# “CARTOGRAFIADO DE QTLs Y GENES CANDIDATOS ASOCIADOS A METABOLITOS DETERMINANTES DE LA CALIDAD DE FRUTO EN MELOCOTÓN”

Tesis presentado por:

**Gerardo Sánchez**

para optar al título de Doctor en Biotecnología

**Director:** Dr. Antonio Granell
**Co-Director:** Dr. Antonio José Monforte, Dra. María Luisa Badenes

-2013-

*A quien me ha esperado para despedirnos (por el momento),*

*a mi padre, in memoriam*

**Agradecimientos**

Durante mi estancia en Valencia recibí la ayuda de muchas personas a quienes quisiera agradecer. Muchos lo han hecho con cuestiones relacionadas al trabajo, otros más relacionadas al alma, incluso algunos lo han hecho a los dos niveles. No importa de que forma ha sido, esta tesis no podría haberla realizado sin este apoyo. Seguro que no podré nombrarlos a todos, aunque sienta una sincera gratitud, igualmente lo intentaré.

En primer lugar quisiera agradecer a mis directores de tesis: Antonio Granell, Marisa Badenes y Antonio Monforte por permitirme desarrollar mi tesis en sus laboratorios, ha sido una experiencia única. De todos ellos he aprendido algo, muchas veces más importante que lo estrictamente laboral, pero especialmente recordaré el apoyo de Toni G. quien siempre ha encontrado tiempo para hablar de ciencia.

Esta tesis fue realizada a partir de la financiación recibida por parte del Instituto Nacional de Tecnología Agropecuaria (INTA) de Argentina y por tanto quiero agradecer a quienes en su momento han confiado en mi propuesta de capacitación: Norberto Ángel, Luís Arroyo y Héctor Martí. También estoy agradecido a todos mis compañeros de la EEA San Pedro, especialmente a Gabriel y Cecilia que me han ayudado con la recolección de muestras mientras estaba en Valencia.

También quisiera agradecer a Benoit (del INRA) por su atención durante mi estadía en Toulouse y a Quino y Mónica del Instituto de la Grasa (Sevilla) por una excelente colaboración. Al departamento de poscosecha del IVIA, a Sandra por dejarnos "cacharrear" ahí y a Cristina con quien nos hemos pegado unos palizones de trabajo pero más importante fue todo lo que aprendimos juntos. Y por eso te digo gracias.

Quisiera agradecer también a todas las personas que he conocido en el IBMCP y en el IVIA, de una u otra forma me han ayudado. Particularmente a Gabino por

prestarme sus melocotoneros y a Miguel Ángel por sus consejos en las transformaciones. A todo el servicio de metabolómica donde he pasado más de una tarde y he recibido la colaboración de las personas que ahí trabajan. También estoy agradecido a mis compañeros de los laboratorios donde realicé mi tesis, todos han sido de gran ayuda, pero especialmente agradezco a Aurora y Sole del Lab 0.08, quienes hicieron que muchos días dentro del lab. fueran más que divertidos (con cubatas incluidos) sin contar las noches. Aurora: gracias por la "valiosa información" compartida en nuestras charlas, a la cual he sabido sacar provecho. Del Lab 2.10, particularmente quiero agradecer a José Luís y Asun por su buena predisposición para ayudar siempre.

A toda la ONU, miembros fundadores, pasados y presentes con quienes hemos compartido más de una tapa y mucho más que una caña, y quiero agradecerlo. A Alfredo que ya ni recuerdo todo lo que nos hemos divertido. A Alessandro con quien hemos compartido varias aficiones y algunos vicios. A Liliam, Pilar, Camilo, Eszter, Sandra, Carolina, Rim y Lorena (que se nos fue pronto pero no se alejó) con quienes muchas veces hicieron que lo mejor del día sea compartir la comida o las cervezas post-laboratorio. Gracias a todos aquellos de Arg. que entendieron que 11.000km no son suficientes separar a las personas y a todos los que lo empezarán a comprender de ahora en más.

Finalmente, quiero agradecer a toda mi familia, a Claudía, mi hermana, cuyas llamadas han sido una alegría y sus visitas (con su flia. incluida!!!) nos ha permitido pasar un tiempo juntos que siempre recordaré. Por último quiero agradecer a los primeros, a mis padres: María del Carmen y Carlos, de quienes he aprendido más en esta vida. Seguir sus ejemplos basadas en el respeto, la perseverancia y la pasión me ha permitido desenvolverme allá donde he ido.

