Validation of genes and markers for fruit quality trait prediction in strawberry: WP1 Vitamin C content volatile content



F1 (Pseudo-test cross) mapping population of 95 progeny lines

X



'232'

Larger fruits, High yield



'1392'

High firmness, Better flavor, Higher L-AA, Darker color QTL controlling L-AA content in strawberry and their applications in breeding



Copyright: University of Waikato, New Zealand

OBJECTIVES

- 1. Identification of QTL for fruit quality traits in the mapping population '232' \times '1392'.
 - SATURATED LINKAGE MAP: A genetic linkage map is a representation of the relative positions of loci on the chromosomes
 - Quantification of traits in the progeny: Quantitative traits, Controlled by many genes (QTL)
- 3. Identification and validation of candidate genes for the QTL
- Analysis of markers linked to the trait in other genotypes: VALIDATION FOR MARKER ASSISTED BREEDING

3 QTLs control aprox. 45% of the variation of L-AA content in strawberry



BULK segregant analysis (BSA) + RNA-seq as strategy for identification of candidate genes





Identification of differentially expressed genes between 2 bulks of 8 lines with contrasting ascorbic acid content

RNA-seq

RNA-seq in L-AA contrasting bulks

Total analyzed genes33,404Differentially expressed211

Among them we have identified a number of genes from L-AA biosynthesis and recycling pathways



Validation of RNA-seq results by qRT-PCR





Correlation RNA-seq and QRT-PCRs in selected genes

Validation of RNA-seq results by qRT-PCR

Not-significant differential expression by RNA-seq

LG7: 14,595,359 .. 14,597,190 bp





LG VII-M1



R²: 12 %

Co-localization between QTL LAA V-M1 and Mannose-6-P isomerase

LG V-M1



LAAV-M1 2008 0,0 _____ -1 EMFvi108-188 -0 ChFaM044-178/188 LAAV-M1 2007 pwiV-M1 2008 14,3 -1 ChFaM108-177 19,7 -0 ChFaM022-198 19,8 -0 cct/aca-390 -1 EMFn010-200/220 20.4 LAAV-M1 2009 -0 BFACT005-173 33,1 R²: 15.1-21.9 % -0 UDF034-155 50,1

QTL markers: LAA V-M1 2007: ChFaM108 (5,745,728 bp) LAA V-M1 2008: EMFvi108 (178,094 bp) LAA V-M1 2009: BFACT5 (17,799,509 bp)

FaGalUR is not differentially expressed but one allele is more abundant than other between pools



FaGalUR

Expression of *FaM6PI* in vegetative tissues and during fruit ripening

ID Strawberry genome Cluster			AV	AV AB AT AR		AR	ЦО	DA	
M6PI	gene10529	6	8,40	8,91	5,80	4,99	но	КA	
			RV	RB	RT	RR	20.41	1 00	
			10,68	5,52	1,08	0,47	20,41	1,00	

Relative expression of FaM6PI



Ongoing functional validations

Cloning FaM6PI



Agro-infiltration of M6PI in strawberry fruits



5-6 days attached to the plant





Identification of FaFAD1 as the gene controlling γ -decalactone content in strawberry fruit

Headspace solid phase micro-extraction + gas chromatography + mass spectrometry (HS-SPME-GC-MS)



The gene controlling γ -decalactone maps to the bottom of LG III-2



Approach to identify the gene underlying the QTL

RNA-seq of bulked pools

Identification of Genes differentially expressed in fruits of contrasting bulks with or $w/o \gamma$ -decalactone

Two pools of fruits from 10 progeny lines with and without γ -decalactone

Relative concentration of g-dec in fruits

wi	1				
line	2007 2008		2009		line
93-01	4.456	1.287	2.849		93-03
93-12	3.114	1.185	1.875		93-07
93-19	3.101	4.300	2.865		93-14
93-36	3.022	3.884	3.182		93-17
93-43	4.237	3.394	3.074		93-18
93-54	3.675	3.376	3.902		93-49
93-61	2.403	1.511	3.293		93-68
93-64	3.413	1.117	2.878		93-69
93-78	1.926	1.171	2.644		93-80
93-92	2.094	1.457	2.658	_	93-89
		-	_		

High γ -dec pool

ine20072008200993-030.0000.0190.01493-070.0020.0010.00593-140.0100.0070.01993-170.0010.0030.00293-180.0010.0080.00693-690.0050.0010.00793-800.0050.0140.01393-890.0020.0040.008

No γ -dec pool

Requirements of candidate gene/s:

- 1. Map to chromosome 3
- Low or no expression in fruits without γ-dec
- 3. Gene Function

RNA-seq in strawberry pools with / without γ-decalactone

Total of 33,458 genes predicted in the two pools 617 differentially expressed genes between the pools:

- 403 up-regulated
- 214 down-regulated

List of the top 25 significantly up-regulated genes in the high γ -dec pool compared to the No γ -dec pool.

