growth (He et al., 2019). Here, it was expressed in SpUn and to a lesser degree in SpCa, but was very low in aluminum, and is perhaps another example of calcium dependency, in this instance correlated with ROS stress response.

There are many multifunctional specialist metabolites, such as flavonoids, that assist with general membrane stability through ROS homeostasis, hormone crosstalk, regulation of stress response transcription, and protein modification (Arora et al., 2000). The cumulative effect of

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

their oxidative capacity via direct and indirect methods likely has a significant effect (Bartwal et al., 2013) in any acclimation these trees experience due to long-term treatments. More ROS response is found in spring samples as fits with the trend of increased overall activity in spring. The extent and complexity of ROS response can also be seen in the *A. saccharum* DEGs and in the variety gene families expanded and contracted in both *A. saccharum* and *A. negundo* relative to each other. Both *Acer* have expanded redox gene families, but *A. saccharum* has more, and there is no close functional similarity in various types seen between the two species. There are a few DEGs that are members of expanded families, but they are mostly just functionally similar, rather than specific overlaps between these two datasets.

Intersection of redox, hormones, and growth in response to aluminum

Expanded genes families with members also highly expressed in AllAl include several that involve hormones. Highly elevated in both seasons and expanded relative to other *Acer*, ent-kaurenoic acid is part of gibberellin biosynthesis and also brassinosteroid synthesis (Helliwell et al., 2001). It is possible there are hormone related growth reductions in the highest levels of aluminum. 1-aminocyclopropane-1-carboxylate (ACC) oxidase is a precursor of ethylene biosynthesis. Indole-3-acetic acid (IAA) -amido synthase controls auxin homeostasis by creating IAA-amino acid conjugates. It suppresses expansin, a cell wall modification gene which loosens cell walls in preparation for growth (Daspute et al., 2017; Ding et al., 2008; Z.-B. Yang et al., 2014). ACC-oxidase and IAA-amido synthase, known participants in aluminum-based root inhibition, are both upregulated in AllAl, and expansin is upregulated in FaUn along with other cell wall and growth related DEGs, such as a variant MDIS1-interacting receptor that increases cell wall integrity maintenance. In both seasons there was also a very strong upregulation of cell cycle regulator E2FE, which prevents increases in growth via cell size by halting endoreduplication. This could be initiated by hormones originating from oxidative stress, but it is possible the underlying issue is DNA damage caused by aluminum toxicity. Endoreduplication is one way plants maintain growth without replicating damaged DNA, and such damage is indicated by several Holliday junction resolvases which repair double stranded breaks (Adachi et al., 2011). Cadmium toxicity reduces endoreduplication, so if aluminum has the same effect,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

perhaps E2FE is part of the suppression mechanism. It is interesting that all these genes, ACC oxidase, IAA-amido synthase, expansion, ATR, and even the endoreduplication response have all been studied in the root, and here, in stem cells, there is some evidence of implication in non-root reductions of growth reported in aluminum treated trees.

Aluminum treatments and the acetyl co a / aldehyde dehydrogenase pathway in spring

Acetyl-coenzyme A synthetase (ACS) activates acetate to acetyl-coenzyme A, a key component of fatty acid metabolism and a source of acetyl moieties for many post-translational modifications and signals, such as lysine acetylation of histones and many of the differentially expressed proteins seen (X. Wu et al., 2011). Acetyl-coenzyme A is the product of several alternate, compartmentalized pathways. Pyruvate dehydrogenase complex (PDC), contained in the peroxisome, is the primary pathway for fatty acid synthesis, while ACS functions in the chloroplast. ACS produces fatty acids and Leucine, but not organic acids citrate, malate, and fumarate and other amino acids that tend to result from PDC (Binder, 2010; Fu et al., 2020). As an alternative pathway, it is important for acetate homeostasis, necessary for proper growth and development. ACS is in many ways redundant to PDC, but it has a different source of acetate, derived from ethanol. Ethanol is converted by alcohol dehydrogenase to acetaldehyde, and aldehyde dehydrogenase (ALDH) converts this to acetate. ACS and ALDH have similar patterns of expression in spring aluminum samples. This particular ACS is a member of a gene family novel to the *Acer* species and is considerably varied from *A. negundo*. The variation and expansion of the gene family combined with expression indicates there is a possibility it is under positive selection.

HBEF expression data profiles trees existing in long-term, chronically stressful conditions. While these are not controlled greenhouse studies on aluminum response or calcium deficiency, and it is possible the trees are responding to more than one stress condition or other variables, replicated samples reveal complex expression indicating multiple forms of stress response in aluminum treatments and also somewhat in unamended plots. Differentially expressed genes, often related to signaling, transport, redox, hormones, and growth, are not seen as extensively in

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

calcium treatments, and furthermore, there is enhancement of other processes, such as senescence, and disease response in calcium. With this data, we provide molecular support to the many studies correlating maple decline with calcium-poor soils exacerbated by acidity, including a 30-year study showing extensive regenerative failure (Cleavitt et al., 2017). Given the lower level of foliar aluminum seen in dominant sugar maple at HBEF, and large amount of variability in aluminum response even between conspecific genotypes, further transcriptomic studies of root tissue would be beneficial for a better understanding of aluminum resistance mechanisms. A number of the differentially expressed genes also seem to be novel or highly diverged, and would benefit from further functional analysis in *Acer*. Higher elevations may become a climatic refugia, and identification of genotypes better adapted to base cation depleted soils may become increasingly important.

Conclusion

In this study we present new chromosomal-length *Acer* genomes for *A. saccharum* and *A. negundo.* We conducted an expression analysis of *A. saccharum* subjected to long-term aluminum and calcium treatments, and identified many genes related to the abiotic stress response and calcium deficiency. Differential expression results from stem tissue were complex, but larger trends were revealed. Aluminum and unamended treatments had upregulated stress response indicating potential damage caused by Al. The necessity of adequate Ca was reflected in calcium treatments by an absence of the abiotic stress response seen at unamended levels, and an upregulation of normal processes, such as seasonal senescence and disease response activity. Gene families related to aluminum stress tolerance were compared between the two species, and showed moderate expansion of families in *A. saccharum*. Broader gene family comparisons revealed expansion in traits associated with invasiveness in *A. negundo*. Release of these genomes and a complementary expression analysis of trees in the HBEF NuPert study shed further light on mechanisms of tolerance to acidic soil conditions and potential adaptations of increasing importance due to climate change.

Methods

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Sequencing

Leaf material for *A. saccharum* was collected from the University of Maryland College Park Campus (GPS location 38°59'16.0"N 76°56'32.8"W). *A. negundo* was sourced from the ForestGEO Forest Dynamics plot (tag # 91603) located in Smithsonian Environmental Research Center, Edgewater, MD. Leaves were dark treated for at least one day prior to collection and shipped to Amplicon Express Inc. (Pullman, WA, USA) for DNA extraction. Both species were sequenced with Pacific BioSciences Sequel v2.1 using a total of fourteen SMRTcells, with the SMRTbell® Express Template Prep Kit v1.0, insert size 40 Kb, library size selection of 66 Kb for *A. saccharum* and 55 Kb for *A. negundo*. The resulting short read sequencing consisted of two lanes of Illumina HiSeq 2500 150PE with insert sizes ranging from 500 to 600 bp (Genomics Resource Center, University of Maryland School of Medicine Institute for Genome Sciences, Baltimore, MD, USA). For Hi-C, DNA leaf material was collected from the same individuals and extracted with the Phase Genomics Proximo Hi-C Plant kit (Phase Genomics, Seattle, WA, USA). Sequencing consisted of two lanes of mid output Illumina NextSeq 500 75PE with an average of 150 M reads per species (Center for Genome Innovation (CGI) at the University of Connecticut, Storrs, CT, USA). For the RNA-Seq used in annotation, leaf samples were collected from one individual per species at two years of age grown at the Michigan State University greenhouse. Sequencing consisted of Illumina HiSeq at 100PE.

Stem tissue from *A. saccharum* was sampled from nine trees across two seasons from the Nutrient Perturbation study plots at HBEF (North Woodstock, NH, USA). A single sapling (dbh < 21 cm) was selected from three plots in each of the three treatments (calcium, control, aluminum). A total of 18 libraries (nine in Oct 2017 (Fall); nine in May 2018 (Spring)) were sampled (Table S2). Tissue was sampled from five years of growth measured by spring wood internodes. Sections were frozen in liquid nitrogen in the field and RNA were extracted after grinding tissues in liquid nitrogen. Extractions were run on Agilent Bioanalyzer Tapestation for quantification and RNA integrity. Two samples were unsuccessful (one aluminum and one

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

calcium) from the spring sampling. Libraries were prepared using Illumina TruSeq Stranded mRNA and sequenced with Illumina NextSeq 500, 150PE (CGI).

Draft genome assembly

In advance of assembly, genome size was estimated with Jellyfish v2.2.6 (21-mers) and GenomeScope (Vurture et al., 2017) using the trimmed Illumina short-reads. The *A. saccharum* reads were test assembled using Illumina short-read, raw and corrected long-reads, and a hybrid of both. Draft assemblies were evaluated in relation to expected genome size, contiguity (N50 and number of contigs), conserved seed plant orthologs, and genomic/transcriptomic read alignments.

The best draft assemblies leveraged the deep PacBio sequencing (*A. saccharum* 103x, *A. negundo* 141x) and prioritized assembling repetitive regions of the genome and resolving the heterozygosity, found in both species. During this phase of the process, two draft assemblies of comparable results were used to investigate scaffolding potential. One was created with FALCON (pb-assembly v0.0.6) a *de novo,* diploid-aware assembler for PacBio (Chin et al., 2016), and the other was done using Canu (v1.6) error-corrected PacBio reads as input (Koren et al., 2017) to Flye v2.3.7, an efficient haploid assembler that leverages repeat graphs with read alignment techniques to resolve areas of repetition (Kolmogorov, 2016/2019).

