

Genomes for Plant and Crop Science Research

Julia Voelker, Mervyn Shepherd, Catherine Nock, Graham J. King, Ramil P. Mauleon

Southern Cross Plant Science

Faculty of Science and Engineering

Southern Cross University



Genomics for better informed (pre) breeding for crops of interest

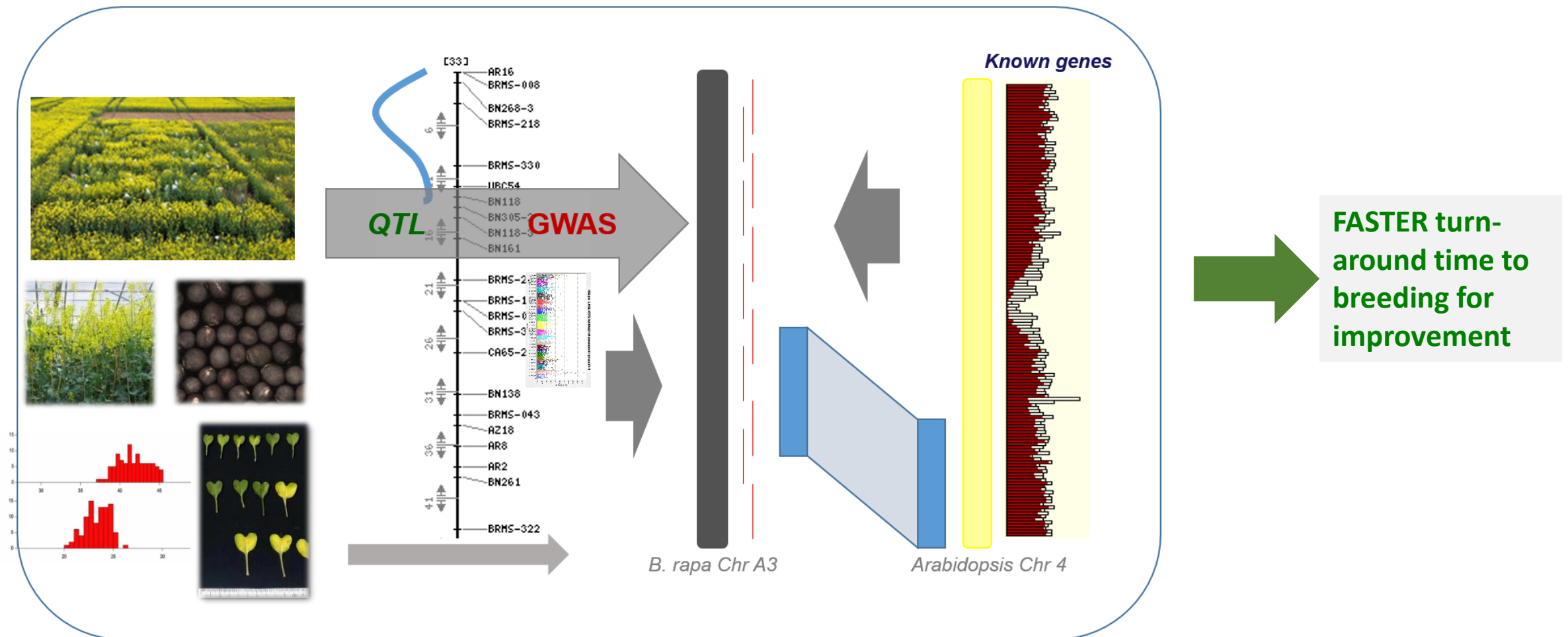
Deep + broad understanding of genome organization

Phenotypic traits

Linkage

Genome

Functional & comparative genomics



CROP OMICS FOR GLOBAL NICHE CROPS

pre-breeding and breeding



Macadamia

Pre-breeding genetics/genome anchoring

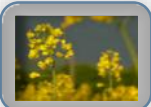


Horticulture
Innovation
Australia



Tea Tree

Pre-breeding /breeding program + CRC-P



**Yellow sarson,
Mustard Oil**

Pre-breeding / breeding selection



Australian
Mustard Oils

**MULTINATIONAL BRASSICA
GENOME PROJECT (MBGP)**

Brassica rapa R-o-18 (yellow sarson)



SCU_BraROA_2.3

Organism name: [Brassica rapa subsp. trilocularis \(field mustard\)](#)

Infraspecific name: Ecotype: Yellow sarson

Isolate: R-o-18

BioSample: [SAMN16250067](#)

