

# Genomic resources for *Macadamia tetraphylla* and an examination of its historic use as a crop resource in Hawaii

Bjarne Bartlett<sup>1</sup>, Alyssa Cho<sup>2</sup>, Daniel Laspisa<sup>1</sup>, Michael A. Gore<sup>3</sup>, Michael Kantar<sup>2\*</sup>

<sup>1</sup> Department of Molecular Biosystems and Bioengineering, University of Hawai‘i at Mānoa, Honolulu, HI 96822, USA

<sup>2</sup> Department of Tropical Plant and Soil Science, University of Hawai‘i at Mānoa, Honolulu, HI 96822, USA

<sup>3</sup> Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853, USA

\*Correspondence to [mbkantar@hawaii.edu](mailto:mbkantar@hawaii.edu)

## Abstract

*Macadamia tetraphylla* is a wild relative of the economically valuable crop *Macadamia integrifolia*. Genomic knowledge of crop wild relatives is central to determining their possible role in breeding programs to mitigate biotic and abiotic stress in the future. The goal of this project was to develop a genomic resource for macadamia agriculture in Hawai‘i through constructing a transcriptome of *M. tetraphylla* and testing for hybridity in University of Hawai‘i at Mānoa breeding material. The transcriptome assembly of *M. tetraphylla* revealed large differences in gene expression attributable to tissue type. Advanced breeding lines (HI862 and HI879) appear to be hybridized with the crop wild relative *M. tetraphylla*. Additionally, a putative *M. tetraphylla* tree sampled from a remnant orchard planting at the Waimānalo research station on Oahu did not match anecdotal accounts of the orchard as it appeared to be of hybrid ancestry.

**Keywords:** macadamia nut, hybrid, remnant orchard

**Abbreviations:** CWR - crop wild relative

# Introduction

*Macadamia nut* (*Macadamia integrifolia*;  $2n = 2x = 28$ ) is an important global crop originating in Australia (Mueller, 1858; Lin et al., 2022]. Recently, there has been increased interest in developing genomic resources for this crop (e.g. genome – Nock et al., 2020) due to its importance as a high-value, nutrient-dense food (Rengel et al., 2015; Navarro et al., 2016) that has undergone relatively little commercial breeding. In Hawai‘i alone, there has been an increase in resources dedicated to the crop due to its large land uses (17,100 acres in 2019) and high value (\$42 million USD) [6]. These efforts have identified agronomic problems in the current orchards that can be attributed to the need for genetic improvement (Gutierrez-Coarite, et al., 2021). There is, however, limited genetic diversity within University of Hawai‘i *M. integrifolia* germplasm (Hardner, 2016). Many factors threaten the livelihoods of the Hawai‘i macadamia industry including pest and disease pressure; however, there is no adapted *M. integrifolia* tolerance for the most important pressures, requiring the use of a different donor species.

In species with limited genetic diversity, the field of plant breeding commonly uses related species as a donor to improve specific traits (Castaneda-Alvarez et al., 2016). In fact, the genetic diversity from crop wild relatives have been valued at over \$100 billion U.S. dollars annually [(PwC, 2013]. The most important relative of *M. integrifolia* is *Macadamia tetraphylla* ( $2n = 2x = 28$ ), a species that is occasionally cultivated but more frequently used as a donor for useful biotic stress tolerance traits (Mueller, 1858; Nin et al., 2022). To best operationalize the use of unadapted plant material for specific localities, genomic information is required (Wambugu and Henry, 2022). While there is a history of breeding in Hawai‘i gaps in staffing have caused records to be in doubt; therefore, there is also a need to reveal genetic ancestry in current promising breeding lines. Recently, a genome of *M. tetraphylla* was released (Nin et al., 2022), providing a useful resource to characterize Hawaiian *M. tetraphylla*. Therefore, the goals of this study were to i) characterize the *M. tetraphylla* transcriptome and ii) explore the ancestry of promising University of Hawai‘i at Mānoa breeding lines.