Alternate heterozygous contigs (haplotigs) were separated from the primary assemblies using Purge Haplotigs v1.0.4 (Roach et al., 2018). To determine coverage, PacBio reads were aligned to the primary assemblies with Minimap2 v2.15-r911-dirty (H. Li, 2018). RepeatModeler v1.0.8 was used to create the repeat libraries for this analysis.

Hi-C scaffolding

Long-range scaffolding of the FALCON/Purge Haplotigs assembly with Hi-C reads followed processes recommended for the following suite of tools (*Genome Assembly Cookbook*, 2019). HiC reads were aligned with BWA mem -5SP and PCR duplicates removed with samblaster

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

v0.1.24 (Faust & Hall, 2014). Scripts from the Juicer v1.5.6 (Durand, Shamim, et al., 2016) pipeline were modified to identify the Sau3AI restriction enzyme. The resulting Hi-C alignment file was provided to 3D-DNA v180419 (Dudchenko et al., 2017) for scaffolding, leveraging different parameters for each species according to the differing draft assembly characteristics. In particular, --diploid was added to *A. saccharum* to address remaining under-collapsed heterozygosity. JuiceBox was used to visualize Hi-C mapping against each scaffold created by the different parameter tests to visually detect which incorporated the majority of the contigs into the expected 13 pseudo-chromosomes (Dudchenko et al., 2018; Durand, Robinson, et al., 2016).

Genome annotation

RepeatModeler v2.01 (Flynn et al., 2020) and RepeatMasker v4.0.6 (Smit et al., 2013) were used to softmask the assembly. Trimmed RNA reads were aligned to the assembly with Hisat2 v2.1.0 (Kim et al., 2015). GenomeThreader v1.7.1 (Gremme et al., 2005) was used to align protein sequences (derived from *de novo* transcriptome assembly; -gcmincoverage 80, -dpminexonlen 20 -startcodon -finalstopcodon -introncutout). Structural gene prediction was executed with BRAKER2 v2.0.5 (Hoff et al., 2016). The process converted RNA-Seq alignments to exon support in GeneMark-ET v4.32 (Lomsadze et al., 2014), and combined this output with protein alignments for two rounds of training with AUGUSTUS v3.2.3 (Stanke et al., 2006, 2008; Camacho et al., 2009) to predict coding regions in the genome. Extensive filtering was performed on the predicted gene space using gFACS v1.1 (Caballero & Wegrzyn, 2019). Evaluation of structural annotations were conducted with BUSCO and the PLAZA CoreGF rosids database v2.5 (Van Bel et al., 2019; Veeckman et al., 2016). Transcriptomic alignments were used to identify where they fully overlapped BRAKER-based models or provided additional support to those that did not pass previous filtering criteria. Transcriptome assemblies were conducted *de novo* with Trinity v2.20 (Grabherr et al., 2011). EnTAP v0.8.0 (Hart et al., 2018) was used to frame-select, functionally annotate, and identify potential contaminants for filtering, including bacteria, archaea, opisthokonta, rhizaria, and fungi. The resulting translated protein sequences were clustered with USEARCH v9.0.2132 (Edgar, 2010) at an alignment identity of 0.90. Transcriptomic alignments created with GMAP v2019-06-10 (T. D. Wu &

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Watanabe, 2005) and gFACs were compared against the genome annotation using GffCompare v0.11.5 (Pertea, 2018).

Functional annotation was performed using EnTAP v0.9.1 (Hart et al., 2020), a pipeline that integrates both similarity search and other annotation sources including gene family (eggNOG), protein domains (Pfam), gene ontology, and pathway assignment (KEGG). The following public databases were included: NCBI RefSeq Complete, EMBL-EBI UniProt, and *Arabidopsis* (TAIR11) .

Comparative genomics

The translated gene space of 22 plant species were used for gene family analysis. *Acer* included *A. saccharum*, *A. negundo*, and *A. yangbiense*. The remaining species were selected from high quality public annotations. Fourteen broadleaf trees (*Betula pendula, Carica papaya*, *Citrus clementina* and *sinensis*, *Eucalyptus grandis*, *Juglans hindsii* and *regia*, *Pistacia vera*, *Populus trichocarpa*, *Prunus persica*, *Quercus lobata* and *robur*, and *Tectona grandis*), plus one woody angiosperm, *Vitis vinifera*, were included, along with one gymnosperm, *Ginkgo biloba*. *Oryza sativa* Kitaake was the representative monocot, along with *Amborella trichopoda* and *Nymphaea colorata*, representing other more basal lineages, and *Arabidopsis thaliana*, being the primary plant model system. OrthoFinder v2.3.7 (Emms & Kelly, 2015, 2018) was used to generate orthogroups. Resulting gene counts per orthogroup for *A. negundo*, *A. saccharum*, and the mean of the combined three *Acer* were each compared to the mean of other species to identify potentially expanded, contracted, missing, and novel gene families. The initial delineation of expansion and contraction was set at 2-fold above the standard deviation. Full absence was verified with alignment of the *Arabidopsis* protein sequence against the assembly. CAFE v5 (Mendes et al., 2020) was used to identify rapidly evolving gene families. Values from the newick species tree produced by OrthoFinder were multiplied by 100 to prevent issues with rounding in CAFE, and the tree was made ultrametric using OrthoFinder. The poisson root frequency distribution was run three times on gene families filtered by size to ensure convergence of a lambda value representing birth and death rate. The selected lambda value was

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

then used to run the large family set. Functional enrichment of the resulting families was obtained from gProfiler v:e99_eg46_p14_f929183 (Raudvere et al., 2019) using the annotation of the representative (longest) gene when aligned to *Vitis vinifera.* This well annotated woody angiosperm represents an ideal source having no recent WGD (Tang et al., 2008).

Whole genome duplication and synteny analysis

Characterization of putative paralogs, including whole genome duplication was done as previously described (Qiao et al., 2019) using DupGen_finder to separate whole genome, tandem, proximal, transposed, and dispersed duplicates. Categorization can be helpful in speculating on the origin of duplication such as transposable element activity, localized replication error, larger segmental translocations, or ploidy events. The whole genome duplicate frequency distribution was plotted by Ks value for analysis of peaks. Microsynteny of the small peak of recent supposed whole genome duplication seen in *A. saccharum* was further analyzed with MCScanX-based collinearity scripts (Nowell et al., 2018), as well as overall macrosynteny between the three *Acer*.

Differential expression analysis

HBEF NuPert plots were used as a basis for this study as they were designed to reflect acidity and calcium levels over time as described by Berger et al. 2001. They consist of 12 *A. saccharum* dominant plots near reference watershed 6, with four receiving annual $AlCl_2$ treatments 12 times from 1995 to 2015, and $CaCl_2$ treatments were applied to 4 other plots for 4 years, followed by applications of slow-release wollastonite in 1999 and 2015 (Table S7). Three samples were collected from aluminum, calcium, and unamended control plots as described in the Sequencing section. Trimmed reads for each of the sixteen successful HBEF-sourced libraries were aligned to the *A. saccharum* reference genome with Hisat2 v2.1.0, and read counts were extracted with htseq-count v0.11.2. The R Bioconductor package, DESeq2 v1.26.0 (Love et al., 2014), was used for the expression analysis with the calcium as the control in pairwise comparisons of unamended to calcium and aluminum to unamended, representing increasing levels of aluminum, and then aluminum to calcium, contrasting the extremes. P-adjusted values greater than 0.1 were

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

filtered. Pairwise comparisons, specific to each season (fall and spring) as well as combined resulted in a total of nine sets. Gene Ontology (GO) enrichment was conducted with g:Profiler (database version e99_eg46_p14_f929183; Raudvere et al., 2019) using alignments of differentially expressed protein models to both *Vitis vinifera* (Phytozome v12.1) and *Arabidopsis thaliana* (TAIR11) as baseline annotations.

A. thaliana was used as a representative model for pathway analysis in the Genemania application for Cytoscape v3.8.0. and *V. vinifera* (NCBI taxon ID:29760) (Franceschini et al. 2013) was used similarly with STRINGDB v.11 in Cytoscape v2.7.1. Differentially expressed proteins for each pairwise comparison were used to visualize the fold-change values in context of the supported pathways. Networks were constructed with a confidence score of 0.4 and 20 maximal additional interactions (Shannon et al. 2003) and additional networks for protein models reported to be responsive to calcium deficiency and aluminum biotoxicity were imported and merged.

Data Availability

Sequencing for both genomes, assemblies, annotations, and RNA-Seq are available in BioProject PRJNA748028. RNA-Seq used in annotation of *A. saccharum* is available in PRJNA413418. Full details on assembly, annotation, gene family analysis, and differential expression analysis can be found at https://gitlab.com/PlantGenomicsLab/AcerGenomes.

Acknowledgements

NGS was funded by the National Science Foundation (NSF DEB-2029997; NSF EF-1638488). Hi-C library preparation and sequencing was funded by the Ronald Bamford Fund Endowment for Ecology and Evolutionary Biology to the Department of Ecology and Evolutionary Biology, University of Connecticut. Support for the HBEF RNA-Seq was provided by the University of Connecticut Center for Biological Risk.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The authors would like to thank the Institute for Systems Genomics (ISG) at UConn for sequencing support, the Microbial Analysis Resources and Services (MARS) within the Center for Open Research Resources & Equipment (COR^{2}E) at UConn for RNA extraction support, the Computational Biology Core for HPC services, and the Hubbard Brook Experimental Forest for access to the NuPert plots. Thanks to Laura Figueroa Corona for assistance with the range map, and Sumaira Zaman, Alison Scott, and Nasim Rahmatpour for sharing computational approaches and scripts. Thank you to Dr. Yaowu Yuan for advice on both data analysis and the manuscript. The *A. negundo* individual sequenced is a part of the Smithsonian Environment Research Center Forest Dynamics Plot, which is part of the Smithsonian's ForestGEO network. We thank ForestGEO and Director Stuart Davies for encouraging tree genome research in the network and for facilitating and funding (NSF DEB-1046113; NSF DEB-1545761) as well as previous working group discussions on the topic.