BioProject: [PRJNA649364](#)

Submitter: Brassica rapa R-o-18 genome sequencing consortium

Date: 2021/03/30



Prof. Graham J. King

GenBank accession: [GCA_017639395.1](#)

Molecule	Total Length	Scaffold Count	Ungapped Length	Scaffold N50	Spanned Gaps	Unspanned Gaps
All	346,506,900	295	345,222,260	32,457,817	855	0
Chromosome A10	22,576,415	1	22,526,320	22,576,415	57	0
Chromosome A01	34,445,504	1	33,995,968	34,445,504	89	0
Chromosome A02	36,095,996	1	36,038,788	36,095,996	73	0
Chromosome A03	32,457,817	1	32,448,697	32,457,817	20	0
Chromosome A04	23,305,641	1	23,284,797	23,305,641	36	0
Chromosome A05	32,871,474	1	32,825,490	32,871,474	76	0
Chromosome A06	30,241,088	1	30,224,719	30,241,088	32	0
Chromosome A07	27,715,701	1	27,692,758	27,715,701	19	0
Chromosome A08	25,318,967	1	25,286,307	25,318,967	31	0
Chromosome A09	60,281,870	1	60,145,694	60,281,870	162	0
unplaced	21,196,427	285	20,752,722	233,017	260	0

Macadamia integrifolia HAES 741




HOME **ISSUES** ABOUT SERIES ARTICLE TYPES PUBLISH

[← Previous Article](#)

[Next Article >](#)

Chromosome-Scale Assembly and Annotation of the Macadamia Genome (*Macadamia integrifolia* HAES 741)

 Catherine J. Nock, Abdul Baten, Ramil Mauleon, Kirsty S. Langdon, Bruce Topp, Craig Hardner, Agnelo Furtado, Robert J. Henry and Graham J. King

G3: GENES, GENOMES, GENETICS October 1, 2020 vol. 10 no. 10 3497-3504;
<https://doi.org/10.1534/g3.120.401326>



Dr. Catherine Nock

ID: 34480

Organism Overview : [Organelle Annotation Report \[1\]](#)

