# Development and validation of *Acacia koa* and *A*. *koaia* nuclear SSRs using Illumina sequencing

Shaneka S. Lawson<sup>1\*</sup>, Aziz Ebrahimi<sup>2</sup>

<sup>1</sup> USDA Forest Service, Northern Research Station, Hardwood Tree Improvement and Regeneration Center (HTIRC), Purdue University, Department of Forestry and Natural Resources, 715 West State Street, West Lafayette, IN 47907 <sup>2</sup> Purdue University, Forestry and Natural Resources, Pfendler Hall, 715 West State Street, West Lafayette, Indiana 47907

\* Corresponding author: Shaneka S. Lawson, Email: sslawson@fs.fed.us

# Abstract

Koa (*Acacia koa*) and sub-species koaia (*A. koaia*) are two of more than 1,200 species from the genus *Acacia* within the *Leguminosae* (also designated Fabaceae) family. In the past, koa and koaia forests were found throughout the Hawaiian archipelago but populations have dramatically decreased. Comprehensive analyses of simple sequence repeats (SSRs) have not been published previously. Here we use genome sequencing and bioinformatics tools to report development of 100,000+ nuclear SSR (nuSSR) markers for use in koa and koaia genome studies with transcriptome SSR information was included for comparison. Over 10,000 high-value SSRs (40-60 % GC content) were isolated with 3,600+ further validated by ePCR. SSRs generated in this work can assist current efforts to sustainably increase in koa and koaia populations.

**Key words**: Acacia koa, Acacia koaia, ePCR, Hawaii, SSRs, tropical hardwood tree

# Introduction

Koa (*Acacia koa*) and koaia (*A. koaia*) are native Hawaiian hardwood trees sacred to the Hawaiian culture and members of the *Leguminosae* family. The primary factor limiting molecular marker use in these species is absence of publicly available genome sequences and lack of high density linkage maps. Simple sequence repeats (SSRs) can be used in construction of genetic linkage maps and identification of quantitative trait loci (QTL), in marker-assisted selection (MAS), structure analysis, and as primers to amplify genomic regions between SSR loci. Some of the earliest uses of these technologies in plants are Morgante and Olivieri (1993) and Young (1996). These markers provide effective means for investigating genetic diversity and accelerating genome studies. Construction and development of SSRs provide a valuable resource for researchers and are essential for subsequent studies of breeding and genetic diversity. The majority of marker-based research employed microsatellites for analysis of genetic differentiation within koa and koaia populations (Fredua-Agyeman et al. 2008) and amplified fragment length polymorphisms (AFLPs) or quantitative reverse-transcriptase PCR (qRT-PCR) for pathogenicity studies (Shiraishi et al. 2012; Rushanaedy et al. 2012). Fredua-Agyeman et al. (2008) and Adamski et al. (2012, 2013) first reported use of SSRs in koa and presented 31 primer pairs for further utilization in genetic diversity and disease resistance studies. Recent publications on SSR development for other Acacia spp. have generated a genetic linkage map for A. mangium (Butcher and Moran 2000), polymorphic microsatellite loci for the hybrid A. mangium × A. auriculiformis (Ng et al. 2005), A. brevispica (Otero-Arnaiz et al. 2005), A. mellifera (Senegalia) (Ruiz-Guajardo et al. 2007), A. saligna (Labill.) (Millar and Byrne 2007), A. karina, A. stanleyi, and A. jibberdingensis (Nevill et al. 2010), A. dealbata (Guillemaud et al. 2015), A. mangium, A. auriculiformis, and A. mangium × A. auriculiformis (Le et al. 2016), comparisons to 454 and pyrosequencing in A. harpophylla F. Muell. Ex Benth (Lepais and Bacles 2011), A. atkinsiana (Levy et al. 2014), A. montana (Hopley et al. 2015), and paternity analyses in A. saligna (Saligna) (Millar et al. 2008).

We present development of a wealth of nuSSR markers for koa and koaia to supplement the shortage of molecular marker data currently available. Past efforts in this arena required use of magnetic beads and the 454 sequencing platform however, following Staton *et al.* (2015), we demonstrate a newer method to improve upon numbers of markers generated. We have described significantly more SSR primer options than those presented in studies of other *Acacia spp*. The nuSSRs identified here were characterized by motif and provide amplification points within both genomes. Future koa and koaia work can use these data for in-depth genomic studies like marker-assisted selection (MAS) or linkage maps as numbers of publically available SSRs are sparse.

# Material and Methods

## Plant materials

Phyllode samples were collected on Hawaii Island from 10 mature healthy koa, 10 koaia, and a single aberrant koaia phenotype koaia-A. These tissues were taken from the most outward facing branch in the middle of the tree and immediately frozen on dry ice before being shipped to the Plant Tissue and Genomics laboratory at Purdue University (https://ag.purdue.edu/fnr/Pages/labtissue.aspx) for nucleic acid extraction.

### DNA / RNA extraction

DNA was extracted from phyllodes using the DNeasy Plant Maxi kit (Qiagen<sup>®</sup>) according to manufacturer's instructions except the following two modifications: (1) frozen phyllodes from each species were ground into powder, pooled, and suspended in 500  $\mu$ L of chilled 100 % ethanol for 15 min before adding the lysis buffer and (2) extracted DNA was eluted in 20  $\mu$ L of sterile water. DNA quantity and quality were determined by measuring the absorbance at 260 nm and the 260/280 nm ratio using a UV spectrophotometer. RNA was extracted from phyllodes using the RNeasy<sup>®</sup> Plant Mini Kit (QIAGEN<sup>®</sup>, Germany) and quantified using the Nanodrop 8000 (Thermo Fisher Scientific Inc., USA) after addition of 1  $\mu$ l DNase<sup>®</sup> (Promega Corporation, USA). RNA quality was ascertained by a RNA 6000 nano chip (Agilent Technologies, USA) with final quality confirmed by electropherogram.

### DNA sequencing, Read quality and Mapping

Genomic *A. koa* and *A. koaia* DNA samples were sequenced at the Purdue Genomics Core Facility (https://www.purdue.edu/ hla/sites/genomics/) using MiSeq (Illumina®) after paired-end library generation (Illumina® TruSeq DNA PCR-Free Library Preparation Kit). Trimmomatic software was used to trim low quality data. Quality control was carried out using FastQC. FastQC (v.0.11.2) and the FASTX toolkit (v.0.013.2) were used to assess sample quality and to execute quality trimming. Bases with Pfred33 scores below 30 were removed. Reads with greater than 50 bases (99 % of total reads) were kept for further downstream analysis. Bowtie2 (v.2.2.6) was used to map quality trimmed reads against the de novo assembled transcriptome for koa using default parameters.

# RNA library construction, De novo assembly and annotation

The cDNA library was generated using the Illumina RNA TruSeq kit (Illumina<sup>®</sup>, USA) and 2  $\mu$ g of high quality RNA. Library

quality was determined with the Agilent Bioanalyzer 2100 (Agilent Technologies, USA) before reverse transcription and sequencing on the Illumina HiSeq2000 platform. *De novo* assembly using Transcriptome Assembly By Short Sequences (Trans-ABySS, v.1.5.3) was employed to generate a reference transcriptome. The *de novo* transcriptome was assembled using Trinity (v.3.0) was performed on the short RNA sequences to generate a reference transcriptome containing reference contigs.

### SSR development and validation

SSR primers for koa and koaia were designed according to the Perl pipeline described in Staton *et al.* (2015). A total of 130,000 SSR motifs were found for the koa and koaia genomes and 5,300 for the koa transcriptome. Primers were sorted by repeat number (>10 repeats) and GC content (40-60 %). Designed primers were validated using ePCR and parameters set forth by Shyu *et al.* (2002) (Figure 1).

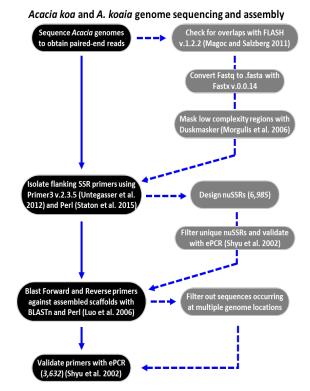


Figure 1

Flow chart. Steps in sequencing the *A. koa* and *A. koaia* genomes. Major steps (solid line); Supporting steps (dotted line)

# Estimate of divergence between flanking region sequences

Raw read data were converted from FASTQ to .fasta format following previously established protocols from Staton *et al.* (2015). Forward and reverse reads were evaluated for overlap in the expected orientation and FLASH (v.1.2.2) software was used to reconstruct the original fragment (Magoc and Salzberg 2011). SSR flanking regions of koa and koaia were obtained from genomic analyses and compared to Glycine max data downloaded from the soybean genome database (http:// www.plantgdb.org/GmGDB/). Flanking region SSRs were aligned and pairwise distances were computed with MEGA6 (v.6.0) software (Tamura *et al.* 2013). Based on pairwise distances a phylogeny tree was constructed with NTSYS (v.2.2) software (Rohlf, 2000).

# Results

### SSR motifs in koa and koaia genomes

Nearly 1.4 % of koa and 0.9 % of koaia sequences exhibited SSRs during analyses of 7.1 million koa and 2.9 million koaia genome sequences (Table 1). A total of 102,656 SSR motifs from koa and 28,275 SSR motifs from koaia were identified after assembly of the koa and koaia genomes. There were 6,985 nuSSRs elucidated from both species. (Table 2).

### Table 1

*A. koa* and *A. koaia* metadata. Read information from initial NextGeneration Sequencing and assembly.