**INDICE**

**RESUMEN**…… **8**

**ABSTRACT**…… **11**

**RESUM**…….... **14**

**CAPÍTULO 1: Introducción general**……. **18**

1.1 El Melocotonero……... **20**

1.2 Importancia de la producción de melocotones……... **21**

1.3 Mejoramiento del melocotonero……... **22**

1.4 Caracteres cuantitativos…….. **24**

1.5 Compuestos volátiles del melocotón…… **26**

1.6 Control genético del aroma del melocotón…….. **38**

1.7 El melocotonero como un nuevo modelo de fruto……. **42**

1.8 Antecedentes y estructuración del presente trabajo……... **45**

1.9 Referencias…….……... **48**

**OBJETIVOS**…….…….. **55**

**CAPÍTULO 2: A non-targeted approach unravels the volatile network in peach fruit**…….…… **58**

**CAPÍTULO 3: An integrative "Omics" approach identifies new candidate genes to impact aroma volatiles in peach fruit**…… **93**

**CAPÍTULO 4: The peach volatiloma modularity is reflected at the genetic and environmental response levels**…… **149**

**CAPÍTULO 5: Discusión general** ……. **188**

**CONCLUSIONES**…… **200**

**ANEXOS**…….…… **204**

Información suplementaria Capítulo 2…….. **205**

Información suplementaria Capítulo 3…….. **225**

Información suplementaria Capítulo 4…….. **240**

## RESUMEN

La mejora de la calidad organoléptica del fruto es probablemente uno de los mayores desafíos a los que se enfrenta la industria del melocotón en la actualidad. Los compuestos orgánicos volátiles (VOCs) son un grupo heterogéneo de sustancias que están íntimamente relacionados con la calidad del melocotón, ya que definen íntegramente el aroma y en combinación con otros compuestos como azucares y ácidos orgánicos configuran el sabor de los frutos. Entre los VOCs de melocotón, los compuestos de tipo lactona son los que otorgan las características típicas al aroma y el sabor del fruto mientras que otros compuestos como algunos ésteres y terpenoides contribuyen en menor medida, aportando ciertas notas específicas que se describen como "frutadas" o "florales". Otro grupo de compuestos formado por aldehídos y alcoholes, conocidos genéricamente como volátiles derivados de lípidos, tendrían un impacto menos deseable sobre el aroma del fruto ya que parecen otorgar notas asociadas a la inmadurez del fruto.

Describir en detalle los componentes de este grupo complejo de compuestos así como también establecer las bases genéticas y moleculares que controlan su síntesis resultan pasos cruciales para emprender con mejores garantías la mejora genética de la calidad del melocotón. En la presente tesis se abordó de forma holística el estudio de la producción de compuestos volátiles mediante tecnologías "omicas" complementarias: metabolómica, transcriptómica y genómica.

Inicialmente se estableció una plataforma metabolómica de alto rendimiento para la identificación y cuantificación de los VOCs del fruto. Esta plataforma permitió, además del escrutinio de una gran proporción de los VOCs del fruto, establecer los grupos de compuestos que comparten una regulación común. Mediante una combinación de análisis basados en correlación se descubrió que la síntesis de volátiles con impacto positivo en el aroma (como las lactonas y

algunos esteres) ocurre conjuntamente. De igual manera ocurre con compuestos con impacto negativo (compuestos derivados de lípidos) y otros compuestos con menor impacto (por ejemplo terpenoides y fenólicos). Sugiriendo que el volatiloma del melocotón está organizado en diferentes módulos.

El descubrimiento de esta red de regulación subyacente a la síntesis de volátiles fue posteriormente aprovechado para la identificación de genes candidatos mediante una aproximación de genómica funcional. Se analizó en paralelo la expresión de genes mediante miroarrays y el contenido de volátiles en una serie temporal de maduración de dos genotipos parentales de nuestra población de mejora. El análisis en conjunto de los patrones de co-regulación permitió la identificación de diferentes genes cuya expresión se mostró altamente correlacionada con los niveles de lactonas, esteres, terpenoides, fenólicos y de compuestos derivados de lípidos entre otros, los cuales fueron propuestos como posibles genes candidatos involucrados en la síntesis de estos aromas. Uno de los genes candidatos identificado por estar asociado a la acumulación de lactonas y esteres, fue clonado y mediante análisis funcional se demostró que posee una actividad ω-6 oleato desaturasa implicada en la generación de ácido linoleico, un posible precursor de aromas del melocotón.

Complementariamente, se describió el control genético de la producción de volátiles mediante un análisis de QTL a gran escala. Una población F1 fue analizada combinando la plataforma metabolómica desarrollada y una plataforma de genotipado masivo basada en marcadores del tipo SNPs. La alta organización del volatiloma en módulos co-regulados se reflejó en la identificación de loci controlando varios de los miembros pertenecientes a los diferentes grupos de co-regulación. De esta forma, se identificó tres loci que controlan la síntesis de lactonas y algunos esteres en los grupos de ligamiento 4, 5 y 6 (LG4, LG5 y LG6). Así mismo, se determinó que la síntesis de monoterpenos está controlada por un locus único ubicado en el extremo

superior del LG4. Los resultados obtenidos sugieren que es posible mejorar el aroma de los frutos mediante mejora asistida por marcadores moleculares.

La integración de los resultados junto con la interpretación en relación al conocimiento existente permitió describir el volatiloma del melocotón, su red regulatoria y describir regiones génicas controlando su síntesis así como también proponer genes candidatos involucrados probablemente en la síntesis de los principales aroma del melocotón.