Macadamia integrifolia

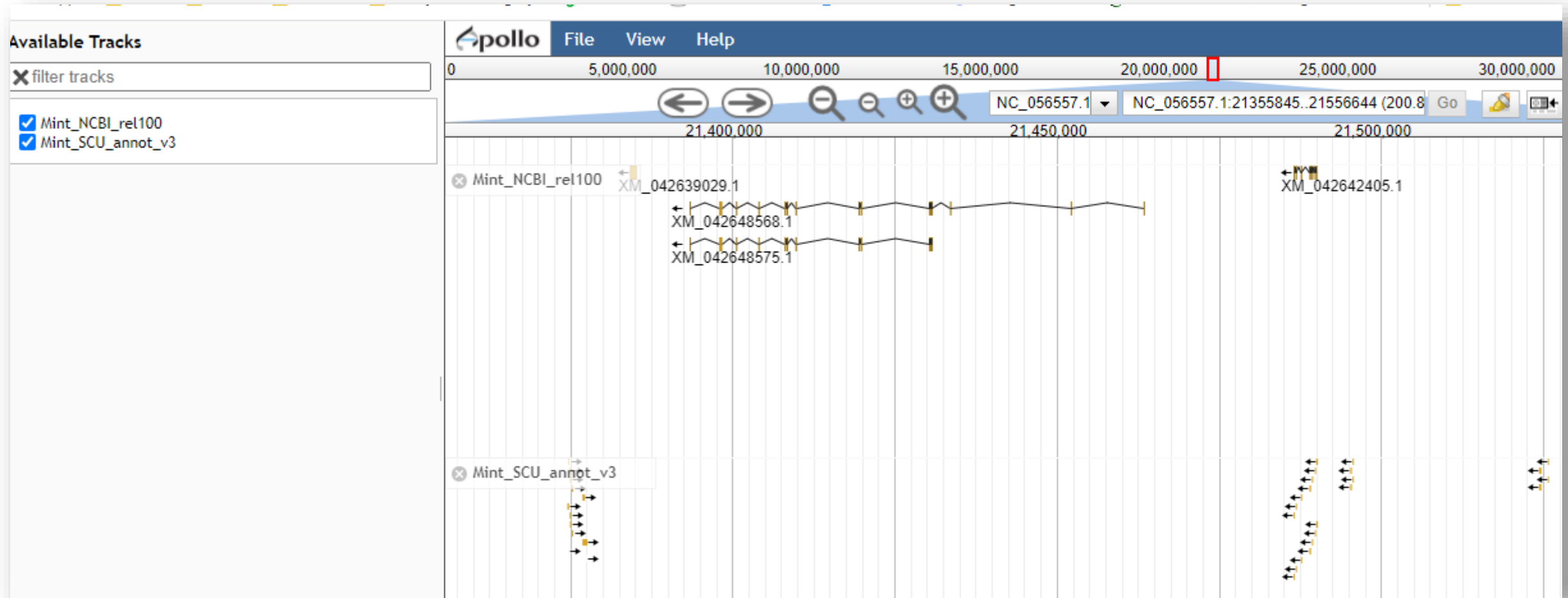
Macadamia integrifolia overview

Lineage: Eukaryota[7993]; Viridiplantae[842]; Streptophyta[758]; Embryophyta[752]; Tracheophyta[744]; Spermatophyta[733]; Magnoliopsida[717]; Proteales[3]; Proteaceae[2]; Macadamia[1]; Macadamia integrifolia[1]
Macadamia integrifolia genome reference project

Summary

Submitter:	Southern Cross University
Assembly level:	Chromosome
Assembly:	GCA_013358625.1 SCU_Mint_v3 scaffolds: 2,644 contigs: 23,449 N50: 60,651 L50: 3,405
BioProjects:	PRJNA748012
Whole Genome Shotgun (WGS):	INSDC: JAAEEG000000000.1
Statistics:	total length (Mb): 744.937 protein count: 46649 GC%: 39.2926
NCBI Annotation Release:	100

Community-driven improvement of macadamia genome annotation...



Improving gene models with data from new studies through the macadamia research community

Melaleuca alternifolia (tea tree)



Data Release

A high-quality draft genome for *Melaleuca alternifolia* (tea tree): a new platform for evolutionary genomics of myrtaceous terpene-rich species

Julia Voelker*, Mervyn Shepherd, Ramil Mauleon

DOI: [10.46471/gigabyte.28](https://doi.org/10.46471/gigabyte.28) Views: 300 Downloads: 41

0 0 0

[Article Review History](#)

Published online : 9 August 2021



PhD candidate Julia Voelker and Dr. Mervyn Shepherd

SCU_Malt_1

Organism name: [Melaleuca alternifolia \(tea tree\)](#)
Isolate: SCU01
BioSample: [SAMN17927736](#)
BioProject: [PRJNA702189](#)
Submitter: Southern Cross University
Date: 2021/09/10
Assembly level: Scaffold
Genome representation: full
RefSeq category: representative genome
GenBank assembly accession: [GCA_019926035.1](#) (latest)
RefSeq assembly accession: n/a
RefSeq assembly and GenBank assembly identical: n/a
WGS Project: [JAGKPW01](#)
Assembly method: MaSuRCA v. 3.4.0
Expected final version: no
Genome coverage: 350.0x
Sequencing technology: PacBio Sequel; Illumina HiSeq

IDs: 10874731 [UID] 28792958 [GenBank]



Apollo usage example:

The manual annotation of the terpene synthase gene family in *Melaleuca alternifolia*

Julia Voelker

PhD candidate at the Faculty of Science and Engineering

Australian Myrtaceae

Melaleuca alternifolia (tea tree)



Eucalyptus grandis



Poyt448 Peter Woodard / CC BY-SA
(<https://creativecommons.org/licenses/by-sa/3.0>)

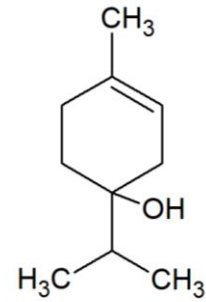
Australian Myrtaceae

Melaleuca alternifolia (tea tree)

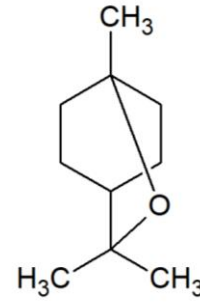


Terpene
synthases
(TPS)