	Raw Reads		Clipped Reads		De novo assembly (ABySS)				
	Total Reads	Max Length	Quality Reads	% Passing QC	Sequences	Min Length	Max Length	% Complete Genes	
A. koa	534,840,500	101	522,220,850	97.62	589,165	500	91,206	83	
A. koaia	179,313,476	101	174,303,956	97.21	308,744	500	25,333	87	
Average	357,076,988	101	348,262,403	97	448,955	500	58,270	85	

#### Table 2

SSR motif information obtained from paired-end A. koa and A. koaia sequences.

	Motif Information	Species I	nformation
Mei	tadata	Acacia koa	Acacia koaia
	Total number of sequences analyzed	7,123,041	2,906,140
	Number of sequences with >1 SSR	98,119	27,083
	Total numbers of SSRs	102,656	28,275
	Number of nuclear SSRs	5,566	1,419
	Number of SSRs with primers	8,482	2,587
Din	ucleotides		
1	AT TA	73,381	3,910
2	GC CG	11	-
3	AC CA TG GT	10,139	1,128
4	AG GA CT TC	4,888	1,267
Trir	nucleotides		
1	GGC GCG CGG GCC CCG CGC	35	29
2	ATG TGA GAT CAT ATC TCA	190	99
3	AGT GTA TAG ACT CTA TAC	724	103
4	AGG GAG GGA CCT CTC TCC	180	103
5	AAT ATA TAA ATT TTA TAT	3,747	1,214
6	CCA CAC ACC TGG GTG GGT	170	49
7	AGC GCA CAG GCT CTG TGC	26	20
8	AAG AGA GAA CTT TTC TCT	774	454
9	AAC ACA CAA GTT TTG TGT	247	184
10	ACG CGA GAC CGT GTC TCG	10	54
Mo	tif Length Excluding Compound SSRs		
	2 bp	88,419	23,516
	3 bp	6,103	2,309
	4 bp	2,568	1,031
SSR	s with Primers Excluding Compound SSRs		
	2 bp	7,797	2,277
	3 bp	583	248
	4 bp	102	62

Characterization of SSRs in koa and koaia genomes Dinucleotide repeats were most abundant, with AT/TA motifs accounting for 71.4 % of all koa SSRs and 13.8 % of all koaia SSRs within the genome (Figure 2). The 2 bp dinucleotide repeats represented 86.1 % of all koa and 83.2 % of all koaia SSRs. The AAT|ATA|TAA|ATT|TTA|TAT| trinucleotide repeat was the most represented of the 3 bp repeats and accounted for 61.4 % of koa and 52.6 % of koaia SSRs. Tetranucleotide motifs accounted for 2.5 and 3.6 % of koa and koaia SSRs, respectively. Primers were generated for 7,797 koa and 2,277 koaia dinucleotide and 583 koa and 248 koaia trinucleotide motifs. Interestingly, with 9.5 and 10.7 %, a greater percentage of primers could be designed for koa and koaia trinucleotide motifs. This higher success rate corresponded to 1 % for 2 bp motifs and 5.6 and 3.7 % for 4 bp motifs in koa and koaia, respectively.

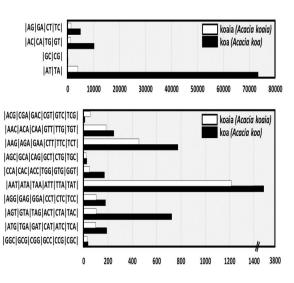


Figure 2

nuSSR counts. Numbers of dinucleotide and trinucleotide nuSSRs found in *A. koa* and *A. koaia* genomic sequences.

### SSR development and validation

Approximately 130,931 SSRs were designed in this study, 78 % of them from koa. Greater than 10,000 nuSSR markers were successfully designed for this study however, after removal of duplicates and overlaps, 6,985 remained. As expected, the koa genome had 77 % of the SSRs with primers. The koaia genome resulted in 2,587, or 23 %, of SSRs with primers. Validation by ePCR restricted high-quality primers to 3,632 markers with specific amplicons and multiple binding site locations. (Table 3). Of these, 90.2 % of high-quality primers were from the koa genome and 9.8 % from the koaia genome (Table S1). Comparison of nuSSRs and EST-SSRs generated from the koa transcriptome indicated that significantly lower numbers of analyzed sequences (7,123,041 versus 667,025) can produce similar SSR numbers for most motif patterns. Surprisingly, the trinucleotide motif pattern |AGC|GCA|CAG|GCT|CTG|TGC| was identified 82 and 77 % more within the koa transcriptome than the koa or koaia genomes (Table 4; Table S1).

Table 3

Validated SSRs. nuSSRs with the greatest numbers of repeats from A. koa and A. koaia with e-PCR. Full list of primers. (see Supplemental Table 1)