# Materials and Methods

A transcriptome assembly of *M. tetraphylla* was constructed from existing public data along with new RNAsequencing (RNAseq) data generated from wild-collected, feral *M. tetraphylla* and putative hybrid breeding lines in Hawai‘i with RNA from leaf tissue extracted using RNAeasy Plant Mini Kit (Qiagen, Valencia, California). RNA was then sent to Novogene Bioinformatics Technology Co., Ltd., for sequencing, which was conducted on an Illumina NovaSeq with 150 bp paired-end reads. Next-generation sequencing data (PRJNA587821) was obtained for *M. tetraphylla* for five different tissue types and assembled using Trinity (Haas et al., 2013). Coding regions within transcripts were identified with TransDecoder V.5.5.0 (Haas et al., 2012). Annotation of transcripts was done with BLASTX (Camacho et al., 2009) using the UniProt database to identify homologous sequences (UniProt Consortium, 2021). The HMMER (Eddy, 2011) and PFAM (Mistry et al., 2021) databases were used to identify protein domains. Data for all analyses are reported in **Table 1**. A read set was generated using combined RNAseq data from 5 tissues (bark, proteoid root, flowering inflorescence, young inflorescence, and leaves; SRR10424518-SRR10424522) that were trimmed, and quality checked using the same methods as for the Trinity assembly. The combined read set was mapped to the *M. tetraphylla* genome

assembly (ASM2298504v1) using TopHat2 (default parameters). The resulting alignment file was filtered for unmapped reads (samtools view -F 0x04), sorted, and indexed using samtools. Expression of transcripts was determined by assessing FPKM (fragments per kilobase million), a transcript was considered expressed if FPKM > 10. Annotations were generated from the sorted alignment file using Cufflinks 2.2.1 (default parameters), which generated the GTF file used in the analysis. The alignment file and annotations were manually inspected in IGV for validation. The tool sppIDer (Langdon et al., 2018) was used to assess hybridity.

