PROPOSAL TO
THE AMERICAN CHESTNUT FOUNDATION
RESEARCH PROGRAM

A. PROJECT TITLE: Genotyping by Sequencing of chestnut trees in Oregon.

B. SUMMARY (not more than 100 words)

This proposal addresses the 2024 TACF priority research area of "Discovery, study, and evaluation of large surviving American chestnuts". American chestnut trees planted in pre-blight times outside their natural range are a potential source of uninfected genotypes for introduction into The American Chestnut Foundation's backcross, testing and research programs. Mixed species plantings might also serve as an additional source of hybrid progeny for research. Large Several American chestnut trees from the late 1800's and early 1900's landscape plantings have been identified in the northwest. For this project, leaf samples from chestnut saplings and large trees from an old, mixed species orchard on the Willamette River in northwestern Oregon will be collected and DNA sequencing conducted for SNP discovery and genotyping to confirm species identify of large trees and determine if any hybrid saplings and their parents can be identified.

C. PRINCIPAL INVESTIGATOR AND INSTITUTIONAL AFFILIATION

John E. Carlson, Professor Emeritus
Department of Ecosystem Science and Management
The Pennsylvania State University
University Park, PA. 16802
(814) 883-5636
jec16@psu.edu

D. DURATION OF PROJECT:

January 1, 2025 to December 31 ,2025 or 12 months from the time funding is received.

E. TOTAL AMOUNT REQUESTED:

Total Project Request: **$6,250** to Penn State University.

No matching funds are available for this project.

F. SHORT AND LONG-TERM GOALS OF THE PROJECT:

The long-term goal of this project is to contribute to the discovery of surviving American chestnut trees, along with wind-pollinated hybrid seed, by expanding the search to historical mixed species plantings in areas outside the natural range of American chestnut and the blight.

The short-term goal is to obtain 20X depth of DNA (RADseq) sequence data from leaf tissues collected from at least 30 healthy chestnut trees growing in remnants of an over 100-year-old mixed-species orchard along the Willamette River in northwest Oregon.

G. NARRATIVE (no more than five pages)

Background

This proposal addresses the 2024 TACF priority research area of "*Discovery, study, and evaluation of "Large Surviving American" (LSA) chestnuts*". American chestnut trees cultivated in pre-blight times far outside their natural range may provide an additional source of uninfected genotypes for introduction into the ACF backcross and testing program, or as an additional source of progeny for research. In the extensive Willamette River Valley region of northwestern Oregon, where I (John Carlson) now reside, there are many large fruit and nut tree farms and orchards. In the last half of the nineteenth century, orchardists traveling to Oregon brought seed, fruit and cuttings of their favorite cultivars from eastern US farms, including American chestnut, to start farms in the Willamette Valley. The establishment of these nurseries and farms often followed large-scale clearing of the forests in the region. Most of the pioneer era fruit and nut farms in the Willamette Valley have since been converted to other uses, or the trees replaced with new varieties. American chestnut trees may have been replaced with chestnut species or hybrids, or with other fruit and nut crops, after the chestnut blight swept through the eastern US. However some survive.

In addition to trees brought by orchardists, in the late 1800s exotic (nonnative) trees, including American chestnut, were also introduced as street and landscape trees, in parks and in private gardens in Oregon, including Portland and the Willamette Valley region. TACF members are familiar with examples of old, stately American chestnut trees in Oregon from reports by Carolinas Chapter President Doug Gillis in past editions of eSprout and the Chestnut journal. In his article "American Chestnut Trees in the Pacific Northwest", in the winter 2017 issue of Chestnut, Gillis reported on six American Chestnut trees over 150 years old, some six feet in diameter and nearly 100 feet tall, that he visited in Tumwater, Washington; Sherwood, Oregon; and Portland, Oregon. Fortunately, these trees were planted in pairs and were still producing nuts. In 2021, Doug reported that the larger tree in the Sherwood, Oregon, pair had succumbed to old age, but was survived by the champion American chestnut tree in Gladstone, Oregon, at 99 feet tall, 19 feet in circumference and 6 feet in diameter, with a spread of 95 feet.