Type         Land         CARCINE NIMID         Constraint         Land         Tun         Size           km         CARCINCTAGTGTCAGTCCA         GGTGAACTGGTGCAGGATACC         af         40         21         10         61.0         58.7         148           km         GGTGAACTGGTGCAGTAGC         GGCGTCTCACCATGCAGATCC         af         39         23         101         61.0         59.9         159           km         GTTGTGACATAGGCATACC         ACCCATTACACACTTTAACCC         tg         39         61         141         56.7         55.5         164           km         GTGTGACATAGGCCAAACCCAACACCACACAGATCC         at         35         21         101         59.9         59.1         139           km         TCTAAGCTTCATATTCC         GTAACACCAAAGAAGGATCCAACAGCC         at         36         71         143         60.1         60.2         113           km         TCGACGAAAGAAGGATCCAAGG         GCGCTCATGCAACACCCACCAAGAGC         at         35         31         101         63.9         99.9         59.1         137           km         TGGCGAAAGACACACAAAGAGGAGC         CCTCATGACACACACACACAAGAGAGGA         35         24         94         55.7         58.9         116           km <th>Туре</th> <th>Forward Primer</th> <th>Reverse Primer</th> <th>Motif</th> <th>#</th> <th>Start</th> <th>End</th> <th>Forward</th> <th>Reverse</th> <th>Fragment</th>	Туре	Forward Primer	Reverse Primer	Motif	#	Start	End	Forward	Reverse	Fragment
kaa         GGTGACTGGGTGACTAGC         GGGTGTCCATTTAAGGTGC         at         39         23         101         61.0         59.9         192           kaa         GTTGTGACTAGGCCAATACC         ACCCATTACACACTTTAACCC         tg         39         63         141         56.7         55.5         164           kaa         GTGTGCACTTGGTGGTGTGG         ACCCATTACACCACCACAGATCC         at         37         27         101         57.5         163           kaa         TCGAAGCTACCACCAACGGCC         GCTAGTGGGTAAACATGACCACCACCAGATCC         at         36         29         101         57.5         59.1         139           kaa         TGCGAAAGAGGATCTAGG         GCGCTCATGACACAAGCC         at         36         30         102         60.1         60.2         113           kaa         TGCGAAAGAGGATCTAGG         GCGCTCATAGCACAAGCC         at         35         31         101         60.3         59.9         150           kaa         TGGAGACAACACACATAGG         GCGCTCATAGCACAACCACCATAGACACAC         at         35         31         101         60.3         59.9         153           kaa         TGGAGACAACACATAGAGGAG         CATGACCAACACACACACACACACACACACACACACACAC	Type			Moth	Repeats		LIIU			Size
kas         GTIGTACALTACGCCATIACC         ACCCATTACACCACITIATACCC         ig         39         63         141         55.5         164           kas         GTGTACACTACGCTATACC         ACAGTAGACACATTATACCC         at AGAATAAGCACCAAGATCC         at         38         25         101         60.6         131           kas         TCTAAGCTCATCATCATCACTCC         AGAATAAGCACCAAGATCC         at         38         25         101         60.1         59.2         101         59.9         101         59.9         103         59.1         119           kas         TCTAAGCTCACGTACCCAAAGACCC         at         36         71         143         60.1         59.2         162           kas         TCCAAAGAGAAGGAGTCAGC         GCCCCTATAATCCAAAGACACACCC         at         35         31         101         59.2         59.1         131           kas         CTCATAATGGAAGAGATTAAG         at         35         31         101         59.3         19.1         131           kas         CTCATAAGCTAAGACAAATTAACACACACACCA         at         35         24         94         59.1         137           kas         TCGAGCAACACATAGAGAGA         at         35         25         29.1         25.8         112	koa			at						
kaa         GTGTTCGGTTCGGTTGTGG         ATGTGTGATGTGGCCTGTGG         at         38         25         101         60.0         60.6         131           kaa         TCGAAGCTCCATCCATCCCC         AGAGATAAGCACCACAGAGTCC         at         37         27         101         57.5         56.6         163           kaa         TCGAACGCTCAATCCCAAAGCGC         GCTAGTGGTGAGTAAGCAACCATCACCCT         at         36         29         101         57.5         59.1         139           kaa         TCCGAACACCCAAAGCGGACTCAGG         GCCACTCATGAACAACACC         at         36         30         102         60.0         60.2         113           kaa         ATTAGCAGCAAGACGGTCC         GACCTCATATTCACAACACC         at         35         31         101         60.3         99.9         150           kaa         TCGATCAGGGACGTACAACAGAGGAGC         CATGACCACACATAGAGGAG         CATGATGGGTGCAACACACATAGAGGAG         35         31         101         60.3         99.9         150           kaa         TGGAGGACACACATAGAGGAGGA         CATGCCCTCTGTGACACACAG         at         35         24         94         59.4         52.2         112           kaa         TGGAGGCACACATAGAGGGACGACACA         ATGAGCGCGTGACCACAGTGGAGGAGGACC         CATGGAGCACACATAGGGAGGCA	koa			at						
kaa         TCTAACCITCATCCCATCCC         AGAGATAAGCCCACAGAGATCC         at         37         27         101         57.5         57.6         163           kaa         TCCATCAGGTGACGTAATTTGC         GGTAAACCTAACCTAACCTAACTCAACTCACTC         at         36         29         101         59.9         103         59.1         139           kaa         TACCAGTGACGTAACCCCAACGC         GCTAGTGGGTGAAACCAACTGAATCC         at         36         30         102         60.1         59.2         162           kaa         ATTACACGACAAGCGGGC         GCATCATTTCACAACACACCC         at         35         31         101         60.3         59.9         154           kaa         TCGATCAGGTGAGTCAATTCC         GGTAAACCAACACACACACCC         at         35         31         101         60.3         59.9         158           kaa         TGGATCAGAAGAGAGTTAAGG         GGTAACACTATCAACCTACTGC         at         35         28         152         59.8         59.2         112           kaa         TGGAGACACACATAGAGAG         CCTGTGGAGAAAATTGTGCC         at         35         28         19.8         178           kaa         TGGAGACACACATAGAGAGGCC         CCTGGGAAAGAGAGTCC         at         35         21         28         99.3 </td <td>koa</td> <td>GTTGTGACATAGGCCAATACC</td> <td>ACCCATTACACACTTTATACCC</td> <td>tg</td> <td>39</td> <td>63</td> <td>141</td> <td>56.7</td> <td></td> <td>164</td>	koa	GTTGTGACATAGGCCAATACC	ACCCATTACACACTTTATACCC	tg	39	63	141	56.7		164
kas         TGATCAGCTGACGTTANTTEC         GGTAACCCTATCACCTCACCTCACTCC         at         36         29         101         99.9         99.1         192           kas         TACCGATAACCCAAACGCC         GGTAGTGGGTAGACAAGTACC         at         36         71         143         60.1         59.2         162           kas         TGCACAAAGAGGACTCAGG         GCCCTCATGACACACACC         at         35         31         101         57.5         59.1         141           kas         TCGACACAGAGAGGTC         GACCTCATATCACACACACG         at         35         31         101         60.3         59.9         136           kas         TCGATCACGTGAAGGAGTCATAGC         GGTAACCTATCACCACACG         at         35         31         101         69.3         59.9         136           kas         TCGATCACGTGAAGGTATAGC         GGTAACCTATCACCCACACG         at         35         24         94         55.7         58.9         116           kas         TGAGAGACACCACATAGAGAGG         CACTGCCACTCATTGACGCACACG         at         35         24         94.4         57.2         122         84.8         184         184         86.8         84.5         122         84.8         184         184         86.6         141<	koa	GTGTTCGGTTCGGTTGTTGG	ATGTGTGATGTGGCCTGTGG	at	38	25	101	60.0	60.6	131
kas         TACCGATLANCCAAACGECC         GCTAGTGGGTGAAACAATGATCC         at         36         71         143         60.1         60.2         162           kas         TGGCGAAAGAAGGATCCAAG         GCGCTCATGACAACAAGC         at         36         30         102         60.1         60.2         113           kas         ATTAGCACAAAGAGGATC         GCGCTCATGACACCAAGCG         at         35         31         101         60.3         99.9         156           kas         TCGATAAGCGAGGATCTAAGG         CTCGAGCAAATACACACG         at         35         31         101         60.3         99.9         151           kas         TGGAGACACACATAGGGAGG         CACTGCACAATACACACG         at         35         24         94         55.7         58.9         115           kas         TGGAGACACACATAGGAGG         CACTGCACAAATTGCTGACC         at         35         24         94         59.4         58.0         115           kas         TGGAGACACACATTGGATGAGGA         CACTGGACAAATTGCTGACC         at         35         25         10.2         58.3         58.9         115           kas         TCGAGACACACATTGGGTGATGGG         CACTGGACACAATTGGGTGAGGAGACATCC         at         35         25         10.2	koa			at	37		101			163
kaa         TGCCCAAAGAAGGATCTAGG         GCCCCTATATCGCAAACAAGC         at         36         30         102         61.1         62.2         113           kaa         ATTAGCAGCAAGCAGGCTC         GACCTCATATTCACAAGTCGG         at         35         31         101         61.5         59.1         141           kaa         CCCTTATATGGTTGGTCTAGTC         CTCGCGGCAAATACACAG         at         35         31         101         69.3         59.9         136           kaa         TGGATGAGTGAAAGGATTAAG         GGTAAACATTACACCTCACCTCACCTCC         at         35         21         96         98.8         59.2         112           kaa         TGGAGACACACATAGAGAGG         CACTGGCAAAGATACACTTGCCC         at         35         24         94         55.7         153         101         99.3         60.5         143           kaa         ACTGCACACAAGAGAGGGCC         CCGTGGAAAGGAGAGGGCC         at         35         29         99.4         55.1         12           kaa         ACTGCACAGAGAGAGGGCTCAGGTAAAGGTGAGGTAACAGCAGAAGC         at         35         30         100         84.8         59.1         143           kaa         ACTGCACAGAGAGAGGGCTCAGGTAAAGG         GGTAAAGGTGAGAGAACACTACACACA         35         32	koa	TCGATCAGCTGACGTTAATTTGC	GGTAACACTATCAACCTCACTGC	at	36	29				139
kma         ATTAGCAGCAMAGGGTCC         GACCTCATATTCACAMATCCACG         at         35         31         101         97.5         99.1         141           kma         CCCTTATATGGTICGGTCTAGTCC         CTGGGGCAATACACACG         at         35         31         101         60.3         99.9         136           kma         TGGGTTGAGAGGGATCTAAGG         AATGCCTCATGCACCTCC         at         35         31         101         60.3         99.9         137           kma         TGGGTGTGAGAGGGATCTAAGG         AATGCCTCATGCACCTCC         at         35         24         94         55.7         58.9         116           kma         TGAGGGGTGTTAAGGGGGC         CATGCCCTCTTGTATGCC         at         35         24         94         59.4         57.2         128           kma         ACTGCACACAATACTGGGGGC         ACTGGCCACACATGCATGCAGGGCC         at         35         23         10.0         58.5         59.2         128           kma         ACTGCACACACATATGGGGGC         ACGGGCAAGCACAGTAGGGAGGCAGCACCAT         at         35         30         100         58.6         59.1         168           kma         ACGCACACATAGGGGGCAGGGG         TGGGCGAGTAGGGAGGGAGGGAGC         at         35         30         100	koa	TACCGATAACCCAAACGGCC	GCTAGTGGGTAGAACAATGAATCC	at	36	71	143	60.1		