**Table 1:** Overview of data files.

| Label      | Name of data file                        | File type | Data repository and identifier                                                                                                                                                                                                                                                                                                                                                                                                                  |
|------------|------------------------------------------|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Data set 1 | SRR10424522_leaves.fasta                 | .fasta    | NCBI Sequence Read Archive:<br><a href="https://identifiers.org/ncbi/insdc.sra:SRR10424522">https://identifiers.org/ncbi/insdc.sra:SRR10424522</a><br>Figshare:<br><a href="https://doi.org/10.6084/m9.figshare.21084853.v1">https://doi.org/10.6084/m9.figshare.21084853.v1</a><br>(Bartlett et al., 2022a)<br>Zenodo:<br><a href="https://doi.org/10.5281/zenodo.7135297">https://doi.org/10.5281/zenodo.7135297</a> (Bartlett et al., 2022b) |
| Data set 2 | SRR10424521_younginflorescence.fasta     | .fasta    | NCBI Sequence Read Archive:<br><a href="https://identifiers.org/ncbi/insdc.sra:SRR10424521">https://identifiers.org/ncbi/insdc.sra:SRR10424521</a><br>Figshare:<br><a href="https://doi.org/10.6084/m9.figshare.21084853.v1">https://doi.org/10.6084/m9.figshare.21084853.v1</a><br>(Bartlett et al., 2022a)<br>Zenodo:<br><a href="https://doi.org/10.5281/zenodo.7135297">https://doi.org/10.5281/zenodo.7135297</a> [21]                     |
| Data set 3 | SRR10424520_floweringinflorescence.fasta | .fasta    | NCBI Sequence Read Archive:<br><a href="https://identifiers.org/ncbi/insdc.sra:SRR10424520">https://identifiers.org/ncbi/insdc.sra:SRR10424520</a><br>Figshare:<br><a href="https://doi.org/10.6084/m9.figshare.21084853.v1">https://doi.org/10.6084/m9.figshare.21084853.v1</a><br>(Bartlett et al., 2022a)<br>Zenodo:<br><a href="https://doi.org/10.5281/zenodo.7135297">https://doi.org/10.5281/zenodo.7135297</a> [21]                     |
| Data set 4 | SRR10424519_proteoidroot.fasta           | .fasta    | NCBI Sequence Read Archive:<br><a href="https://identifiers.org/ncbi/insdc.sra:SRR10424519">https://identifiers.org/ncbi/insdc.sra:SRR10424519</a><br>Figshare:<br><a href="https://doi.org/10.6084/m9.figshare.21084853.v1">https://doi.org/10.6084/m9.figshare.21084853.v1</a><br>(Bartlett et al., 2022a)<br>Zenodo:<br><a href="https://doi.org/10.5281/zenodo.7135297">https://doi.org/10.5281/zenodo.7135297</a> [21]                     |
| Data set 5 | SRR10424518_bark.fasta                   | .fasta    | NCBI Sequence Read Archive:<br><a href="https://identifiers.org/ncbi/insdc.sra:SRR10424518">https://identifiers.org/ncbi/insdc.sra:SRR10424518</a><br>Figshare:<br><a href="https://doi.org/10.6084/m9.figshare.21084853.v1">https://doi.org/10.6084/m9.figshare.21084853.v1</a><br>(Bartlett et al., 2022a)<br>Zenodo:<br><a href="https://doi.org/10.5281/zenodo.7135297">https://doi.org/10.5281/zenodo.7135297</a> [21]                     |
| Data set 6 | MT1_879.vcf.gz                           | .vcf.gz   | Figshare:<br><a href="https://doi.org/10.6084/m9.figshare.21084853.v1">https://doi.org/10.6084/m9.figshare.21084853.v1</a><br>(Bartlett et al., 2022a)                                                                                                                                                                                                                                                                                          |