Gene	Locus position	Predicted Function			log2 fold change	Test statistic	p-value	q-value
gana24070 v1 0 hybrid	brid IG1.16175288-16176946 Cinnamyl alcohol dehydrogenase-like		3 70			1.9E+209	2 78E 15	2 06E 12
gene24970-v1.0-liyblid	LG1.10175288-10170940	Alde kate reductese	3.79	0.00	-1.6E+306	-1.6E+306	0.000572	2.00E-12
gene22143-v1.0-llyblid	LG4.24174475-24170874	Microsomal data 12 alasta dasaturasa	1.90	0.00	-3.10	3.44	0.000373	0.021971
gene24414-v1.0-nybrid	LG3.31112417-31114043	Microsoffal delta-12 ofeate desaturase	323.92	11.41	-4.04	33.00	0	0
gene1/831-v1.0-hybrid	LGI:12556/15-12561321	Isoflavone 2 -hydroxylase-like	26.20	1.52	-4.10	13.80	0	0
gene11616-v1.0-hybrid	unanchored	NA	1.07	0.10	-3.49	6.49	8.34E-11	2.37E-08
gene12565-v1.0-hybrid	LG7:19409089-19409688	S-norcoclaurine synthase-like	8.52	1.00	-3.09	6.15	7.56E-10	1.61E-07
gene09812-v1.0-hybrid	LG6:9651860-9658187	Thaumatin-like protein	9.64	1.18	-3.03	8.53	0	0
gene29430-v1.0-hybrid	unanchored	Salicylic acid-binding protein 2-like	4.75	0.60	-2.98	4.62	3.92E-06	0.000362
gene09059-v1.0-hybrid	LG2:20948366-20950293	Hypothetical protein	1.10	0.14	-2.98	3.89	0.000102	0.005647
gene12485-v1.0-hybrid	LG1:6048496-6049132	Auxin-binding protein abp19a-like	2.15	0.28	-2.96	3.50	0.000468	0.018845
gene09427-v1.0-hybrid	LG5:10292998-10293780	Probable glutathione s-transferase-like	94.21	13.04	-2.85	17.49	0	0
gene16882-v1.0-hybrid	LG4:17649473-17650304	Probable glutathione s-transferase-like	5.13	0.76	-2.75	5.74	9.30E-09	1.60E-06
gene05671-v1.0-hybrid	LG6:28779898-28789864	Beta-glucanase	9.32	1.45	-2.69	10.13	0	0
-	LG6:25256895-25257433	Hypothetical protein	2.32	0.37	-2.67	3.24	0.001216	0.038671
gene08424-v1.0-hybrid	unanchored	Pathogenesis-related protein 4	7.04	1.15	-2.61	6.47	1.01E-10	2.79E-08
-	LG7:2164010-2164526	NA	4.04	0.67	-2.60	4.09	4.29E-05	0.002708
-	unanchored	NA	77.57	13.29	-2.55	3.37	0.000739	0.026493
gene13265-v1.0-hybrid	LG7:21691207-21692009	Par-1a protein	3.39	0.59	-2.53	4.10	4.17E-05	0.002653
-	LG4:2760209-2761669	Inhibitor of trypsin and hageman factor	4.34	0.76	-2.51	3.41	0.000656	0.024349
gene28799-v1.0-hybrid	LG3:9340369-9341925	Nectarin-3-like	2.84	0.52	-2.45	4.74	2.12E-06	0.000217
gene32603-v1.0-hybrid	LG4:2356895-2357929	Sieve element-occluding protein	1.82	0.34	-2.40	3.48	0.000503	0.019820
gene02395-v1.0-hybrid	LG3:5514062-5518181	Cytochrome p450	106.92	20.98	-2.35	19.15	0	0
gene21028-v1.0-hybrid	LG7:17655017-17664058	Psbp domain-containing protein chlorop-like	1.80	0.36	-2.33	4.37	1.24E-05	0.000952
gene17832-v1.0-hybrid	LG1:12556715-12561321	Hypothetical protein	19.76	4.11	-2.26	5.07	3.9E-07	4.78E-05
-	LG3:10391258-10391650	NA	14.68	3.06	-2.26	3.78	0.000156	0.007817