Author Contributions

SM, NS, JW, PS, and UU conceived and designed various aspects of the study; JW and NS coordinated and managed the study; SM, UU, NS, PS, AT performed the sampling and experiments; AT performed transcriptomic analysis; SM performed the assembly, annotation, and related comparative genomic and expression analysis under the advisement of JW; SM, JW wrote the manuscript; UU, NS, and AT contributed sections. All authors read and approved the final manuscript.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

Adachi, S., Minamisawa, K., Okushima, Y., Inagaki, S., Yoshiyama, K., Kondou, Y., Kaminuma, E., Kawashima, M., Toyoda, T., Matsui, M., Kurihara, D., Matsunaga, S., & Umeda, M. (2011). Programmed induction of endoreduplication by DNA double-strand breaks in Arabidopsis. *Proceedings of the National Academy of Sciences*, *108*(24), 10004–10009. https://doi.org/10.1073/pnas.1103584108

Arora, A., Byrem, T. M., Nair, M. G., & Strasburg, G. M. (2000). Modulation of Liposomal Membrane Fluidity by Flavonoids and Isoflavonoids. *Archives of Biochemistry and Biophysics*, *373*(1), 102–109. https://doi.org/10.1006/abbi.1999.1525

Bal, T. L., Storer, A. J., Jurgensen, M. F., Doskey, P. V., & Amacher, M. C. (2015). Nutrient stress predisposes and contributes to sugar maple dieback across its northern range: A review. *Forestry: An International Journal of Forest Research*, *88*(1), 64–83. https://doi.org/10.1093/forestry/cpu051

Bartwal, A., Mall, R., Lohani, P., Guru, S. K., & Arora, S. (2013). Role of Secondary Metabolites and Brassinosteroids in Plant Defense Against Environmental Stresses. *Journal of Plant Growth Regulation*, *32*(1), 216–232. https://doi.org/10.1007/s00344-012-9272-x

Begum, H. H., Osaki, M., Watanabe, T., & Shinano, T. (2009). Mechanisms of Aluminum Tolerance in Phosphoenolpyruvate Carboxylase Transgenic Rice. *Journal of Plant Nutrition*, *32*(1), 84–96. https://doi.org/10.1080/01904160802531035

Berger, T. W., Eagar, C., Likens, G. E., & Stingeder, G. (2001). Effects of calcium and aluminum chloride additions on foliar and throughfall chemistry in sugar maples. *Forest Ecology*

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and Management, *149*(1), 75–90. https://doi.org/10.1016/S0378-1127(00)00546-6

Binder, S. (2010). Branched-Chain Amino Acid Metabolism in Arabidopsis thaliana. *The Arabidopsis Book / American Society of Plant Biologists*, *8*. https://doi.org/10.1199/tab.0137

Caballero, M., & Wegrzyn, J. (2019). gFACs: Gene Filtering, Analysis, and Conversion to Unify Genome Annotations Across Alignment and Gene Prediction Frameworks. *Genomics, Proteomics & Bioinformatics*, *17*(3), 305–310. https://doi.org/10.1016/j.gpb.2019.04.002

CABI. (2021). *Acer Negundo*. Invasive Species Compendium. https://www.cabi.org/ISC

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, *10*(1), 421. https://doi.org/10.1186/1471-2105-10-421

Cardoso, T. B., Pinto, R. T., & Paiva, L. V. (2019). Analysis of gene co-expression networks of phosphate starvation and aluminium toxicity responses in Populus spp. *PLOS ONE*, *14*(10), e0223217. https://doi.org/10.1371/journal.pone.0223217

Castro, B., Citterico, M., Kimura, S., Stevens, D. M., Wrzaczek, M., & Coaker, G. (2021). Stress-induced reactive oxygen species compartmentalization, perception and signalling. *Nature Plants*, *7*(4), 403–412. https://doi.org/10.1038/s41477-021-00887-0

Chen, Z., Lu, X., Xuan, Y., Tang, F., Wang, J., Shi, D., Fu, S., & Ren, J. (2019). Transcriptome analysis based on a combination of sequencing platforms provides insights into leaf pigmentation in Acer rubrum. *BMC Plant Biology*, *19*(1), 240. https://doi.org/10.1186/s12870-019-1850-7

Chin, C.-S., Peluso, P., Sedlazeck, F. J., Nattestad, M., Concepcion, G. T., Clum, A., Dunn, C.,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

O'Malley, R., Figueroa-Balderas, R., Morales-Cruz, A., Cramer, G. R., Delledonne, M., Luo, C., Ecker, J. R., Cantu, D., Rank, D. R., & Schatz, M. C. (2016). Phased diploid genome assembly with single-molecule real-time sequencing. *Nature Methods*, *13*(12), 1050–1054. https://doi.org/10.1038/nmeth.4035

Choudhury, F. K., Rivero, R. M., Blumwald, E., & Mittler, R. (2017). Reactive oxygen species, abiotic stress and stress combination. *The Plant Journal*, *90*(5), 856–867. https://doi.org/10.1111/tpj.13299

Cleavitt, N. L., Battles, J. J., Johnson, C. E., & Fahey, T. J. (2017). Long-term decline of sugar maple following forest harvest, Hubbard Brook Experimental Forest, New Hampshire. *Canadian Journal of Forest Research*. https://doi.org/10.1139/cjfr-2017-0233

Conner, J., & Liu, Z. (2000). LEUNIG, a putative transcriptional corepressor that regulates AGAMOUS expression during flower development. *Proceedings of the National Academy of Sciences*, *97*(23), 12902–12907. https://doi.org/10.1073/pnas.230352397

Contreras, R. N., & Shearer, K. (2018). Genome Size, Ploidy, and Base Composition of Wild and Cultivated Acer. *Journal of the American Society for Horticultural Science*, *143*(6), 470–485. https://doi.org/10.21273/JASHS04541-18

Crowley, D., Barstow, M., Rivers, M., & Harvey-Brown, Y. (2020). *The Red List of Acer*. BGCI.

Daspute, A. A., Sadhukhan, A., Tokizawa, M., Kobayashi, Y., Panda, S. K., & Koyama, H. (2017). Transcriptional Regulation of Aluminum-Tolerance Genes in Higher Plants: Clarifying the Underlying Molecular Mechanisms. *Frontiers in Plant Science*, *8*. https://doi.org/10.3389/fpls.2017.01358

Dawson, T. E., & Ehleringer, J. R. (1993). Gender-Specific Physiology, Carbon Isotope

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Discrimination, and Habitat Distribution in Boxelder, Acer Negundo. *Ecology*, *74*(3), 798–815. https://doi.org/10.2307/1940807

De Angeli, A., Zhang, J., Meyer, S., & Martinoia, E. (2013). AtALMT9 is a malate-activated vacuolar chloride channel required for stomatal opening in Arabidopsis. *Nature Communications*, *4*(1), 1804. https://doi.org/10.1038/ncomms2815

Ding, X., Cao, Y., Huang, L., Zhao, J., Xu, C., Li, X., & Wang, S. (2008). Activation of the Indole-3-Acetic Acid–Amido Synthetase GH3-8 Suppresses Expansin Expression and Promotes Salicylate- and Jasmonate-Independent Basal Immunity in Rice. *The Plant Cell*, *20*(1), 228–240. https://doi.org/10.1105/tpc.107.055657

Donaldson, J. G., & Jackson, C. L. (2011). ARF family G proteins and their regulators: Roles in membrane transport, development and disease. *Nature Reviews. Molecular Cell Biology*, *12*(6), 362–375. https://doi.org/10.1038/nrm3117

Du, Q., Wang, L., Zhou, D., Yang, H., Gong, C., Pan, W., & Zhang, D. (2014). Allelic variation within the S-adenosyl-L-homocysteine hydrolase gene family is associated with wood properties in Chinese white poplar (Populus tomentosa). *BMC Genetics*, *15*(1), S4. https://doi.org/10.1186/1471-2156-15-S1-S4

Dudchenko, O., Batra, S. S., Omer, A. D., Nyquist, S. K., Hoeger, M., Durand, N. C., Shamim, M. S., Machol, I., Lander, E. S., Aiden, A. P., & Aiden, E. L. (2017). De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science*, *356*(6333), 92–95. https://doi.org/10.1126/science.aal3327

Dudchenko, O., Shamim, M. S., Batra, S. S., Durand, N. C., Musial, N. T., Mostofa, R., Pham, M., Hilaire, B. G. S., Yao, W., Stamenova, E., Hoeger, M., Nyquist, S. K., Korchina, V.,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Pletch, K., Flanagan, J. P., Tomaszewicz, A., McAloose, D., Estrada, C. P., Novak, B. J., … Aiden, E. L. (2018). The Juicebox Assembly Tools module facilitates de novo assembly of mammalian genomes with chromosome-length scaffolds for under $1000. *BioRxiv*, 254797. https://doi.org/10.1101/254797

Durand, N. C., Robinson, J. T., Shamim, M. S., Machol, I., Mesirov, J. P., Lander, E. S., & Aiden, E. L. (2016). Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Systems*, *3*(1), 99–101. https://doi.org/10.1016/j.cels.2015.07.012