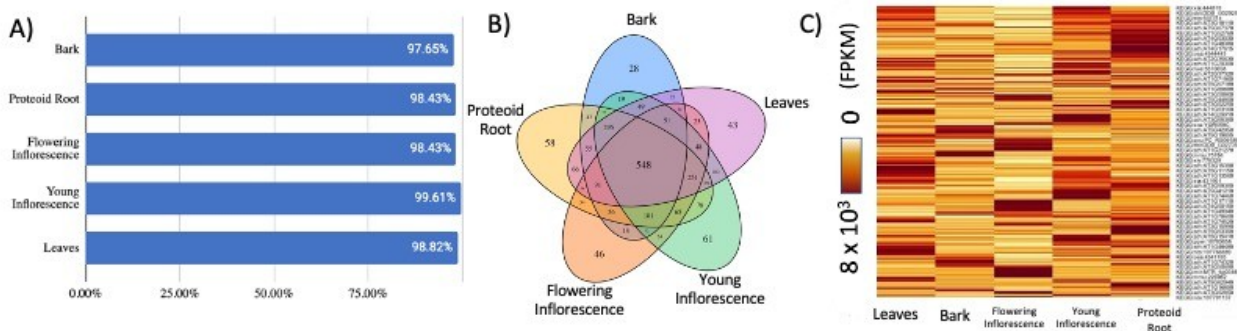
|             |                                  |         |                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|-------------|----------------------------------|---------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
|             |                                  |         | Zenodo:<br><a href="https://doi.org/10.5281/zenodo.7135297">https://doi.org/10.5281/zenodo.7135297</a> (Bartlett et al., 2022b)<br>European Variation Archive:<br><a href="https://identifiers.org/ebi/bioproject:PRJEB56438">https://identifiers.org/ebi/bioproject:PRJEB56438</a> )                                                                                                                                                           |
| Data set 7  | MT5_OT.vcf.gz                    | .vcf.gz | Figshare:<br><a href="https://doi.org/10.6084/m9.figshare.21084853.v1">https://doi.org/10.6084/m9.figshare.21084853.v1</a><br>(Bartlett et al., 2022a)<br>Zenodo:<br><a href="https://doi.org/10.5281/zenodo.7135297">https://doi.org/10.5281/zenodo.7135297</a> (Bartlett et al., 2022b)<br>European Variation Archive:<br><a href="https://identifiers.org/ebi/bioproject:PRJEB56438">https://identifiers.org/ebi/bioproject:PRJEB56438</a> ) |
| Data set 8  | MT9_OTPH.vcf.gz                  | .vcf.gz | Figshare:<br><a href="https://doi.org/10.6084/m9.figshare.21084853.v1">https://doi.org/10.6084/m9.figshare.21084853.v1</a><br>(Bartlett et al., 2022a)<br>Zenodo:<br><a href="https://doi.org/10.5281/zenodo.7135297">https://doi.org/10.5281/zenodo.7135297</a> (Bartlett et al., 2022b)<br>European Variation Archive:<br><a href="https://identifiers.org/ebi/bioproject:PRJEB56438">https://identifiers.org/ebi/bioproject:PRJEB56438</a>   |
| Data set 9  | MT10_862.vcf.gz                  | .vcf.gz | Figshare:<br><a href="https://doi.org/10.6084/m9.figshare.21084853.v1">https://doi.org/10.6084/m9.figshare.21084853.v1</a><br>(Bartlett et al., 2022a)<br>Zenodo:<br><a href="https://doi.org/10.5281/zenodo.7135297">https://doi.org/10.5281/zenodo.7135297</a> (Bartlett et al., 2022b)<br>European Variation Archive:<br><a href="https://identifiers.org/ebi/bioproject:PRJEB56438">https://identifiers.org/ebi/bioproject:PRJEB56438</a>   |
| Data set 10 | MT1_1.fastq.gz<br>MT1_2.fastq.gz | .fq.gz  | NCBI Sequence Read Archive:<br><br><a href="https://identifiers.org/ncbi/insdc.sra:SRR21236275">https://identifiers.org/ncbi/insdc.sra:SRR21236275</a>                                                                                                                                                                                                                                                                                          |
| Data set 11 | MT10_1.fq.gz<br>MT10_2.fq.gz     | .fq.gz  | NCBI Sequence Read Archive:<br><br><a href="https://identifiers.org/ncbi/insdc.sra:SRR21236272">https://identifiers.org/ncbi/insdc.sra:SRR21236272</a>                                                                                                                                                                                                                                                                                          |
| Data set 12 | MT9_1.fq.gz<br>MT9_2.fq.gz       | .fq.gz  | NCBI Sequence Read Archive:<br><br><a href="https://identifiers.org/ncbi/insdc.sra:SRR21236273">https://identifiers.org/ncbi/insdc.sra:SRR21236273</a>                                                                                                                                                                                                                                                                                          |
| Data set 13 | MT5_1.fq.gz<br>MT5_2.fq.gz       | .fq.gz  | NCBI Sequence Read Archive:<br><br><a href="https://identifiers.org/ncbi/insdc.sra:SRR21236274">https://identifiers.org/ncbi/insdc.sra:SRR21236274</a>                                                                                                                                                                                                                                                                                          |

## Results

### Differential gene expression across tissue type

The *de novo* *M. tetraphylla* transcriptomes showed a high level of completeness (**Figure 1a**) using Benchmarking Universal Single-Copy Orthologous (BUSCO) for bark (97.65% complete), proteoid root (98.43% complete), flowering inflorescence (98.43% complete), young inflorescence (99.61% complete), and leaves (98.82% complete). For validation and annotation, the assembled sequence was compared to the KEGG and PFAM databases (Mistry et al., 2021; Ogata et al., 1998). Similar to a previous report (Niu et al., 2022), a total of 46,967 transcripts were called representing 34,444 unigenes and 12,523 isoforms. As expected, there was tissue specific expression of transcripts and pathways (FPKM > 10 **Figure 1b, c**).

