



International Center for Tropical Agriculture
Since 1967 / *Science to cultivate change*

Sequencing of common bean genotypes provides a comprehensive resources for genetic introgression studies and molecular breeding

Juan David Lobaton, Tamara Miller

Juanita Gil, Alvaro Soler,

Jorge Duitama, Paul Gepts and

Bodo Raatz

June 2016

Lima, Peru

j.d.lobaton@cigar.org

b.raatz@cgiar.org



Hopes grow for climate-proof beans

By Mark Kinver
Environment reporter, BBC News

© 25 March 2015 | Science & Environment

A breakthrough in the development of temperature-resilient beans could help sustain a vital source of protein for millions of people around the globe.

The news offers a glimmer of hope for many farmers in developing nations.

Climate projections suggest that 50% of the countries' bean production will be lost by 2050 if farmers do not have access to the new variety of bean.

The discovery was made by plant breeders at CGIAR, a global agricultural research group.

One of the plant breeders involved in the research, Steve Beebe - a senior bean researcher at the International Center for Tropical Agriculture (CIAT) - explained that bean plants were, as a rule, very sensitive to excessive heat.



Beans are the staple diet for almost half a billion people around the globe.



High-iron beans: a new "superfood" for Colombia

by Neil Palmer | Jun 13, 2016



New high-iron beans have been officially released in Colombia.

They contain as much as 60% more iron than normal beans, and are intended to address the problem of iron deficiency - the world's most persistent nutritional ailment.



Pulse Profile: Steve Beebe, the Bean Breeder

by Paul Blake | Jun 7, 2016



Bioinformatic analysis of genotype by sequencing (GBS) data with NGSEP

Claudia Perea, Juan Fernando De La Hoz, Daniel Felipe Cruz, Juan David Lobaton, Paulo Izquierdo, Juan Camilo Quintero, Bodo Raatz and Jorge Duitama ✉

BMC Genomics 2016 17(Suppl 5):498 | DOI: 10.1186/s12864-016-2827-7 | © The Author(s) 2016

Published: 31 August 2016

Abstract

Background

The recent development and availability of different genotype by sequencing (GBS) protocols provided a cost-effective approach to perform high-resolution genomic analysis of entire populations in different species. The central component of all these protocols is the digestion of the initial DNA with known restriction enzymes, to generate sequencing fragments at predictable and reproducible sites. This allows to genotype thousands of genetic markers on populations with hundreds of individuals. Because GBS protocols achieve parallel genotyping through high throughput sequencing (HTS), every GBS protocol must include a bioinformatics pipeline for analysis of HTS data. Our bioinformatics group recently developed the Next Generation Sequencing Eclipse Plugin (NGSEP) for accurate, efficient, and user-friendly analysis of HTS data.

Results

Here we present the latest functionalities implemented in NGSEP in the context of the analysis of GBS data. We implemented a one step wizard to perform parallel read alignment, variants identification and genotyping from HTS reads sequenced from entire populations. We added different filters for variants, samples and genotype calls as well as calculation of summary statistics overall and per sample, and diversity



NGSEP

NGSEP (Next Generation Sequencing Experience Platform)

Brought to you by: dfcruz00, jduitama, juandelahoz

Summary | Files | Reviews | Support | Wiki | Tickets | Discussion | Blog | Code

★ Add a Review

↓ 224 Downloads (This Week)

📅 Last Update: 3 days ago



Download

net.sf.ngsep_3.0.1.201608051808.jar

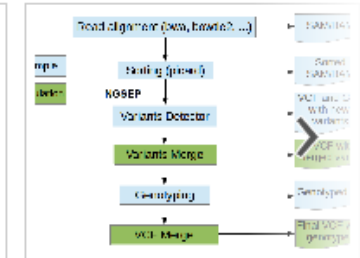
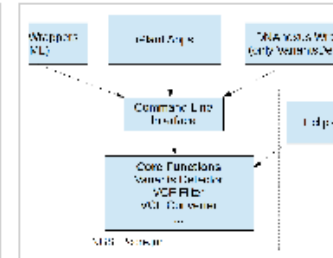
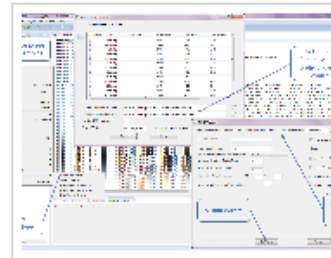
🐦 Tweet

👍 G+1 0

👍 Me gusta 1



Browse All Files



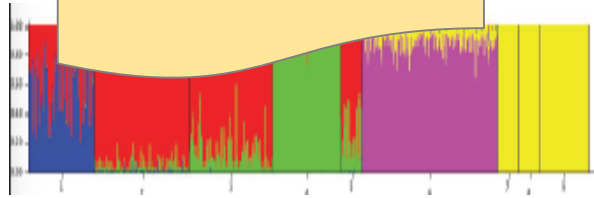
Description

NGSEP now is much more than the plugin so we decided to rebrand it as Next Generation Sequencing Experience Platform. NGSEP is an integrated framework for analysis of high throughput sequencing reads with several deployment alternatives for different kinds of users. The main functionality of NGSEP is the variants detector, which allows to make simultaneous discovery of SNVs, indels, and CNVs. We also provide a user interface for bowtie2, alignments sorting, merging of variants from different samples, functional annotation, filtering, format conversion, and several additional features.

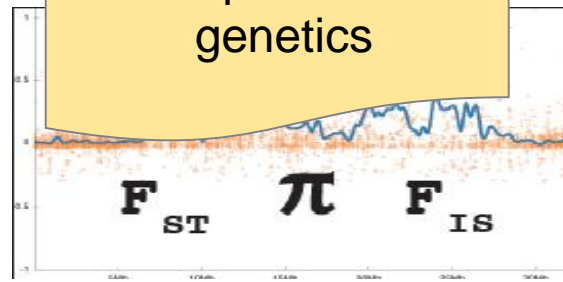
Note: Since we offer different distribution alternatives for different kinds of users, the quick download (green) button may not lead you to the version of the product that you need. Please visit our web page to choose the distribution of NGSEP that better suits your needs.

Variants detection
SNPs , small /large indels,
SNVs and
translocations

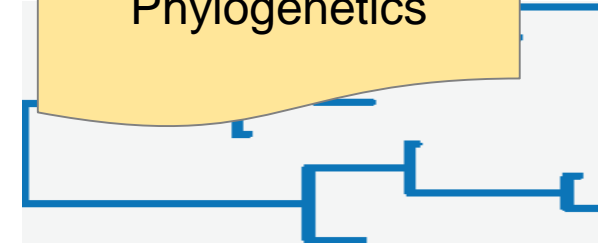
Population
structure



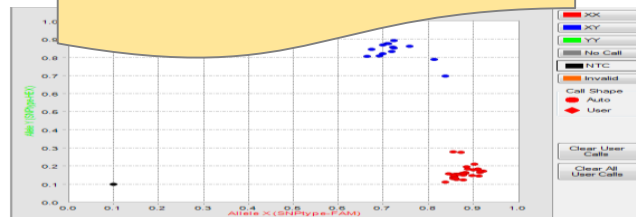
Population
genetics



SNPs
Phylogenetics



Genotyping
Chips



SNP
selection

Yale_AFR298	T	T	C	CA	C	T	T	T	G	C	T	T	A	G	C	C	C
Yale_AND696	T	T	C	CA	C	T	T	A	G	C	T	T	A	G	C	C	C
Yale_G10474	T	T	C	CA	C	T	T	A	G	C	T	T	A	G	C	C	C
Yale_G35346	G	A	A	AT	T	A	C	A	A	A	C	C	G	A	T	A	A
Yale_G40001	T	A	C	CA	C	T	C	T	G	C	T	T	A	A	C	C	C
Yale_G5686	T	T	C	CA	C	T	T	T	G	C	T	T	A	G	C	C	C
Yale_MD23-24	T	T	C	CA	C	T	T	T	G	C	T	T	A	G	C	C	C
Yale_SEA5	T	T	C	CA	C	T	T	T	G	C	T	T	A	G	C	C	C
Yale_VAX1	T	T	C	CA	C	T	T	T	G	C	T	T	A	G	C	C	C