Durand, N. C., Shamim, M. S., Machol, I., Rao, S. S. P., Huntley, M. H., Lander, E. S., & Aiden, E. L. (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. *Cell Systems*, *3*(1), 95–98. https://doi.org/10.1016/j.cels.2016.07.002

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, *26*(19), 2460–2461. https://doi.org/10.1093/bioinformatics/btq461

Emms, D. M., & Kelly, S. (2015). OrthoFinder: Solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biology*, *16*(1), 157. https://doi.org/10.1186/s13059-015-0721-2

Emms, D. M., & Kelly, S. (2018). OrthoFinder2: Fast and accurate phylogenomic orthology analysis from gene sequences. *BioRxiv*. https://doi.org/10.1101/466201

Estravis-Barcala, M., Mattera, M. G., Soliani, C., Bellora, N., Opgenoorth, L., Heer, K., & Arana, M. V. (2020). Molecular bases of responses to abiotic stress in trees. *Journal of Experimental Botany*, *71*(13), 3765–3779. https://doi.org/10.1093/jxb/erz532

Faust, G. G., & Hall, I. M. (2014). SAMBLASTER: Fast duplicate marking and structural

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

variant read extraction. *Bioinformatics*, *30*(17), 2503–2505. https://doi.org/10.1093/bioinformatics/btu314

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., & Smit, A. F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proceedings of the National Academy of Sciences*, *117*(17), 9451–9457. https://doi.org/10.1073/pnas.1921046117

Fu, X., Yang, H., Pangestu, F., & Nikolau, B. J. (2020). Failure to Maintain Acetate Homeostasis by Acetate-Activating Enzymes Impacts Plant Development. *Plant Physiology*, *182*(3), 1256–1271. https://doi.org/10.1104/pp.19.01162

Genome Assembly Cookbook. (2019). DNA Zoo. https://www.dnazoo.org/methods

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen, Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. https://doi.org/10.1038/nbt.1883

Gremme, G., Brendel, V., Sparks, M. E., & Kurtz, S. (2005). Engineering a software tool for gene structure prediction in higher organisms. *Information and Software Technology*, *47*(15), 965–978. https://doi.org/10.1016/j.infsof.2005.09.005

Halman, J. M., Schaberg, P. G., Hawley, G. J., Hansen, C. F., & Fahey, T. J. (2015). Differential impacts of calcium and aluminum treatments on sugar maple and American beech growth dynamics. *Canadian Journal of Forest Research*, *45*(1), 52–59. https://doi.org/10.1139/cjfr-2014-0250

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Halman, J. M., Schaberg, P. G., Hawley, G. J., Pardo, L. H., & Fahey, T. J. (2013). Calcium and aluminum impacts on sugar maple physiology in a northern hardwood forest. *Tree Physiology*, *33*(11), 1242–1251. https://doi.org/10.1093/treephys/tpt099

Hart, A. J., Ginzburg, S., Xu, M. (Sam), Fisher, C. R., Rahmatpour, N., Mitton, J. B., Paul, R., & Wegrzyn, J. L. (2018). EnTAP: Bringing Faster and Smarter Functional Annotation to Non-Model Eukaryotic Transcriptomes. *BioRxiv*, 307868. https://doi.org/10.1101/307868

Hart, A. J., Ginzburg, S., Xu, M. (Sam), Fisher, C. R., Rahmatpour, N., Mitton, J. B., Paul, R., & Wegrzyn, J. L. (2020). EnTAP: Bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. *Molecular Ecology Resources*, *20*(2), 591–604. https://doi.org/10.1111/1755-0998.13106

Hasan, M. K., Cheng, Y., Kanwar, M. K., Chu, X.-Y., Ahammed, G. J., & Qi, Z.-Y. (2017). Responses of Plant Proteins to Heavy Metal Stress—A Review. *Frontiers in Plant Science*, *8*. https://doi.org/10.3389/fpls.2017.01492

He, F., Li, H.-G., Wang, J.-J., Su, Y., Wang, H.-L., Feng, C.-H., Yang, Y., Niu, M.-X., Liu, C., Yin, W., & Xia, X. (2019). PeSTZ1, a C2H2-type zinc finger transcription factor from Populus euphratica, enhances freezing tolerance through modulation of ROS scavenging by directly regulating PeAPX2. *Plant Biotechnology Journal*, *17*(11), 2169–2183. https://doi.org/10.1111/pbi.13130

He, F., Pan, Q.-H., Shi, Y., & Duan, C.-Q. (2008). Biosynthesis and Genetic Regulation of Proanthocyanidins in Plants. *Molecules*, *13*(10), 2674–2703. https://doi.org/10.3390/molecules13102674

Helliwell, C. A., Chandler, P. M., Poole, A., Dennis, E. S., & Peacock, W. J. (2001). The

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

CYP88A cytochrome P450, ent-kaurenoic acid oxidase, catalyzes three steps of the gibberellin biosynthesis pathway. *Proceedings of the National Academy of Sciences*, *98*(4), 2065–2070. https://doi.org/10.1073/pnas.98.4.2065

Hendrix, S., Keunen, E., Mertens, A. I. G., Beemster, G. T. S., Vangronsveld, J., & Cuypers, A. (2018). Cell cycle regulation in different leaves of Arabidopsis thaliana plants grown under control and cadmium-exposed conditions. *Environmental and Experimental Botany*, *155*, 441–452. https://doi.org/10.1016/j.envexpbot.2018.06.026

Higo, A., Kawashima, T., Borg, M., Zhao, M., López-Vidriero, I., Sakayama, H., Montgomery, S. A., Sekimoto, H., Hackenberg, D., Shimamura, M., Nishiyama, T., Sakakibara, K., Tomita, Y., Togawa, T., Kunimoto, K., Osakabe, A., Suzuki, Y., Yamato, K. T., Ishizaki, K., … Araki, T. (2018). Transcription factor DUO1 generated by neo-functionalization is associated with evolution of sperm differentiation in plants. *Nature Communications*, *9*(1), 5283. https://doi.org/10.1038/s41467-018-07728-3

Hoekenga, O. A., Maron, L. G., Piñeros, M. A., Cançado, G. M. A., Shaff, J., Kobayashi, Y., Ryan, P. R., Dong, B., Delhaize, E., Sasaki, T., Matsumoto, H., Yamamoto, Y., Koyama, H., & Kochian, L. V. (2006). AtALMT1, which encodes a malate transporter, is identified as one of several genes critical for aluminum tolerance in Arabidopsis. *Proceedings of the National Academy of Sciences*, *103*(25), 9738–9743. https://doi.org/10.1073/pnas.0602868103

Hoff, K. J., Lange, S., Lomsadze, A., Borodovsky, M., & Stanke, M. (2016). BRAKER1: Unsupervised RNA-Seq-Based Genome Annotation with GeneMark-ET and AUGUSTUS. *Bioinformatics*, *32*(5), 767–769.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

https://doi.org/10.1093/bioinformatics/btv661

Hu, B., Wang, N., Bi, X., Karaaslan, E. S., Weber, A.-L., Zhu, W., Berendzen, K. W., & Liu, C. (2019). Plant lamin-like proteins mediate chromatin tethering at the nuclear periphery. *Genome Biology*, *20*(1), 87. https://doi.org/10.1186/s13059-019-1694-3

Huang, J., Zhao, X., & Chory, J. (2019). The Arabidopsis Transcriptome Responds Specifically and Dynamically to High Light Stress. *Cell Reports*, *29*(12), 4186-4199.e3. https://doi.org/10.1016/j.celrep.2019.11.051

Huggett, B. A., Schaberg, P. G., Hawley, G. J., & Eagar, C. (2007). Long-term calcium addition increases growth release, wound closure, and health of sugar maple (Acer saccharum) trees at the Hubbard Brook Experimental Forest. *Canadian Journal of Forest Research*, *37*(9), 1692–1700. https://doi.org/10.1139/X07-042

Jaillon, O., Aury, J.-M., Noel, B., Policriti, A., Clepet, C., Casagrande, A., Choisne, N., Aubourg, S., Vitulo, N., Jubin, C., Vezzi, A., Legeai, F., Hugueney, P., Dasilva, C., Horner, D., Mica, E., Jublot, D., Poulain, J., Bruyère, C., … French-Italian Public Consortium for Grapevine Genome Characterization. (2007). The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*, *449*(7161), 463–467. https://doi.org/10.1038/nature06148

Jansen, S., Broadley, M. R., Robbrecht, E., & Smets, E. (2002). Aluminum hyperaccumulation in angiosperms: A review of its phylogenetic significance. *The Botanical Review*, *68*(2), 235–269. https://doi.org/10.1663/0006-8101(2002)068[0235:AHIAAR]2.0.CO;2

Jasinski, S., Fabrissin, I., Masson, A., Marmagne, A., Lécureuil, A., Bill, L., & Chardon, F. (2021). ACCELERATED CELL DEATH 6 Acts on Natural Leaf Senescence and

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Nitrogen Fluxes in Arabidopsis. *Frontiers in Plant Science*, *11*. https://doi.org/10.3389/fpls.2020.611170

Jockusch, S., & Kräutler, B. (2020). The red chlorophyll catabolite (RCC) is an inefficient sensitizer of singlet oxygen – photochemical studies of the methyl ester of RCC. *Photochemical & Photobiological Sciences*, *19*(5), 668–673. https://doi.org/10.1039/D0PP00071J

Kar, D., Pradhan, A. A., & Datta, S. (2021). The role of solute transporters in aluminum toxicity and tolerance. *Physiologia Plantarum*, *171*(4), 638–652. https://doi.org/10.1111/ppl.13214

Kidokoro, S., Yoneda, K., Takasaki, H., Takahashi, F., Shinozaki, K., & Yamaguchi-Shinozaki, K. (2017). Different Cold-Signaling Pathways Function in the Responses to Rapid and Gradual Decreases in Temperature. *The Plant Cell*, *29*(4), 760–774. https://doi.org/10.1105/tpc.16.00669

Kim, D., Langmead, B., & Salzberg, S. L. (2015). HISAT: A fast spliced aligner with low memory requirements. *Nature Methods*, *12*(4), 357–360. https://doi.org/10.1038/nmeth.3317

Kolmogorov, M. (2019). *Fast and accurate de novo assembler for single molecule sequencing reads: Fenderglass/Flye* [C++]. https://github.com/fenderglass/Flye (Original work published 2016)

Koren, S., Walenz, B. P., Berlin, K., Miller, J. R., Bergman, N. H., & Phillippy, A. M. (2017). Canu: Scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research*, *27*(5), 722–736. https://doi.org/10.1101/gr.215087.116

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Krizek, B. A. (2015). AINTEGUMENTA-LIKE genes have partly overlapping functions with AINTEGUMENTA but make distinct contributions to Arabidopsis thaliana flower development. *Journal of Experimental Botany*, *66*(15), 4537–4549. https://doi.org/10.1093/jxb/erv224

Kunkel, T. A. (2000). DNA replication fidelity. *Annu. ReV. Biochem*, 497–529.