kas         CCCTTATATTGGTTCGGTCTAGTCC         CTGCGGGCAAATACACACG         at         35         31         101         60.3         59.9         136           kas         TCGATCAGTGACGTTAATTGCC         GGTAAACCTATCAACTCAACTCACTGC         at         35         31         101         69.3         59.9         136           kas         TGGATCAGTGAAAGGATCTAAGG         GGTAAACCTATCAACCCACCC         at         35         26         96         98.8         59.2         112           kas         TGGAGACAACATAGAGGGC         CATCGCCCTCATGTAACGACACACG         at         35         24         94         55.7         58.9         116           kas         TGGAGACACACATAGGGC         CATCGCCCATTGTATCCC         at         35         24         94         59.4         59.2         123           kas         TGGAGACACACATTGGACGC         CATGGCGAAATTGTATCCC         at         35         21         92.3         60.5         143           kas         TCGAGACACAGTCAGTCAACG         GGTGAACTGGGTGAATCACTTACCC         at         35         31         00         86.8         59.1         168           kas         TCGAGACACATTCATCGG         GGTGAACTGAGTAACG         GGTGAACTGAGTAACG         CTTTCCAAGAACACATTACATCC         at         35 <td>koa</td> <td>TGCGCAAAGAAGGGATCTAGG</td> <td>GCGCCTCATGACAACAAGC</td> <td>at</td> <td>36</td> <td></td> <td>102</td> <td>60.1</td> <td>60.2</td> <td></td>	koa	TGCGCAAAGAAGGGATCTAGG	GCGCCTCATGACAACAAGC	at	36		102	60.1	60.2	
kaa         TCGATCAGCTGACGTTAATTEC         GGTAACACTATCACCCACTCACC         at         35         31         101         99.9         99.1         137           kaa         TGGGTTGAGAGAGGGATCTAAGG         AATGCCTCATGACACCACACC         at         35         24         94         95.7         58.0         115           kaa         TGGGAGACACACATAGAGAGGAGC         CACTGCCCTCTTGTAGTCC         at         35         24         94         95.7         58.0         115           kaa         TGGACACACATACATGATGACGC         CACTGGAGGAATTATTGCCC         at         35         24         94         99.4         57.2         123           kaa         AATTAGCACACATATGATGCC         CACTGGAGGAATGATGTGGC         at         35         32         102         58.3         59.8         179           kaa         AATTAGCACACAATTAGTGCC         CCTGGAAGAATGGGTGCAGGAGGATCC         at         35         30         100         58.6         60.0         117           kaa         TCGAGGGAATCAGTTTAGG         GGTTATACCACATACAGTAGGGGGCAGACACC         at         35         30         100         58.5         112           kaa         ACCTTACACATATAGGTGGAACACTTACCAGACACATACAGATACAGATAGAGGAAGACATTACGTGGCAATCACAC         at         21         28         99<	koa	ATTAGCAGCAAGACGGTCC	GACCTCATATTTCACAAGTCCGG	at	35	31	101	57.5	59.1	141
kan         TGCGTTGAAAGGGATCTAAGG         AATGCCTCATGACAGCAAGC         at         35         25         96         98.8         59.2         112           kan         TGGAGACACACATGAGAGGG         CACTGCCCTCTTGTATGCC         at         35         24         94         55.7         58.9         116           kan         TGAGACACACATGGGAGC         CACTGCACAAATGTGGCC         at         35         24         94         55.7         58.0         116           kan         ACTGCACACAATTGTGACGC         CACTGGAAAAATGTGGCC         at         35         24         94         59.8         59.2         122           kan         ACTGCACACAATTGTGACTGC         ACGTGGAAGAAATGCATTGCACATTC         at         35         23         102         61.6         61.0         117           kan         TCGGAGGAAATCAGGTAACG         GGGTGAACTGAGGTAAGC         at         35         35         10         86.5         59.1         168           kan         TCGGAGGTGAATTCCGGGT         CGTTACCACAATACCTGGAGGAACGAGCC         at         35         51         125         58.8         57.4         132           kania         TCGGACATCATGTTATACTGCG         TCATGCACAACAACTCCC         at         19         48         86         85.7 <td>koa</td> <td>CCCTTATATTGGTTCGGTCTAGTCC</td> <td>CTGCGGGCAAATACACACG</td> <td>at</td> <td>35</td> <td>31</td> <td>101</td> <td>60.3</td> <td>59.9</td> <td>136</td>	koa	CCCTTATATTGGTTCGGTCTAGTCC	CTGCGGGCAAATACACACG	at	35	31	101	60.3	59.9	136
kaa         TGGAGACACACATAGAGAGG         CACTGCCTTGTTATACC         at         35         24         94         55.7         58.9         116           kaa         TGATGCACACATAGAGAGGC         CACTGGACAATCATTGCCC         at         35         82         152         95.1         58.9         115           kaa         ACTCGACACATTCGACTGC         ACGTGGAGAAATCTTTGCCC         at         35         24         94         95.7         58.9         112           kaa         AATTACCAGCACATTGGACTGC         CCGTGGAAGAAATTGGTGC         at         35         24         94         95.7         58.8         172           kaa         ACTGCAGACACATTCACG         GCGGAAAGGTGCAGATACG         at         35         29         102         84.8         172           kaa         TCGAGACACATTACGG         TGGTGAAAGGTGTGAGTACATACG         at         35         35         102         61.6         101         117           kaa         TCGCAGACAATATCTACTCTG         GGTGAATGCATTACACATACAATACATACATACATATAGAGACAAAAGACAAACAA	koa	TCGATCAGCTGACGTTAATTTGC	GGTAACACTATCAACCTCACTGC	at	35	31	101	59.9	59.1	137
kaa         TGATCGGTGTTAATGGGC         CACTGGACAATCATTGCCC         at         35         82         15.2         99.1         57.0         115           kaa         ACTGCACACAATTCTGACGC         ACGTGGAAGAATTGTGCC         at         35         24         94         99.4         57.2         123           kaa         AATTGCAGCAAGAGGGTC         CCGTGGAAGGAATGCATCC         at         35         22         10         58.3         59.8         179           kaa         TCGGACACACTCTTACCG         TCGGGCAAATCGGTGCAAGG         at         35         23         10.0         8.6         6.0.0         117           kaa         TCGGGCTGAAATCGGTGGCAGCGG         GCTTATCCACATATCTTGCGGCG         TGGACGAATAGGGGGCAGGGC         at         35         51         125         8.8         60.0         145           kaa         ACCCTCACATATCTTTCGCGG         GCTTACCACATACCTCCTGGGG         TGGACACATATGGGGCAGGC         at         21         28         91         95.5         55.1         122         8.84         55.7         133           kaaia         TCGACACATATCGTGCGGCGGGTGGTGTGTGTGG         TGGAGACACATTAGGGCACACCC         tag         16         8.84         55.7         133           kaaia         TGGGCGTGTGTGTGTGTGG         GAGATCCACA	koa	TGCGTTGAGAAGGGATCTAAGG	AATGCCTCATGACAGCAAGC	at	35	26	96	59.8	59.2	112
kaa         ACTGCACACANTTCGACTGC         ACCTGGAAGAAATTGTGGCC         at         35         24         94         994         994         572         123           kaa         AATTAGCAGCAAGACGGTCC         CCTGTGAAGAGATGTCC         at         35         24         94         994         994         594         572         123           kaa         TICGAGACAGCATTCCCTACCG         CCTGTGAAGGTAACAGCATTCC         at         35         29         102         843         598         179           kaa         TCGGAGCAAGTCAGTAACG         GGTGAACTGGGTGAATCATTCC         at         35         30         100         856         591         168           kaa         TCCGAGCAATTCGTTCGG         GGTGACTGGGGTGAAATCCTTTACCC         at         35         35         100         856         591         112           kaai         TCGAGAATTCATTCGGCTGG         TGATGAACACATTAGAGCC         at         21         28         91         95         55.1         112           kaai         TCCAAATTCATGCTTACTGC         TCATCCAACCACATCACCCC         TCCTTCCAAGAACACACCC         at         12         50         59.5         131           kaaia         TGCAGCAATTCATTCGG         TCAAAGCACATCACCCCC         TCCTACACATCTCACCCCC         TCCTCCAAC	koa	TGGAGACACACATAGAGAGG	CACTGCCCTCTTGTTATGCC	at	35	24	94	55.7	58.9	116
kaa         AATTAGCAGCAAGAGCGTCC         CCTGTGAAGGTAACAGCATTCC         at         35         32         102         83         59.8         179           kaa         TTGGAGAGCATCCTTACCG         TCGGTCAAAGGTGAGCATCC         at         35         29         99         99         50.3         60.5         143           kaa         TCGGGCAAGCAGCAGTAAC         GGTGAACTGGGTGACGACTAC         at         35         32         102         84.3         59.8         101           kaa         TCGGCGGAAATCAGTTCTGG         GGTTAACCACATCAATCACTACATCAATACATCC         ta         35         30         100         85.6         61.0         117           kaa         ACCCTACCACTATCACTTTTAGC         GGTGAACTGGGGTGACTACACATCATACATCC         ta         35         55         15         55.1         112           kain         TGCAACAATTAGTTCGTATACAACAC         TCCTCCAAGAACCAATGCACCC         ctTGAATCCAACACATGATGCC         ctTGAATCCAACACATCCACCCC         17         48         85         55.7         133           kain         TGGGCGTGTGTGCTCAGG         GAATCCAACACCATACCCC         tag         16         28         64         64.4         59.4         171           kain         TGGGCGTGTGTGTCACCAGG         GGTGTGTGTGTGGC         ctTGCTGACACACACCCAATCCCACC </td <td>koa</td> <td>TGATGCGTGTTTAAGTGGGC</td> <td>CACTGGACAAATCATTTGCCC</td> <td>at</td> <td>35</td> <td>82</td> <td>152</td> <td>59.1</td> <td>58.0</td> <td>115</td>	koa	TGATGCGTGTTTAAGTGGGC	CACTGGACAAATCATTTGCCC	at	35	82	152	59.1	58.0	115
kaa         TTCGAGAGCATCCTTACCGG         TCGGTCAAAGGTCAGTACG         at         35         29         99         99.3         60.5         143           kaa         TCCGGCTGAAGTCAGGTAAGG         GGTGAAATCAGGTCAGGTAAGG         at         35         32         102         61.6         61.0         117           kaa         TCCGGCTGAAGTCAGGTAAGG         GGTGAAATCAGTACCAGTAACTC         ta         35         30         100         85.6         59.1         168           kaa         ACCCTTCACCTATTCTTCTTACG         GGTGAAATCAATACTCGGGAGCC         ta         35         55         125         58.8         59.1         122           koaia         CCCGAATTCACTGACTTATTCGG         TCAAGTCACAATACCTGGAGGACC         at         21         28         91         95.5         55.1         112           koaia         TGGGGGTGGTGTCTGTCTGAC         TGAAGTCACAACACCCC         at         19         48         86         58.4         57.4         132           koaia         TGGGGGTGGTGTGTGTCTGCTC         TGAAGTCACAACACCCACCCCC         tgTGTGAAAACACATAGACACCCC         17         41         45.8         59.6         137           koaia         TGTGGCGGTGGTGTGTGTGTGTC         ta         16         28         60         59.1 <t< td=""><td>koa</td><td>ACTGCACACAATTCTGACTGC</td><td>ACGTGGAAGAAATTGTGTGC</td><td>at</td><td>35</td><td>24</td><td>94</td><td>59.4</td><td>57.2</td><td>123</td></t<>	koa	ACTGCACACAATTCTGACTGC	ACGTGGAAGAAATTGTGTGC	at	35	24	94	59.4	57.2	123