The fatty acid desaturase FaFAD1 and γ -decalactone in strawberry fruit

- FaFAD1 contains the Δ12 Oleic Acid Desaturase conserved domain, 6 putative trans-membrane domains and the 3 characteristic His-rich motifs.
- Also shares important Aa residues with Fatty Acid Hydroxylases





FaFAD1 NtFAD OeFAD2-1 DiFAD2 CaFAD2-2 AtFAD2 PpFAD1B-6 RcFAH12 LfFAH12	* MGAGCNMSVVT MGAGGRLSVPA MGAGGRMSVPT MGAGGCMSVSD MGAGGRMSVPT MGAGGRMSAPP MGGGGRMSTVI MGAGGRMSTVI MGAGCRIMVTP	20 GKTGEKK TKAEEKK TKSEQK TKSEQKN SSKKSET TRKNAET TSNNSEKKGGS: SSKKSET	* QRLPVSTPPFT NPLEKVPTSKPPFT NPLKRVPYLKPPFT NPLKRVPSEKPPFT DTTKRVPCEKPPFS DNPKRVPYSKPPFS SHLKRPPHTKPPFT EALKRGPCEKPPFT	40 * LSQLKKAIPPHCF WGDIKKAIPPHCF LGDVKKAIPPHCF SDLKKVIPPHCF VGDLKKAIPPHCF LGQIKKAIPPHCF LGDLKRAIPPHCF LGDLKRAIPPHCF VKDLKKAIPCHCF	60 C SLIRSFSYVIYDLVLVSL SLVRSFSYVVYDLILVSV SLRSFSYVVYDLFLVFL SLFRSFSYVVYDLSLASL SLTRSFSYVVSDLIISFL SIPRSFSYLISDIIIASC SVIRSFSYVFYDLTIAFL SFVRSFSYVAYDVCLSFL SIPRSFSYLLTDIILVSC	80 YHIATSYFHLIFHPL YYIATSYFHLIFSPYC YYIATSYFHLIFSPFC YYIATSYFHLIFSPFC YHIATAYFHHLFPFI YHIATAYFHHFPPIS YYVATNYFSILFOPIS YSIATNFFPYISSPI YYVATNYFSILFOPIS	* -YIAWPYYW -YLAWPIYW -YLGWSVYW -CIAWPYYW -SLAWLPYW -FLAWPYYW -FLAWPYYW -YVAWLVYW TYLAWPLYW	100 ILQGCTLTGV ICQGCVCTGI ILQGCVCTGV ITQGCVLTGV ACQGCVLTGV YVQGCVLTGV LFQGCILTGL VCQGCVLTGI	WVIAHECGHHAFS WVIAHECGHHAFS WVIAHECGHHAFS WVIAHECGHHAFS WVIAHECGHHAFS WVIAHECGHHAFS WVIGHECGHHAFS WVIGHECGHHAFS	YCLLDDVIG YCWVDDTVG YCWVDDTVG YCWVDDTVG YCWVDDTVG YCWLDDTVG YCWLDDTVG YCWLDDTVG YCULDDTVG YCLADDIVG YCLADDIVG	7 : 118 : 125 : 125 : 124 : 125 : 125 : 125 : 125 : 125 : 125	355455596
		c	o		TM-Helix1		D TA	1-Helix2	H-BOX1		-	
	140	*	160	* 180	* 200	*	220 5	*	240	* 260	-	
FaFAD1 NtFAD OeFAD2-1 DiFAD2 CaFAD2-2 AtFAD2 PpFAD1B-6 RcFAH12 LfFAH12	: LHSAL VPYFS : LHSALVPYFS : LHSTLVPYFS : LHSSLVPYFS : LHSALVPYFS : FHSFL VPYFS : VHSALVPYFS : FHSFL VPYFS	WKYSHRRHHSN WKYSHRRHHSN WKYSHRRHHSN WKYSHRRHHSN WKYSHRRHHSN WKYSHRRHHSN WKYSHRRHHSN WKYSHRRHHSN WKYSHRRHHSN	IGSMERDEAFVPKP IGSLERDEVFVPKP IGSLERDEVFVPKP IGSLERDEVFVPKP IGSLERDEVFVPKQ IGSLERDEVFVPKQ IGSLERDEVFVPKQ IGSLERDEVFVPKS	KSKISWYNKYENNP KSQLGWYSKYLNNP KSQLGWYSKYLNNP KSQLAWYSQYLNNP RSKITWYSKYLNNP KSAIKWYGKYLNNP KSAIGWYSKYLNNP KSAIGWYSKYSNNP