**Figure 1.** A) The *de novo* *M. tetraphylla* transcriptomes showed high completeness based on Benchmarking Universal Single-Copy Orthologous (BUSCO) quality assessments. B) Number of genes expressed in different tissues by KEGG Pathway in *M. tetraphylla*. Total number of genes with FPKM > 10 across proteoid root, flowering inflorescence, young inflorescence, and leaves. C) Gene expression of KEGG pathways represented as FPKM (fragments per kilobase million) in *M. tetraphylla* for leaves, bark, flowering inflorescence, young inflorescence, and proteoid root.

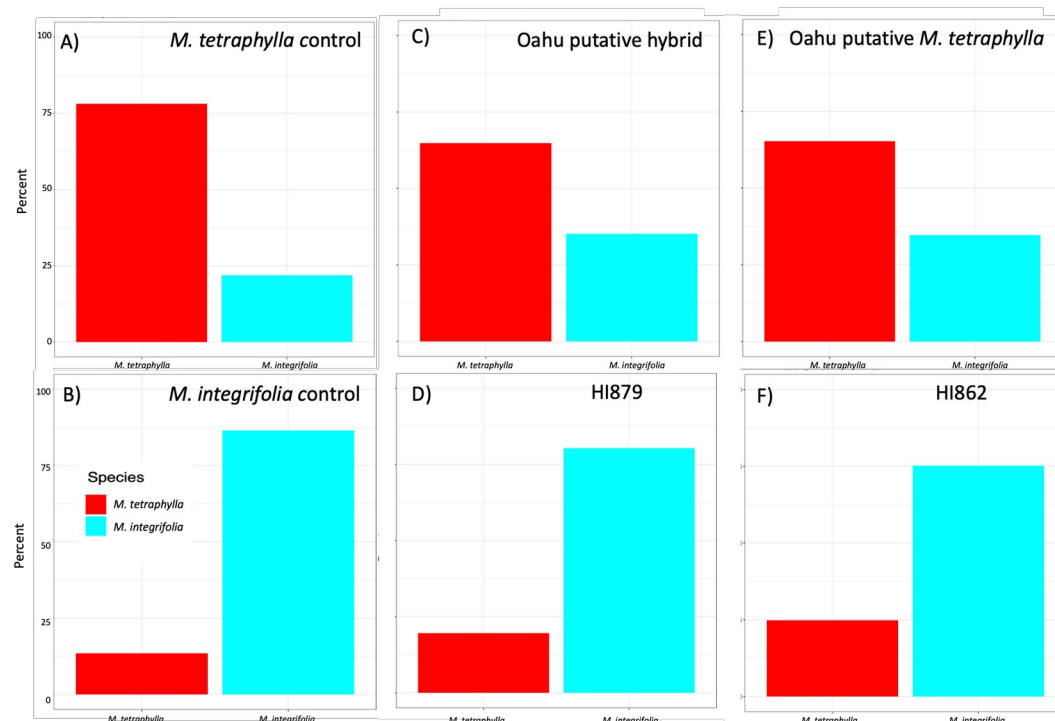


### Hybridity of Hawai'i Collected Individuals

Combining the information from the *de novo* assembly and the previous sequence showed that the new assembly has potential to inform *M. integrifolia* breeding. We identified control samples from previously published data to calibrate putative hybridization (**Figure 2A, B**). We sampled a tree (Oahu Putative Hybrid from the Tantalus Lookout Puu Ualakaa State Park - **Figure 2C**) from a remnant orchard planting that showed ancestry percentages supporting a hybrid origin. However, its percentage of *M. integrifolia* ancestry was estimated to be ~35%, suggesting that this putative hybrid is an F<sub>1</sub> with some segregation distortion or a late generation hybrid. Surprisingly, the *M. tetraphylla* from a remnant planting at Waimānalo (**Figure 2E**) also was inferred to be a hybrid, again likely an F<sub>1</sub> or a later generation hybrid. Additionally, we found close to the expected ancestry percentage (25%) for two putative advanced BC<sub>1</sub> ((*M. tetraphylla* x *M. integrifolia*) x *M. integrifolia*) breeding lines (HI862 - 19.6% and HI879 - 24.8%) from the