kaa         TCGGGCTGAGTCAGGTAACG         GGTGAACTGGGTGCAGTAGC         at         35         32         102         61.6         61.0         117           kaa         TCTGGCTGAATCAGTTCTGG         GCTTATCCACATCCATACC         ta         35         30         100         85.6         61.0         117           kaa         ACCTTACCATTATCTTTAGG         GGTGAACTGGGAAACC         ta         35         30         100         85.6         61.0         145           kaai         TCCAACATTACCATTATTTAGG         GTGAAACCATTACC         ta         21         28         91         95.5         55.1         112           koaia         CCGAATCATTACTCAGCT         TCCTTCACCCATAAATTCGG         TCAAAGACACACCC         tat         21         28         91         95.5         55.1         112           koaia         CTGGAATCCATACTTAAATTCGG         TCAAAGCACACCC         tat         71         48         85         58.4         55.7         133           koaia         TCGAAACATTACCC         TGTGTGTGTGTGGTGGTGGCGTGCC         ta         16         21         23         93         94.6         134           koaia         TCAAACACACACCACCACC         tag         TGTGGGGTGTGTGGTGGTGGCGGC         ta         16         91	koa	AATTAGCAGCAAGACGGTCC	CCTGTGAAGGTAACAAGCATTCC	at	35	32	102	58.3	59.8	179
kea         TCTGGCTGAAATCAGTTICTGG         GCTTIATCCACAATCCAATACAGT         ta         35         30         100         58.6         59.1         168           kea         ACCCTTCACCTATATCTTCTTTAGG         GGTGATGAGAAATCCTTACCC         ta         35         55         125         58.8         600         145           keani         ACCCACAATATCTTCGGCGG         GGTGACGATGAGACGAATGCGGGGGC         ta         21         28         91         95.5         55.1         112           keani         CCGACATCAGTCTTACACACC         TCCTTCCCAGAACGAACGAAGG         ett         21         28         91         95.5         55.1         133           keani         TGGGCGTGTGTGTGTCTTTACG         TCATAGACACATTAGACACCCC         tg         17         64         152         57.4         132           keani         TGGGCGTGTGTGTGTGTGC         GAGATCCACACCCCCCC         TGTGTGTGTGTGTGTGC         tg         16         41         57.3         133           keani         TGGGCGTGTGTGTGTGG         GAGATCCACACCCCCCCC         TGTGTGTGTGTGTGGTGC         ta         16         28         60         9.1         161           keani         TCCAAACACTCACCACCCCCCCC         TGTGTGTGTGTGTGGGGC         ta         15         61         91         91.	koa	TTCGAGAGCATCCTTACCGG	TCGGTCAAAGGTTGAGGTGC	at	35	29	99	59.3	60.5	143
kana         ACCCTTCACCTTATTCTTTACG         AGTGACTGAGAAATCCTTTACCC         ts         35         155         125         58.8         60.0         145           komin         TGCAACAATTATTCTGGCTGG         TGATGAAAATCCTTTACCC         ts         31         35         55         12.5         88.8         60.0         145           komin         CAGGCATTCATTCGGCTGG         TGATGAACACATTGAGGCGC         att         21         28         91         95.5         55.1         112           komin         CCGGAATTCATGCTTAATATTCGG         TTAAGTCACAGAACG         att         19         48         86         58.4         55.7         133           komin         TGTGCAGAAACATGTATATCG         TTAAGTCACAGCCAACCCG         att         17         82         16         63.8         59.6         137           komin         TCTAGCACAACTCCACCCCCCC         TGTGTGTGTTCATGGGTGTGG         att         16         91         12.5         51.9         178           komin         CCCAACACTCTGTCGCCG         TGTGTGTGTTCATGAGTGTGTG         att         16         91         12.2         57.3         134           komin         CCCAACACTCTGTCGCCG         TGTGTGTGTGTGTGTGTGG         att         16         91         10.9         0.0	koa	TCGGGCTGAGTCAGGTAACG	GGTGAACTGGGTGCAGTAGC	at	35	32	102	61.6	61.0	117
koaia         TGCAACAATTACTTCGGCTGG         TGATGAACACATTAGGAGGC         ant         21         28         91         99.5         55.1         112           koaia         CAGGCATTCATTCTTTACACACC         TCCTTCCCAAGAACGAACGAAGG         ett         21         28         91         99.5         55.1         112           koaia         CCAGCATTCATTCTTACACACC         TCCTTCCCAAGAACGAACGAAGG         ett         21         93         93.8         55.7         133           koaia         CTGACATTCATTCTTATTTTCG         TGAAGCACCAATACCC         tg         17         84         86         55.7         133           koaia         TGTGCGAGAACTATGATCCC         TGTGTGTGTTGTGC         GAACTCCAACCACCACC         tg         17         64         132         57.4         59.1         16           koaia         TCTAACACCACTACCCACC         tg         17         64         132         57.4         59.1         18           koaia         CTAACACACTATCGTCACTAG         TGTGTGTGTGTGTGGC         ca         16         28         60         9.1         27.5         57.3         57.9         178           koaia         CTAACACACATTGTGTCAGG         TGTGTGTGTATAGGACGGG         ca         15         70         100         60	koa	TCTGGCTGAAATCAGTTTCTGG	GCTTTATCCACAATCCAATACATGC	ta	35	30	100	58.6	59.1	168
	koa	ACCCTTCACCTTATTCTTCTTTACG	AGTGACTGGAGAAATCCTTTACCC	ta	35	55	125	58.8	60.0	145
koaia         CCTGAATTICCATGCITAATATTCGG         TIAAAGTCACGAGACGAACCATCC         at         19         48         86         58.4         55.7         133           koaia         TTGGGGGTGTGTCTTGC         GAOATCCAACCCATACCCG         tg         17         76         115         63.8         55.7         133           koaia         TTGGGGGTGTGTGTCTGC         GAOATCCAACACCACC         tg         17         64         12         57.4         59.1         167           koaia         TCCAAACCACCACCACC         tg         17         64         12         57.4         59.1         167           koaia         CCAACACCACCACCACCAC         tg         16         28         64         60.4         59.3         134           koaia         TCCAAACACTCACACATGCG         TGTGTGTGTGAGAGGTGGGGAGGG         ca         16         59         125         57.3         57.9         178           koaia         GTATGAGAGAGTAGAGGCC         GGCATCCCAATTAGAGTGGGAGGG         ca         15         70         100         60.3         157           koaia         CCCAACATCCAATACCACCC         GTGTGTGTGGGGGGGGG         ca         15         71         101         57.7         58.2         137           koaia	koaia	TGCAACAATTACTTCGGCTGG	TGATGAACACATTAGGAGGC	aat	21	28	91	59.5	55.1	112
koaia         TTGGGCGTGTGTGTGTCTGCC         GAATCCAACCCATACCCG         tg         17         82         116         63.8         59.6         137           koaia         TGTGCAGAAAACATGTATGCC         CTCATTCAACGCCACACC         tag         17         64         132         57.4         59.6         137           koaia         TGTGCAGAAAACTGTATGCC         CTCATTCAACGCCACCAC         tag         17         64         132         57.4         60.4         59.3         134           koaia         TCAAACACCACTCACCACGC         TGTGTGTGTTGTAGAGTGTTGGG         ca         16         28         60         69.1         27         57.3         57.9         178           koaia         CCAAACACTCAACCACGG         GGCATGCCTATTAGGAGGG         ca         16         29         12         57.3         57.9         178           koaia         CCAAACACTATAGTGACAC         TATGGGTGGTGGGGGG         ca         15         61         91         57.5         58.2         137           koaia         CCAAACAACCAAACC         TATGGGTGGTGGGGGGG         ca         15         71         101         60.0         57.6         125           koaia         ACACAACCAACACAC         GATGGGATGGGGGGGG         ca         15 <td< td=""><td>koaia</td><td>CAGGCATTCAGTCTTTACACACC</td><td>TCCTTCCCAAGAACGAAAGG</td><td>ctt</td><td>21</td><td>30</td><td>93</td><td>59.8</td><td>57.4</td><td>132</td></td<>	koaia	CAGGCATTCAGTCTTTACACACC	TCCTTCCCAAGAACGAAAGG	ctt	21	30	93	59.8	57.4	132
konia         TGTEGCAGAAACATGTATGCC         CTCATTCAAGCCCACC         unig         17         64         132         57.4         59.1         167           konia         CCAAACACCACCAACCGC         TGTCTGTGTTGTGTGTGGGTGGCG         ca         16         32         64         69.3         134           konia         CCCAAACCCACCACAGCGC         TGTCTGTGTGTGTGTGGGGTGGGC         ca         16         28         60         99.1         59.4         127           konia         CCCAACACTCCACCAGGG         TGTGTGTGTCATGAGTGTGTG         at         16         28         60         99.1         59.4         127           konia         CCCAACATCGACCACATGG         TGTGTGTGTGTCAATGAGTGGTGG         at         16         9         125         57.3         157           konia         CACACATCCAACCACCCAATGGAGCC         GGCATCCATATGGAGCGGGG         at         15         71         101         60.3         157           konia         ACACAAGCACATACCAACC         TATGGGTGGTGGGAGGG         at         15         71         101         60.3         57.5         168           konia         TAGACTACCCACCG         AGGTAGATTCATCGGGTGG         at         40         97         57.5         58.8         171           k	koaia	CCTGAATTCCATGCTTAATATTCGG	TTAAAGTCACAGAACAGATCCC	at	19	48	86	58.4	55.7	133
koain         CTAACACCACCTACCCACCGC         TGTCIGTGTTGAGTGTGGC         ca         16         32         64         60.4         59.3         134           koain         TCCAAACACTCACCACAGGC         TGTGTGTGTTGAGTGTTTGG         ca         16         32         64         60.4         59.3         134           koain         TCCAAACACTCACCACAGGC         TGTGTGTGTCAGATCTTGG         ca         16         92         80         59.1         57.4         175           koain         GTATGAGAGATGTAGGGCCC         GGCATCATAGTGAGGC         ta         15         61         91         51.3         57.9         178           koain         GTATGAGAGATGTAGGGCCC         GGCATCCATATAGCACCG         GGCATCCATATGGAGGGC         ta         15         70         100         60.0         57.6         125           koain         CCCAACTCAAACCACATACCACCC         GGTGCTTCTATCAGGGGGGGC         ta         15         71         101         57.7         58.2         137           koain         CCCAACTACACACCCC         GGTGATTCTATCGGGGGGGG         ta         14         41         15         54.9         57.5         165           koain         CAAACCACACACACCCC         AGTGATTCATCGGGGGGG         ta         14         48	koaia	TTGGGCGTGTGTGTGTCTGTGC	GAGATCCACACCCATACCCG	tg	17	82	116	63.8	59.6	137
keaia         TCCAAACACTCAACAAGGG         TGTGTGTGTCTAGAGGTCTTGGG         ca         16         28         60         59.1         59.4         127           keaia         CCCACAATTCGTGCTAGG         TCATGGCACAATGGGG         at         16         91         123         57.3         59.4         127           keaia         GATGGAGGATGTGTGCTCAAG         TCATGGCCCCAATGGAGG         at         16         91         123         57.3         57.9         178           keaia         GATGGAGGAGTAGGGCCC         GGCATCCCTAATGGAGCTGG         at         15         61         91         123         57.3         57.6         125           keaia         ACACAACCAATACCAACC         TATGGGTGTGGGGGGG         ac         15         71         101         60.0         57.7         58.2         137           keaia         ACCAACCAATACCAACC         GGTGATTCTATCGAGGGGG         ac         14         115         143         58.0         57.5         56           keaia         CTAAGCCACTTCAAGGACC         AGGTGAATCTCAACGGGGGGG         ac         14         69         97         57.7         58.8         147           keaia         CTAAGCCACTTCAAGCACC         AGGGGAATCTCACGGGG         ac         14         71         <	koaia	TGTTGCAGAAACATGTATGCC	CTCATTCAACGCCACCACC	tatg	17	64	132	57.4	59.1	167