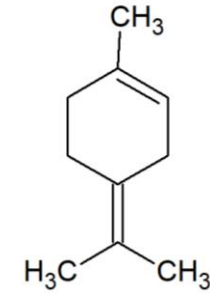
Terpenoids



terpinen-4-ol



1,8-cineole



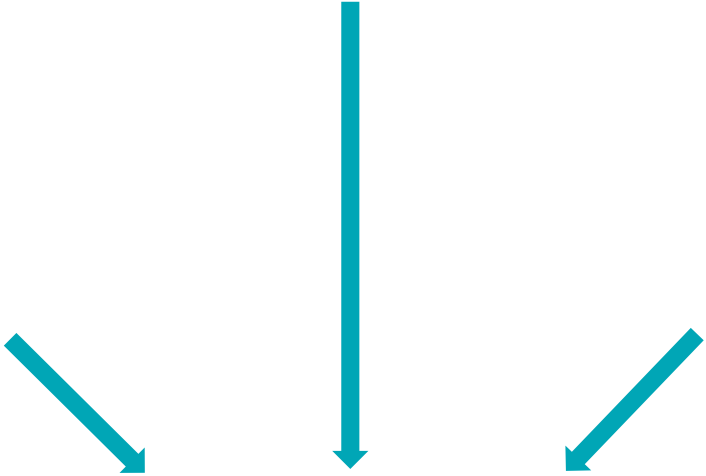
terpinolene

Automatic Annotation
Fgenesh++
GenemarkET+
Protein and RNAseq alignment evidence

Assembly
MaSuRCA: hybrid assembly using
PacBio and Illumina reads

Gene structure evidence
Exonerate: TPS protein alignments to genome
STAR: RNAseq alignments to genome