## ABSTRACT

The improvement of fruit quality is probably the main challenge facing the peach industry nowadays. The volatile organic compounds (VOCs) are a heterogeneous group of substances that are intimately related to peach quality since they entirely define its aroma and in combination with others compounds like sugars and organic acids shape fruit flavor. Among peach VOCs, the lactone-type compounds are those conferring the characteristic peach notes for both aroma and flavor, while others compounds such as esters and terpenoids contributes in a lesser extent by adding specific notes often described as "fruity" and "floral". Another group of compounds consisting of aldehydes and alcohols, known generically as lipid-derived volatiles, should have a less desirable impact on the flavor of the fruit since they seem to provide notes associated with immature fruit.

Describing in detail the components of this complex group of compounds as well as establishing the genetic and molecular bases that control their synthesis is a prerequisite before undertaking any peach breeding project aimed to increase peach quality with more chances of success.

In the current thesis, the study of the volatile compounds production was addressed holistically by means of a series of "omics" complementary technologies: metabolomics, transcriptomics and genomics.

Initially a high-throughput metabolomics platform for identification and quantification of the VOCs from the fruit was established. Besides the screening of a large proportion of the VOCs from the fruit, this platform enabled, the definition of groups of compounds that share a common regulation.

By a combination of a series of analyzes based on data correlation it was found that the synthesis of volatiles with positive impact in aroma (such as lactones and certain esters) occurs in a coordinated manner. The same happens with negative impact compounds (lipid derived compounds) and other compounds

with less impact (e.g. terpenoids and phenolics). All this suggesting that the peach volatilome is organized into a series of interconnected modules.

The discovery of this regulatory network underlying volatile synthesis was then exploited to propose a number of candidate genes by a functional genomics approach. The gene expression levels were analyzed by miroarrays and the volatile contents were profiled in parallel along maturation time course series in the two parental genotypes of our breeding population.

The combined analysis of the co-regulation patterns allowed the identification of different genes whose expression was highly correlated with the levels of lactones, esters, terpenoids, phenolic compounds and lipid derivatives among others, which were proposed as possible candidate genes involved in the synthesis of these aromas.

One of the candidate genes, identified as being associated with the accumulation of lactones and esters, was cloned and by functional analysis showed to be an oleate ω-6 desaturase involved in the generation of linoleic acid, a potential precursor peach aromas.

Complementarily, the genetic control of the production of volatile was described by a broad scale QTL analysis. An F1 peach population segregating for fruit aroma was analyzed combining the metabolomics platform developed previously and a high-throughput genotyping platform based on SNPs markers detection. The high organization of the volatiloma in co-regulated modules was reflected in the identification of loci controlling several members belonging to different groups of co-regulation. In this way, three loci that control the synthesis of lactones and some esters were identified in linkage groups 4, 5 and 6 (LG4, LG5 and LG6). Similarly, it was found that the synthesis of monoterpenes is controlled by a single locus located in the upper end of LG4.

The results obtained indicate the presence of important regions in peach genome affecting fruit volatile production and suggest that it should be possible to improve the flavor of the fruits by molecular marker-assisted breeding.

The integration of the results along with the interpretation in light of the existing knowledge allowed the description of the peach volatilome, its regulatory network and the identification of genomic regions controlling their synthesis as well as propose candidate genes involved in the synthesis of the major peach aromas, which could be also used as targets for classical and biotech breeding for flavor and aroma in peach.

## RESUM

La millora de la qualitat organolèptica del fruit és tal vegada una de les qüestions pendents més importants que troba les companyies productores de bresquilles en l'actualitat. Els compostos anomenats volàtils orgànics (VOCs) són un grup heterogeni de substàncies que són íntimament relacionades amb la qualitat de la bresquilla en definir la seua aroma d'una manera integra i a la vegada conformen el sabor dels fruits amb la participació dels sucres i els àcids. Entre els VOCs de la bresquilla són els compostos de tipus lactona els que confereixen les característiques típiques a l'aroma i al sabor del fruit, mentre que altres compostos com són alguns èsters i terpenoides contribueixen en menor mesura aportant notes específiques que sovint són descrites a la literatura com "afruitades" o "florals" Un altre grup de compostos que es format per aldehids i alcohols, i que són coneguts genèricament com volàtils derivats dels lípids, tindrien en canvi un impacte negatiu sobre l' aroma del fruit en conferir-hi notes de fruit verd.

El fer una descripció detallada dels components d'aquest complexe grup de compostos, així com establir les bases genètiques i moleculars que controlen la seua síntesi són fites fonamentals per començar un programa de millora genètica de la qualitat de la bresquilla. En aquesta tesi s'abordà de manera holística estudiar la formació dels compostos volàtils mitjançant un conjunt de tècniques "òmiques" de naturalesa complementària: la metabolòmica, la transcriptòmica i la genòmica.

Vàrem començar establint una plataforma metabolòmica d'alta eficàcia per tal d'identificar i quantificar els VOCs del fruit. Aquesta plataforma va permetre a més de avaluar una gran part dels VOCs del fruit, establir grups de compostos que comparteixen una regulació comú. Mitjançant una combinació de tècniques d'anàlisi de dades basades en la correlació entre les dades trobarem que la síntesis del volàtils amb un impacte positiu en l'aroma (com són les lactones i

alguns èsters) ocurreix conjuntament. Tanmateix passa amb alguns compostos que tenen un impacte negatiu (compostos derivats de lípids) i altres compostos menys importants (com són els terpenoides i derivats fenòlics). Tot això suggereix que el volatiloma de la bresquilla està organitzat en un sistema modular.