KSAIGWYSKYSNNP	GRLITIIFTLTLGWPLYLA GRVMSLTVTLTLGWPLYLA GRVMTLVITLTLGWPLYLA GRILTLVITLTLGWPLYLA GRIISMLATLTLGWPLYLA GRIMLTVQFVLGWPLYLA GRFLTLTIQLTLGWPLYLA GRVLTLAATLLLGWPLYLA	ENVSGREYDRFACHYI INVSGREYDRFACHYI INVSGREYDRFACHYI ENVSGREYDRFACHYI ENVSGREYDFACHFI ENVSGREYDGFACHFI ENVSGREYDGFACHFI ENVSGREYDGFACHYI ENVSGREYDGFASHFI	PYSPIFSDR DPYGPIYNDR DPHGPIYNDR VPNSPIYNDR ENAPIYNDR ENAPIYNDR HPYGPIFSDR DPYGPIFSER FPHAPIFKDR	SLQIYISDL RLQIFLSDA RLQIYISDV RLQIYISDV RLQIYLSDA RLQIWLSDV RLQIFLSDA RLQIFLSDA RLQIYISDA	GILVTASVLYRLAN GVLGAGYLLYRIAI CVLATSYILYRVAI GVLATYLLYRVAI GIATMSYFLYRLA GIATMSYFLYRLA GILAVCFGLYRYA GVLAVVYGLYRLA GIFATTFVLYQATN GILAVCYGLYRYA	IAKGITWLVV VKGLAWLVV AQGLVWLTVV IAKGLAWLICV VLKGVSWVIVV VAKGLAWLVV VAKGLAWVVQV IAKGLAWVVRI ISQGLTAMICV	: 248 : 255 : 254 : 255 : 255 : 255 : 255 : 255 : 255 : 255 : 255	355455596
	TM-Helix3	H-BOX2		6	TM-Helix4 📢 🔫				TM-Helix5		-	
FaFAD1 NtFAD OeFAD2-1 DiFAD2 CaFAD2-2 AtFAD2 PpFAD1B-6 RcFAH12 LfFAH12	* GVPLLITNGFI GVPLLIVNGFI GVPLLIVNGFI GVPLLIVNGFI GVPLLIVNAFI GVPLLIVNGFI GVPLLIVNCFI GVPLLIVNFFI	280 VLITYLQHTHP VLITYLQHTHP VLITYLQHTHP VLITYLQHTHP VLITYLQHTHP VLITYLQHTHP VLITYLQHTHP VLITYLQHTHP VLITYLQHTHP VLITYLQHTHP	* SLPHYDSSEWDWLR SLPHYDSSEWDWLR ALPHYDSSEWDWLR SLPHYDGSEWDWLR SLPHYDSSEWDWLR SLPHYDSSEWDWLR AIPRYGSSEWDWLR SLPHYDSSEWDWLR	300 * GALSTVDRDCGVTH GALATVDRDYGILNM GALATVDRDYGVLNM GALATVDRDYGVLNM GALATVDRDYGVLNM GALATVDRDYGILNM GALATVDRDYGILNM GALATVDRDYGVLNM GALATVDRDYGVLNM	320 P VFHNITDTHVVHHLFSTI VFHNITDTHVVHHLFSTM VFHNITDTHVAHHLFSTM VFHNITDTHVAHHLFSTM VFHNITDTHVAHHLFSTM VFHNITDTHVAHHLFSTM VFHNITDTHVAHHLFSTM VFHNIADTHVAHHLFATT	340 PHYNAMEATTAVKPILO PHYNAMEATKAVKPILO PHYHAMEATKAIKPILO PHYHAMEATKAIKPILO PHYNAMEATKAIKPILO PHYNAMEATKAIKPILO PHYHAMEATKAIKPILO PHYNAMEATFAIKPILO	* EDYYGFDGTP EDYYGFDGTP EDYYGDGTP EDYYGFDGTP EDYYGFDGTP EDYYGDDRTP EDYYGDDRTP EDYYGFDGTP EDYYRFDGTP	360 IYKALWREAR YKAMWREAK FYKAMWREAK WYVAMWREAR WYVAMWREAK YFKAMFREAK FYKALWREAK	* 38 ECLYVEPDEDAPNI ECLYVEPDEASQ-G ECLYVEPDE-P-N ECLYVEPDG-EE-F ECLYVEPDREGD-F ECLYVEPDREGDK- ECLYVEPDEGAP-I ECLYVEPDTERG-F	O KGVFWYRNKL KGVFWYRNKF IKGVFWYRNKF GGVEWYRNKI KGVYWYNNKL KGVFWYRNKL KGVYYYNNKL	: 375 : 383 : 381 : 382 : 382 : 382 : 383 : 383 : 387 : 384	