koain         CCCACCAATTCTGTCCTAGG         TCATGGCTCCAATTGAGTGG         at         16         91         123         57.3         57.9         178           koain         GTATGAGAGGATGTAGGCCC         GGCATCCTATTAGGACCTGG         at         15         61         91         92.1         57.3         57.9         178           koain         GTATGAGAGGATGTAGGCCCC         GGCATCCTATTAGGACCTGG         ta         15         61         91         92.1         60.3         157           koain         CCCAACTCCAATCCCATCCC         TATGTGGTGGGGGGC         ac         15         70         100         60.0         57.6         125           koain         CCCAAGCAATCCAATCCAACCC         GGTGATTGTATGTAGGGCG         ca         15         23         53         58.4         58.0         168           koain         ACACAGCACATAGAGACCCA         AGGTAAGTGTAGAGCCC         caGTAAATGTGTACCATCCC         ca         15         23         53         58.4         167           koain         AAACATACCCACACACCCA         AGGTAATGTAGCATGCGGGG         ca         14         69         97         57.7         58.8         147           koain         AAATGAGCCACCTTGAGGCC         TATGAGGATGTAGCGCG         CAAATGAGACATAACATGAGCATTAGAGATGAGACATAACATGTAGCGC <td>koaia</td> <td>CTAACACCACCTACCCACGC</td> <td>TGTCTGTGTTTGAGTGTGTGC</td> <td>ca</td> <td>16</td> <td>32</td> <td>64</td> <td>60.4</td> <td>59.3</td> <td>134</td>	koaia	CTAACACCACCTACCCACGC	TGTCTGTGTTTGAGTGTGTGC	ca	16	32	64	60.4	59.3	134
koaia         GTATGAGAGGATGTAGGGCCC         GGGATGCCTATATAGGACCTGG         ta         15         61         91         99.1         60.3         157           koaia         CACGACATCCCATATGTACCCG         TATGTGGTGTACAGGGGCG         ac         ti         5         70         100         60.0         105         76.1         125           koaia         CCAACTAACCACCAACC         TATGTGGTGGTGGGGGGG         ac         15         71         101         60.0         57.7         58.2         137           koaia         ACACAACCAACTAACCACC         GGTGCTTCTATCTACTACCTC         ca         15         71         101         57.7         58.2         137           koaia         ACACAACCACACACCACCC         GGTGCTTCTATCTACACCC         ca         15         23         53         58.4         57.5         165           koaia         AAACATACCACACACCACCACCC         AGGTGATTCTATCGGGTGG         ac         14         41         15         13         58.8         17.7           koaia         CTAAGCCCACTTCATCTAGCATC         TATAGGGTGGATCACA         ac         14         71         9         60.4         55.2         157           koaia         CATAGCTAGCATTAGAGATTAGAGATTAGAGATGAGC         TATAGGGGTGGATCACAA         ac	koaia	TCCAAACACTCACACATGCG	TGTGTGTGTGTTCATGAGTCTTTGG	ca	16	28	60	59.1	59.4	127
keaia         CACGACATCCCTATGTACCCG         TATGTGTGTACATGGGTGGC         ac         15         70         100         60.0         57.6         125           keaia         CCCAACTCAAACCC         TATATGGGTGGGGGGGG         ac         15         70         100         60.0         57.6         125           keaia         CCCAACTCAAACGCAAACC         TATATGGGTGGGGGGGGG         ac         15         71         101         57.7         52.2         137           keaia         TCTAGGACACTAACACC         GGTGTAAGTGTACAGTCC         ac         14         115         143         58.0         57.5         165           keaia         ACACATACCACACACACCC         AGGTGATTCATGGGGTGGGG         ac         14         40         97         57.7         58.8         147           keaia         TCAGCCATTTCCTCCCC         TATAGGGTGGGGG         ac         14         40         97         57.8         59.8         147           keaia         TCAGCCTTTCCTCCCGATGC         AATAGACGACCGGTTCACGGGG         ac         14         71         99         60.4         56.2         157           keaia         TCAGCCTTTCACGATGGC         TGACATAGCATAGCATAGACATAGAGC         TGACAACAATAGCACGGGC         at         44         88         1	koaia	CCCACCAATTCTGTCCTAGG	TCATGGCTCCAATTGAGTGG	at	16	91	123	57.3	57.9	178
keaia         CCCAACTCAAACCC         TATATGGGTTGGTGGGAGGG         ac         15         71         101         57.7         58.2         137           keaia         ACACAAGCAATACCAACC         GGTGCTTCTATCTAGTGCGC         ca         15         23         53         58.4         59.0         168           keaia         TCTAAGCACTTGAAGCACCA         GGTGCTTCTATCTAGTGCGC         ca         15         23         53         58.4         59.0         168           keaia         TCTAAGCACTTGAAGCACCA         GAGTATGTAGACTTAGGGTGGG         ca         14         115         143         58.0         57.5         165           keaia         ACACATACCACACACCACCA         AGGTGATTCATGCGGTGGG         ca         14         69         97         57.7         58.8         147           keaia         CTAAGCCACACACACACCC         AGGTGATTCATGCAGTGCGGG         ca         14         62         90         58.8         157           keaia         CTAAGCCACTTCATGACCT         AATAGAGCAGCGGTGGCGGC         ca         14         71         99         60.4         56.2         157           keaia         GATGAACTGACATTAGCATTAGCAGC         TGAATGTACACTAGCATAGGCG         TGAATGTAACACTAGCATAGGCG         TGAAATGAACATAGCATAGCATAGGCGC         TGAAATGAACATAGC	koaia	GTATGAGAGGATGTAGGGCCC	GGCATGCCTATTAGGACCTGG	ta	15	61	91	59.1	60.3	157
keain         ACACAAGCACATACCAACCC         GGTGCTTCTATCTATGTACGTCC         ca         15         23         53         58.4         59.0         168           keain         TCTAGGACACTAGCAACCC         GGTGATGTGAAATAGCACCC         ac         14         115         143         58.0         57.5         165           keain         ACACATCCCCACACCC         AGGTGAAGTGGAAATAGCACC         ac         14         61         5         57.7         59.8         147           keain         CTAAGCCCACTTTCATGCGC         TTCAATGAGCACGTGGGG         ac         14         60         97.7         59.8         147           keain         CTAAGCCCACTTTCATGCGC         TTCAATGAGCACGTGCG         ac         14         71         90.04         55.2         157           keain         GATGAACTGTAACTAGCATTAGGGC         TGGAAACATACACTGTACGC         at         14         71         90.04         55.2         157           keain         GATGAACTGTAACTAGCATTAGGGCC         TGGAAACATACACTGTACGC         at         14         81         800         57.1         58.8         140           keain         CCATTACTGTAGGGCC         TGGAAACATACACTGTAACGGCC         at         4         41         83         56.7         59.8         140	koaia	CACGACATCCCTATGTACCCG	TATGTGTGTACATGGGTGGC	ac	15	70	100	60.0	57.6	125
keaia         TCTAGGACACTTAGAGACCACC         AGGTTAAGGTGTAGATATGACCC         ac         14         115         143         58.0         57.5         165           keaia         AAACATACCCACACACCG         AGGTGATTCATGGGTGGGG         ac         14         04         97         57.7         98.8         147           keaia         CAACATACCCACACACCGC         AGGTGATTCATGGGGTGGGG         ac         14         62         9         57.5         165           keaia         TCAGCCACTTACCCC         ATATGATGCGGTGGGGG         ac         14         62         9         58.9         147           keaia         TCAGCCTTTCCTCCCC         ATATGACGATCGGGC         ta         14         71         99         60.4         56.2         157           keaia         GATGAACTATAGCATTAGAGG         TGACAATAGACATGGCG         ta         14         78         100         51.3         125           keaia         TGACTTACAGCTGTCAGGCC         TGACAATAGACATGCACATTCAAGGGC         ta         14         41         8         56.7         59.8         140           keaia         TGACTTIACCTGTCAGGTGC         TCAAATGCACACTCTAACAGCA         ta         14         51         93         57.8         59.2         107	koaia	CCCAACTCAAACGCAAACC	TATATGGGTTGGTGGGAGGG	ac	15	71	101	57.7	58.2	137
koain         AAACATACCCACACCCCG         AGGTGATTICTATCGGGTIGGG         ac         14         69         97         57.7         59.8         147           koain         CTAACCCCATTTCATCCCC         TTTCAATGACATGTGGCG         ac         14         69         97         57.7         59.8         147           koain         CTAACCCCATTTCATCCCC         TTTCAATGACATGTTCACG         ta         14         71         99         60.4         55.2         157           koain         GATGAATCGTAATACCATTAGACG         TGAAAACATACACTATAGCGC         ta         14         71         99         60.4         55.2         157           koain         CCACHACTGAAACACATAGACATTAGAGGC         TGAAAACATACACATGTACACTTAGGGC         ta         14         88         100         57.1         55.8         140           koain         CCATHATCTACGTGCAGGCC         TGAAAACAACACACCACCCTCTAACAGC         att         4         51         35         57.8         59.2         107	koaia	ACACAAGCACATACCAACCC	GGTGCTTCTATCTATGTACGTGC	ca	15	23	53	58.4	59.0	168
konia         CTAAGCCCACTITCATGCCC         TTCAATGAGCATGCTGGCG         ac         14         62         90         58.9         59.8         173           konia         TCAGCCTTTCCTTCGATGC         AAATAGAGGACGTTTCAGG         ta         14         71         99         60.4         55.2         157           konia         GATGAACTAGCATTAGCATTAGAGT         TGAGAACATACATGCACGTTGCG         ata         14         58         100         57.1         56.3         125           konia         CATATCATCHTCTATGATGACAGGC         GGCAAATGTCTAACTAGCGC         ata         14         41         43         56.7         59.8         140           konia         TGACHTACCTGTCGAGGC         TCAAATGCACACCCTCTAACAGC         ata         14         51         93         57.8         59.2         107	koaia	TCTAGGACACTTAGAGACCACC	AGGTTAAGGTGTAGAATATGACCC	ac	14	115	143	58.0	57.5	165
koaia         TCAGCCTTICCTTCCGATGC         AAATAGACGACCGTTTCACG         ta         14         71         99         60.4         56.2         157           koaia         GATGAACTGTAACGATTAGAGG         TGAGAACAATACACTGTAGCTG         ata         14         78         100         57.1         55.3         125           koaia         CATATACATCTATACTCTAATGAGCC         TGACAAATAGACATGTCAATTTCAATGT         44         44         8         56.7         59.8         104           koaia         TGACTTIACCTGTCAGCTGC         TCAAATGCAACCTCTAACAGC         aat         14         51         93         57.8         59.2         107	koaia	AAACATACCCACACACCCG	AGGTGATTTCTATCGGGTTGGG	ac	14	69	97	57.7	59.8	147
konia         GATGAACTGTAACATAGCATTAGAGG         TGAAACAATACATGTACGTCG         ata         14         88         100         97.1         55.3         125           konia         CCATTATCTATCHTCATGAGGCC         GTGCAAATGTCAACATGTAGGCC         att         44         41         83         56.7         59.8         140           konia         TGACATTACCTGTCAGCTCC         TCAAATGCACACCTTAACAGC         att         44         51         95         57.8         59.2         107	koaia	CTAAGCCCACTTTCATGCCC	TTTCAATGAGCATGCTGGCG	ac	14	62	90	58.9	59.8	173
konia         CCATTATCTATCTTTCATAGAGGCC         GTGCAAATGTTCAAATTTAGGGCC         att         14         41         83         56.7         59.8         140           konia         TGACTTTACCTGTCAGCTGC         TCAAATGCACACCTCTAACAGC         aat         14         51         93         57.8         59.2         107	koaia	TCAGCCTTTCCTTCCGATGC	AAATAGACGACCGTTTCACG	ta	14	71	99	60.4	56.2	157
koaia TGACTTTACCTGTCAGCTGC TCAAATGCACACCTCTAACAGC aat 14 51 93 57.8 59.2 107	koaia	GATGAACTGTAACTAGCATTAGAGG	TGAGAACAATACATGTACGTCG	ata	14	58	100	57.1	56.3	125
	koaia	CCATTATCTATCTTTCATAGAGGCC	GTGCAAATGTTCAAATTTAGGGCC	att	14	41	83	56.7	59.8	140
koaia AATTCTAACCTGATATCGGAGC ATGCTTCTTTGGGCGAGAGG aag 14 57 99 55.5 60.4 114	koaia	TGACTTTACCTGTCAGCTGC	TCAAATGCACACCTCTAACAGC	aat	14	51	93	57.8	59.2	107
	koaia	AATTCTAACCTGATATCGGAGC	ATGCTTCTTTGGGCGAGAGG	aag	14	57	99	55.5	60.4	114