Andean Ref G19833



AFR_298



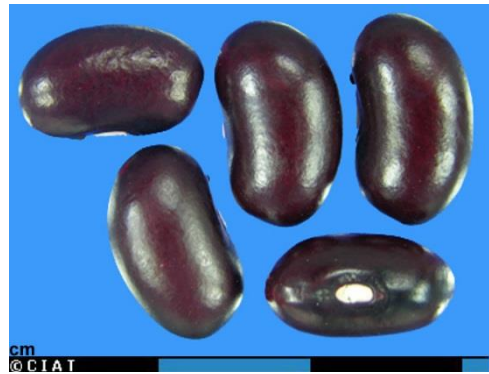
AND_665



G5686



Mesoamerican MD_23-24



SEA_5



VAX_1



G10474



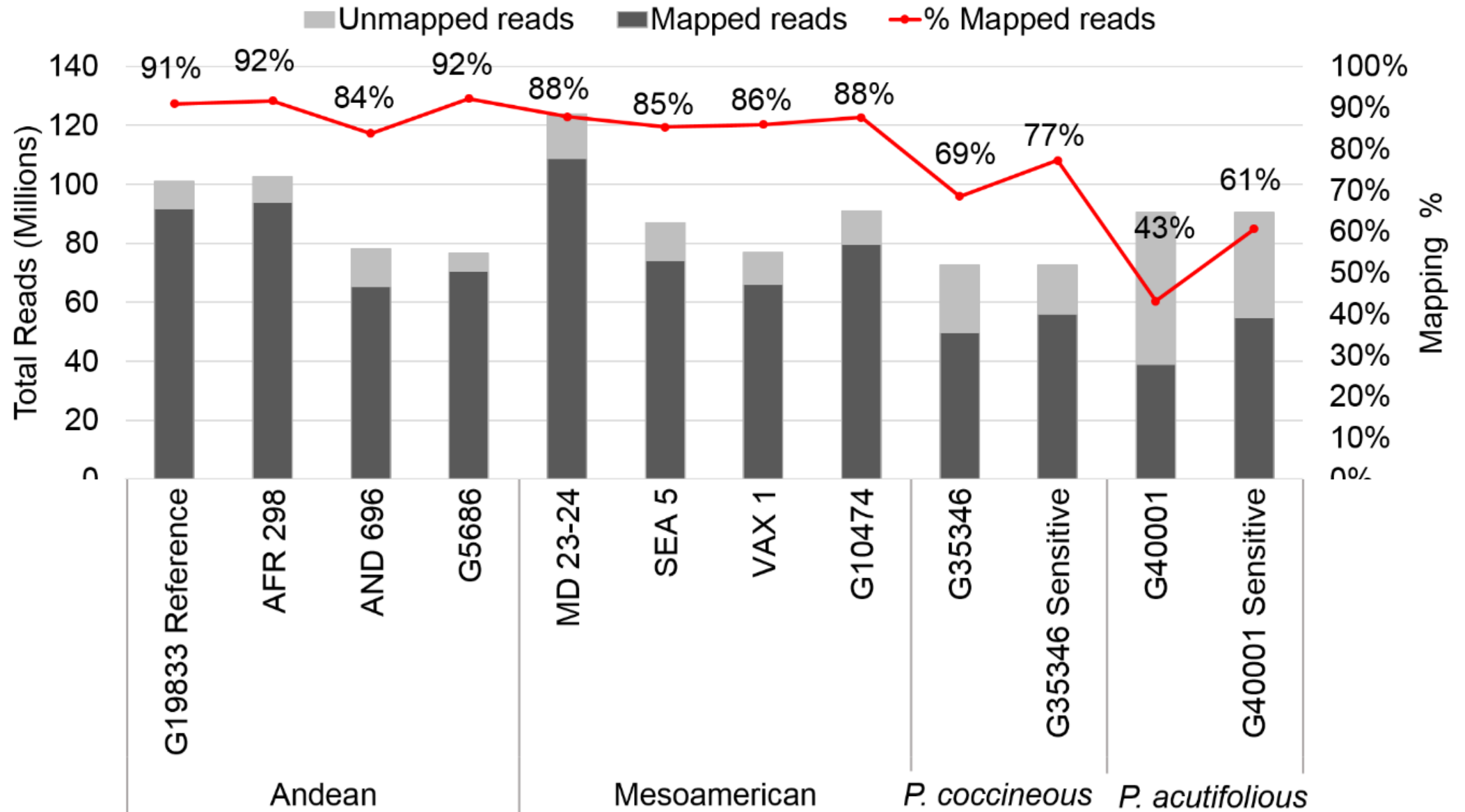
P. coccineus G35346



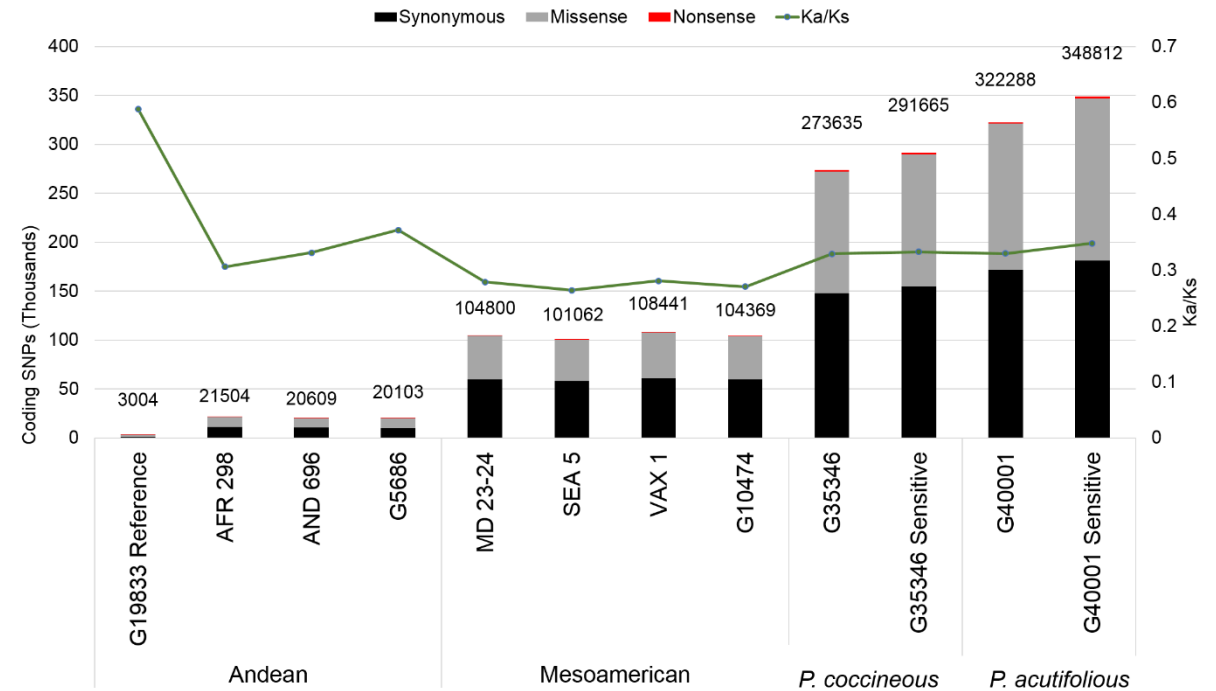
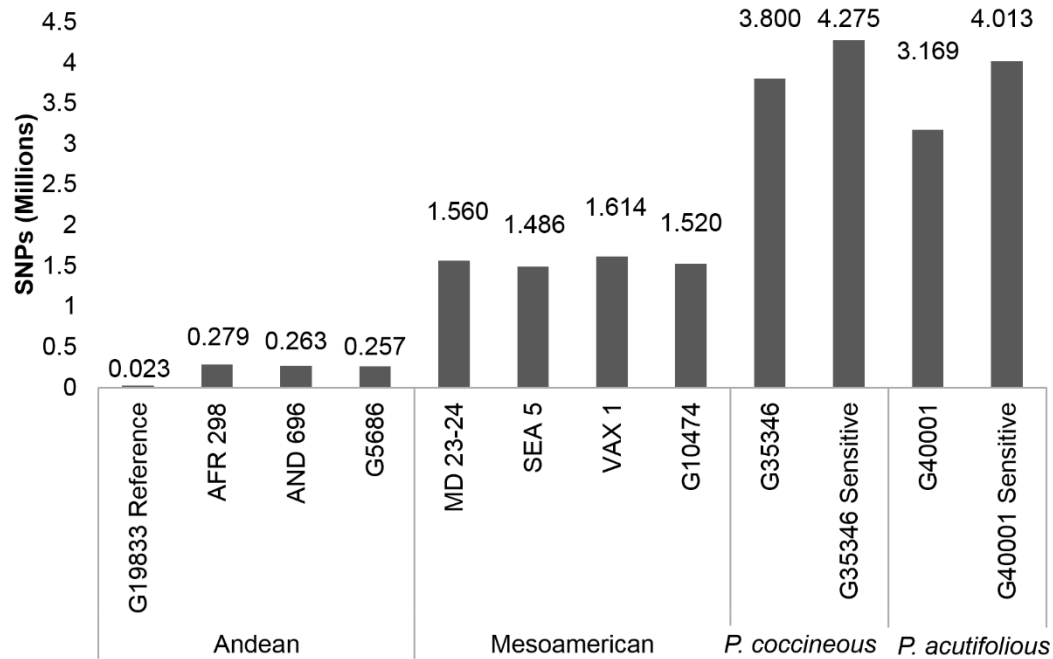
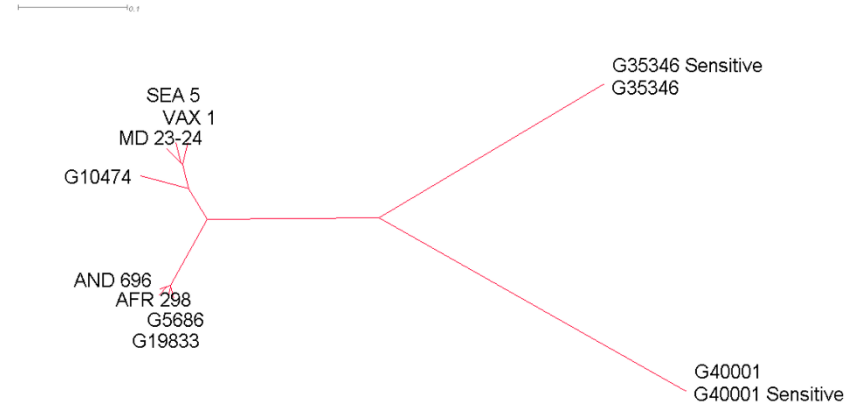
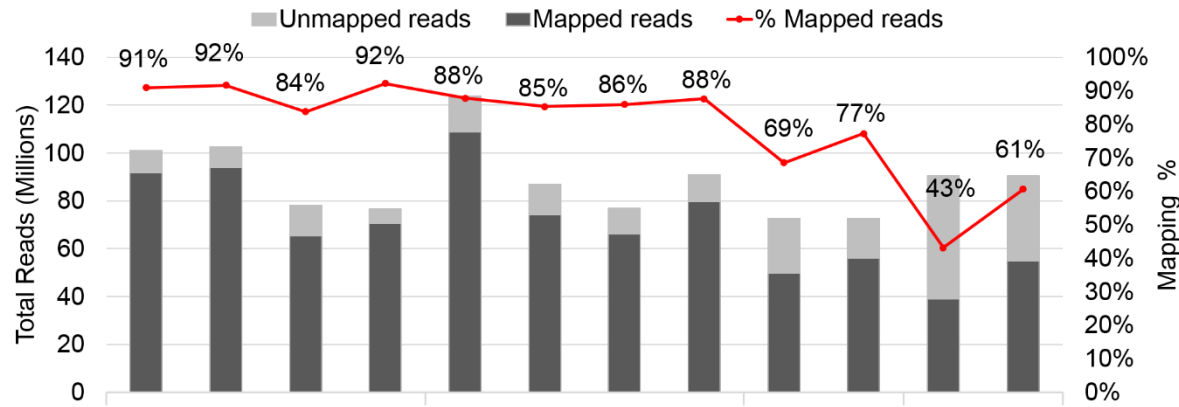
P. acutifolius G40001



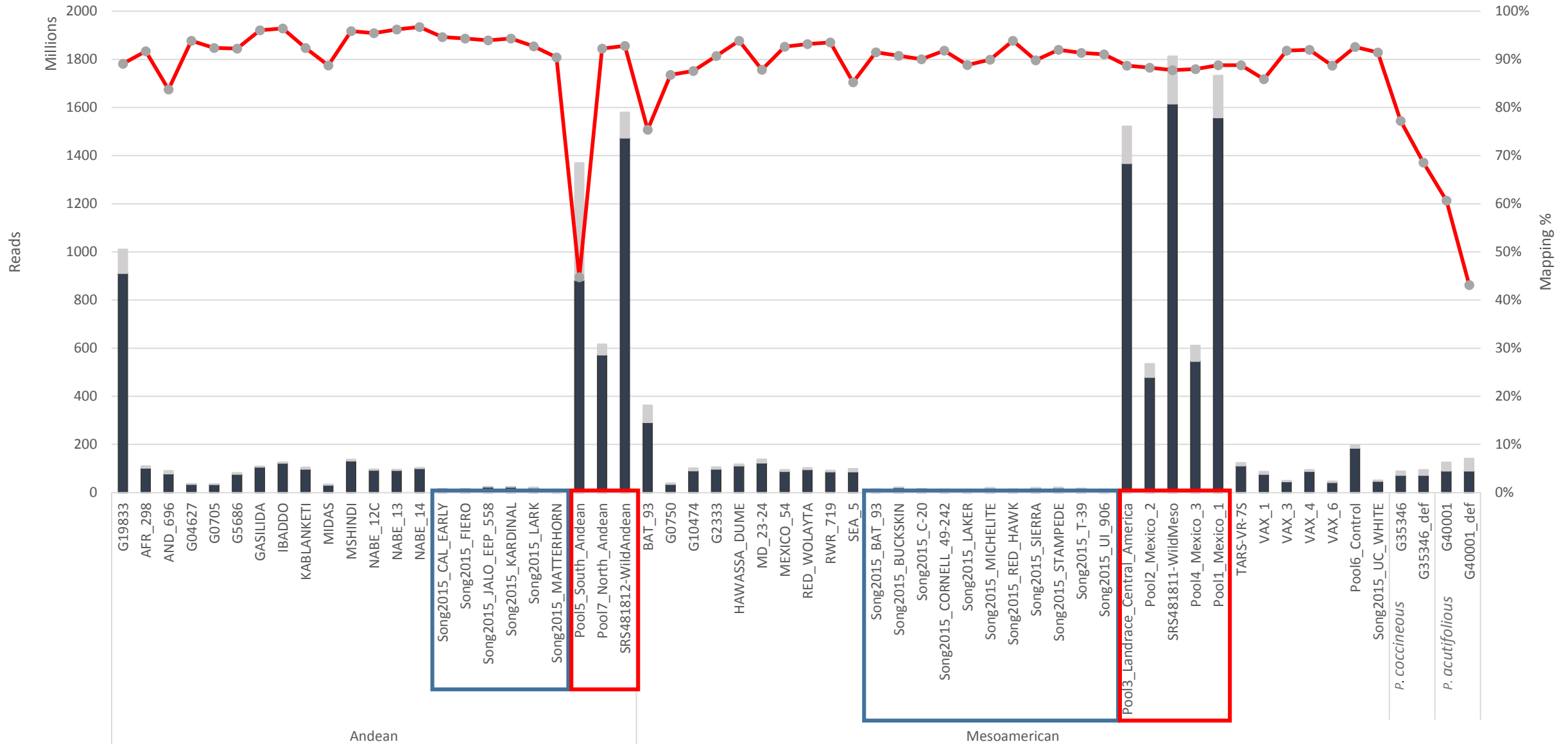
a.