Lamarque, L. J., Lortie, C. J., Porté, A. J., & Delzon, S. (2015). Genetic differentiation and phenotypic plasticity in life-history traits between native and introduced populations of invasive maple trees. *Biological Invasions*, *17*(4), 1109–1122. https://doi.org/10.1007/s10530-014-0781-3

Lamarque, L. J., Porté, A. J., Eymeric, C., Lasnier, J.-B., Lortie, C. J., & Delzon, S. (2013). A Test for Pre-Adapted Phenotypic Plasticity in the Invasive Tree Acer negundo L. *PLOS ONE*, *8*(9), e74239. https://doi.org/10.1371/journal.pone.0074239

Lan, T., You, J., Kong, L., Yu, M., Liu, M., & Yang, Z. (2016). The interaction of salicylic acid and Ca2+ alleviates aluminum toxicity in soybean (Glycine max L.). *Plant Physiology and Biochemistry*, *98*, 146–154. https://doi.org/10.1016/j.plaphy.2015.11.019

Leitch, I. J., Johnston, E., Pellicer, J., Hidalgo, O., & Bennett, M. (2019). *Angiosperm DNA C-values Database (release 9.0, Apr 2019)*. https://cvalues.science.kew.org/search

Li, H. (2018). Minimap2: Pairwise alignment for nucleotide sequences. *Bioinformatics*, *34*(18), 3094–3100. https://doi.org/10.1093/bioinformatics/bty191

Li, J., Stukel, M., Bussies, P., Skinner, K., Lemmon, A. R., Lemmon, E. M., Brown, K., Bekmetjev, A., & Swenson, N. G. (2019). Maple phylogeny and biogeography inferred from phylogenomic data. *Journal of Systematics and Evolution*, *57*(6), 594–606.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

https://doi.org/10.1111/jse.12535

Li, L.-S., Ying, J., Li, E., Ma, T., Li, M., Gong, L.-M., Wei, G., Zhang, Y., & Li, S. (2021). Arabidopsis CBP60b is a central transcriptional activator of immunity. *Plant Physiology*, *kiab164*. https://doi.org/10.1093/plphys/kiab164

Li, N., Meng, H., Xing, H., Liang, L., Zhao, X., & Luo, K. (2017). Genome-wide analysis of MATE transporters and molecular characterization of aluminum resistance in Populus. *Journal of Experimental Botany*, *68*(20), 5669–5683. https://doi.org/10.1093/jxb/erx370

Li, Q., Yu, H., Cao, P. B., Fawal, N., Mathé, C., Azar, S., Cassan-Wang, H., Myburg, A. A., Grima-Pettenati, J., Marque, C., Teulières, C., & Dunand, C. (2015). Explosive Tandem and Segmental Duplications of Multigenic Families in Eucalyptus grandis. *Genome Biology and Evolution*, *7*(4), 1068–1081. https://doi.org/10.1093/gbe/evv048

Li, X., Mao, X., Xu, Y., Li, Y., Zhao, N., Yao, J., Dong, Y., Tigabu, M., Zhao, X., & Li, S. (2021). Comparative transcriptomic analysis reveals the coordinated mechanisms of Populus × canadensis 'Neva' leaves in response to cadmium stress. *Ecotoxicology and Environmental Safety*, *216*, 112179. https://doi.org/10.1016/j.ecoenv.2021.112179

Liang, S. C., Hartwig, B., Perera, P., Mora-García, S., Leau, E. de, Thornton, H., Alves, F. L. de, Rapsilber, J., Yang, S., James, G. V., Schneeberger, K., Finnegan, E. J., Turck, F., & Goodrich, J. (2015). Kicking against the PRCs – A Domesticated Transposase Antagonises Silencing Mediated by Polycomb Group Proteins and Is an Accessory Component of Polycomb Repressive Complex 2. *PLOS Genetics*, *11*(12), e1005660. https://doi.org/10.1371/journal.pgen.1005660

Likens, G. E., Driscoll, C. T., Buso, D. C., Siccama, T. G., Johnson, C. E., Lovett, G. M., Fahey,

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

T. J., Reiners, W. A., Ryan, D. F., Martin, C. W., & Bailey, S. W. (1998). The biogeochemistry of calcium at Hubbard Brook. *Biogeochemistry*, *41*(2), 89–173. https://doi.org/10.1023/A:1005984620681

Likens, G. E., & Lambert, K. F. (1998). The Importance of Long-Term Data in Addressing Regional Environmental Issues. *Northeastern Naturalist*, *5*(2), 127–136. https://doi.org/10.2307/3858583

Liu, J., Magalhaes, J. V., Shaff, J., & Kochian, L. V. (2009). Aluminum-activated citrate and malate transporters from the MATE and ALMT families function independently to confer Arabidopsis aluminum tolerance. *The Plant Journal*, *57*(3), 389–399. https://doi.org/10.1111/j.1365-313X.2008.03696.x

Liu, J., & Zhou, M. (2018). The ALMT Gene Family Performs Multiple Functions in Plants. *Agronomy*, *8*(2), 20. https://doi.org/10.3390/agronomy8020020

Lomsadze, A., Burns, P. D., & Borodovsky, M. (2014). Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Research*, *42*(15), e119–e119. https://doi.org/10.1093/nar/gku557

Long, R. P., Horsley, S. B., Bailey, S. W., Hallett, R. A., & Hall, T. J. (2019). Sugar maple decline and lessons learned about Allegheny Plateau soils and landscapes. *In: Stout, Susan L., Ed. SILVAH: 50 Years of Science-Management Cooperation. Proceedings of the Allegheny Society of American Foresters Training Session; 2017 Sept. 20-22; Clarion, PA. Gen. Tech. Rep. NRS-P-186. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station: 80-97.*, 80–97. https://doi.org/10.2737/NRS-GTR-P-186-Paper8

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Love, M., Anders, S., & Huber, W. (2014). Differential analysis of count data–the DESeq2 package. *Genome Biol*, *15*(550), 10–1186.

Luo, J., Zhou, J., Li, H., Shi, W., Polle, A., Lu, M., Sun, X., & Luo, Z.-B. (2015). Global poplar root and leaf transcriptomes reveal links between growth and stress responses under nitrogen starvation and excess. *Tree Physiology*, *35*(12), 1283–1302. https://doi.org/10.1093/treephys/tpv091

Ma, Q., Sun, T., Li, S., Wen, J., Zhu, L., Yin, T., Yan, K., Xu, X., Li, S., Mao, J., Wang, Y., Jin, S., Zhao, X., & Li, Q. (2020). The Acer truncatum genome provides insights into nervonic acid biosynthesis. *The Plant Journal*, *104*(3), 662–678. https://doi.org/10.1111/tpj.14954

Mach, J. M., Castillo, A. R., Hoogstraten, R., & Greenberg, J. T. (2001). The Arabidopsis-accelerated cell death gene ACD2 encodes red chlorophyll catabolite reductase and suppresses the spread of disease symptoms. *Proceedings of the National Academy of Sciences*, *98*(2), 771–776. https://doi.org/10.1073/pnas.98.2.771

Mano, J., Belles-Boix, E., Babiychuk, E., Inzé, D., Torii, Y., Hiraoka, E., Takimoto, K., Slooten, L., Asada, K., & Kushnir, S. (2005). Protection against Photooxidative Injury of Tobacco Leaves by 2-Alkenal Reductase. Detoxication of Lipid Peroxide-Derived Reactive Carbonyls. *Plant Physiology*, *139*(4), 1773–1783. https://doi.org/10.1104/pp.105.070391

Mekawy, A. M. M., Assaha, D. V. M., Munehiro, R., Kohnishi, E., Nagaoka, T., Ueda, A., & Saneoka, H. (2018). Characterization of type 3 metallothionein-like gene (OsMT-3a) from rice, revealed its ability to confer tolerance to salinity and heavy metal stresses. *Environmental and Experimental Botany*, *147*, 157–166.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

https://doi.org/10.1016/j.envexpbot.2017.12.002

Mendes, F. K., Vanderpool, D., Fulton, B., & Hahn, M. W. (2020). CAFE 5 models variation in evolutionary rates among gene families. *Bioinformatics*, *btaa1022*. https://doi.org/10.1093/bioinformatics/btaa1022