El descobriment d'una xarxa de regulació per sota la síntesi dels volàtils ha estat aprofitat posteriorment per a identificar gens candidat mitjançant una aproximació de genòmica funcional. Es va analitzar paral·lelament l'expressió de gens mitjançant la utilització de microarrays i el contingut de volàtils en sèries temporals de maduració de dos genotips parentals de la nostra població de millora. L'anàlisi combinat dels patrons de co-regulació va permetre identificar alguns gens la expresió dels quals estava altament correlacionada amb els nivells de lactones, èsters, terpenoides, fenòlics i de compostos derivats dels lípids, tots els quals varen ser proposat com a gens candidats a estar involucrats en l'acumulació d'aquests composts aromàtics. Un gen candidat que es va identificar per la seua associació amb l'acumulació de lactones i esters va ser clonat i mitjançant anàlisi funcional demostrarem que tenia una activitat ω-6 oleat desaturasa implicada en la generació d'acid linoleic, un possible precursor d'aromes de bresquilla.

Complementàriament, varem descriure el control genètic de la producció de volàtils mitjançant una anàlisi de QTL a gran escala. Una població F1 va ser analitzada combinant la plataforma metabolòmica desenvolupada prèviament i una plataforma de genotipat massiu basada en marcadors de tipus SNPs. L' alta organització del volatiloma en mòduls co-regulats es reflectí amb la identificació de loci que control·laven diferents membres pertanyents als diferents grups de co-regulació. D'aquesta forma, s'identificà tres loci que controlen la síntesi de lactones i alguns èsters als grups de lligament 4, 5 y 6 (LG4, LG5 y LG6). Tanmateix es determinà que l'acumulació de monoterpens està controlada per un locus únic localitzar a l'extrem superior del LG4. Els

resultats obtinguts suggereixen que és possible millorar l'aroma de les bresquilles mitjançant la millora assistida per marcadors moleculars.

La integració dels resultats combinada amb la interpretació segons el coneiximent a l'abast va permetre descriure el volatiloma de la bresquilla, la seua xarxa reguladora i descriure regions gèniques que controlen la seua acumulació, aixi com també proposar gens candidat per participar en aquest procés d'acumulació.

# Capítulo 1:
# INTRODUCCION GENERAL

# INTRODUCCIÓN GENERAL

## 1.1 El Melocotonero

La familia *Rosaceae* con sus 3000 especies distribuidas en 110 géneros es una de las familias más importantes por el valor económico de algunas de las plantas que la conforman y su amplia distribución en el mundo (Takhtajan, 1997). El género *Prunus* es probablemente el que posee mayor número de especies (unas doscientas) para el consumo humano, algunas con alto valor económico, dentro las cuales destacan: el melocotonero [*P. persica* (L.), el albaricoquero [*P.armeniaca* L.], el almendro [*Prunus dulcis* (Mill.) D.A. Webb], Batsch], el ciruelo europeo [*P. domestica*], el ciruelo japonés [*P. salicina*], el cerezo [*P. avium*] y el guindo [*P. cerasus*].

El melocotonero es un frutal originario de China donde las referencias de su cultivo se remontan hasta unos 3000 años de antigüedad (Martínez-García et al., 2013). Los romanos lo llevaron a Persia (actual Irán) a través de las rutas comerciales, llegando a ser conocido como fruta pérsica, de ahí su nombre botánico: *Prunus persica* (Gordó, 2012). Hacia el año 330 a.C. se introdujo en Grecia, y durante la Edad Media su cultivo se extendió por toda Europa. Durante el periodo de colonización (siglo XVI) los españoles lo llevaron a México desde donde se extendió a otros países limítrofes (Martínez-García et al., 2013). En países americanos más alejados, como es el caso de la Argentina, el melocotonero fue introducido posteriormente por europeos hacia finales del siglo XVIII. Sin embargo la consolidación definitiva se produce después de la finalización de la primera guerra mundial donde ocurre una migración masiva, principalmente de italianos y españoles, que además de los materiales vegetales traen consigo su experiencia en el cultivo (Pagliaricci and Ángel, 2012).

La historia relacionada con este cultivo así como sus requerimientos fisiológicos explican las zonas actuales donde se cultiva el melocotonero en el mundo.

## 1.2 Importancia de la producción de melocotones

En cuanto a producción mundial, el melocotonero es la tercera especie frutal después del manzano y del peral (Arús et al., 2012). En el hemisferio norte los principales países productores son China, Italia, España, EE.UU., Grecia y Francia (Fig. 1, FAO, 2011). España es el tercer productor a nivel mundial después de China, Italia. En el hemisferio sur los mayores productores son: Chile, Argentina, Brasil y Sudáfrica (FAO, 2011).