In November of 2017, I encountered chestnut trees during a walk along the Willamette River where the city of Lake Oswego, Oregon, originated. The path is paved and was originally 'Old River Road', an historic corridor for travelers to Willamette Falls, a long-standing attraction. The river road path starts at a river landing that was used by native Americans for over 10,000 years. This landing played a big role in the area's industrial origins, first as a stopping point for explorers, fur traders, and pioneers in the early 19th century, and then as a launch for boats transporting goods up and down the river. At this landing, the Oregon Iron Company built a furnace that produced iron from 1867 to 1894, which is now a national historical landmark. Now the 'Oswego Furnace', it was the first, and is the only surviving, stone blast furnace on the Pacific Coast. This site became the George Rogers Park in 1945. Also, within the current park boundaries, the Crown Willamette Paper Company erected a large cement log hoist in 1905, which operated for several decades, along with a house for the manager of the company's tugboat operation. The tug master's house is still occupied and is recognized as a local landmark (see Figure 1). There are several large chestnut trees, which appear to be *C. dentata*, on the tug master's property (Figure 1), likely planted circa 1905. These large trees produce substantial nut crops annually.

Additional groups of chestnut trees extend along on the banks of Willamette River beyond the George Rogers Park. Adjacent to the park, the first president of the Portland General Electric company, Parker Farnsworth Morey, established a large mansion and farm in the late 1800's. The 600-acre tract of land was originally forested, but heavily logged for charcoal production for the Oregon Iron Company furnace. Morey planted hundreds of forest trees, both local species and trees not native to the area, some of which remain as large specimen trees. Morey also established fruit and nut tree orchards, including chestnut trees. Quite a few of these trees still grow along the Willamette River near the neighborhood of Glenmorrie (see Figure 2) and on the campus of the Marylhurst University (Figure 3). After retiring and moving to Oregon in 2022, I confirmed that the chestnut trees remain at these sites and George Rogers Park. The trees now range from saplings to the huge original trees. American, Chinese, and Japanese chestnut are all represented (if identified correctly). The range of tree sizes and the abundance of nuts, suggest the remaining chestnut trees from the park and from the Morey orchard appear to be naturalizing.

Figure 1. The_George Rogers Park in Lake Oswego, Oregon, including historical markers, the tug master's house (and garage), the Oregon Iron Company furnace built in 1867, and large chestnut trees planted when the tug master's house was built in 1905. (my photos, Nov 24, 2022)

Figure 2. Chestnut trees and a 'carpets' of burs and leaves along Old River Road on the Willamette River, near Lake Oswego city's oldtown. (My photos from Nov. 19, 2017.)

Figure 3. Large chestnut tree (one of several) on the campus of Marylhurst University thought to be from the original fruit and nut trees planted on the estate of Parker Farnsworth Morey, circa 1905 (center image). Left and right images are close-ups of bur-laden trees from the same tree. (my photos from November 24, 2022.)

I propose to confirm the genetic identifies of these historic trees and their offspring using a Genotyping-by-Sequencing (GBS) approach in which genomic DNA sequence is generated for each tree, the species of the trees confirmed by mapping to reference genomes, levels genetic variation among the trees determined at the single nucleotides (SNP) level, and relationships among adult and juvenile trees will be uncovered through phylogenetic analyses.

Research Objectives:

- Identify surviving American chestnut trees in historical (over 100-year-old) mixed-species orchard and landscaping tree plantings outside the natural range of chestnut and the chestnut blight (near Lake Oswego, OR, on the Willamette River in northwest Oregon).
- Use a Genotyping-by-Sequencing (GBS) approach to genotype the chestnut trees resulting from the historical plantings.
- Use the GBS data to determine if younger chestnut plants in the area, that appear to be volunteers, are hybrid and if so if the parent trees in the area can be identified.
- If hybrids are identified, and TACF is interested, request permission from the city parks department for permission to collect seed to send TACF for restoration research.
- Create entries into TreeSnap for each genotyped tree, including saplings. Populate the TreeSnap entries with information on size, growth, nut production, health, phenology, etc.

Research Approach:

1. Record GPS locations and photos of chestnut trees of various ages and species along the Willamette River (Old River Road and George Rogers Park) near Lake Oswego, Oregon.