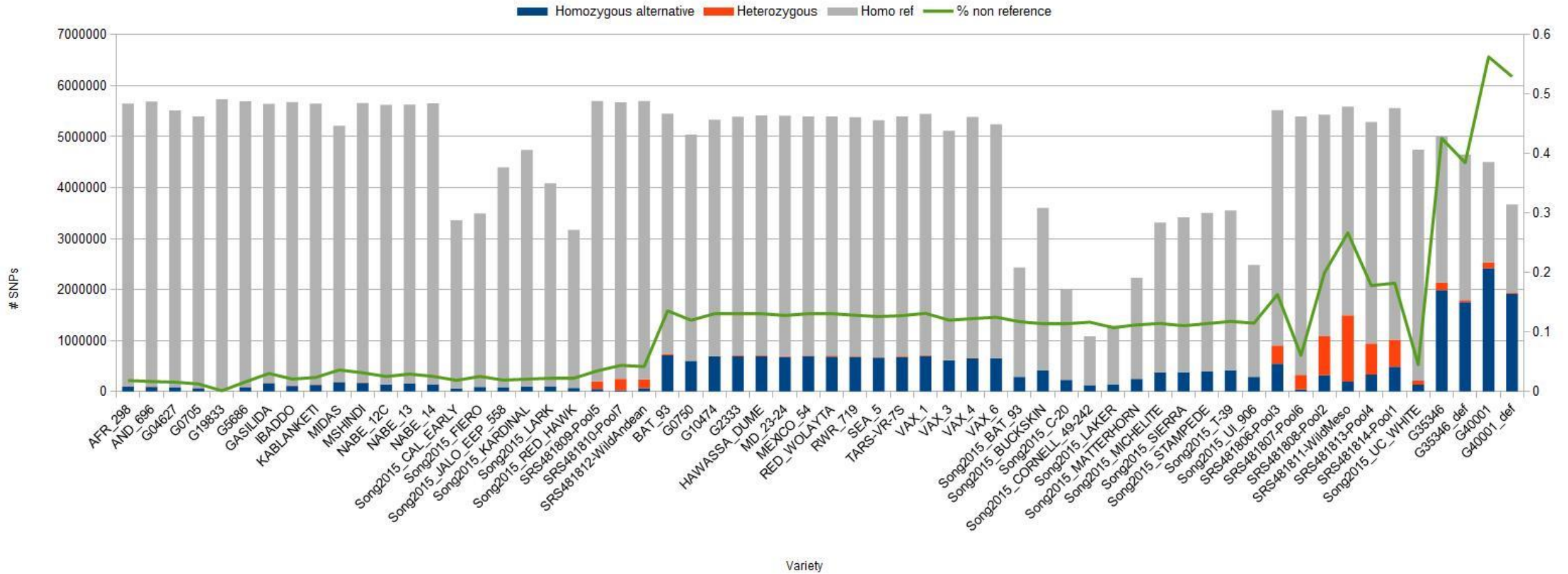


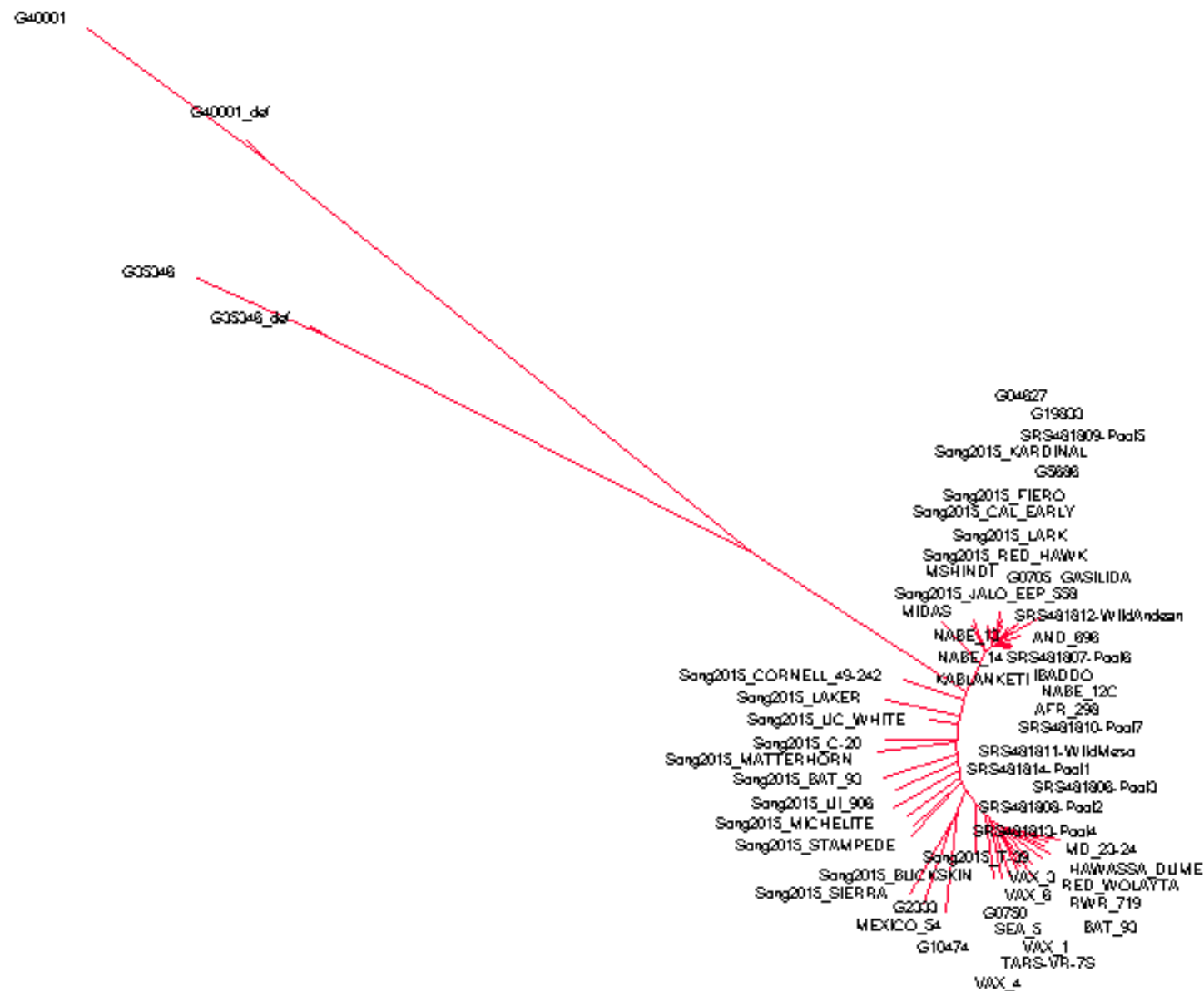
Reads, Overall SNPs and coding SNPs



Total reads and mapping percentage to the Andean reference

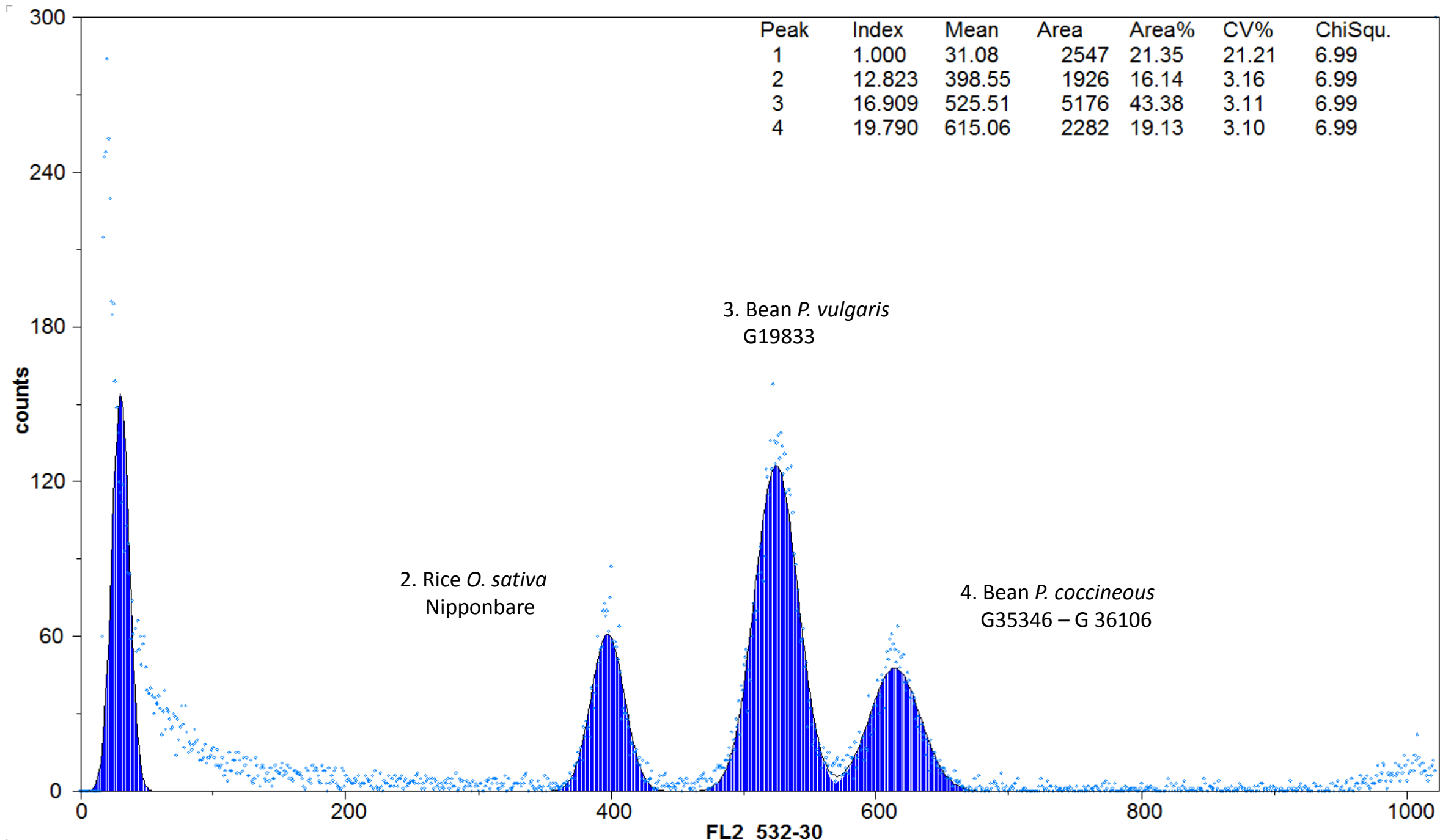


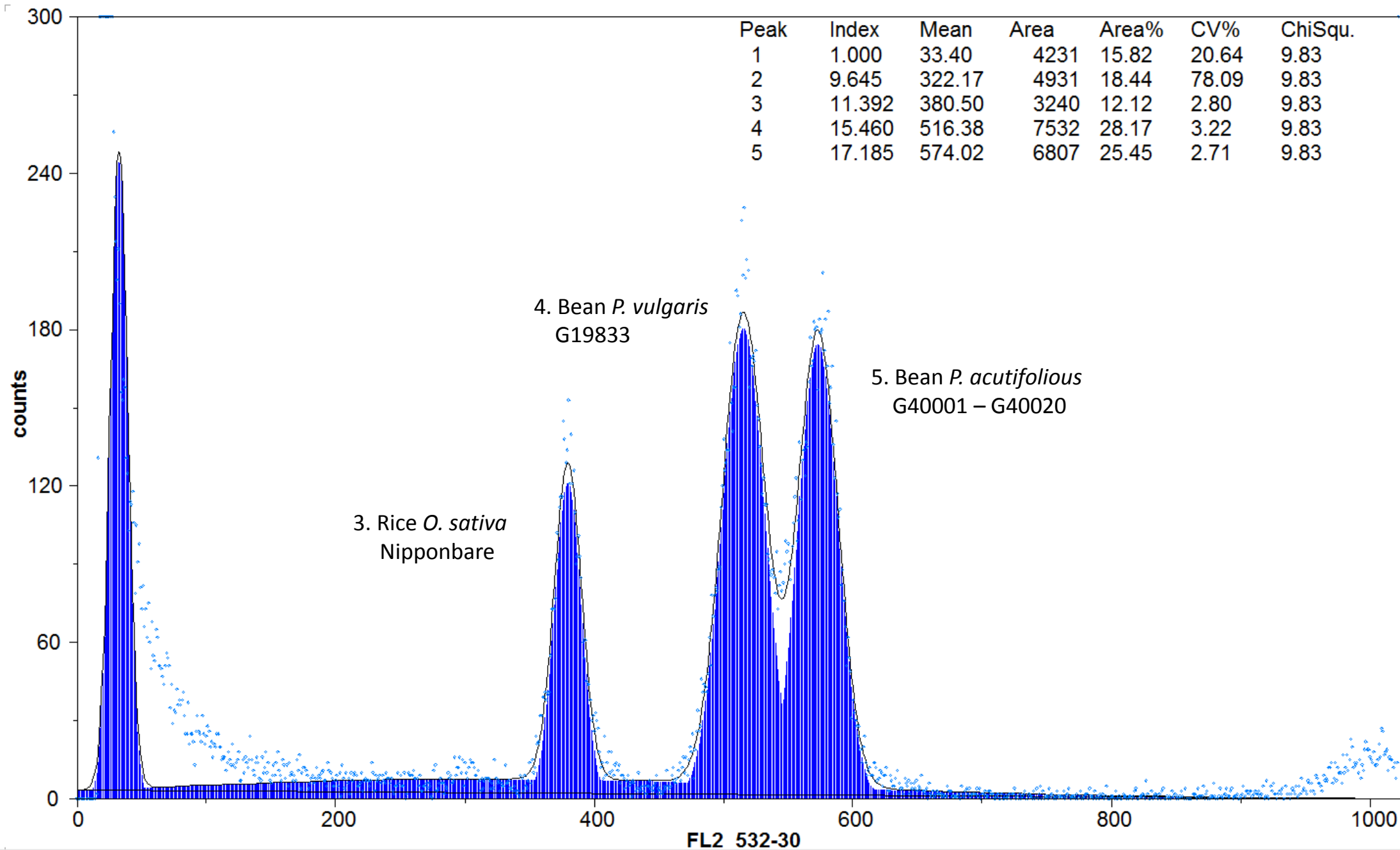




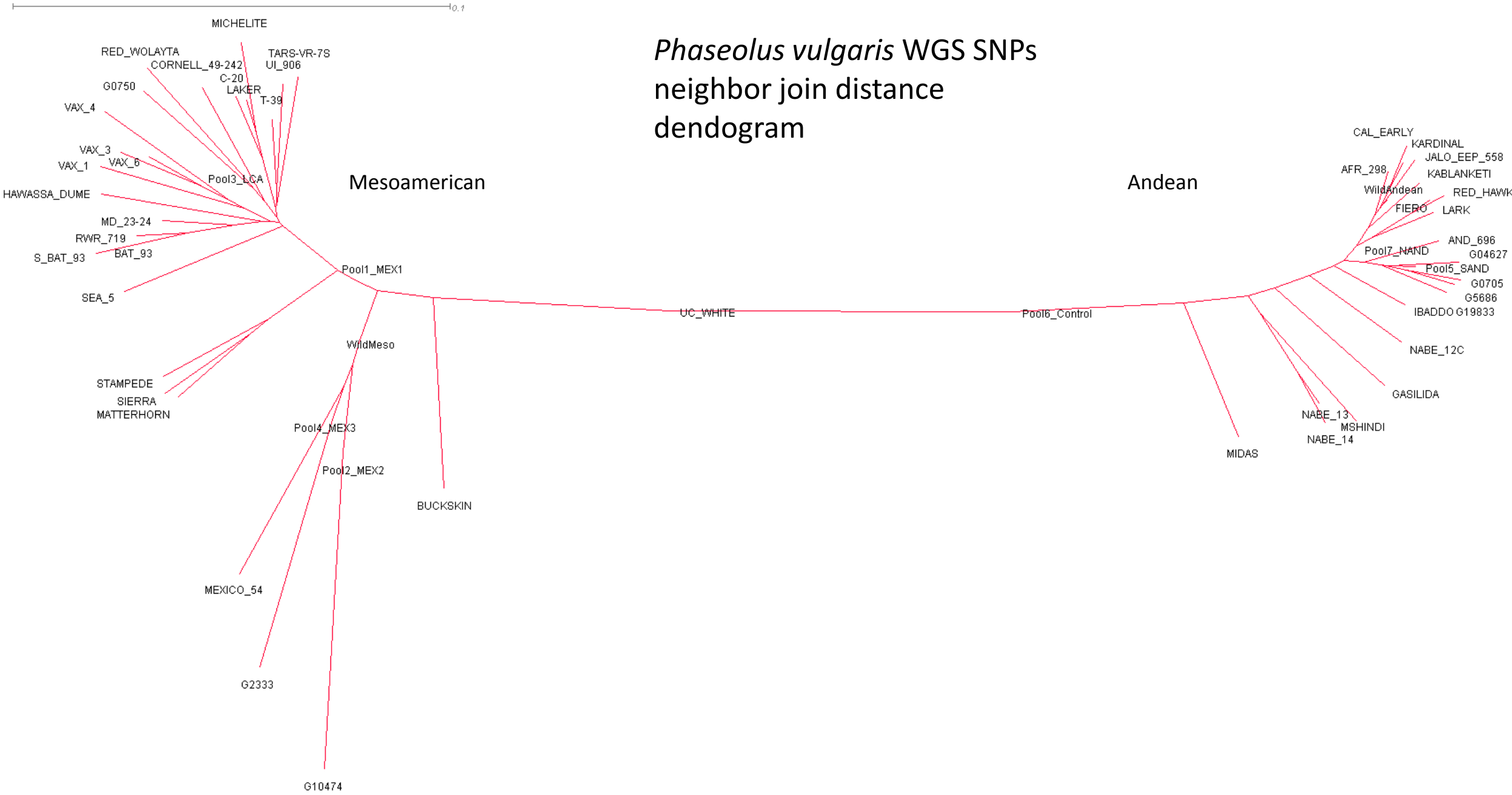
Phaseolus sp. WGS SNPs
neighbor join distance
dendrogram