Ming, R., Hou, S., Feng, Y., Yu, Q., Dionne-Laporte, A., Saw, J. H., Senin, P., Wang, W., Ly, B. V., Lewis, K. L. T., Salzberg, S. L., Feng, L., Jones, M. R., Skelton, R. L., Murray, J. E., Chen, C., Qian, W., Shen, J., Du, P., … Alam, M. (2008). The draft genome of the transgenic tropical fruit tree papaya (Carica papaya Linnaeus). *Nature*, *452*(7190), 991–996. https://doi.org/10.1038/nature06856

Mohamed, H. K., Pathak, S., Roy, D. N., Hutchinson, T. C., McLaughlin, D. L., & Kinch, J. C. (1997). RELATIONSHIP BETWEEN SUGAR MAPLE DECLINE AND CORRESPONDING CHEMICAL CHANGES IN THE STEM TISSUE. *Water, Air, and Soil Pollution*, *96*(1), 321–327. https://doi.org/10.1023/A:1026429122530

Moore, J.-D., & Ouimet, R. (2021). Liming still positively influences sugar maple nutrition, vigor and growth, 20 years after a single application. *Forest Ecology and Management*, *490*, 119103. https://doi.org/10.1016/j.foreco.2021.119103

Naslavsky, N., & Caplan, S. (2005). C-terminal EH-domain-containing proteins: Consensus for a role in endocytic trafficking, EH? *Journal of Cell Science*, *118*(18), 4093–4101. https://doi.org/10.1242/jcs.02595

Nisa, M.-U., Huang, Y., Benhamed, M., & Raynaud, C. (2019). The Plant DNA Damage Response: Signaling Pathways Leading to Growth Inhibition and Putative Role in Response to Stress Conditions. *Frontiers in Plant Science*, *10*.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

https://doi.org/10.3389/fpls.2019.00653

Nkongolo, K., Theriault, G., & Michael, P. (2018). Nickel-induced global gene expressions in red maple (Acer rubrum): Effect of nickel concentrations. *Plant Gene*, *14*, 29–36. https://doi.org/10.1016/j.plgene.2018.04.003

Nowell, R. W., Almeida, P., Wilson, C. G., Smith, T. P., Fontaneto, D., Crisp, A., Micklem, G., Tunnacliffe, A., Boschetti, C., & Barraclough, T. G. (2018). Comparative genomics of bdelloid rotifers: Insights from desiccating and nondesiccating species. *PLOS Biology*, *16*(4), e2004830. https://doi.org/10.1371/journal.pbio.2004830

Pellicer, J., Hidalgo, O., Dodsworth, S., & Leitch, I. J. (2018). Genome Size Diversity and Its Impact on the Evolution of Land Plants. *Genes*, *9*(2), 88. https://doi.org/10.3390/genes9020088

Pereira, J. F., & Ryan, P. R. (2019). The role of transposable elements in the evolution of aluminium resistance in plants. *Journal of Experimental Botany*, *70*(1), 41–54. https://doi.org/10.1093/jxb/ery357

Pertea, G. (2018). *GffCompare*. http://ccb.jhu.edu/software/stringtie/gffcompare.shtml

Porté, A. J., Lamarque, L. J., Lortie, C. J., Michalet, R., & Delzon, S. (2011). Invasive Acer negundo outperforms native species in non-limiting resource environments due to its higher phenotypic plasticity. *BMC Ecology*, *11*(1), 28. https://doi.org/10.1186/1472-6785-11-28

Pyšek, P., Skálová, H., Čuda, J., Guo, W.-Y., Suda, J., Doležal, J., Kauzál, O., Lambertini, C., Lučanová, M., Mandáková, T., Moravcová, L., Pyšková, K., Brix, H., & Meyerson, L. A. (2018). Small genome separates native and invasive populations in an ecologically

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

important cosmopolitan grass. *Ecology*, *99*(1), 79–90. https://doi.org/10.1002/ecy.2068

Qiao, X., Li, Q., Yin, H., Qi, K., Li, L., Wang, R., Zhang, S., & Paterson, A. H. (2019). Gene duplication and evolution in recurring polyploidization–diploidization cycles in plants. *Genome Biology*, *20*(1), 38. https://doi.org/10.1186/s13059-019-1650-2

Racero, M., & J, F. (2020). *Caracterización funcional del canal aniónico ALMT10*. https://digital.csic.es/handle/10261/229928

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g:Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research*, *47*(W1), W191–W198. https://doi.org/10.1093/nar/gkz369

Renner, S. S., Beenken, L., Grimm, G. W., Kocyan, A., & Ricklefs, R. E. (2007). The Evolution of Dioecy, Heterodichogamy, and Labile Sex Expression in Acer. *Evolution*, *61*(11), 2701–2719. https://doi.org/10.1111/j.1558-5646.2007.00221.x

Rizza, A., Boccaccini, A., Lopez‐Vidriero, I., Costantino, P., & Vittorioso, P. (2011). Inactivation of the ELIP1 and ELIP2 genes affects Arabidopsis seed germination. *New Phytologist*, *190*(4), 896–905. https://doi.org/10.1111/j.1469-8137.2010.03637.x

Roach, M. J., Schmidt, S. A., & Borneman, A. R. (2018). Purge Haplotigs: Allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics*, *19*(1), 460. https://doi.org/10.1186/s12859-018-2485-7

Robineau, T., Batard, Y., Nedelkina, S., Cabello-Hurtado, F., LeRet, M., Sorokine, O., Didierjean, L., & Werck-Reichhart, D. (1998). The Chemically Inducible Plant Cytochrome P450 CYP76B1 Actively Metabolizes Phenylureas and Other Xenobiotics1.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Plant Physiology, *118*(3), 1049–1056. https://doi.org/10.1104/pp.118.3.1049

Roddy, A. B., Théroux-Rancourt, G., Abbo, T., Benedetti, J. W., Brodersen, C. R., Castro, M., Castro, S., Gilbride, A. B., Jensen, B., Jiang, G.-F., Perkins, J. A., Perkins, S. D., Loureiro, J., Syed, Z., Thompson, R. A., Kuebbing, S. E., & Simonin, K. A. (2019). The Scaling of Genome Size and Cell Size Limits Maximum Rates of Photosynthesis with Implications for Ecological Strategies. *International Journal of Plant Sciences*, *181*(1), 75–87. https://doi.org/10.1086/706186

Rosario, L. C. (1988). *Acer negundo*. Fire Effects Information System (FEIS). U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Fire Sciences Laboratory. https://www.fs.fed.us/database/feis/plants/tree/aceneg/all.html

Sade, H., Meriga, B., Surapu, V., Gadi, J., Sunita, M. S. L., Suravajhala, P., & Kavi Kishor, P. B. (2016). Toxicity and tolerance of aluminum in plants: Tailoring plants to suit to acid soils. *BioMetals*, *29*(2), 187–210. https://doi.org/10.1007/s10534-016-9910-z

Sakamoto, A. N. (2019). Translesion Synthesis in Plants: Ultraviolet Resistance and Beyond. *Frontiers in Plant Science*, *10*, 1208. https://doi.org/10.3389/fpls.2019.01208

Schaberg, P. G., DeHayes, D. H., & Hawley, G. J. (2001). Anthropogenic Calcium Depletion: A Unique Threat to Forest Ecosystem Health? *Ecosystem Health*, *7*(4), 214–228. https://doi.org/10.1046/j.1526-0992.2001.01046.x

Schaberg, P. G., Tilley, J. W., Hawley, G. J., DeHayes, D. H., & Bailey, S. W. (2006). Associations of calcium and aluminum with the growth and health of sugar maple trees in Vermont. *Forest Ecology and Management*, *223*(1), 159–169. https://doi.org/10.1016/j.foreco.2005.10.067

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Schlaich, N. L. (2007). Flavin-containing monooxygenases in plants: Looking beyond detox. *Trends in Plant Science*, *12*(9), 412–418. https://doi.org/10.1016/j.tplants.2007.08.009

Shi, X., Sun, H., Chen, Y., Pan, H., & Wang, S. (2016). Transcriptome Sequencing and Expression Analysis of Cadmium (Cd) Transport and Detoxification Related Genes in Cd-Accumulating Salix integra. *Frontiers in Plant Science*, *7*. https://doi.org/10.3389/fpls.2016.01577

Shu, K., Zhou, W., & Yang, W. (2018). APETALA 2-domain-containing transcription factors: Focusing on abscisic acid and gibberellins antagonism. *New Phytologist*, *217*(3), 977–983. https://doi.org/10.1111/nph.14880

Šmarda, P., Bureš, P., Horová, L., Leitch, I. J., Mucina, L., Pacini, E., Tichý, L., Grulich, V., & Rotreklová, O. (2014). Ecological and evolutionary significance of genomic GC content diversity in monocots. *Proceedings of the National Academy of Sciences*, *111*(39), E4096–E4102. https://doi.org/10.1073/pnas.1321152111

Smit, A., Hubley, R., & Green, P. (2013). *RepeatMasker Open-4.0*. http://www.repeatmasker.org

Sork, V. L., Fitz-Gibbon, S. T., Puiu, D., Crepeau, M., Gugger, P. F., Sherman, R., Stevens, K., Langley, C. H., Pellegrini, M., & Salzberg, S. L. (2016). First Draft Assembly and Annotation of the Genome of a California Endemic Oak Quercus lobata Née (Fagaceae). *G3: Genes, Genomes, Genetics*, *6*(11), 3485–3495. https://doi.org/10.1534/g3.116.030411

Stafford, H. A. (1988). Proanthocyanidins and the lignin connection. *Phytochemistry*, *27*(1), 1–6. https://doi.org/10.1016/0031-9422(88)80583-1

Stanke, M., Diekhans, M., Baertsch, R., & Haussler, D. (2008). Using native and syntenically