Manual TPS annotation with Apollo



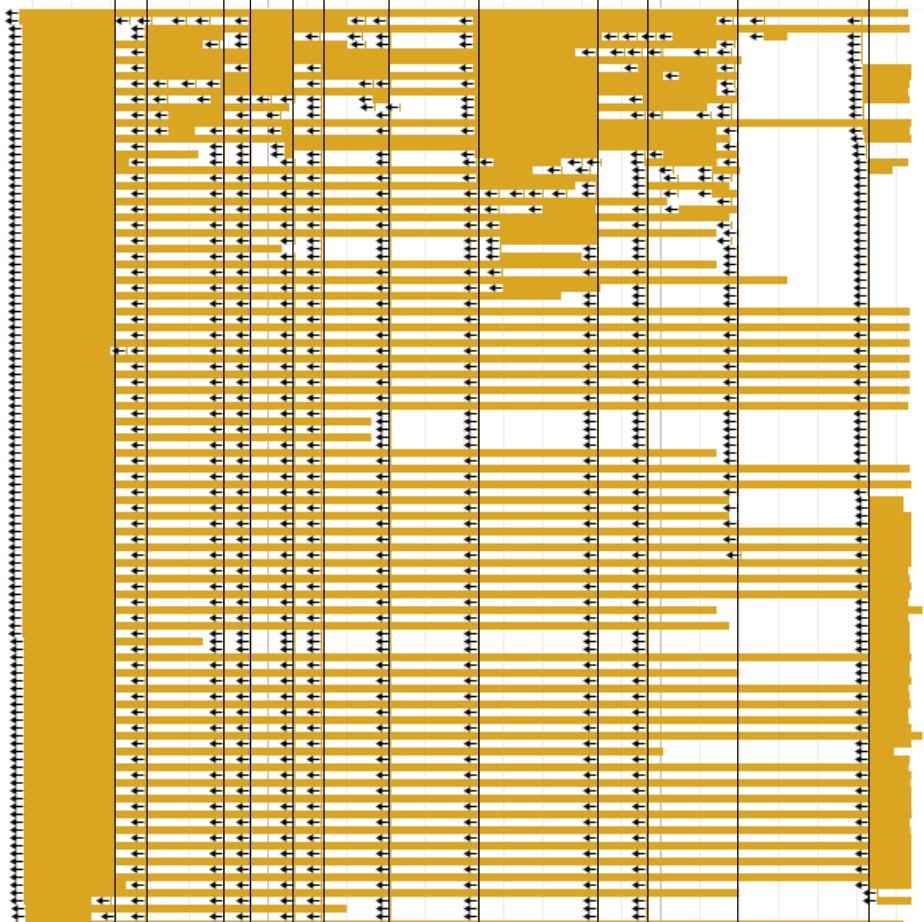
573,750 575,000 576,250 577,500

User-created Annotations scf7180000029212_mRNA_102-00001

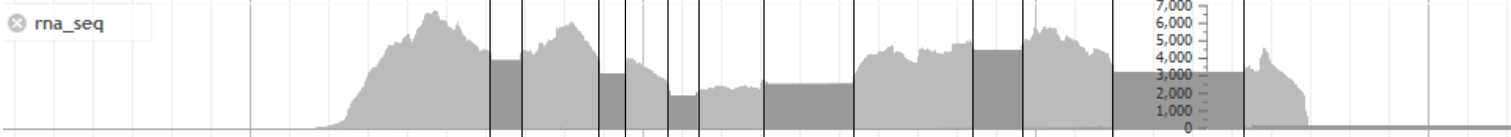
fgenes scf7180000029212_mRNA_102 protein supported,protein: gi 702517665 |ref|XP_010042263.1| PREDICTED: (-)-germacrene D synthase-like [Eucalyptus grandis]

genemark 45730_t 45731_t

exonerate



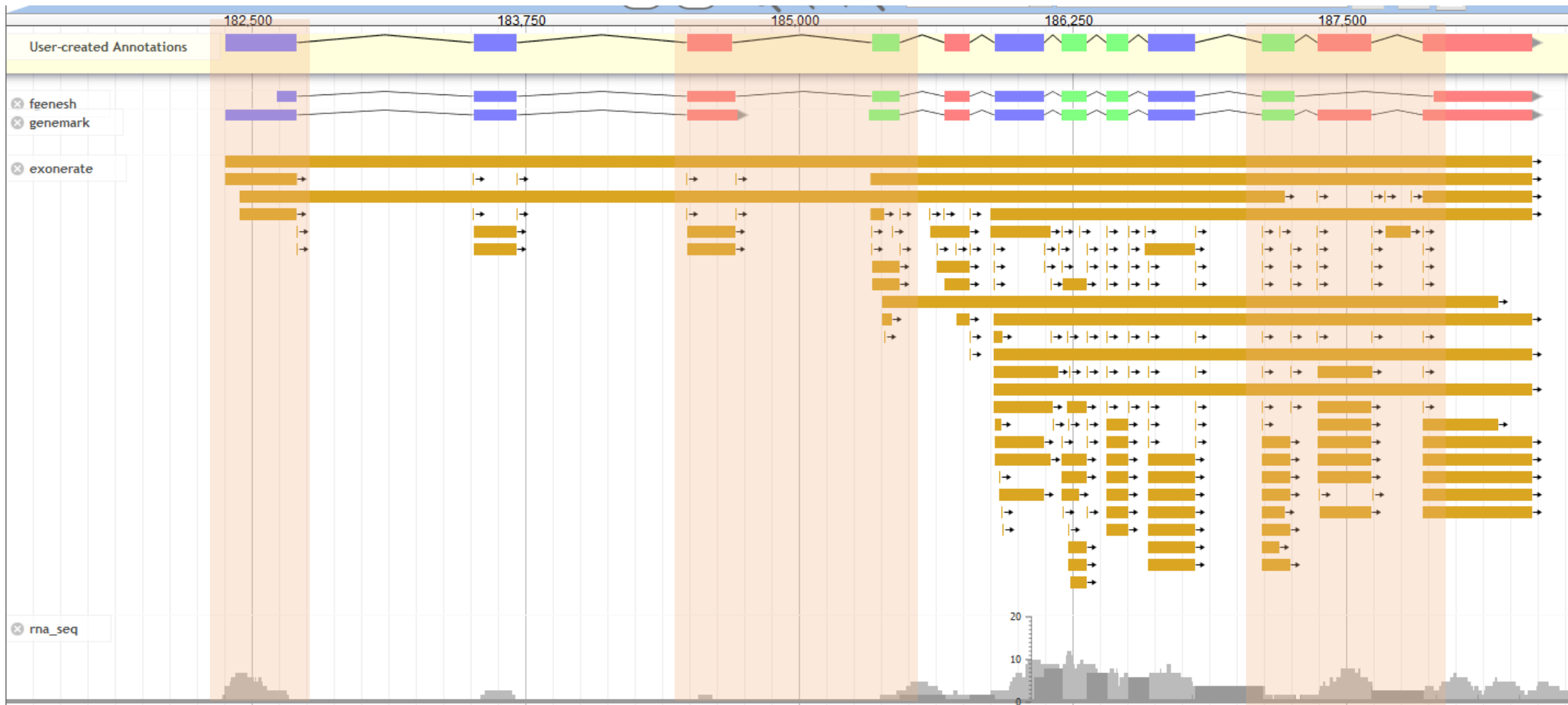
height reached Max height reached Max height reached Max height reached