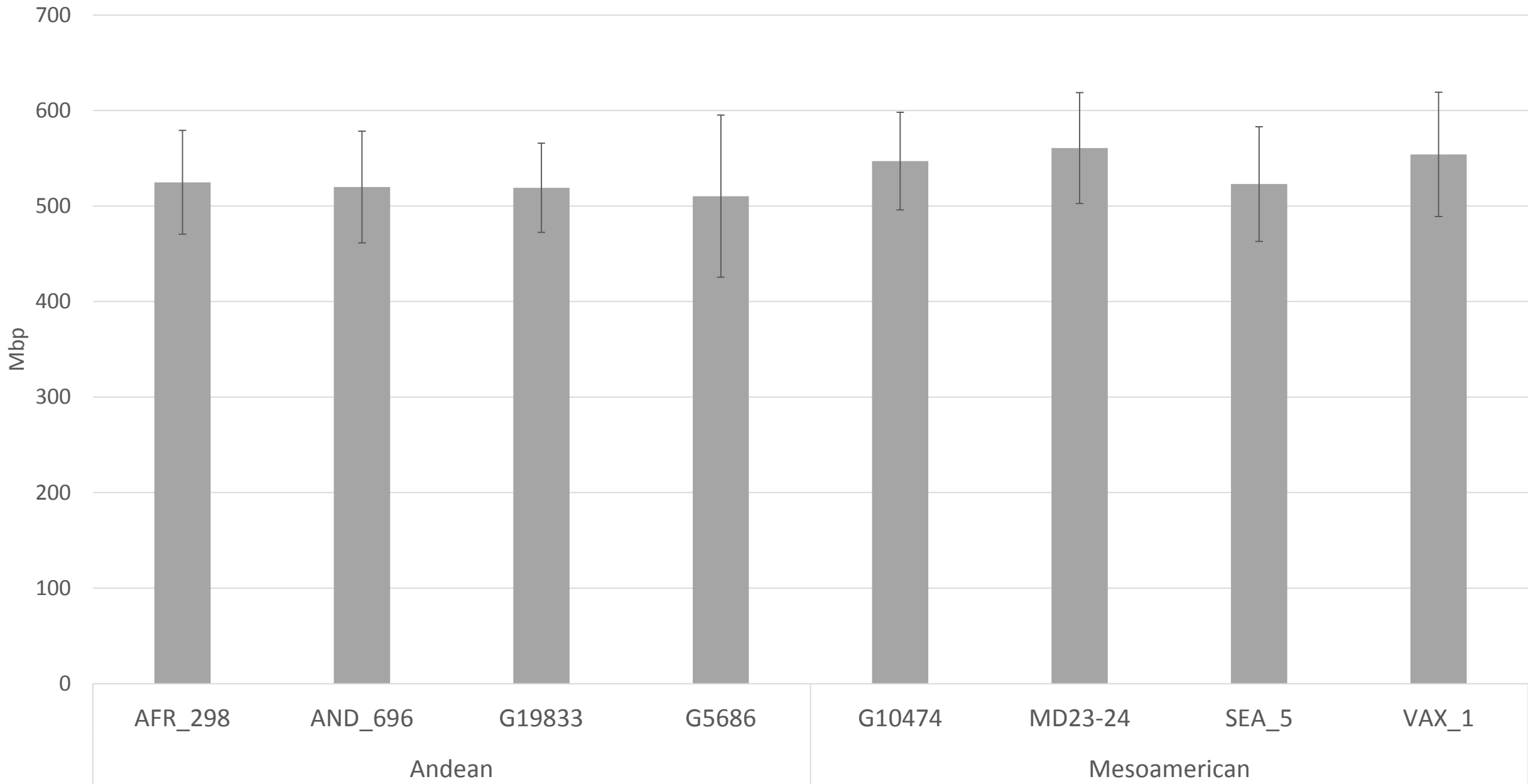




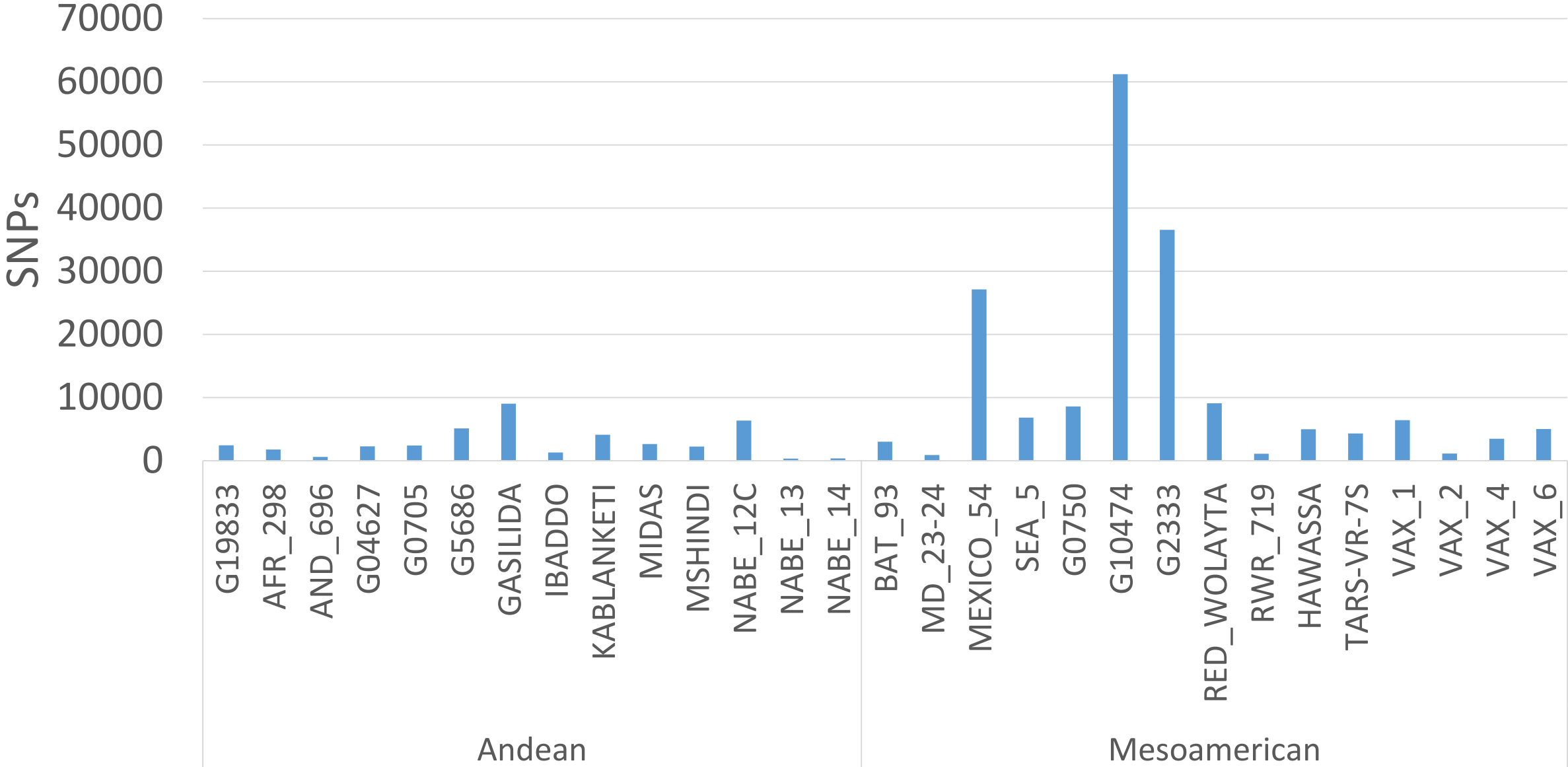
Phaseolus vulgaris WGS SNPs
neighbor join distance
dendrogram

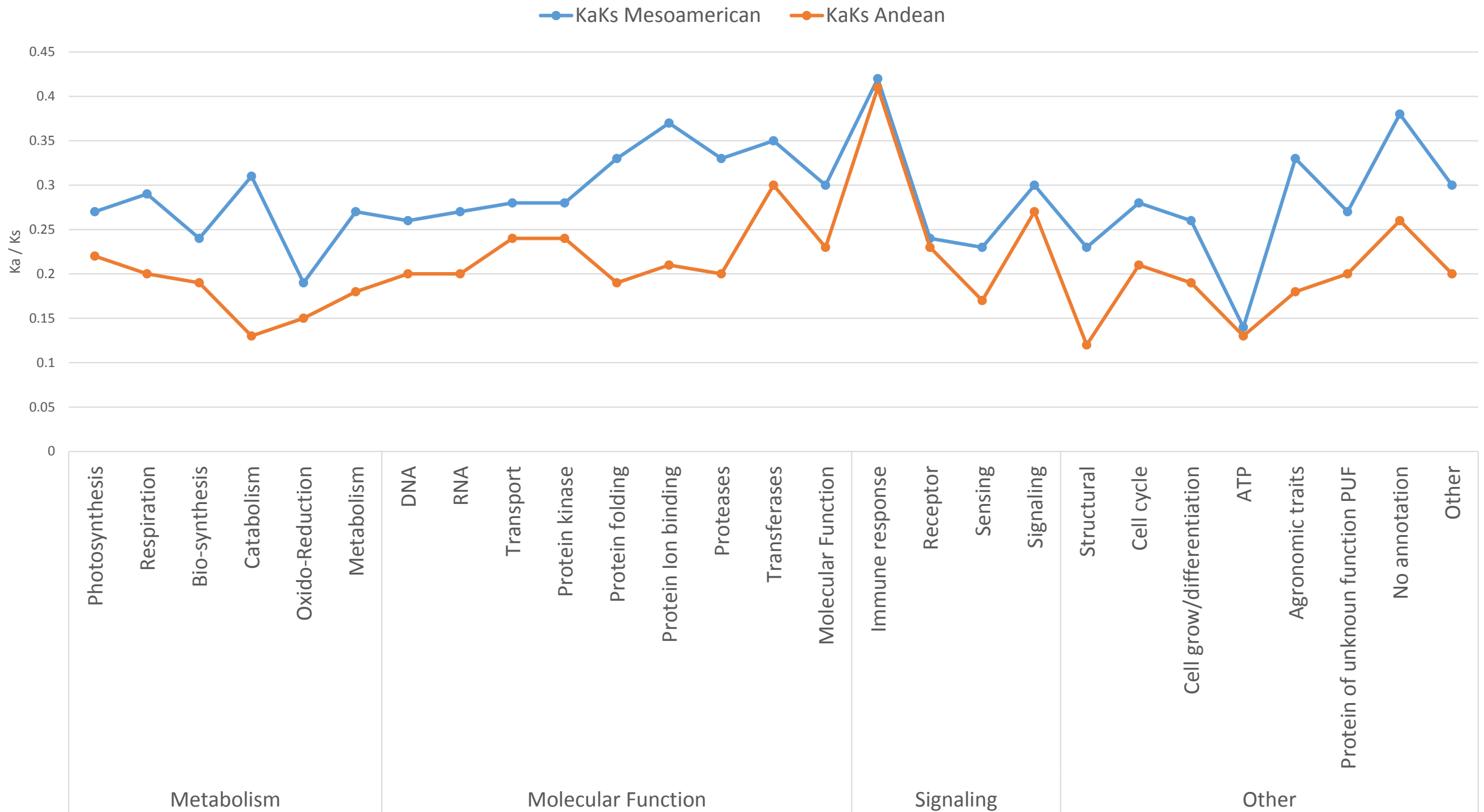


Genome size



Unique SNPs per sample

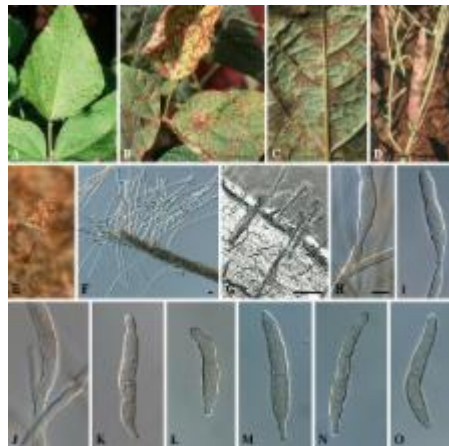




SNPs Markers for Angular Leaf spot (ALS) resistance

ALS is caused by the fungus *Phaeoisariopsis griseola*, is considered a serious disease of beans in many regions

Marker **PF5** in red was the previous SCAR marker for ALS resistance. With WGS we were able to saturate the ALS QTL Region with a total of 23 SNPs markers. In this set 5 SNPs had been selected for the CIAT breeding program (highlighted in blue).