Uno de los aspectos que caracterizan al sector productivo del melocotón en España es su creciente competitividad en los mercados internacionales, principalmente dentro de Europa. España, desde el año 2006, es el primer exportador de melocotón de Europa, logrando en 2010 una exportación de 585.000 toneladas por un valor aproximado de 500 millones de euros (FAO, 2010), siendo Francia, Alemania y Portugal sus principales mercados (Iglesias and Casals, 2011).

En ese mismo año Argentina exportó más de 10.000 toneladas de melocotones, por un valor cercano a los 7 millones de euros; colocándola como el 16º exportador a nivel mundial y segundo en Sudamérica detrás de Chile (FAO, 2010). La producción en el caso de Argentina tiene como destino final la exportación para consumo en fresco, fundamentalmente a Brasil (Pagliaricci and Ángel, 2012).

La comercialización del melocotón experimenta el fenómeno conocido como primicia, refiriéndose a que el valor del producto es mayor al inicio de la campaña debido a que los consumidores aprecian más las frutas que hace tiempo que no consumen. Por esta razón, si bien los productores de ambos

hemisferios no compiten directamente debido a la diferente estacionalidad de la producción de melocotón, la importación proveniente del hemisferio opuesto antes del inicio de la campaña local, puede romper la primicia y afectar negativamente el valor del producto durante la estación de cosecha.

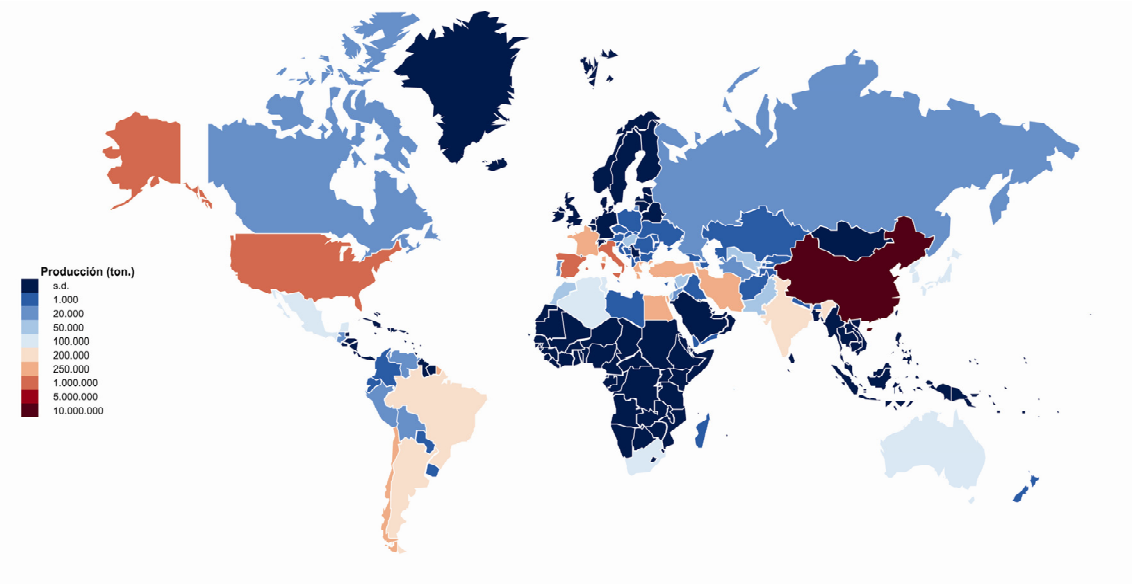



**Figura 1. Producción de melocotones en el mundo.** Con diferentes colores se muestra la producción de melocotones de cada país (en toneladas, ton.) según los datos de la FAO (2011). Los países con una producción menor a 1000 ton. o sin datos registrados por la FAO (s.d.) se muestran en azul oscuro.

## 1.3 Mejoramiento del melocotonero

Tradicionalmente los programas de mejora del melocotón se centraron fundamentalmente en la obtención de genotipos élite de alta productividad, resistentes a plagas y patógenos, que se adaptaran a diferentes zonas agroecológicas y que produjeran frutos de gran tamaño y buen aspecto (Daorden, 2012). Como resultado, muchos de estos programas han obtenido cultivares de excelentes características agronómicas. No obstante, el mejoramiento selectivo hacia caracteres agronómicos puede ir en detrimento de la calidad organoléptica del fruto como fue demostrado en el caso de la fresa y el tomate donde algunos aromas se perdieron en el proceso de mejora (Klee and

Giovannoni, 2011; Olbricht et al., 2008). En el caso del melocotón, la disminución de la calidad del fruto ha sido también percibida por los consumidores y además es la mayor causa de insatisfacción de los mismos (Bruhn et al., 1991). Un probable consecuencia de esto puede ser el bajo consumo de melocotón (2 Kg de fruta per capita por año) en comparación con otras frutas como el plátano (9 Kg) y la manzana (16 Kg), y que se ha identificado en algunos países como EE.UU. (Crisosto, 2006). En el caso de España, si bien en la actualidad el consumo se encuentra estabilizado alrededores de los 4 Kg per capita por año, ha sufrido una disminución de 46% desde 1989 (Iglesias, 2010). Estudios pioneros han establecido que el aroma junto con la firmeza y el color del fruto son los atributos principales por los cuales los consumidores juzgan la calidad del melocotón (Bruhn, 1995). En este sentido, parece imprescindible incorporar la obtención de genotipos "más aromáticos" dentro de los objetivos de los programas de mejora de cultivares elite de mayor calidad.