#### Table 4

# Koa transcriptome. Koa EST-SSR information for comparison to genomic nuSSRs

	Motif Information	Transcriptome
Me	tadata	Acacia koa
	Total number of sequences analyzed	667025
	Number of sequences with >1 EST-SSR	5245
	Total numbers of EST-SSRs	5397
	Number of EST-SSRs	42
	Number of EST-SSRs with primers	2426
Dir	nucleotides	
1	AT TA	1195
2	GC CG	1
3	AC CA TG GT	859
4	AG GA CT TC	2426
Tri	nucleotides	
1	GGC GCG CGG GCC CCG CGC	22
2	ATG TGA GAT CAT ATC TCA	92
3	AGT GTA TAG ACT CTA TAC	11
4	AGG GAG GGA CCT CTC TCC	66
5	AAT ATA TAA ATT TTA TAT	88
6	CCA CAC ACC TGG GTG GGT	70
7	AGC GCA CAG GCT CTG TGC	112
8	AAG AGA GAA CTT TTC TCT	212
9	AAC ACA CAA GTT TTG TGT	94
10	ACG CGA GAC CGT GTC TCG	45
Мо	tif Length Excluding Compound SSRs	
	2 bp	4481
	3 bp	812
	4 bp	62
SSI	Rs with Primers Excluding Compound SSRs	
	2 bp	2067
	3 bp	339
	4 bp	20

# Evaluation of phylogeny using flanking region sequences

Flanking regions and a distance table were used to draw a phylogenetic tree for visualization of the relationship between these species. Phylogeny results indicated koa and koaia were closely related to each other however the aberrant koaia-A sorted separately. *G. max,* another legume family member, was also examined and is distinct from the *Acacia spp.* in this study (Figure 3).

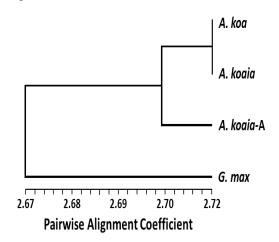


Figure 3

Phylogeny. Pairwise alignment of koa, koaia, koaia-A, and the soybean reference genomes.

# Discussion

The nuSSR abundance within the koa genome was 45 % greater than that of koaia, a somewhat expected result as 2.5 times more koa sequence was retrieved and analyzed than koaia. A total of 28,275 nuSSRs in koaia resulted in 9.73 SSRs per kb of genome sequence while koa had approximately 14.4 SSRs per kb of sequence. Differences in nuSSR abundance have been reported previously for other somewhat related species such as Glycine spp. (Ozyigit et al. 2015). Model species such as Arabidopsis (Arabidopsis thaliana) and tomato (Solanum lycopersicum) have considerably higher numbers of SSRs (529,000 and 285,000) per kb of sequence (Cheng et al. 2016). The model tree poplar (Populus trichocarpa) has 667,900 SSRs per kb of sequence (Sonah et al. 2011). Therefore, the lower volume of nuSSRs within the koa and koaia genomes sequenced here may explain the limited information available regarding markers for these two species. EST-SSRs were included for comparison

Here, we developed the first koa and koaia database containing high numbers of nuSSR and EST-SSR markers with unique flanking sequences. SSRs are as useful as SNPs for subsequent parental analyses and basic population genetics despite the introduction of single-nucleotide polymorphism (SNP) analyses (Cappa *et al.* 2016). Recent results in apple (Zhang *et al.* 2012) and ten other hardwood tree species (Staton *et al.* 2015) showed AT/TA motifs were the most common dinucleotides found within the sequences. These results mirrored those observed in this study where the AT/TA dinucleotide motif was most abundant. We obtained 700 tri- and tetranucleotide nuSSR primers for koa and 300 for koaia. In general, longer SSRs amplify more alleles among cultivars and species, whereas shorter SSRs do so only among species. (Smulders *et al.* 1997).

As NextGeneration sequencing technologies continue to provide high-throughput high-resolution data, analyses of expression data can be essential to experiments on non-model organisms. NextGeneration sequencing (NGS) provides a relatively inexpensive method for analysis of genome and transcriptomic sequences for non-model species such as koa, Despite its low cost, very little NGS has been published for koa thus, efforts to supply a wealth of nuSSRs and EST-SSRs for use in future studies are highly desirable.

Our results showed that *A. koa* and *A. koia* are sorted with each other. Recently a new study based on genotyping by sequencing data of various population of *A. koa* and *A. koaia* revealed that both species are genetically similar. So. Our result based on flanking region of SSRs confirmed their study. We provide, for public use, a multitude of primers for use in genomic and transcriptomic studies.

# Conclusions

Koa and koaia grow on several islands within the Hawaiian archipelago. Limited regeneration and the perseverance of disease and invasive species has hindered population growth thus, increased efforts to identify genomic and genetic methods of selection are prudent. Additional sampling and genome analyses may help uncover molecular variation and other nuances within the genetic structures of koa and koaia. Development of SSRs is an inexpensive and invaluable method to research and evaluate *Acacia spp*. genomes. This study identified 11,019 SSRs unique primer pairs confirmed with ePCR. These SSRs will aid future investigations of koa and koaia genomes and support future breeding and genetic improvement program efforts.

# Acknowledgements

The authors thank Drs. Nicholas LaBonte and Mark Coggeshall for their comments and critiques of a previous version of this manuscript. We acknowledge the State of Hawai'i Department of Land and Natural Resources (DLNR) – Division of Forestry and Wildlife (DOFAW) and/or State Parks, as well as the USDA Forest Service (FS) as land managers and administrators for the Hawai'i Experimental Tropical Forest. Funding provided by the US Forest Service and the Atherton Family Foundation. Mention of a trademark, proprietary product, or vendor does not constitute a guarantee or warranty of the product by the US Department of Agriculture and does not imply its approval to the exclusion of other products or vendors that also may be suitable.