2. Collect healthy leaf samples (preferably early in season) from at least 30 chestnut trees representing different ages (sizes) and species phenotypically; Store at -20C.
3. Deliver leaf samples to the Floragenex company in Beaverton, Oregon, for DNA isolations, DNA sequencing, data QC, and SNP calling.
4. The SNP data from the Floragenex company will be analyzed to:
    a. Develop unique genotypes (marker sets) for each tree from the SNP data.
    b. Conduct cluster analyses with the SNP data to determine if species and/or family groupings can be identified, including paternity.
5. We will produce a final report and provide a TACF journal publication if results warrant. All data and results will be provided to TACF for

RADseq generation of DNA sequence data to determine genotypes

I propose to collect healthy leaves (preferably early in season) for DNA sequencing from at least 30 chestnut trees at the site along the Willamette River previously described above. For sequencing we propose to use the 'RADseq' approach which is a reduced-representation DNA sequencing technique that reduces the complexity of the data produced by targeting sequencing at a specific restriction enzyme cut sites in the genome (Andrews et al, 2016). This approach can produce DNA sequences for sites in genomes that are more likely to be shared, or at common positions, in individuals within a species. Thus, RADseq data can be more efficiently used, with less computing power, in applications such as genotyping, population genetics, and phylogenetics than whole genome sequencing data. The genomics company Floragenex (https://www.floragenex.com/) in Oregon has provided a competitive quote for generating RADseq data from up to 56 chestnut leaf samples collected in this project. Floragenex has extensive experience with trees, including large projects for the US Forest Service and the Morton Arboretum. Their comprehensive service includes extracting DNA from the samples, constructing RADseq libraries, producing app. 20X depth of sequence data on the Illumina NovaSeq 6000 S4 sequencing platform followed by quality checks and filtering of the data to remove poor or short reads, and finally 'calling' (identifying robust) SNP sites. The SNP calls and sequence data will be delivered in STACKs format for genetic analyses (Catchen et al, 2013; Rivera-Colón et al, 2022).

Genotyping and species identification by mapping SNPS to reference genomes to confirm

SNP sequences called for each tree will be aligned to the *Q. dentata*, *Q. mollissima*, and *C. crenata* reference genomes using the multiple alignment tools at the Galaxy Bioinformatics website at Penn State (https://dev.galaxy.psu.edu/). The alignment results will reveal the overall amount and patterns of sequence variation among the sampled trees relative to the reference, providing a unique genotype for each tree (Parchman et al, 2018). The species assignment for each genotype will also be determined from the summary statistics of similarity to the three reference genomes (i.e. relative number of polymorphisms) from global alignment tallies, from the Stacks program's locus-specific statistical tests, and with phylogenetic analysis programs (e.g. Hipp et al, 2014; Hipp et al, 2020).

Similarity Analysis with SNP data to determine paternity and sibship relationships

Statistical bioinformatics programs will be used to display relatedness of individuals graphically that are statistically identified based on overall similarities/differences in sequence data. SNP calls from the RADseq data will be analyzed by Principal Components Analysis (PCA), to determine to what extent individuals might cluster based on shared SNPs (e.g., Hsu, et al, 2022)

Clusters may identify families (sibs and parents) or simply group trees into the same species, or both, depending on the range of variation existing among the chestnut trees growing at the site. If hybrids exist among the cohort of younger chestnut trees sampled, these individuals could be identified as outliers between species-specific clusters.

All of the raw data, SNP call files, and results from analysis of the genotypes will be provided to the TACF for further analysis as interest dictates.

References

Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G. and Hohenlohe, P.A., 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, *17*(2), pp.81-92.

Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A. and Cresko, W.A., 2013. Stacks: an analysis tool set for population genomics. *Molecular ecology*, *22*(11), pp.3124-3140.

Rivera-Colón, A.G. and Catchen, J., 2022. Population genomics analysis with RAD, reprised: Stacks 2. In *Marine genomics: methods and protocols* (pp. 99-149). New York, NY: Springer US.

Parchman, T.L., Jahner, J.P., Uckele, K.A., Galland, L.M. and Eckert, A.J., 2018. RADseq approaches and applications for forest tree genetics. *Tree Genetics & Genomes*, *14*, pp.1-25.