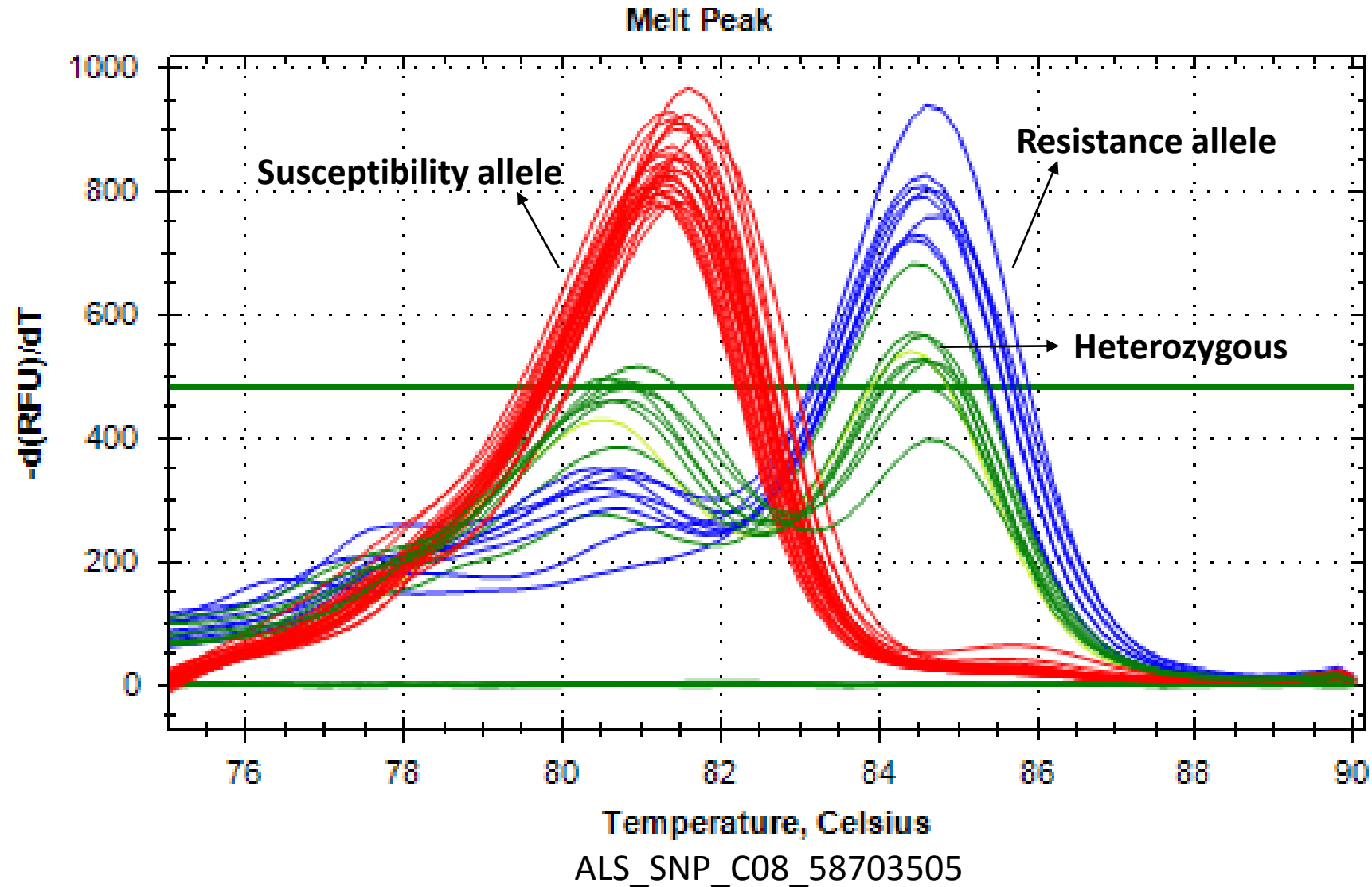
Mbp	Marker	Chromosome 08		
54				
54.2				
54.4	ALS_PG1			
54.6				
54.8				
55				
55.2				
55.4	ALS_PG2			
55.6				
55.8				
56				
56.2				
56.4	ALS_PG3			
56.6				
56.8	ALS_PG4			
57				
57.2	ALS_PG5			
57.4				
57.6	ALS_PG6	CB_00128	CB_00129	ALSChr08_CT_57798588
57.8	PF5	ALS_PG7	CB_00130	
58				
58.2	CB_00131	ALS_PG8		
58.4	CB_00132	ALS_PG9	CB_00133	
58.6	CB_00134	sc267437		
58.8	CB_00135	CB_00136	ALS_PG10	
59	CB_00137	CB_00138		
59.2	ALS_PG11			
59.4				
59.6				
59.8				
60				

G5686

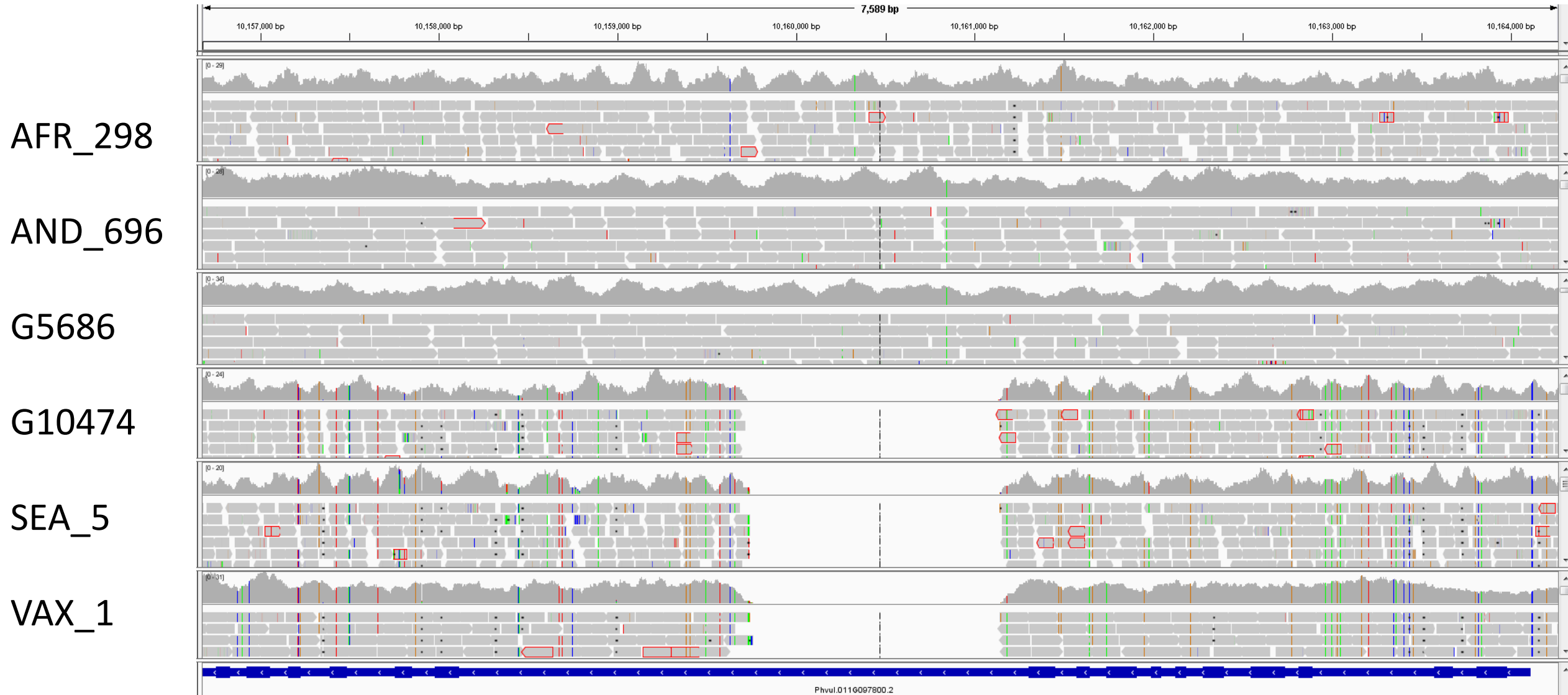


SNP marker tagging ALS resistance on chromosome 8.

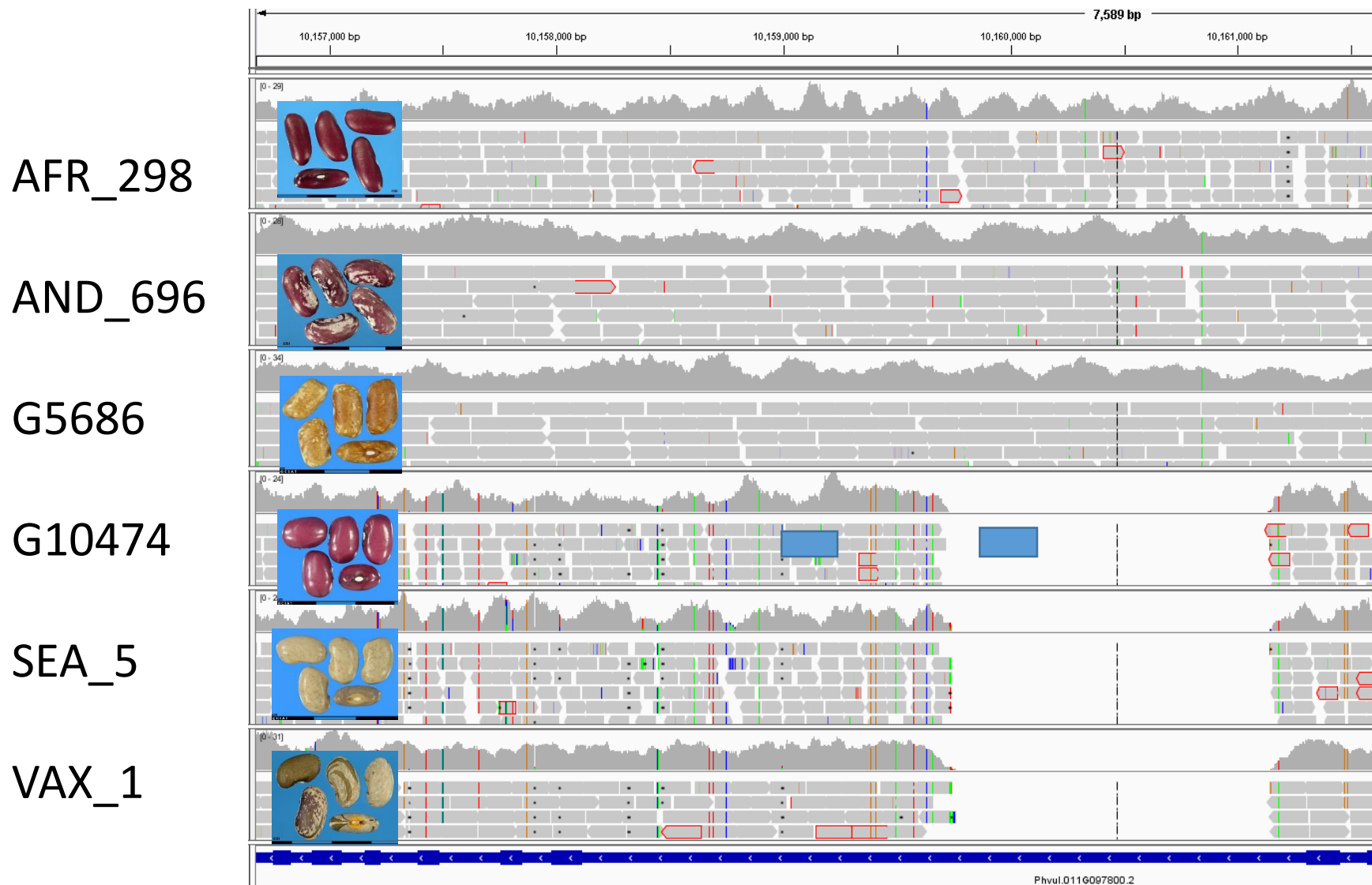
SNP selected from WGS and GBS data specific for G10474 and MD23-24 (WGS), MAB348 and MAB349 (GBS). All Andean and other Meso (WGS) have the reference allele.