Si bien, como se comentó anteriormente, muchas de las variedades disponibles actualmente fueron obtenidas teniendo en cuenta casi exclusivamente un criterio de productividad, tamaño y aspecto del fruto, la tendencia actual de los programas de mejora se dirigen a aumentar el periodo de disponibilidad de la fruta así como también mejorar la calidad organoléptica del fruto (Daorden, 2012). Con respecto al último punto, los esfuerzos se centran principalmente en obtener genotipos con diferentes relaciones entre el contenido de sólidos solubles y la acidez. Este parámetro está directamente relacionado con el sabor, el carácter dulce y la percepción de acidez del fruto, todos estos íntimamente relacionados con la aceptación de los melocotones (Kader, 1999). De esta forma, se han obtenido variedades dulces, semi-dulces, equilibradas, ácidas y muy ácidas. Si bien, los parámetros mencionados son clave para obtener un fruto de

calidad, los compuestos volátiles orgánicos (VOCs) son los que otorgan los sabores característicos al melocotón además de definir íntegramente el aroma. En este sentido, los programas de mejoramiento italianos y el que se lleva a cabo en el IVIA se han centrado en recuperar características de aromas y sabor que poseen variedades tradicionales locales.

El melocotón como especie arbórea, presenta ciertas características que dificultan las tareas de mejora, como son el poseer un período de floración corto, lo que obliga a concentrar las tareas en un periodo de tiempo estrecho; y un tiempo de generación (obtención de semilla a partir de semilla) largo que varia entre 3 y 5 años (Abbott et al., 2008). Además al ser una especie autocompatible existe una reducida variabilidad genética del germoplasma cultivado, lo cual podría ser subsanado incorporando germoplasma exótico o de especies relacionadas. Todo ello, hace que la mejora clásica del melocotón acarree unos costos operativos altos tanto de mano de obra como de mantenimiento, pues hay que conservar las poblaciones de mejora en condiciones de campo hasta la caracterización del fruto, que es el órgano donde se miden la mayoría de los caracteres a mejorar. Estos altos costos pueden ser minimizados si los programas de mejora utilizan marcadores moleculares durante el cribado para la identificación de los genotipos deseados (mejora asistida por marcadores). Para poder disponer de esta herramienta, primero debe invertirse en la identificación de marcadores asociados a los caracteres a mejorar.

### 1.4 Caracteres cuantitativos

Diseccionar el control genético de caracteres que influyen en la calidad del fruto es el paso inicial para implementar el uso de marcadores moleculares en programas de mejora. Estos caracteres están sometidos frecuentemente a un

control multigénico y por tanto su estudio debe realizarse a través de la identificación de loci de caracteres cuantitativos (QTL, del inglés "*Quantitative Trait Loci*"). Los QTL se identifican basándose en el principio de asociación entre los marcadores moleculares polimórficos y el fenotipo de los individuos de una población y consiste en cuatro etapas fundamentales [para revisión ver (Collard et al., 2005)]: **1)** la obtención de una población segregante, **2)** la identificación de los marcadores polimórficos y su posterior genotipado en la población, **3)** la medición del carácter a analizar en la población, que se conoce como fenotipado y **4)** el análisis estadístico de los datos de genotipado y fenotipado para descubrir la asociación entre marcadores y caracteres. En el caso del melocotón y en particular del control genético del aroma, el análisis de QTL tiene algunas dificultades adicionales. Debido al prolongado periodo juvenil del melocotonero, lo cual es una desventaja *per se*, rara vez suelen desarrollarse poblaciones de mapeo que requieren varias generaciones (como por ejemplo Back Cross o RILs) sino que frecuentemente se acopla el estudio de QTL a los programas de mejora utilizando las poblaciones desarrolladas, que en su mayoría son F1 o como mucho F2 o retrocruzamientos. Se ha demostrado que el melocotonero puede ser cruzado con otras especies de su subgénero (*Amygdalus*) como *P. ferganensis*, *P. mira*, *P. davidiana* o el almendro (*P. dulcis*) produciendo híbridos fértiles. Incluso se ha conseguido cruzar con otros frutales perteneciente a otros subgéneros (*Prunophora* y *Cerasus*) como a los que pertenecen el albaricoque (*P. armeniacana*), ciruelo japones (*P. salicina*) y el cerezo (*P. cerasus*) aunque los híbridos producidos raramente son fértiles. Por tanto hay un gran *pool* de genes disponible para la mejora, sin embargo estas fuentes de variabilidad han sido explotadas ocasionalmente (Abbott et al., 2008) debido a las dificultades asociadas a la eliminación de los caracteres indeseables por los prolongados tiempos de generación. Por tanto, los genotipos que