Hipp AL, Eaton DA, Cavender-Bares J, Fitzek E, Nipper R, Manos PS (2014) A framework phylogeny of the American oak clade based on sequenced RAD data. PLoS One 9:e93975

Hipp, A.L., Manos, P.S., Hahn, M., Avishai, M., Bodénès, C., Cavender-Bares, J., Crowl, A.A., Deng, M., Denk, T., Fitz-Gibbon, S. and Gailing, O., 2020. Genomic landscape of the global oak phylogeny. *New Phytologist*, *226*(4), pp.1198-1212.

Hsu, Y.M., Wang, S.S., Tseng, Y.C., Lee, S.R., Fang, H., Hung, W.C., Kuo, H.I. and Dai, H.Y., 2022. Assessment of genetic diversity and SNP marker development within peanut germplasm in Taiwan by RAD-seq. *Scientific Reports*, *12*(1), p.14495.

H. TIMELINE, SHOWING START AND COMPLETION DATES FOR EACH GOAL

Project Period: January 1, 2025 to December 31, 2025.

Objectives completion dates:

Objective / Milestone	**Start Date**	**Completion Date**
6. Identify locations with multiple chestnut species, including American chestnut trees.	pre-award	completed
7. Collect young leaf samples from at least 30 chestnut trees along the Willamette River in Oregon; Store at -20C.	03-01-2025	04-30-2025
8. Deliver all leaf samples to Floragenex company.	03-30-2025	05-15-2025
9. Floragenex company conducts DNA isolations, sequencing, QC, and SNP calling.	05-15-2025	06-15-2025
10. PI conducts DNA sequence genotype data analyses.	06-15-2025	12-30-2025
11. PI submits final report for project.	01-01-2026	01-31-2026

I. HOW RESULTS WILL BE MEASURED AND REPORTED

Updates on progress in reaching milestones will be provided through regular communications with Dr. Jared Westbrook for feedback analyses most appropriate for optimal informativeness to TACF breeding objectives. Final results will be provided in reports and manuscript.

J. BREAKDOWN OF HOW AND WHEN FUNDS WILL BE SPENT

Funds will be used to collect and deliver leaf samples, and to purchase DNA isolation, sequencing and bioinformatics services from the Floragenex company in Oregon, near where the samples will be collected.

Personnel	
Salaries	$0
Fringe Benefits	$0
Subtotal Salaries, Wages, Fringe	**$0**
Total Personnel	**$0**
Other Direct Costs	
Materials & Supplies: Collection tubes (1.2 ml cryotubes in strips of 8, with storage boxes) from Qiagen company - $209. Styrofoam insulated shipping boxes for sample collection and shipping - $58,	 $209. $58.
Travel	0
Publications	0
Other Direct Costs	
Fee for service paid to Floragenex, Inc. (9590 Southwest Gemini Drive, Beaverton, OR, 97008 (541)343-0747) to produce Genotypes-by-Sequencing for 30 chestnut leaf samples collected in Oregon (i.e., isolating and sequencing DNAs from the leaf samples followed by QC and analysis of the sequence data produced).	$5,983.
Equipment	0
Total Direct Costs	**$6,250.**

Indirect Costs (refer to budget notes above)	**0**
Total Costs	**$6,250.**

L. A CONFLICT OF INTEREST OR COMMITMENT (COI OR COC) STATEMENT.

There are neither conflicts of interest or commitment for the applicant with TACF or with the Floragenex company, not other parties relevant to the proposal.

K. BRIEF CURRICULUM VITAE (CV) FOR PRINCIPAL INVESTIGATOR

John E. Carlson
326 Forest Resources Building,
Department of Ecosystem Science and Management
The Pennsylvania State University
University Park, PA 16802
Phone: (814) 883-0989; E mail: jec16@psu.edu

Education and Training:

BS, 1974, Biochemistry, Honors, University of Pittsburgh, Pittsburgh, PA
MS, 1978, Agronomy & Plant Physiology, University of Illinois, Urbana – Champaign, IL
Ph.D., 1983, Genetics, University of Illinois, Urbana – Champaign, IL
Postdoctoral Fellow, 1982-1983, Plant Pathology, Kansas State University, Manhattan, KS

Professional Positions Held:

1983-1988	Research Scientist, Allelix, Inc., Mississauga, Ontario, Canada
1988-1994	Assistant Professor, University of British Columbia, Vancouver, B.C., Canada
1994-1997	Associate Professor, University of British Columbia, Vancouver, B.C., Canada
1995-1997	Chair**,** Genetics Graduate Program**,** University of British Columbia
1997-2007	Associate Professor, tenured, School of Forest Resources and The Life Sciences Consortium, Pennsylvania State University
2007-2021	Full Professor, tenured, Department of Ecosystem Science and Management, Huck Institutes of the Life Sciences, Penn State University
2000-2021	Director, The Louis W. Schatz Center for Tree Molecular Genetics at Penn State
2010-2013	Visiting Professor, Chonnam National University, South Korea
2019-2024	Visiting Professor, Georg August University of Göttingen, Göttingen, Germany

Awards

2012, Alex and Jessie C. Black Faculty Award for Excellence in Research, College of Agricultural Sciences, Pennsylvania State University

2010, Association of Public and Land-Grant Universities Excellence in Multistate Research Award for USDA Multistate Project "Biological Improvement of Chestnut through Technologies that Address Management of the Species, its Pathogens, and Pests".

Refereed Publications related to chestnut (of 172 total peer-reviewed papers):

Research Journals

Staton ME, Addo-Quaye C, Cannon N, Yu J, Zhebentyayeva T, Huff M, Islam-Faridi N, Fan S, Georgi LL, Nelson CD, Bellis E, Fitzsimmons S, Henry N, Drautz-Moses D, Noorai RE, Ficklin S, Saski C, Mandal M, Wagner T, Zembower N, Bodénès C, Holliday J, Westbrook J, Lasky J, Hebard FV, Schuster SC, Abbott AG, Carlson JE. 2020. A reference genome assembly and adaptive trait analysis of *Castanea mollissima* 'Vanuxem', a source of resistance to chestnut blight in restoration breeding. Tree Genetics & Genomes 16:57, pages 1-23.

Staton M, Zhebentyayeva T, Olukolu B, Fang GC, Nelson D, **Carlson JE**, Abbott AG. 2015. Substantial genome synteny preservation among woody angiosperm species: comparative genomics of Chinese chestnut (*Castanea mollissima*) and plant reference genomes. BMC Genomics 16(1):1-13.

Fang GC, Blackmon BP, Staton ME, Nelson CD, Kubisiak TL, Olukolu BA, Henry D, Zhebentyayeva Y, Saski CA, Cheng CH, Monsanto M, Ficklin S, Atkins M, Georgi LL, Barakat A, Wheeler N, **Carlson JE**, Sederoff R, Abbott AG. 2013. A physical map of the Chinese chestnut (*Castanea mollissima*) genome and its integration with the genetic map. Tree Genetics & Genomes 9:525-537.

Barakat A, Staton M, Cheng C-H, Park J, Buang N, Yassin M, Ficklin S, Yeh C-C, Hebard F, Baier K, Powell W, Schuster S, Wheeler N, Abbott A, **Carlson J,** Sederoff R. 2012. Chestnut resistance to the blight disease: insights from transcriptome analysis. BMC Plant Biology, 12:38, 14 pages.

Barakat A, DiLoreto DS, Zhang Y, Smith C, Baier K, Powell W, Wheeler N, Sederoff R, **Carlson JE**. 2009. Comparison of the transcriptomes of American chestnut (*Castanea dentata*) and Chinese chestnut (*Castanea mollissima*) in response to the chestnut blight infection. BMC Plant Biol, 9:51.

Books and Book Chapters

Merkle SA, Vieitez FJ, Corredoira E, **Carlson JE**. 2020. 10.1 Castanea spp. Chestnut. *In*: R. Litz, F. Alfaro, & J. Hormaza (Eds.), Biotechnology of Fruit and Nut Crops (2nd ed.). CAB International, published January 2020, pages 206 – 237. CABI cat # 9781780648279.