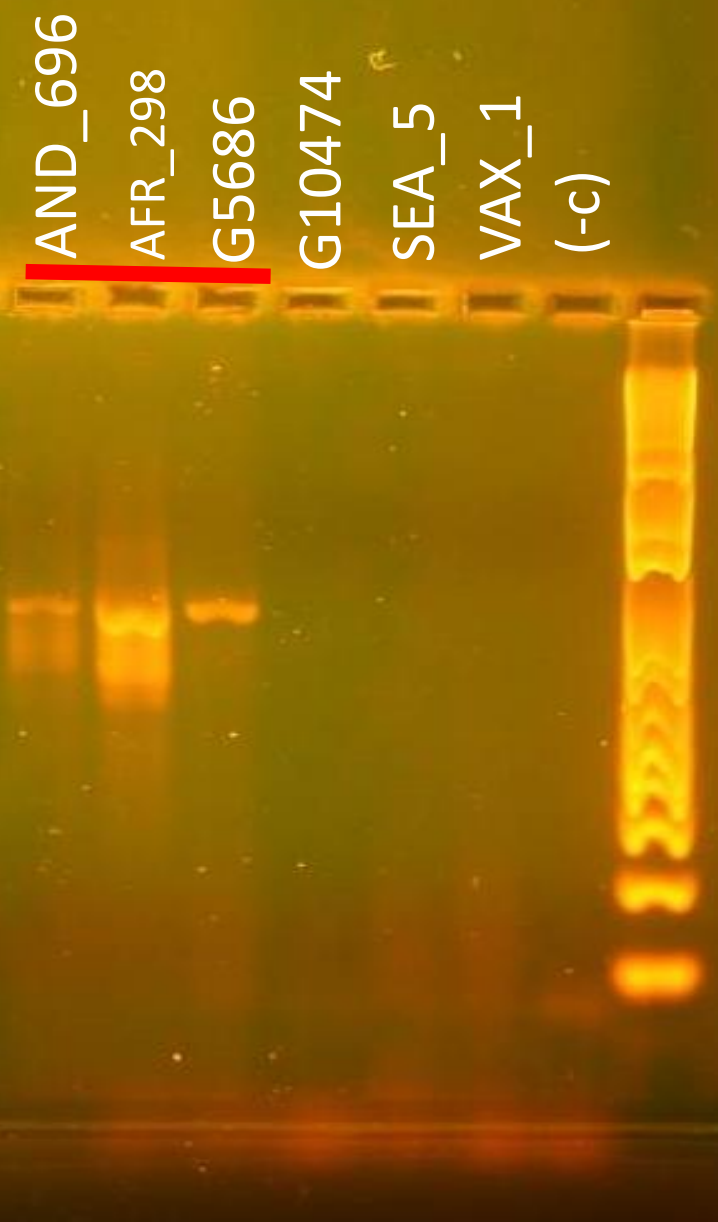
Indel markers for Mesoamerican Andean differentiation



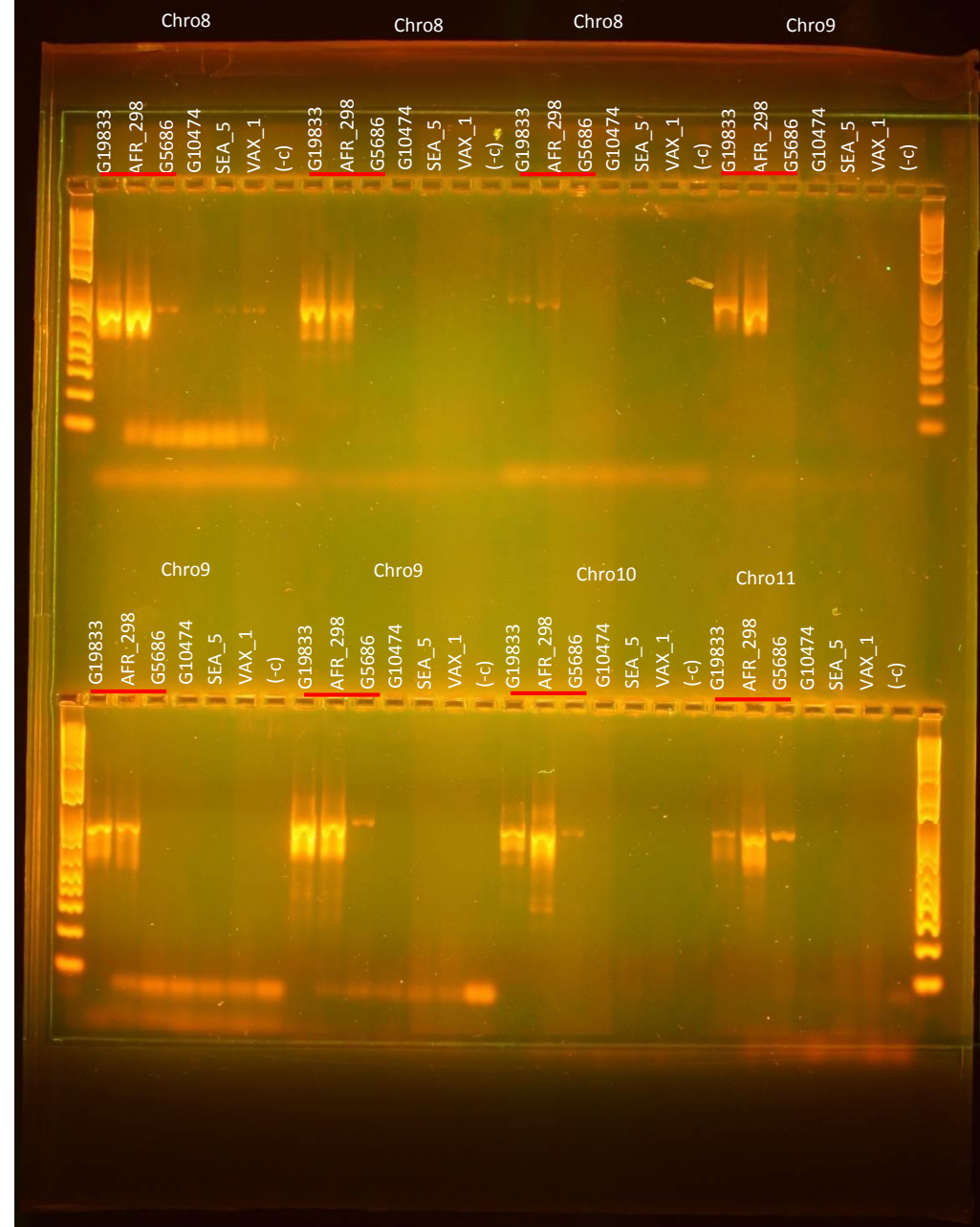
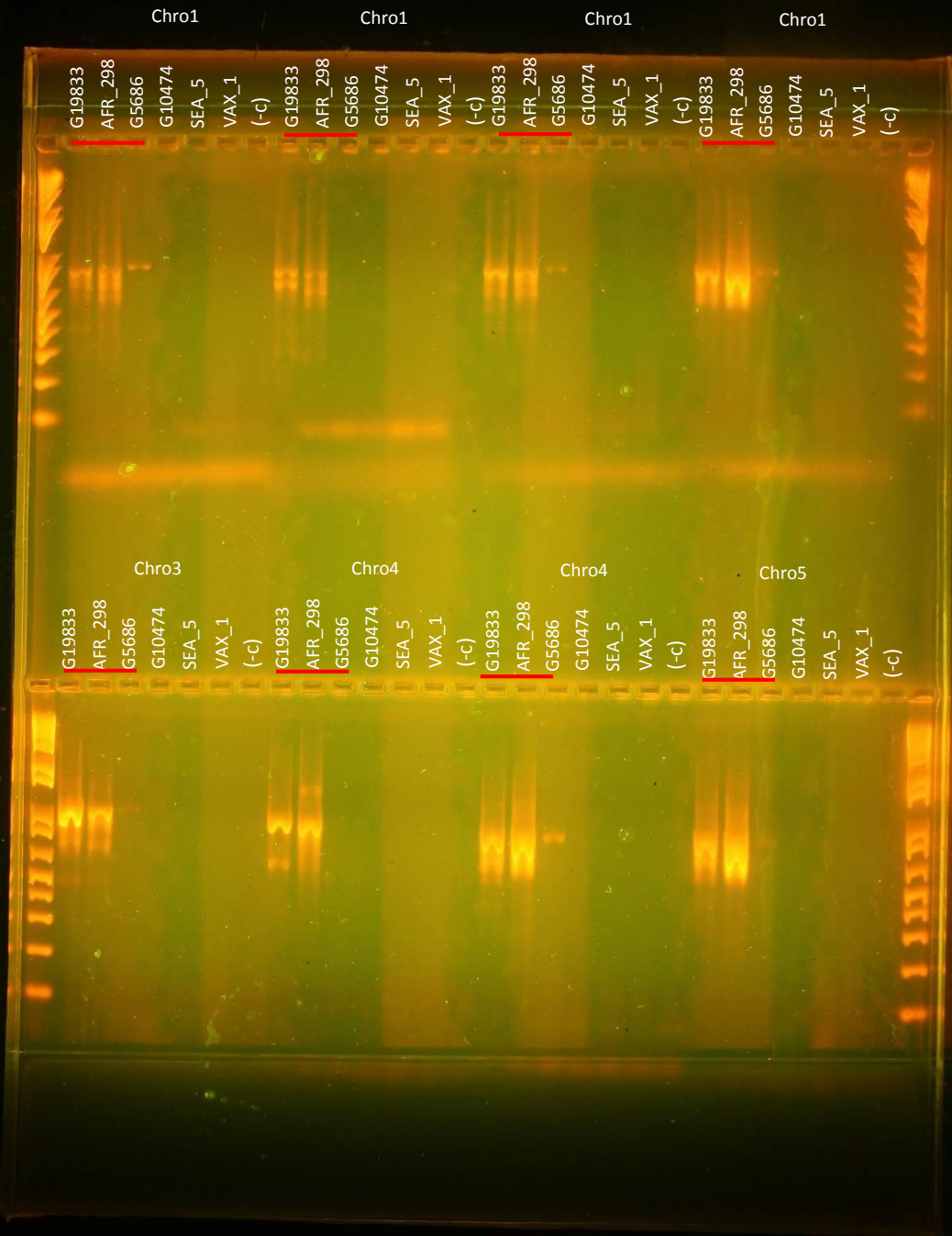
Indel markers for Mesoamerican Andean differentiation 5'-3' EXORIBONUCLEASE 2-RELATED