resultan más prácticos para su uso como parentales en programa de mejora son generalmente cultivares comerciales que muestran una baja variabilidad genética lo cual se traduce en un obstáculo a la hora de identificar marcadores polimórficos segregantes y en algunos casos un bajo número de caracteres contrastantes entre los genotipos utilizados. Por otro lado, en el caso del estudio de QTL de aromas se encuentra el obstáculo adicional de que la fracción volátil del fruto en general la forman muchos compuestos de naturaleza química diversa. En los últimos años las plataformas tecnológicas de alto rendimiento, que genéricamente se conocen como tecnologías "omicas", han mejorado significativamente, permitiendo la adquisición de un gran volumen de datos en un único análisis. En este sentido, el análisis de la fracción volátil del fruto , también conocida como volatiloma, en su conjunto puede ser abordada con tecnologías metabolómicas (Bicchi and Maffei, 2012). Por otro lado, la baja variabilidad genética de los cultivares de melocotón que exigía un escrutinio exhaustivo para la identificación de marcadores polimórficos, en la actualidad puede realizarse mediante un "array" que permite la evaluación en simultaneo de 9000 marcadores del tipo SNPs (del inglés "*Single Nucleotide Polimorfism*") en melocotonero (Verde et al., 2012). Ciertas restricciones relacionadas con la fisiología de la especie pueden ser parcialmente subsanadas. Por ejemplo el tiempo de generación (obtención de semilla a semilla) puede acortarse si se emplea cultivo *in vitro* en vez de la estratificación tradicional del hueso del melocotón.

## 1.5 Compuestos volátiles del melocotón

### Volátiles involucrados en el flavor del fruto

El término inglés "*flavor*" no tiene un equivalente en el castellano, sino más bien engloba la suma del aroma y el sabor de un alimento. La percepción del *flavor*

por parte de los humanos involucra la integración de una gran cantidad de información cuantitativa proveniente de múltiples sistemas sensoriales. En el caso de los frutos comienza con la apariencia de los mismos. Principalmente, el color afecta la evaluación subjetiva del *flavor*, la cual también está influenciada por la textura y la sensación que provoca en la boca (Christensen, 1983; Stommel et al., 2005; Zanor et al., 2009). Desde el punto de vista químico, el *flavor* es la suma de una gran cantidad de metabolitos primarios y secundarios que son percibidos por los sistemas gustativo y olfativo. Existen cinco clases de receptores gustativos que reconocen lo dulce, salado, agrio, amargo y la sensación difícil de definir conocida como umami, presente por ejemplo en los tomates por su alto cantidad de glutamato. La base del sabor de los frutos está proporcionada por azucares (glucosa y fructosa) y ácidos orgánicos (citrato y malato principalmente), actuando en un balance adecuado. Sin embargo, la complejidad inherente al *flavor* característico de cada fruta está dada por el sistema olfativo (Shepherd, 2006). Existiendo más de 350 receptores olfativos con los cuales interaccionan los compuestos volátiles orgánicos (VOCs). Los volátiles del melocotón han sido estudiados con anterioridad, describiéndose un poco más de 100 compuestos incluyendo: lactonas, esteres, terpenos, aldehídos, ácidos carboxílicos y alcoholes entre otros [(Aubert et al., 2005); y referencias incluidas]. Se ha demostrado que la producción de volátiles en melocotón es un proceso dinámico y está influenciado por diversos factores como el genotipo (Wang et al., 2009), el estado de madurez (Aubert et al., 2003; Chapman et al., 1991; Visai and Vanoli, 1997; Zhang et al., 2010), la aplicación de tratamientos poscosecha (Robertson et al., 1990; Sumitani et al., 1994; Zhang et al., 2010) y las prácticas de cultivo empleadas (Jia et al., 2005).

Potencialmente todos los compuestos volátiles pueden interaccionar con los receptores olfativos de los humanos, sin embargo una concentración mínima es

requerida para generar una respuesta. En este sentido, se ha estimado que menos del 5% de los volátiles de los alimentos contribuyen al aroma (Grosch, 2001). En el caso del melocotón, la contribución de los compuestos volátiles al aroma de fruto ha sido estudiada, y se han propuesto unos pocos compuestos claves (Eduardo et al., 2010). Dentro de éstos se encuentran las lactonas (Derail et al., 1999; Eduardo et al., 2010; Horvat et al., 1990) y en particular la γ-decalactona cuyo aroma fue descrito como "típico de melocotón" (Derail et al., 1999). Otros compuestos que también contribuirían al aroma de fruto maduro son los esteres lineales: hexil acetato y (*Z*)-3 hexenil acetato, que aportan notas "frutadas" al aroma general del melocotón (Derail et al., 1999; Eduardo et al., 2010) Mientras que los compuestos terpenoides: Linalool y β-ionona contribuirían con notas "florales" (Eduardo et al., 2010; Greger and Schieberle, 2007; Guillot et al., 2006). Al contrario, los compuestos derivados del catabolismo de lípidos como los aldehídos (*E*)-2-hexenal y (*Z*)-3-hexenal poseen aromas que se definen como "verdes" (Derail et al., 1999) los cuales se asocian a fruta inmadura.