UH Mānoa breeding program (**Figure 2D and F; Supplemental Table 1**). However, the 13.4% *M. tetraphylla* in the *M. integrifolia* control (Figure 2B) may instead indicate that the breeding lines are BC<sub>2</sub> rather than BC<sub>1</sub>. Introgressions from *M. tetraphylla* appeared to be consistent within HI862 and HI879, occurring on chromosome 2, 5, and 9 (**Supplemental Figure 1**).

**Figure 2.** Hybridity of Hawai‘i collected individuals. A) *M. tetraphylla* control (SRR10424522), B) *M. integrifolia* control (PRJNA593881), C) Putative hybrid collected from a remnant orchard planting in Tantalus Lookout Puu Ualakaa State Park on Oahu, D) Putative hybrid breeding line from UH Mānoa, E) *M. tetraphylla* collected from the UH Mānoa Waimānalo research station on Oahu, and F) Putative hybrid breeding line from UH Mānoa.



## Discussion

It is now a routine procedure to construct genomic resources for non-model crop species and their relatives (Wambugu and Henry, 2022). This is an important step in increasing the efficiency of interspecific breeding, especially in species that have long generation time (Valle-Echevarria et al., 2021). Transcriptomes provide insight into gene function and are a potentially rich source of information on crop relatives. Here, we showed considerable tissue-specific expression across functional pathways. Furthermore, we found evidence of hybridization not only in breeding lines, but also unexpectedly for a putative *M. tetraphylla* in remnant orchard material at the Waimānalo research station on Oahu. Although the breeding lines had the expected levels of *M. tetraphylla* when compared to pedigrees, additional generations of backcrossing to *M. integrifolia* could be needed for production in commercial Hawaiian orchards. There are some limitations associated with our data, (1) short read data (~150bp) has limited ability to explore multiple isoforms, (2) we only have single control for each sample, and (3) the limited sampling in remnant seed orchards prevented a more complete characterization of potential genetic

variation for introgressions. Wild relatives of crops have huge potential value, but in order to maximize the ability to use them they will need to be extensively characterized.

## Declarations

**Availability of data and materials:** The data described here can be freely and openly accessed at: <https://doi.org/10.6084/m9.figshare.21084853.v1> -- (Transcriptome Leaf of *M. tetraphylla*, Transcriptome young inflorescence of *M. tetraphylla*, Transcriptome flowering inflorescence of *M. tetraphylla*, Transcriptome proteoid root of *M. tetraphylla*, Transcriptome bark of *M. tetraphylla*, VCF file of *M. integrifolia* × *M. tetraphylla* hybrid Hawaii 879, VCF file of Oahu *M. tetraphylla*, VCF file of Oahu putative *M. integrifolia* × *M. tetraphylla* hybrid, VCF file of *M. integrifolia* × *M. tetraphylla* hybrid Hawaii 862 – [20]) and at Zenodo. <https://doi.org/10.5281/zenodo.7135297> -- (Transcriptome Leaf of *M. tetraphylla*, Transcriptome young inflorescence of *M. tetraphylla*, Transcriptome flowering inflorescence of *M. tetraphylla*, Transcriptome proteoid root of *M. tetraphylla*, Transcriptome bark of *M. tetraphylla*, VCF file of *M. integrifolia* × *M. tetraphylla* hybrid Hawaii 879, VCF file of Oahu *M. tetraphylla*, VCF file of Oahu putative *M. integrifolia* × *M. tetraphylla* hybrid, VCF file of *M. integrifolia* × *M. tetraphylla* hybrid Hawaii 862 [21]). DNA from MT1, MT5, MT9 and MT10 upon reasonable request.