Pereira-Lorenzo S, Costa R, Anagnostakis S, Serdar U, Yamamoto T, Saito T, Ramos-Cabrer AM, Ling Q, Barreneche T, Robin C, Botta R, Contessa C, Conedera M, Martín LM, Martín A, Laranjo J, Villani F, **Carlson JE**. 2016. Chapter 15 - Interspecific hybridization of chestnut. *In*: Polyploidy and Hybridisation for Crop Improvement. Published May 15, 2016 by CRC Press, 440 pages.

Nelson CD, Powell WA, Merkle SA, **Carlson JE,** Hebard FV, Islam-Faridi N, Staton ME, Georgi L. 2014. Biotechnology of Trees: Chestnut. *In*: Ramawat K, Merillon JM, Ahuja MR (eds), Tree Biotechnology, Chapter 1. CRC Press, Boca Raton, Florida, USA, April 1, 2014, pp. 3-35.

Steiner K, **Carlson J**. 2006. Proceedings of a Conference and Workshop, held May 4-6, 2004 at the North Carolina Arboretum, Steiner, KC and JE Carlson, (eds.). US Department of the Interior, National Park Service, National Capital Region, Center for Urban Ecology.

Conference Proceedings:

Carlson JE, Staton ME, Addo-Quaye C, Cannon N, Zhebentyayeva T, Islam-Faridi N, Yu J, Huff M, Fan S, Conrad, AO, Schuster SC, Abbott AG, Westbrook J, Holliday J, Nelson CD, Georgi L, Hebard FV. 2020. An Improved Chinese Chestnut Genome, *In*: Proceedings of the Sixth International Workshop on the Genetics of Host-Parasite Interactions in Forestry - Tree Resistance to Insects and Diseases: Putting Promise into Practice, eds. Nelson CD, Koch L, Sniezko RA, Mt. Sterling, Ohio, USA, 5-10 August 2018. e-Gen. Tech. Rep. SRS–252. Asheville, NC: U.S. Department of Agriculture Forest Service, Southern Research Station. p. 118.

Islam-Faridi N, Majid MA, Zhebentyayeva T, Georgi LL, Fan S, Hebard V, Sisco PH, Westbrook J, **Carlson JE**, Abbott AG, Nelson CD. 2016. FISH Confirmation of a Reciprocal Translocation in Chestnut. *In:* Cytogenetic and Genome Research 2016 Jan 1, Vol. 148, No. 2-3, pages. 144-144, Allschwilerstrasse 10, Ch-4009 Basel, Switzerland: Karger.

Nelson CD, Powell WA, Maynard CA, Baier KM, Newhouse A, Merkle SA, Nairn CJ, Kong L, **Carlson JE**, Addo-Quaye C, Staton ME, Hebard FV, Georgi LL, Abbott AG, Olukolu BA and Zhebentyayeva, T. 2014. The forest health initiative, American chestnut (*Castanea dentata*) as a model for forest tree restoration: biological research program. *In*: Proc of the 5th International Chestnut Symposium, Acta Hort 1019:179-189.

Sisco PH, Sederoff RR, Tomkins JP, **Carlson JE**, Kubisiak TL, Staton ME, Hebard FV, Anagnostakis SL, Powell WA, Smith CP. 2009. The United States National Science Foundation project on developing tools for the study of the Fagaceae: *Castanea, Quercus*, and *Fagus*. *In*: Proc of the 4th International Chestnut Symposium, Beijing, China, Sept. 25-28, 2008, Acta Hort 844:267-274.

Carlson J. 2006. Biological dimensions of the GMO issue. **In**: Proc of Conf on Restoration of American Chestnut to Forest Lands, Steiner, Carlson, (eds.), North Carolina Arboretum, Asheville, May 4-6, 2004, pp 151-158.

Liu S, **Carlson JE**. 2006. Variation in the amount of Chinese vs. American genetic material remaining in resistant backcross progeny. **In**: Proc of Conf Restoration of American Chestnut to Forest Lands, Steiner, Carlson (eds.), North Carolina Arboretum, Asheville, May 4-6, 2004, pp 133-150.

Carlson JE. 2006. Biological dimensions of the GMO issue. In Restoration of American Chestnut to Forest Lands: Proceedings of a Conference and Workshop Held May 4-6, 2004 at the North Carolina Arboretum, Steiner, KC and JE Carlson, (eds.). US Department of the Interior, National Park Service, National Capital Region, Center for Urban Ecology.