H-BOX3

Expression of FaFAD1 is correlated to γ -decalactone production in strawberry fruit



Correlation between *FaFAD1* and γ -decalactone content in the 232 x 1392 population

eQTL analysis detects a *loci* controlling γ -decalactone content, *FaFAD1* and *FaFAH1* expression at the bottom of LG III-2



Agro-infiltration of 'Festival' fruits with FAD1-myc construct + P19



D 1. 5. Validation of markers for fruit volatile content prediction in strawberry

[#	ACCID	Accession	_	#	ACCID	Accession
ſ	-1		232		35	843	Jucunda
ľ	2		1392		36	854	Lanai
ľ	3		93-23		37	777	Macarena
ľ	4		93-34		38	213	Mara des bois
ľ	5		93-10		39	433	Medina
ľ	6		93-47		40	307	Mieze Schindler
ľ	7	480	Amiga		41	77	Milsey
ľ	8	300	Aromas		42	697	Naiad
ľ	9	288	Avalon		43	104	Oso Grande
ľ	10	72	Camarosa		44	856	Palomar
ľ	11	677	Camino Real		45	39	Pandora
ľ	12	715	Candonga		46	809	Pedrone
ľ	13	292	Captiva		47	674	Plarionfre, Chiflón
ľ	14	84	Chandler		48	316	Premial
ľ	15	664	Cifrance		49	844	Reusraths aller Krüester
ľ	16	301	Cijosee, Cireane		50	706	Roxana
ſ	17	701	Commitment,		51	839	Rubigen
ſ	18	778	Coral	-	- 52	673	Ruby
ſ	19	304	Deutsch Evern		53	965	Sabrina
ſ	20	302	Diamante		54	862	Splendor
ſ	21	130	Douglas		55	112	Tioga
ſ	22	956	Driscoll Magdalena		56	245	Toyonoka
ſ	23	955	Driscoll Windsor		57	676	Ventana
ſ	24	126	Elvira		58	294	Ville de Paris
ſ	25	699	Endurance		59	866	Virtude
ſ	26	315	Everest		60	953	Viva Patricia
ſ	27	732	Festival		61	838	Winter Down
ſ	28	879	Fuentepina		62	596	F. vesca (Blanca)
ľ	29	678	Galante		63	660	F. vesca 'Reine des vallées'
ľ	30	733	Galexia		64	180	F. virginiana
ľ	31	131	Gento		65	591	F. moschata 'Capron Royale
	32	308	Gigantella		66	606	F. Moschata (Rusian)
	33	700	Honor		67		CS 9/2 EU
	34	832	Hood		68		CS 13/2 EU

D 1. 5. Validation of markers for fruit volatile content prediction in strawberry

Marker for **γ**-decalactone

68 cultivars and species have been genotyped and and volatile content measured (still on progress for 16 of them)

The developed marker in the gene *FaFAD1* is able to predict γ -decalactone production with a 94.23% accuracy



Analysis of mesifurane content and *FaOMT*

Homoeology Group VII (HG VII)

LG VII-1

LG VII-2

LG VII-3

LG VII-4



The gene FaOMT controls the variation in mesifurane content in strawberry



FaOMT expression in lines not producing mesifurane is barely detectable

An indel in the promoter of *FaOMT* fully co-segregates with mesifurane content

3 potential cisregulatory elements are specific of the functional allele

Co-segregation of the *FaOMT* promoter marker and mesifurane content in 232 x 1392



D 1. 5. Validation of markers for fruit volatile content prediction in strawberry

Marker for Mesifurane

68 cultivars and species have been genotyped and and volatile content measured (still on progress for 16 of them)

The developed marker in the gene *FaOMT* is able to predict mesifurane production with a 93,6 % accuracy