The variant data for this study have been deposited in the European Variation Archive (EVA) at EMBL-EBI under accession number PRJEB56438 (<https://identifiers.org/ebi/bioproject:PRJEB56438> – [22])

*Raw sequence data are available at:*

Yunnan Institute of Tropical Crops: RNA seq of *Macadamia tetraphylla*: leaves 2019; NCBI Sequence Read Archive: <https://identifiers.org/ncbi/insdc.sra:SRR10424522> [23]

Yunnan Institute of Tropical Crops: RNA seq of *Macadamia tetraphylla*9: young inflorescence; NCBI Sequence Read Archive: <https://identifiers.org/ncbi/insdc.sra:SRR10424521> [24]

Yunnan Institute of Tropical Crops: RNA seq of *Macadamia tetraphylla*: flowering inflorescence 2019; NCBI Sequence Read Archive: <https://identifiers.org/ncbi/insdc.sra:SRR10424520> [25]

Yunnan Institute of Tropical Crops: RNA seq of *Macadamia tetraphylla*: proteoid root 2019; NCBI Sequence Read Archive: <https://identifiers.org/ncbi/insdc.sra:SRR10424519> [26]

Yunnan Institute of Tropical Crops: RNA seq of *Macadamia tetraphylla*: bark 2019; NCBI Sequence Read Archive: <https://identifiers.org/ncbi/insdc.sra:SRR10424518> [27]

University of Hawaii: Hawaii 879 2022; NCBI Sequence Read Archive: <https://identifiers.org/ncbi/insdc.sra:SRR21236275> [28]

University of Hawaii: Hawaii 862 2022; NCBI Sequence Read Archive: <https://identifiers.org/ncbi/insdc.sra:SRR21236272> [29]

University of Hawaii: Oahu putative hybrid *Tantalus* 2022; NCBI Sequence Read Archive: <https://identifiers.org/ncbi/insdc.sra:SRR21236273> [30]

University of Hawaii: Oahu *Tetraphylla* Waimanalo 2022; NCBI Sequence Read Archive: <https://identifiers.org/ncbi/insdc.sra:SRR21236274> [31]

**Competing interests:** No competing interests

**Funding:** The authors would like to thank the University of Hawai'i Team Science Grant. Additional funding provided by The College of Tropical Agriculture and Human Resources, University of Hawai'i at Mānoa; USDA Cooperative State Research, Education and Extension (CSREES), Grant/Award Number: HAW00942-H and HAW08039-H.

**Authors' contributions:** BB: Analysis, data generation, drafting, revision, AC: Conceptualization, writing/revision, DL: Analysis, writing/revision, MK: Conceptualization, writing/revision, MG: Conceptualization, writing/revision

**Acknowledgements:** We would like to thank Cornell University for supporting the sabbatical of Dr. Michael A. Gore so he could contribute to this manuscript.

# References

Bartlett, B, Cho, A, Laspisa, D, Gore, M, Kantar,M. 2022. Genomic resources for Macadamia tetraphylla and an examination of its historic use as a crop resource in Hawai'i. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.21084853.v1>

Bartlett, B, Cho, A, Laspisa, D, Gore, M, Kantar,M. (2022). Genomic resources for Macadamia tetraphylla and an examination of its historic use as a crop resource in Hawaii [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.7135297>

Castañeda-Álvarez NP, Khoury CK, Achicanoy HA, Bernau V, Dempewolf H, et al. Global conservation priorities for crop wild relatives. Nature plants. 2016 Mar 21;2(4):1-6.

Eddy SR. Accelerated profile HMM searches. PLoS computational biology. 2011 Oct 20;7(10):e1002195

Gutierrez-Coarite R, Cho AH, Mollinedo J, Pulakkatu-Thodi I, Wright MG. Macadamia felted coccid impact on macadamia nut yield in the absence of a specialized natural enemy, and economic injury levels. Crop Protection. 2021 Jan 1;139:105378.