Fgenes++
GenemarkET+

Exonerate TPS alignments

RNAseq alignments

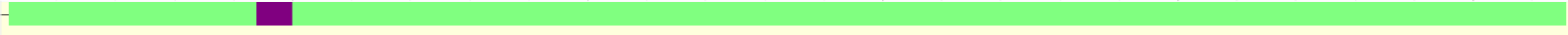




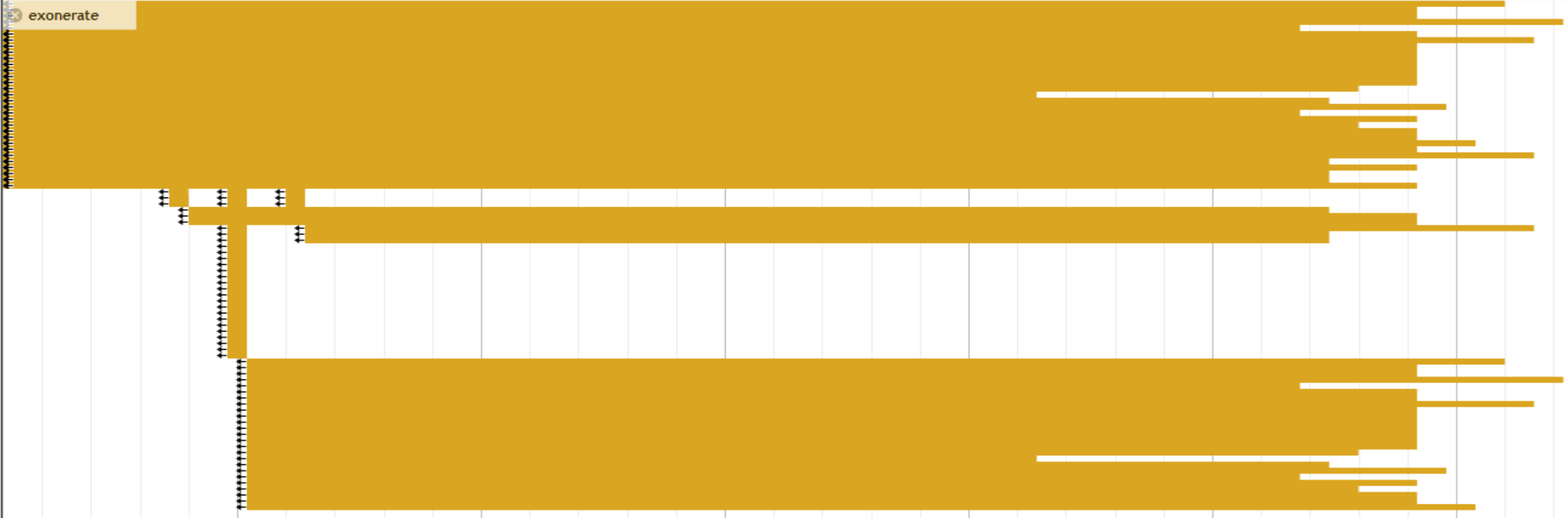
Reference sequence

R K T N K R N A T M A L E S E A L Y L R K I S P P H M P G I I C S A F S Y V A C F L S W R R E R D E E I K R G H F S F
E N E K K C N H G V G I R S L F E K I I P P Y A W M I I C S A F S Y V A C F L S W R R E R D E E I K R G H F S F
AGAAAACGAATAAAAGAAATGCAACCATGGCGTTGGAAATCAGAAGCTTATTTGAGAAAATAATCCCCCATATGCTGGATGATAATTGCTGATCGGC
TCTTTTGCCTTATTTTCTTTTACGTTACCGCAACCTTAGTCTTTCGAATAAACTCTTTTATTAGGGGGGTATACGGACCTACTATTAACGACTAGCCG
L F R I F S I C G H R Q F F S I Q S F L G G M H R S S L K S I P T S H P R N G S S F I F L N S P V N R
L F V F L L F A V M A N S D S A K L F Y D G W I G P H Y N A S R R E A I H G T E K A P S S S S I Q L S M E
S F S Y F F H L W P T P I L L K N S F I I G G Y A Q I I I Q Q D A N L T A Q K R L Q L L H L S K F P C K R

User-created Annotations



fgenesch
genemark 21934_t





Thanks for your attention!