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mapped cDNA alignments to improve de novo gene finding. *Bioinformatics*, *24*(5), 637–644. https://doi.org/10.1093/bioinformatics/btn013

Stanke, M., Schöffmann, O., Morgenstern, B., & Waack, S. (2006). Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics*, *7*(1), 62. https://doi.org/10.1186/1471-2105-7-62

Staton, M., Best, T., Khodwekar, S., Owusu, S., Xu, T., Xu, Y., Jennings, T., Cronn, R., Arumuganathan, A. K., Coggeshall, M., Gailing, O., Liang, H., Romero-Severson, J., Schlarbaum, S., & Carlson, J. E. (2015). Preliminary Genomic Characterization of Ten Hardwood Tree Species from Multiplexed Low Coverage Whole Genome Sequencing. *PLOS ONE*, *10*(12), e0145031. https://doi.org/10.1371/journal.pone.0145031

Stevens, R., Grelon, M., Vezon, D., Oh, J., Meyer, P., Perennes, C., Domenichini, S., & Bergounioux, C. (2004). A CDC45 Homolog in Arabidopsis Is Essential for Meiosis, as Shown by RNA Interference–Induced Gene Silencing. *The Plant Cell*, *16*(1), 99–113. https://doi.org/10.1105/tpc.016865

Suda, J., Meyerson, L. A., Leitch, I. J., & Pyšek, P. (2015). The hidden side of plant invasions: The role of genome size. *New Phytologist*, *205*(3), 994–1007. https://doi.org/10.1111/nph.13107

Sullivan, T. J., Lawrence, G. B., Bailey, S. W., McDonnell, T. C., Beier, C. M., Weathers, K. C., McPherson, G. T., & Bishop, D. A. (2013). Effects of Acidic Deposition and Soil Acidification on Sugar Maple Trees in the Adirondack Mountains, New York. *Environmental Science & Technology*, *47*(22), 12687–12694. https://doi.org/10.1021/es401864w

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Sun, X.-L., Yu, Q.-Y., Tang, L.-L., Ji, W., Bai, X., Cai, H., Liu, X.-F., Ding, X.-D., & Zhu, Y.-M. (2013). GsSRK, a G-type lectin S-receptor-like serine/threonine protein kinase, is a positive regulator of plant tolerance to salt stress. *Journal of Plant Physiology*, *170*(5), 505–515. https://doi.org/10.1016/j.jplph.2012.11.017

Tang, H., Bowers, J. E., Wang, X., Ming, R., Alam, M., & Paterson, A. H. (2008). Synteny and Collinearity in Plant Genomes. *Science*, *320*(5875), 486–488. https://doi.org/10.1126/science.1153917

Theriault, G., Michael, P., & Nkongolo, K. (2016). Comprehensive Transcriptome Analysis of Response to Nickel Stress in White Birch (Betula papyrifera). *PLOS ONE*, *11*(4), e0153762. https://doi.org/10.1371/journal.pone.0153762

Tian, J., & Kong, Z. (2019). The role of the augmin complex in establishing microtubule arrays. *Journal of Experimental Botany*, *70*(12), 3035–3041. https://doi.org/10.1093/jxb/erz123

Tokizawa, M., Kobayashi, Y., Saito, T., Kobayashi, M., Iuchi, S., Nomoto, M., Tada, Y., Yamamoto, Y. Y., & Koyama, H. (2015). SENSITIVE TO PROTON RHIZOTOXICITY1, CALMODULIN BINDING TRANSCRIPTION ACTIVATOR2, and Other Transcription Factors Are Involved in ALUMINUM-ACTIVATED MALATE TRANSPORTER1 Expression. *Plant Physiology*, *167*(3), 991–1003. https://doi.org/10.1104/pp.114.256552

Tommasini, R., Vogt, E., Fromenteau, M., Hörtensteiner, S., Matile, P., Amrhein, N., & Martinoia, E. (1998). An ABC-transporter of Arabidopsis thaliana has both glutathione-conjugate and chlorophyll catabolite transport activity. *The Plant Journal*, *13*(6), 773–780. https://doi.org/10.1046/j.1365-313X.1998.00076.x

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Toyota, M., Spencer, D., Sawai-Toyota, S., Jiaqi, W., Zhang, T., Koo, A. J., Howe, G. A., & Gilroy, S. (2018). Glutamate triggers long-distance, calcium-based plant defense signaling. *Science*, *361*(6407), 1112–1115. https://doi.org/10.1126/science.aat7744

Trávníček, P., Čertner, M., Ponert, J., Chumová, Z., Jersáková, J., & Suda, J. (2019). Diversity in genome size and GC content shows adaptive potential in orchids and is closely linked to partial endoreplication, plant life-history traits and climatic conditions. *New Phytologist*, *224*(4), 1642–1656. https://doi.org/10.1111/nph.15996

Tuskan, G. A., DiFazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., Putnam, N., Ralph, S., Rombauts, S., Salamov, A., Schein, J., Sterck, L., Aerts, A., Bhalerao, R. R., Bhalerao, R. P., Blaudez, D., Boerjan, W., Brun, A., Brunner, A., … Rokhsar, D. (2006). The Genome of Black Cottonwood, Populus trichocarpa (Torr. & Gray). *Science*, *313*(5793), 1596–1604. https://doi.org/10.1126/science.1128691

Valencia, J. D., & Girgis, H. Z. (2019). LtrDetector: A tool-suite for detecting long terminal repeat retrotransposons de-novo. *BMC Genomics*, *20*(1), 450. https://doi.org/10.1186/s12864-019-5796-9

Van Bel, M., Bucchini, F., & Vandepoele, K. (2019). Gene space completeness in complex plant genomes. *Current Opinion in Plant Biology*, *48*, 9–17. https://doi.org/10.1016/j.pbi.2019.01.001

Veeckman, E., Ruttink, T., & Vandepoele, K. (2016). Are We There Yet? Reliably Estimating the Completeness of Plant Genome Sequences. *The Plant Cell*, *28*(8), 1759–1768. https://doi.org/10.1105/tpc.16.00349

Veleba, A., Šmarda, P., Zedek, F., Horová, L., Šmerda, J., & Bureš, P. (2017). Evolution of

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

genome size and genomic GC content in carnivorous holokinetics (Droseraceae). *Annals of Botany*, *119*(3), 409–416. https://doi.org/10.1093/aob/mcw229

Verde, I., Jenkins, J., Dondini, L., Micali, S., Pagliarani, G., Vendramin, E., Paris, R., Aramini, V., Gazza, L., Rossini, L., Bassi, D., Troggio, M., Shu, S., Grimwood, J., Tartarini, S., Dettori, M. T., & Schmutz, J. (2017). The Peach v2.0 release: High-resolution linkage mapping and deep resequencing improve chromosome-scale assembly and contiguity. *BMC Genomics*, *18*(1), 225. https://doi.org/10.1186/s12864-017-3606-9

Vurture, G. W., Sedlazeck, F. J., Nattestad, M., Underwood, C. J., Fang, H., Gurtowski, J., & Schatz, M. C. (2017). GenomeScope: Fast reference-free genome profiling from short reads. *Bioinformatics*, *33*(14), 2202–2204. https://doi.org/10.1093/bioinformatics/btx153

Wang, Y., Zhao, Y., Wang, S., Liu, J., Wang, X., Han, Y., & Liu, F. (2021). Up-regulated 2-alkenal reductase expression improves low-nitrogen tolerance in maize by alleviating oxidative stress. *Plant, Cell & Environment*, *44*(2), 559–573. https://doi.org/10.1111/pce.13956

Wu, T. D., & Watanabe, C. K. (2005). GMAP: A genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics*, *21*(9), 1859–1875. https://doi.org/10.1093/bioinformatics/bti310

Wu, X., Oh, M.-H., Schwarz, E. M., Larue, C. T., Sivaguru, M., Imai, B. S., Yau, P. M., Ort, D. R., & Huber, S. C. (2011). Lysine Acetylation Is a Widespread Protein Modification for Diverse Proteins in Arabidopsis. *Plant Physiology*, *155*(4), 1769–1778. https://doi.org/10.1104/pp.110.165852

Xia, K., Ou, X., Tang, H., Wang, R., Wu, P., Jia, Y., Wei, X., Xu, X., Kang, S.-H., Kim, S.-K., &

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Zhang, M. (2015). Rice microRNA osa-miR1848 targets the obtusifoliol 14α-demethylase gene OsCYP51G3 and mediates the biosynthesis of phytosterols and brassinosteroids during development and in response to stress. *New Phytologist*, *208*(3), 790–802. https://doi.org/10.1111/nph.13513

Xu, M., Cho, E., Burch-Smith, T. M., & Zambryski, P. C. (2012). Plasmodesmata formation and cell-to-cell transport are reduced in decreased size exclusion limit 1 during embryogenesis in Arabidopsis. *Proceedings of the National Academy of Sciences*, *109*(13), 5098–5103. https://doi.org/10.1073/pnas.1202919109

Yamamoto, Y., Kobayashi, Y., Devi, S. R., Rikiishi, S., & Matsumoto, H. (2003). Oxidative stress triggered by aluminum in plant roots. In J. Abe (Ed.), *Roots: The Dynamic Interface between Plants and the Earth: The 6th Symposium of the International Society of Root Research, 11–15 November 2001, Nagoya, Japan* (pp. 239–243). Springer Netherlands. https://doi.org/10.1007/978-94-017-2923-9_23

Yang, J., Wariss, H. M., Tao, L., Zhang, R., Yun, Q., Hollingsworth, P., Dao, Z., Luo, G., Guo, H., Ma, Y., & Sun, W. (2019). De novo genome assembly of the endangered Acer yangbiense, a plant species with extremely small populations endemic to Yunnan Province, China. *GigaScience*, *8*(7). https://doi.org/10.1093/gigascience/giz085

Yang, W.-Y., Zheng, Y., Bahn, S. C., Pan, X.-Q., Li, M.-Y., Vu, H. S., Roth, M. R., Scheu, B., Welti, R., Hong, Y.-Y., & Wang, X.-M. (2012). The Patatin-Containing Phospholipase A pPLAIIα Modulates Oxylipin Formation and Water Loss in Arabidopsis thaliana. *Molecular Plant*, *5*(2), 452–460. https://doi.org/10.1093/mp/ssr118

Yang, Z.-B., Geng, X., He, C., Zhang, F., Wang, R., Horst, W. J., & Ding, Z. (2014).