Hardner C. Macadamia domestication in Hawai 'i. Genetic resources and crop evolution. 2016 Dec;63(8):1411-30.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. Nature Protocols, 8(8), 2013. 1494-1512.

Haas, B., & Papanicolaou, A. (2012). TransDecoder (Find Coding Regions within Transcripts) [WWW Document]. URL <https://transdecoder.github.io>.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009. 15;10:421.

Langdon QK, Peris D, Kyle B, Hittinger CT. sppIDer: a species identification tool to investigate hybrid genomes with high-throughput sequencing. Molecular biology and evolution. 2018 Nov 1;35(11):2835-49.

Lin J, Zhang W, Zhang X, Ma X, Zhang S, et al. Signatures of selection in recently domesticated macadamia. Nature communications. 2022 Jan 11;13(1):1-2.

Mistry J, Chuguransky S, Williams L, Qureshi M, Salazar GA, et al. Pfam: The protein families database in 2021. Nucleic acids research. 2021 Jan 8;49(D1):D412-9.

Mueller, Ferdinand. Philosophical Institute of Victoria. (1858). Transactions of the Philosophical Institute of Victoria from January to December. (Vol. 2). The Institute.

Niu Y, Li G, Ni S, He X, Zheng C, et al. The chromosome-scale reference genome of *Macadamia tetraphylla* provides insights into fatty acid biosynthesis. *Frontiers in genetics*. 2022;13.

Nock CJ, Baten A, Mauleon R, Langdon KS, Topp B, et al. Chromosome-scale assembly and annotation of the macadamia genome (*Macadamia integrifolia* HAES 741). *G3: Genes, Genomes, Genetics*. 2020 Oct 1;10(10):3497-504.

Navarro SL, Rodrigues CE. Macadamia oil extraction methods and uses for the defatted meal byproduct. *Trends in Food Science & Technology*. 2016 Aug 1;54:148-54.

Ogata H, Goto S, Fujibuchi W, Kanehisa M. Computation with the KEGG pathway database. *Biosystems*. 1998 Jun 1;47(1-2):119-28.

PwC, P. (2013). Crop Wild Relatives: A valuable resource for crop development. [www.pwc.co.uk/valuations](http://www.pwc.co.uk/valuations). Available online: <https://pwc.blogs.com/files/pwc-seed-bank-analysis-for-msb-0713.pdf> (accessed on 15 November 2021).

Rengel A, Pérez E, Piombo G, Ricci J, Servent A, et al. Lipid profile and antioxidant activity of macadamia nuts (*Macadamia integrifolia*) cultivated in Venezuela. *Nat Sci*; 2015. 7: 535–547.

The UniProt Consortium, UniProt: the universal protein knowledgebase in 2021, *Nucleic Acids Research*, Volume 49, Issue D1, 8 January 2021, Pages D480–D489, <https://doi.org/10.1093/nar/gkaa1100>

[USDA-NASS] United States Department of Agriculture- National Agricultural Statistics Service. 2019. Pacific region- Hawaii macadamia nuts final season estimates. [https://www.nass.usda.gov/Statistics\\_by\\_State/Hawaii/Publications/Fruits\\_and\\_Nuts/072019MacNutFinal.pdf](https://www.nass.usda.gov/Statistics_by_State/Hawaii/Publications/Fruits_and_Nuts/072019MacNutFinal.pdf) Accessed 11 September 2021

Valle-Echevarria AD, Fumia N, Gore MA, Kantar M. Accelerating crop domestication in the era of gene editing. *Plant breeding reviews*. 2021 Dec 28;45:185-211.

Wambugu PW, Henry R. Supporting in situ conservation of the genetic diversity of crop wild relatives using genomic technologies. *Molecular Ecology*. 2022 Apr;31(8):2207-22.