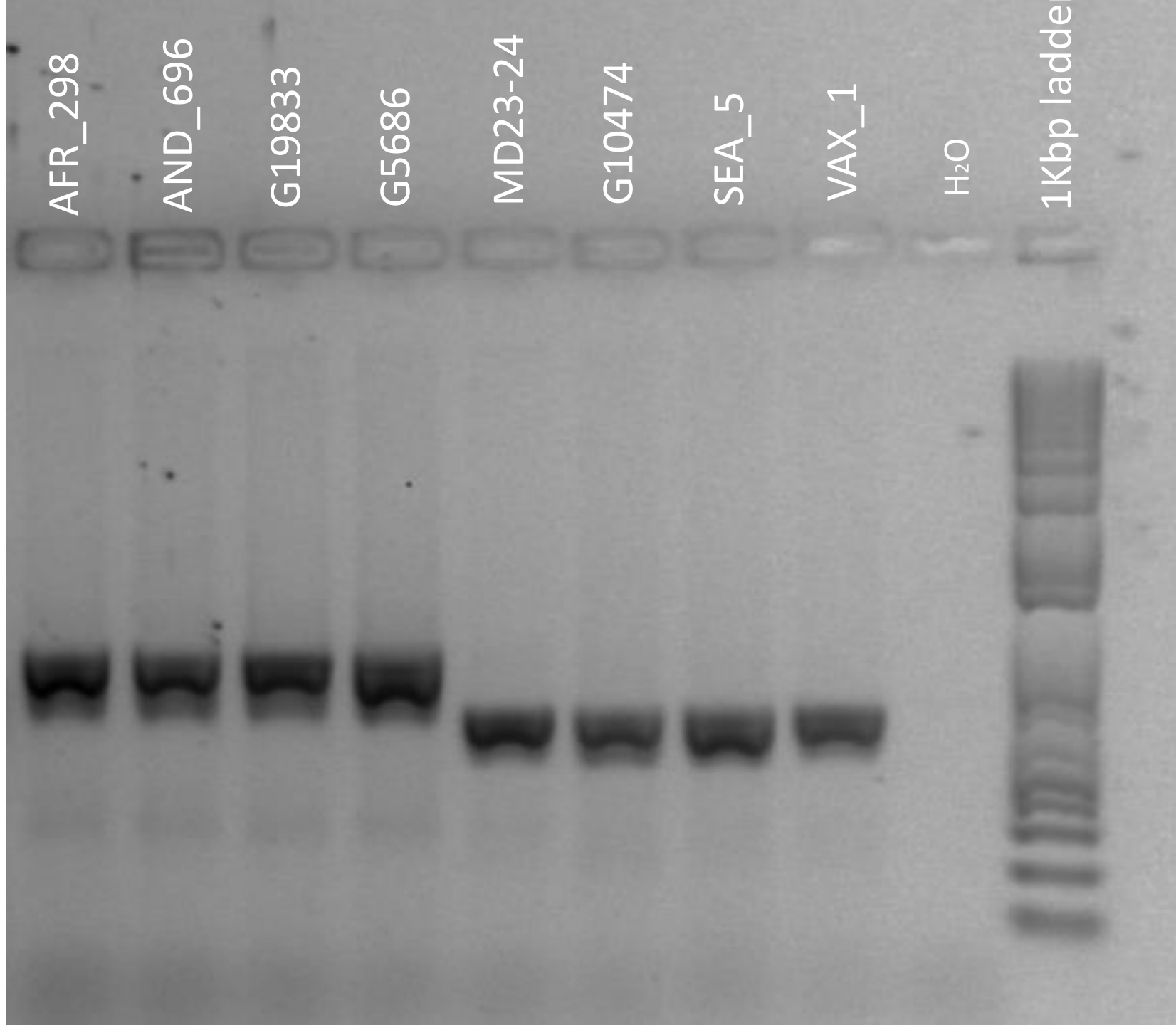


Chro11 10'159701 1500bp



Genome deletions

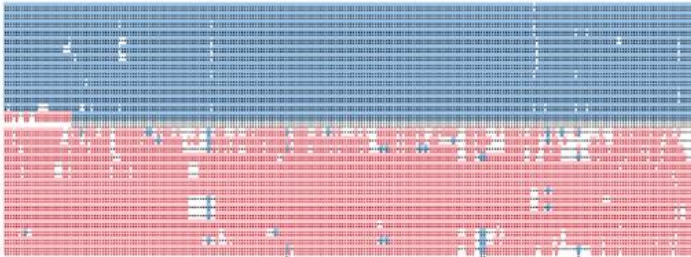




Chr09

Mesoamerican-Andean introgression blocks 50 SNPs window analysis

Chromosome 01



Chromosome 02



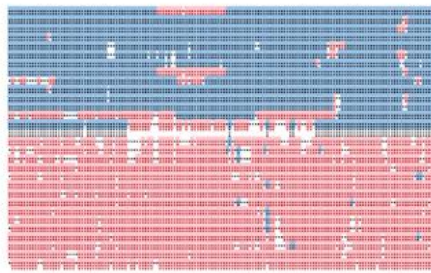
Chromosome 03



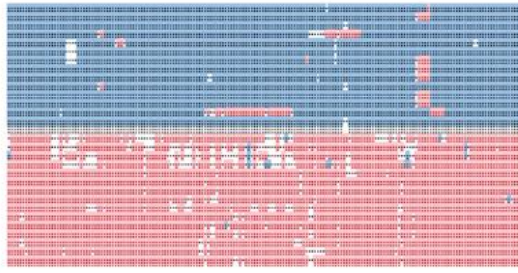
Chromosome 04



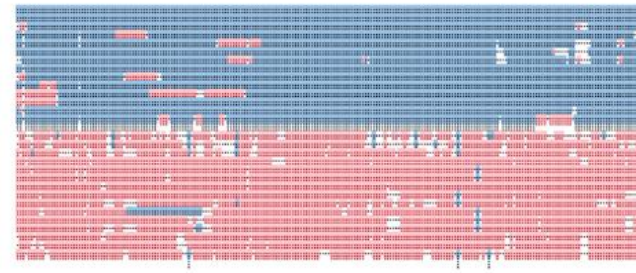
Chromosome 05



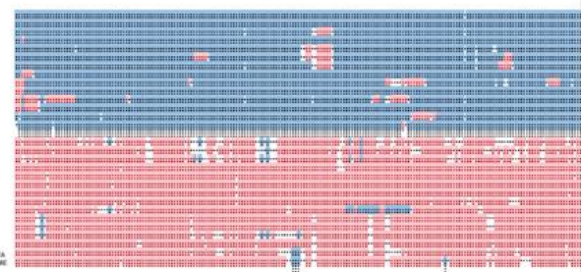
Chromosome 06



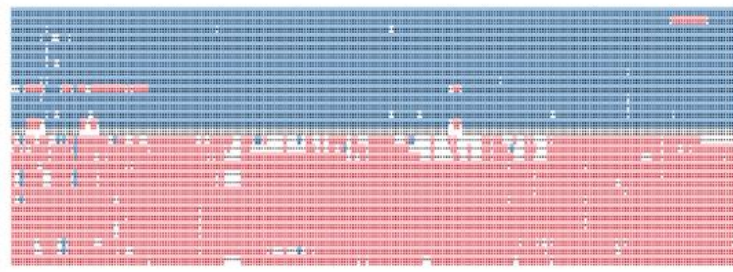
Chromosome 07



Chromosome 08



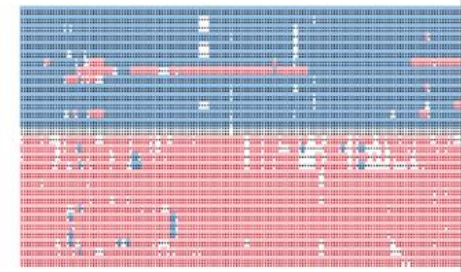
Chromosome 09



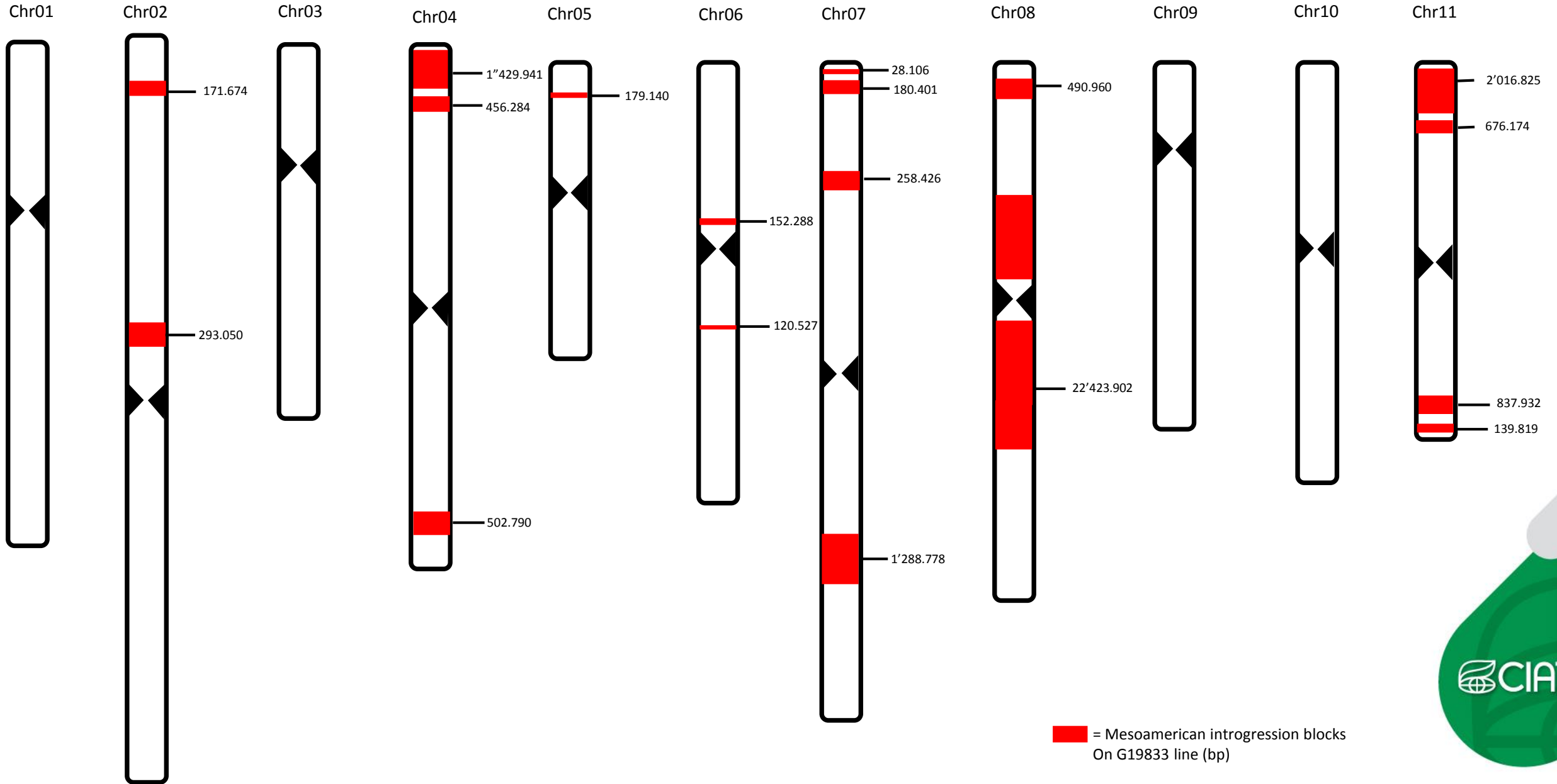
Chromosome 10



Chromosome 11



Mesoamerican introgression regions in Andean G19833 line

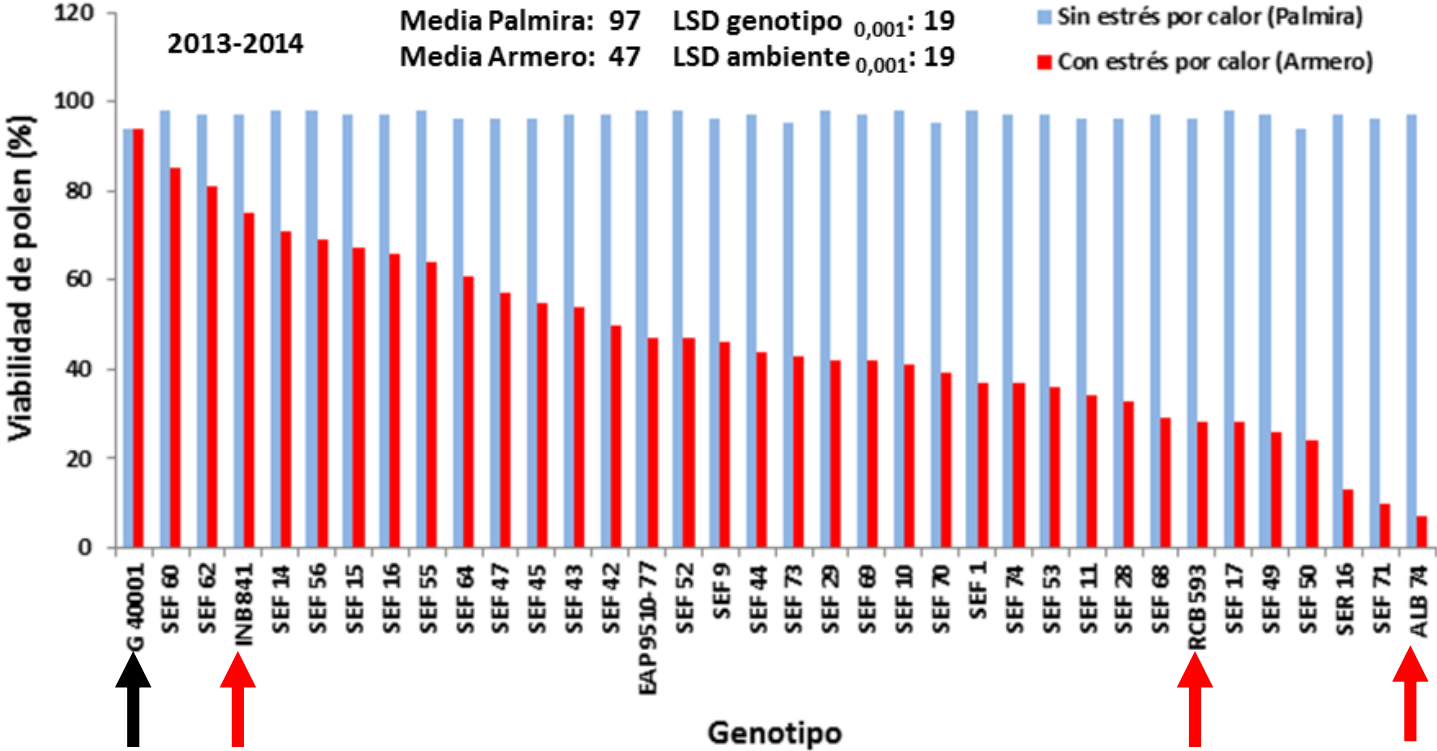


INB 841: A unique introgression line

- Resulted from a cross among several introgression lines from a previous project
- Selected originally under drought
 - Expressed resistance to wilting under intense mid-season drought
 - May have rapid pod elongation
- Has been an excellent parent...progeny express:
 - Heat resistance
 - Resistance to wilting
 - Uniform and stable maturity
- We have RILs of INB 841 x RCB 293
 - Tim has evaluated for heat
 - We evaluated for drought and are repeating



2014: Pollen Viability in Interspecifics with Tepary with and without Heat Stress