# References

Adamski DJ, NS Dudley, CW Morden and D Borthakur (2012) Genetic differentiation and diversity of Acacia koa populations in the Hawaiian Islands. Plant Species Biol. 27(3): 181-190.

https://doi.org/10.1111/j.1442-1984.2011.00359.x

- Adamski DJ, NS Dudley, CW Morden and D Borthakur (2013) Cross-amplification of nonnative Acacia species in the Hawaiian Islands using microsatellite markers from Acacia koa. Plant Biosyst. 147(4): 1088-1091. https://doi.org/10.1080/11263504.2012.749958
- Butcher PA and GF Moran (2000) Genetic linkage mapping in Acacia mangium.
  2. Development of an integrated map from two outbred pedigrees using RFLP and microsatellite loci. Theor Appl Genet. 101(4): 594-605.
  <u>https://doi.org/10.1007/s001220051521</u>
- Cappa EP, J Klapste, MN Garcia, PV Villalba and SN Marcucci Poltri (2016) SSRs, SNPs and DArTs comparison on estimation of relatedness and genetic parameters' precision from a small half-sib sample population of Eucalyptus grandis. Mol Breeding. 36(7): 96-114. https://doi.org/10.1007/s11032-016-0522-7
- Cheng J, Z Zhao, B Li, C Qin, Z Wu, DL Trejo-Saavedra, X Luo, J Cui, RF Rivera-Bustamante, S Li and K Hu (2016) A comprehensive characterization of simple sequence repeats in pepper genomes provides valuable resources for marker development in Capsicum. Scientific Reports. 6(1): 1-12. <u>https://doi.org/10.1038/srep18919</u>
- Fredua-Agyeman R, D Adamski, RJ Liao, C Morden and D Borthakur (2008) Development and characterization of microsatellite markers for analysis of population differentiation in the tree legume Acacia koa (Fabaceae: Mimosoideae) in the Hawaiian Islands. Genome. 51(12): 1001-1015. <u>https://doi.org/10.1139/g08-087</u>
- Guillemaud T, L Broadhurst, I Legoff, M Henery, A Blin, C Ducatillion, N Ferrando and T Malausa (2015) Development of 23 Polymorphic Microsatellite Loci in Invasive Silver Wattle, Acacia dealbata (Fabaceae). Applications in Plant Sciences. 3(5): A1500018. <u>https://doi.org/10.3732/apps.1500018</u>
- Gugger PF, CT Liang, VL Sork, P Hodgskiss, and JW Wright (2017) Applying landscape genomic tools to forest management and restoration of Hawaiian koa (Acacia koa) in a changing environment. Evolutionary Applications. 0(0): 1-12. <u>https://doi.org/10.1111/eva.12534</u>
- Hopley T, LM Broadhurst and MG Gardner (2015) Isolation via 454 sequencing and characterisation of microsatellites for Acacia montana (Fabaceae), Mallee wattle: an endemic shrub from southeastern Australia. Conserv Genet Resour. 7(1): 171-172. <u>https://doi.org/10.1007/s12686-014-0321-6</u>
- Le S, R Wickneswari, CE Harwood, MJ Larcombe, RA Griffin, A Koutoulis, JL Harbard, KS Cyer, LW Yee, TH Ha and RE Vaillancourt (2016) A multiplexed set of microsatellite markers for discriminating Acacia mangium, A. auriculiformis, and their hybrid. Tree Genet Genomes. 12(2): A31. https://doi.org/10.1007/s11295-016-0990-2
- Lepais O and CFE Bacles (2011) Comparison of random and SSR-enriched shotgun pyrosequencing for microsatellite discovery and single multiplex PCR optimization in Acacia harpophylla F. Muell. Ex Benth. Mol Ecol Resour. 11(4): 711-724. <u>https://doi.org/10.1111/j.1755-0998.2011.03002.x</u>
- Levy E, M Byrne, DJ Coates, SJ van Leeuwen, S McArthur, B Macdonald and MG Gardiner (2014) Isolation via 454 sequencing, and characterisation of microsatellite markers for the Pilbara endemic Acacia atkinsiana (Fabaceae). Conserv Genet Resour. 6(3): 585-587.

https://doi.org/10.1007/s12686-014-0146-3

- Magoč T and SL Salzberg (2011) FLASH: Fast length adjustment of short reads to improve genome assemblies. Bioinformatics. 27(21): 2957-2963. https://doi.org/10.1093/bioinformatics/btr507
- Millar MA and M Byrne (2007) Characterization of polymorphic microsatellite DNA markers for Acacia saligna (Labill.) H.L.Wendl. (Mimosaceae). Mol Ecol Notes. 7(6): 1372-1374. <u>https://doi.org/10.1111/j.1471-8286.2007.01890.x</u>
- Millar MA, M Byrne, I Nuberg and M Sedgley (2008) High outcrossing and random pollen dispersal in a planted stand of Acacia saligna subsp. saligna re-

vealed by paternity analysis using microsatellites. Tree Genet Genomes. 4(3): 367-377. <u>https://doi.org/10.1007/s11295-007-0115-z</u>.

Morgante M and AM Olivieri (1993) PCR-amplified microsatellites as markers in plant genetics. Plant J. 3(1):175-182.

#### https://doi.org/10.1111/j.1365-313x.1993.tb00020.x.

Nevill PG, JM Anthony and SL Krauss (2010) Isolation and characterization of microsatellite markers for the banded ironstone endemic Acacia karina (Leguminosae: Mimosaceae) and cross-species amplification with A. stanleyi and A. jibberdingensis. Conserv Genet Resour. 2(1): 321-323.

### https://doi.org/10.1007/s12686-010-9219-0.

Ng CH, SC Koh, SL Lee, KKS Ng, A Mark, M Norwati and R Wickneswari (2005) Isolation of 15 polymorphic microsatellite loci in Acacia hybrid (Acacia mangium x Acacia auriculiformis). Mol Ecol Notes. 5(3): 572-575.

### https://doi.org/10.1111/j.1471-8286.2005.00994.x

- Otero-Arnaiz A, A Schnabel, TC Glenn, NA Schable, C Hagen and L Ndong (2005) Isolation and characterization of microsatellite markers in the East African tree, Acacia brevispica (Fabaceae: Mimosoideae). Mol Ecol Notes. 5(2): 366-368. <u>https://doi.org/10.1111/j.1471-8286.2005.00929.x.</u>
- Ozyigit II, I Dogan and E Filiz (2015) In silico analysis of simple sequence repeats (SSRs) in chloroplast genomes of Glycine species. POJ. 8(1): 24-29. doi: <u>http://www.pomics.com/FILITZ 8\_1\_2015\_24\_29.pdf</u>.
- Rohlf FJ. (2000) NTSYSpc Numerical Taxonomy and Multivariate Analysis System version 2.2. User Guide. Exeter Software. https://www.researchgate.net/publication/246982444\_NTSYS-pc\_-\_Numerical\_Taxonomy\_and\_Multivariate\_Analysis\_System. Accessed 26 August 2017
- Ruiz-Guajardo JC, A Otero-Arnaiz, T Taylor, G Stone, TC Glenn, NA Schable, JT Miller, S Preuss, and A Schnabel (2007) Isolation of polymorphic microsatellite markers in the sub-Saharan tree, Acacia (Senegalia) mellifera (Fabaceae: Mimosoideae). Mol Ecol Notes. 7(6): 1138-1140.

# https://doi.org/10.1111/j.1471-8286.2007.01809.x

- Rushanaedy I, TC Jones, NS Dudley, RJF Liao, R Agbayani, and D Borthakur (2012) Chitinase is a Potential Molecular Biomarker for Detecting Resistance to Fusarium oxysporum in Acacia koa. Trop Plant Biol. 5(3):244-252. <u>https://doi.org/10.1007/s12042-012-9108-7.</u>
- Shiraishi A, JF Leslie, S Zhong, and JY Uchida (2012) AFLP, pathogenicity, and VCG analyses of Fusarium oxysporum and Fusarium pseudocircinatum from Acacia koa. Plant Dis. 96(8): 1111-1117. <u>https://doi.org/10.1094/pdis-06-11-0491</u>

### Shyu C, JA Foster and LJ Forney (2002) Electronic polymerase chain reaction (ePCR) search algorithm. Proceedings of the IEEE Computer Society Bioinformatics Conference (CSB'02). <u>https://doi.org/10.1109/csb.2002.1039361</u>

Smulders MJM, G Bredemeijer, W Rus-Kortekaas, P Arens and B Vosman (1997) Use of short microsatellites from database sequences to generate polymorphisms among Lycopersicon esculentum cultivars and accessions of other Lycopersicon species. Theor Appl Genet. 94(2): 264-272.

### https://doi.org/10.1007/s001220050409

Sonah H, RK Deshmukh, A Sharma, VP Singh, DK Gupta, RN Gacche, JC Rana, NK Singh and TR Sharma (2011) Genome-wide distribution and organization of microsatellites in plants: An insight into marker development in Brachypodium. PLoS One. 6(6): e21298.

### https://doi.org/10.1371/journal.pone.0021298

Staton M, T Best, S Khodwekar, S Owusu, T Xu, Y Xu, R Jennings, R Cronn, A Kathiravetpilla, M Coggeshall, O Gailing, H Liang, J Romero-Severson, S Schlarbaum and JE Carlson (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

# Tamura K, G Stecher, D Peterson, A Filipski and S Kumar (2013) MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30(12): 2725-2729. <u>https://doi.org/10.1093/molbev/mst197.</u>

Young ND. (1996) QTL mapping and quantitative disease resistance in plants. Annu Rev Phytopathol. 34(1): 479-501.

#### https://doi.org/10.1146/annurev.phyto.34.1.479

Zhang Q, B Ma, H Li, Y Chang, Y Han, J Li, G Wei, S Zhao, A Awas Khan, Y Zhou, C Gu, X Zhang, Z Han, S Korban, S Li and Y Han (2012) Identification, characterization, and utilization of genome-wide simple sequence repeats to identify a QTL for acidity in apple. BMC Genom. 13(1): a537. <u>https://doi.org/10.1186/1471-2164-13-537.</u>