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TAA1-Regulated Local Auxin Biosynthesis in the Root-Apex Transition Zone Mediates the Aluminum-Induced Inhibition of Root Growth in Arabidopsis. *The Plant Cell*, *26*(7), 2889–2904. https://doi.org/10.1105/tpc.114.127993

Zeng, L., Tu, X.-L., Dai, H., Han, F.-M., Lu, B.-S., Wang, M.-S., Nanaei, H. A., Tajabadipour, A., Mansouri, M., Li, X.-L., Ji, L.-L., Irwin, D. M., Zhou, H., Liu, M., Zheng, H.-K., Esmailizadeh, A., & Wu, D.-D. (2019). Whole genomes and transcriptomes reveal adaptation and domestication of pistachio. *Genome Biology*, *20*(1), 79. https://doi.org/10.1186/s13059-019-1686-3

Zhang, Y., Guo, J., Chen, M., Li, L., Wang, L., & Huang, C.-F. (2018). The Cell Cycle Checkpoint Regulator ATR Is Required for Internal Aluminum Toxicity-Mediated Root Growth Inhibition in Arabidopsis. *Frontiers in Plant Science*, *9*. https://doi.org/10.3389/fpls.2018.00118

Zhao, D., Hamilton, J. P., Bhat, W. W., Johnson, S. R., Godden, G. T., Kinser, T. J., Boachon, B., Dudareva, N., Soltis, D. E., Soltis, P. S., Hamberger, B., & Buell, C. R. (2019). A chromosomal-scale genome assembly of Tectona grandis reveals the importance of tandem gene duplication and enables discovery of genes in natural product biosynthetic pathways. *GigaScience*, *8*(3). https://doi.org/10.1093/gigascience/giz005

Zhu, W., Li, L., Neuhäuser, B., Thelen, M., Wang, M., Chen, J., Wei, L., Venkataramani, K., Exposito-Alonso, M., Liu, C., Keck, J., Barragan, A. C., Schwab, R., Lutz, U., Ludewig, U., & Weigel, D. (2021). Small peptides modulate the immune function of the ion channel-like protein ACD6 in Arabidopsis thaliana. *BioRxiv*, 2021.01.25.428077. https://doi.org/10.1101/2021.01.25.428077

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure and Table Captions

Figure 1. Distribution of *A. saccharum* (blue) and *A. negundo* (orange) in North America. *A. saccharum* range in blue, *A. negundo* in orange. Leaves indicate location of individuals selected for the reference genomes; *A. saccharum* from the University of Maryland campus, and *A. negundo* from the Smithsonian Environmental Research Center. HBEF (Hubbard Brook Experimental Forest) is the location of the 9 individuals used for RNA-seq, three each from calcium, aluminum, and control plots.
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/fia_little_acer_map_labelled.jpg

Figure 2. Results of assembly testing with *A. saccharum*, comparing fragmentation in terms of total contigs versus assembly length. The dashed line represents the estimated genome size. Gray dots are short-read assemblers, shown as highly fragmented. Blue dots are long-read tests of assembly workflows. Canu refers to the use of reads error-corrected by the Canu pipeline. The red dot is the selected draft assembly, and the green dot shows scaffolding results following Hi-C. Detailed assembly statistics are available in File S1.
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/assemblycomparisongraph_busco.jpg

Figure 3. Ks distribution for WGD synteny blocks with a summary of duplication types in (a) *A. negundo* and (b) *A. saccharum*. Abbreviations for categories of duplication: WGD, whole genome duplication; TD, tandem duplication; PD, proximal duplication; TRD, transposed duplication; DSD, dispersed duplication. (c) Circos plot of the thirteen chromosomes ordered largest to smallest for *A. negundo* (orange bars) and *A. saccharum* (blue bars) with distributions of gene density (green) and transposable element frequency (purple). Syntenic regions are linked in gray with darker shades to visually highlight larger recombinations.

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/acsa_acne_manuscript_circos.png

Figure 4. a) Differential expression study design showing number of samples collected in fall and spring from treatment plots at the Hubbard Brook Experimental Forest, Nutrient Perturbation study. b) Differentially expressed genes (up and downregulated) for each treatment and season comparison. Charts display both significance and relative expression denoted as log-fold change. Dotted lines indicate thresholds of significance (0.1 p-adjusted, 1.5 $\log_2$ fold change). https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/hbefplots.jpg

Figure 5. a) Gene ontology enrichments for *Acer* (all three species combined), *A. negundo*, and *A. saccharum*. Abbreviations for gene family dynamics: E, expanded; N, novel; RC, rapidly contracting. b) Total gene families, shared and unique, among the *Acer*. c) Reconstructed gene tree showing contracted gene families in red and expanded in green. https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/cafetreevennGO.jpg

Figure 6. *A. negundo* gene families with ontology related to DNA damage and repair, and secondary enrichments categorized by color. Circles with multiple colors indicate multiple ontology assignments. Lines indicated known or predicted interactions, or other association via text-mining, co-expression, or protein homology. https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/acne_dnadamage.jpg

Figure 7. Orthogroup sizes for aluminum tolerance gene families are presented by species. Families were selected for inclusion based on documented aluminum tolerance and/or presence in the HBEF RNA-Seq differential expression results. Color represents the proportion of gene membership per species, with darker purple equating to more contracted families relative to the

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

median, and dark green indicating expansion. (H) Family contains HBEF differentially expressed gene; (E) Expanded in *A. saccharum*; (C) Contracting; (M) Missing; (N) Novel; (*) Rapidly expanding; Categorization of tolerance is according to literature describing aluminum stress phenotypes. The undetermined category contains species where tolerance to aluminum or acidic soils has not been reported. [1]*B. pendula* is undetermined due to high variability in tolerance by genotype.
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/genefamily_al_comparisons_manuscript_customnorm.html.jpg

Supplemental

Figure S1. Genome size estimation using k-mer distribution analysis
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/Figure_1_SuppInfo.pdf
Figure S2. Hi-C plots
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/Figure_2_SuppInfo.pdf
Figure S3. PCA plot
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/Figure_3_SuppInfo.pdf
Figure S4. Syntenic comparisons between the three *Acer* genomes
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/figures/Figure_4_SuppInfo.pdf

Table S1. Illumina, PacBio, and Hi-C sequencing data summaries
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/tables/sequencing_tables.pdf
Table S2. HBEF table of trees

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/tables/hbef_trees.pdf

Table S3. HBEF GO enrichment

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/tables/hbef_functional_enrichment.pdf

Table S4. Orthogroup statistics by species

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/tables/Orthogroup_summary_table.pdf

Table S5. OF dynamics GO enrichment

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/tables/orthogroup_functional_enrichment.pdf

Table S6. CAFE GO enrichment

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/tables/cafe_functional_enrichment.pdf

Table S7. HBEF Nutrient Perturbation Treatment Schedule

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/tables/HBEFNuPertTreatmentTable2015.pdf

File S1. Assembly output stats

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/assembly_statistics.xlsx

File S2. Annotation details

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/annotation_statistics.xlsx

File S3. Collinearity analysis of recent Ks peak of WGD frequency

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/Acer_microsynteny.xlsx

File S4. HBEF differentially expressed genes

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/HBEF_DEGs.xlsx

File S5. Orthofinder significant dynamics
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/orthofinder_dynamics.xlsx

File S6. CAFE significant rapid evolution
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/cafe_rapidly_evolving.xlsx

File S7. *A. negundo* vs *Acer* contracted or missing using longest overall gene as annotation
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/acne_contracedmissing_longestoverall.xlsx

File S8. *A. negundo* vs *Acer* contracted or missing using longest *Acer* gene as annotation
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/acne_contractedmissing_longestacer.xlsx

File S9. *A. negundo* vs *Acer* expanded or novel using longest overall gene as annotation
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/acne_expandednovel_longestacer.xlsx

File S10. *A. negundo* vs *Acer* expanded or novel using longest *Acer* gene as annotation
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/acne_expandednovel_longestoverall.xlsx

File S11. *A. saccharum* vs *Acer* contracted or missing using longest *Acer* gene as annotation
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/acsa_contractedmissing_longestacer.xlsx

File S12. *A. saccharum* vs *Acer* contracted or missing using longest overall gene as annotation
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/acsa_contractedmissing_longestoverall.xlsx

File S13. *A. saccharum* vs *Acer* expanded or novel using longest *Acer* gene as annotation
https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/acsa_expandednovel_longestacer.xlsx

bioRxiv preprint doi: https://doi.org/10.1101/2021.07.19.452996; this version posted July 20, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

File S14. *A. negundo* verified missing orthogroups

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/acne_missing_verified_string_mapping.tsv

File S15. *A. saccharum* verified missing orthogroups

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/acsa_missing_verified_string_mapping.tsv

File S16. Orthogroup comparisons for HBEF DEG and Al tolerance genes

https://gitlab.com/PlantGenomicsLab/AcerGenomes/-/blob/master/acer/supplemental/files/hbef_orthogroup_Al_comparisons.xlsx