P. acutifoliosus introgression into *P. vulgaris*

Q40
RepMasked
MINI 300
MAF 0.01
Variants Group select

G40001W_sen Vs G19833W-G10474W

G40001 WGS
G40001 WGS_sen ←
G40001 GBS
G40035 GBS
G40036 GBS

AM

INB

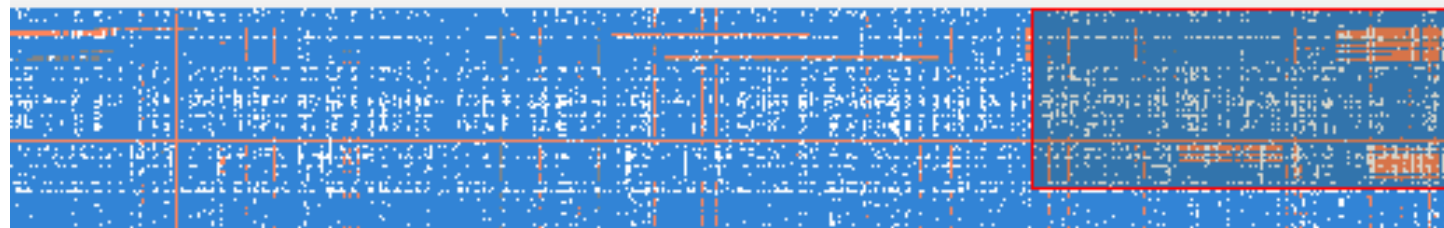


G19833W *P. vulgaris*

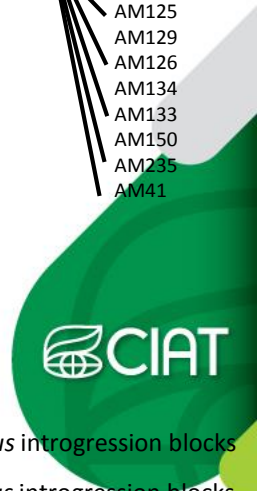
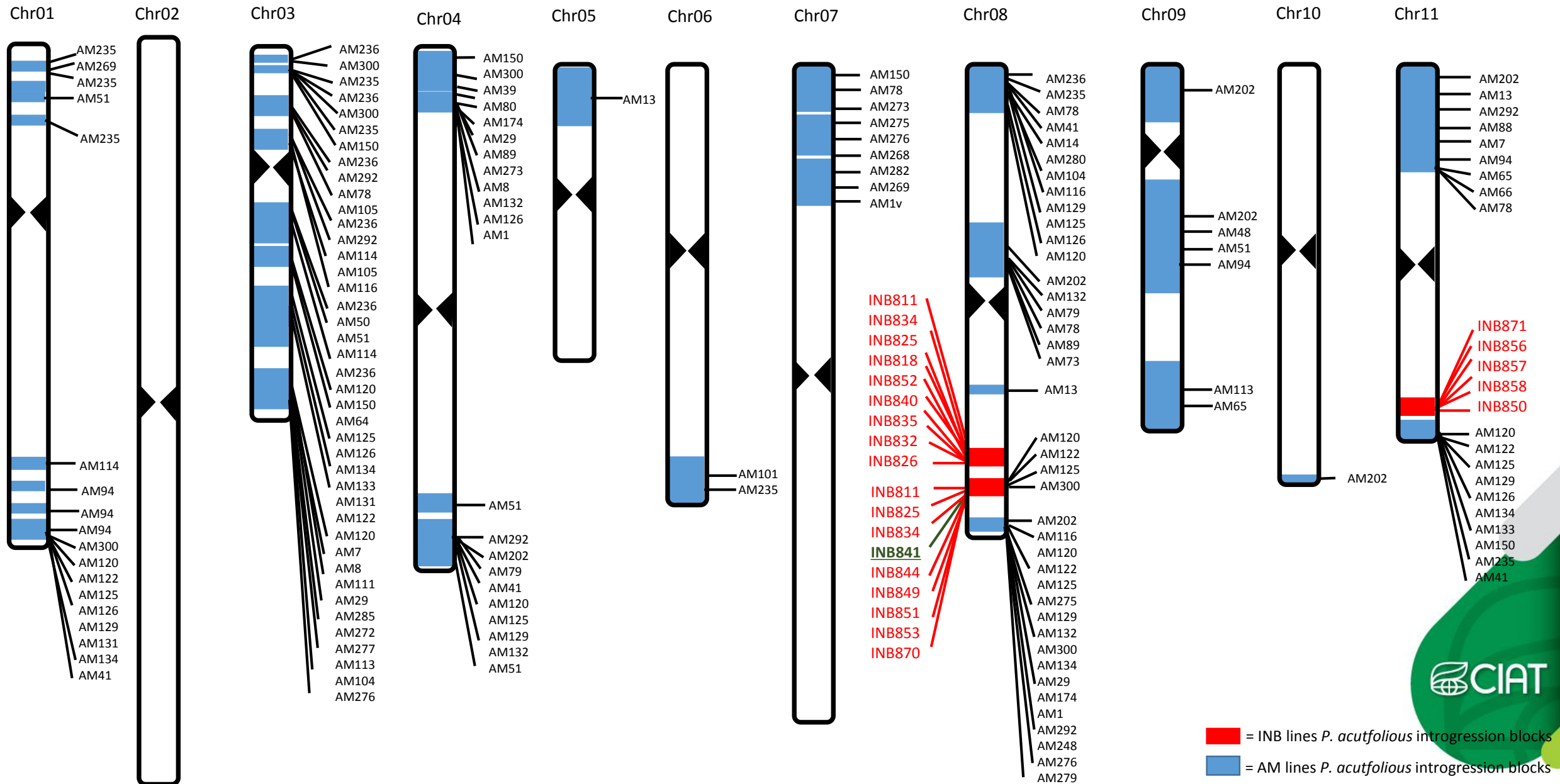


G40001W *P. acutifoliosus*

Overview




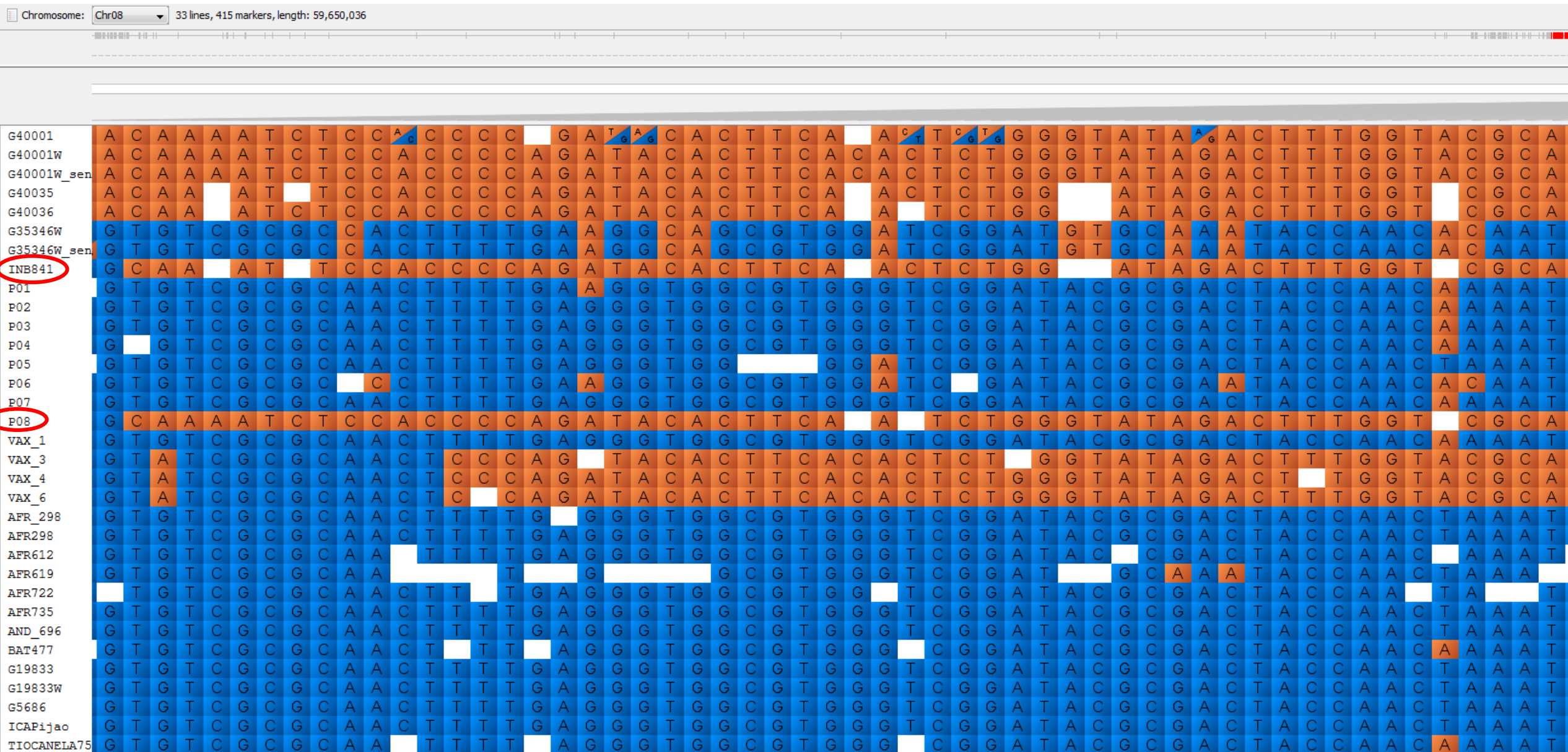
Phaseolus acutifolius introgression regions in *P. vulgaris* at AM and INB interspecific cross populations



Phaseolus acutifolius introgression on INB 841, MAGIC Parent P08 and VAX lines

 *P. acutifolius* alleles

 *P. vulgaris* and *P. coccineus* alleles



SNPs gene location on the *P. acutifoliosus* introgression at INB841 interspecific line (Chr 08 58'724.703 – 59'187.803) 500kbp

#CHROM	POS	Gene pos	code	Protein Domains	#CHROM	POS	Gene pos	code	Protein domains
Chr08	58841149	TA=FivePrimeUTR	PAC:27153086		Chr08	59386906	TA=Synonymous	Phvul.008G291100	BR-signaling kinase 1
Chr08	58842950	TA=Synonymous	Phvul.008G282600	seed imbibition 1 Raffinose synthase	Chr08	59386969	TA=Synonymous	Phvul.008G291100	BR-signaling kinase 1
Chr08	58843028	TA=Synonymous	Phvul.008G282600	seed imbibition 1 Raffinose synthase	Chr08	59387023	TA=Synonymous	Phvul.008G291100	BR-signaling kinase 1
Chr08	58844144	TA=Missense	Phvul.008G282600	seed imbibition 1 Raffinose synthase	Chr08	59387027	TA=Missense	Phvul.008G291100	BR-signaling kinase 1
Chr08	58846878	TA=Missense	Phvul.008G282700	cytochrome c oxidase 15	Chr08	59414987	TA=Intron	PAC:27156131	
Chr08	58861011	TA=Synonymous	Phvul.008G282900	CCCH-type zinc fingerfamily protein with RNA-binding domain	Chr08	59415015	TA=Intron	PAC:27156131	
Chr08	58915717	TA=Missense	Phvul.008G283700	Neutral/alkaline non-lysosomal ceramidase	Chr08	59418163	TA=Synonymous	Phvul.008G291500	isocitrate dehydrogenase 1
Chr08	58955990	TA=Synonymous	Phvul.008G284100	Transducin/WD40 repeat-like superfamily protein	Chr08	59418238	TA=Downstream	PAC:27156131	
Chr08	58977912	TA=Missense	Phvul.008G284300	F-box family protein	Chr08	59421145	TA=Missense	Phvul.008G291600	Phosphoglycerate mutase family protein
Chr08	59021428	TA=Missense	Phvul.008G284900	Nucleoporin, Nup133/Nup155-like	Chr08	59422596	TA=Missense	Phvul.008G291700	C2H2-like zinc finger protein
Chr08	59026184	TA=Synonymous	Phvul.008G284900	Nucleoporin, Nup133/Nup155-like	Chr08	59422817	TA=Synonymous	Phvul.008G291700	C2H2-like zinc finger protein
Chr08	59059770	TA=Intron	PAC:27155480		Chr08	59429661	TA=Missense	Phvul.008G291800	Plant regulator RWP-RK family protein
Chr08	59059940	TA=Intron	PAC:27155480		Chr08	59453220	TA=Intron	PAC:27153741	
Chr08	59060035	TA=Synonymous	Phvul.008G285400	HEAT SHOCK PROTEIN 89.1	Chr08	59459148	TA=Synonymous	Phvul.008G292000	Protein kinase superfamily protein
Chr08	59104120	TA=Missense	Phvul.008G286000	GRAS family transcription factor	Chr08	59459160	TA=Synonymous	Phvul.008G292000	Protein kinase superfamily protein
Chr08	59104256	TA=Synonymous	Phvul.008G286000	GRAS family transcription factor	Chr08	59462633	TA=Intron	PAC:27155306	
Chr08	59112952	TA=Synonymous	Phvul.008G286200	DGCR14-related	Chr08	59462641	TA=Intron	PAC:27155306	
Chr08	59113009	TA=Synonymous	Phvul.008G286200	DGCR14-related	Chr08	59464515	TA=Intron	PAC:27155306	
Chr08	59120580	TA=Synonymous	Phvul.008G286400	Phototropic-responsive NPH3 family protein	Chr08	59466377	TA=Intron	PAC:27155306	
Chr08	59185195	TA=Intron	PAC:27155851		Chr08	59479372	TA=Missense	Phvul.008G292300	homoserine kinase
Chr08	59185202	TA=Intron	PAC:27155851		Chr08	59479376	TA=Synonymous	Phvul.008G292300	homoserine kinase
Chr08	59185207	TA=Intron	PAC:27155851		Chr08	59479390	TA=Missense	Phvul.008G292300	homoserine kinase
Chr08	59186244	TA=Missense	Phvul.008G287400		Chr08	59479407	TA=Missense	Phvul.008G292300	homoserine kinase
Chr08	59186348	TA=FivePrimeUTR	PAC:27155851		Chr08	59484931	TA=Missense	Phvul.008G292400	tonoplast monosaccharide transporter2
Chr08	59214540	TA=Intron	PAC:27154413		Chr08	59484932	TA=Missense	Phvul.008G292400	tonoplast monosaccharide transporter2
Chr08	59303118	TA=Intron	PAC:27154734		Chr08	59484948	TA=Synonymous	Phvul.008G292400	tonoplast monosaccharide transporter2
Chr08	59303144	TA=Intron	PAC:27154734		Chr08	59485055	TA=Missense	Phvul.008G292400	tonoplast monosaccharide transporter2
Chr08	59310941	TA=Missense	Phvul.008G289500	PHE ammonia lyase 1	Chr08	59485104	TA=Synonymous	Phvul.008G292400	tonoplast monosaccharide transporter2
Chr08	59310946	TA=Synonymous	Phvul.008G289500	PHE ammonia lyase 1	Chr08	59492351	TA=Upstream	PAC:27155451	
Chr08	59310959	TA=Missense	Phvul.008G289500	PHE ammonia lyase 1	Chr08	59492372	TA=Upstream	PAC:27155451	
Chr08	59311750	TA=Synonymous	Phvul.008G289500	PHE ammonia lyase 1	Chr08	59492379	TA=Upstream	PAC:27155451	
Chr08	59312560	TA=Synonymous	Phvul.008G289500	PHE ammonia lyase 1	Chr08	59492380	TA=Upstream	PAC:27155451	
Chr08	59348281	TA=Synonymous	Phvul.008G290500	PHE ammonia lyase 1	Chr08	59616761	TA=Intron	PAC:27153582	
Chr08	59374770	TA=Missense	Phvul.008G291000	UDP-Glycosyltransferase superfamily protein	Chr08	59630997	TA=Missense	Phvul.008G293100	SEC7-like guanine nucleotide exchange family protein
Chr08	59374916	TA=Synonymous	Phvul.008G291000	UDP-Glycosyltransferase superfamily protein	Chr08	59631051	TA=Missense	Phvul.008G293100	SEC7-like guanine nucleotide exchange family protein
Chr08	59374932	TA=Missense	Phvul.008G291000	UDP-Glycosyltransferase superfamily protein	Chr08	59650036	TA=Synonymous	Phvul.008G293200	FAR1-related sequence 9



CGIAR Research Program on Grain Legumes

Our Strategy

CGIAR Research Programs

- Agriculture for Nutrition and Health
- Aquatic Agricultural Systems
- Climate Change, Agriculture and Food Security (CCAFS)
- Dryland Cereals
- Dryland Systems
- Forests, Trees and Agroforestry
- Grain Legumes

PARTNERS:

- CIAT
- Ethiopian Institute of Agricultural Research (EIAR), Ethiopia
- Brazilian Agricultural Research Corporation (EMBRAPA)
- General Directorate of



NGSEP team @NGSEP · 6 h

At the Colombian Bioinformatics Congress #CCBCOL3