Además de su contribución a la calidad del fruto, los compuestos volátiles del melocotón son productos muy deseados en la industria alimenticia y de perfumería donde son usados como agentes saborizantes. En particular la γ-decalactona es uno de los compuestos más utilizados para conferir el aroma típico del melocotón (Dufosse et al., 1994) lo cual genera que su demanda mundial, calculada en 10.000 Kg en 1997, esté en expansión (Gatfield, 1999).

**Biosíntesis de los volátiles relacionados al aroma**

Desde el punto de vista químico, los VOCs son sustancias que poseen carbonos con una alta presión de vapor lo que permite que se conviertan fácilmente en vapor o gases. En las plantas, se han identificado más de 1000 compuestos

volátiles con funciones muy diversas (Pichersky et al., 2006). Sin embargo, los que han recibido mayor atención probablemente son aquellos que están involucrados con la percepción del aroma.

La biosíntesis de las moléculas relacionadas con el aroma y sabor puede dividirse en grupos de acuerdo a su origen biogénico (Fig. 2).

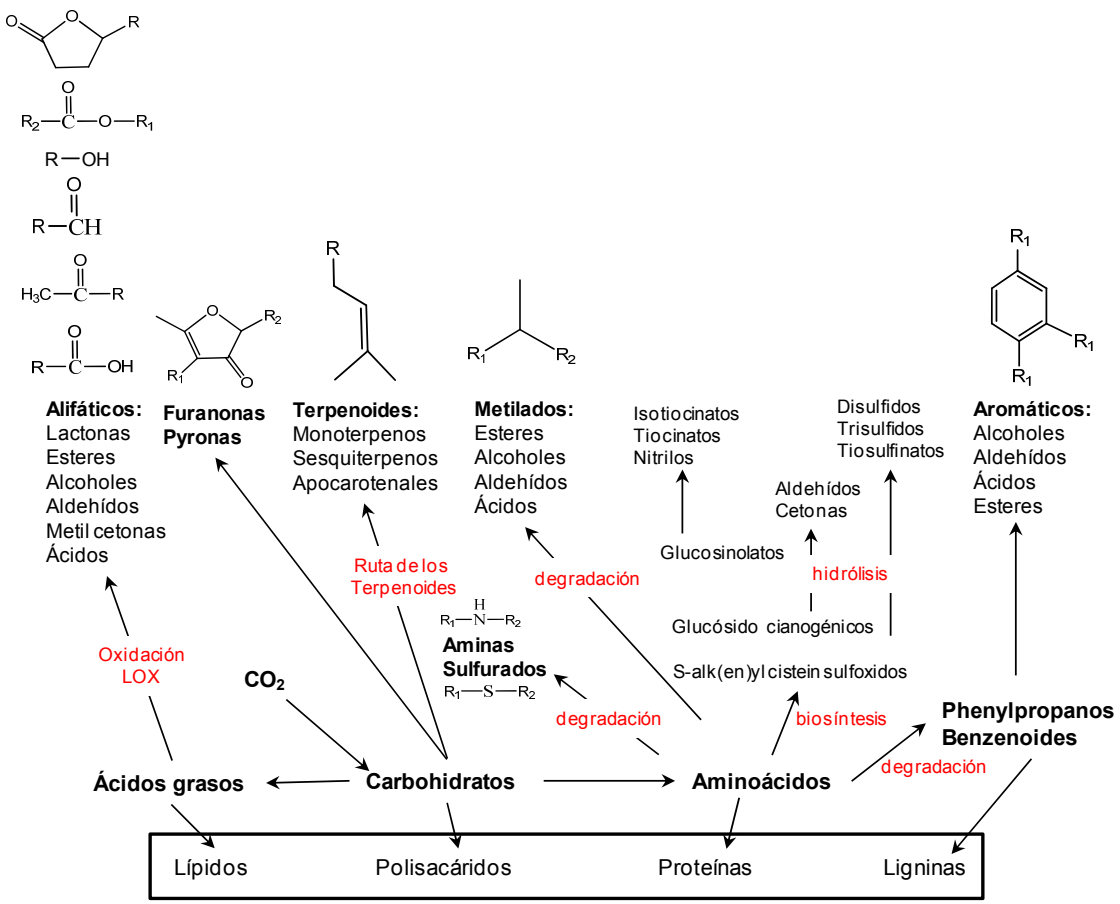



**Figura 2. Origen biogénico de los compuestos volátiles.** Adaptado de Schwab et al. 2008. LOX, ruta de las lipoxigenasas.

**A. Compuestos derivados de los ácidos grasos**. Los volátiles que se forman a partir de los ácidos grasos en las plantas provienen de tres procesos degradativos: la ruta de las lipoxigenasas, la $\alpha$-oxidación y $\beta$-oxidación (Schwab et al., 2008). Los ácidos grasos se almacenan como triacilgliceridos en las

plantas. Si bien, se ha demostrado que las cadenas de ácidos grasos pueden ser oxidadas cuando están formando parte de los acilgliceridos (Buseman et al., 2006), es probable que la mayor proporción de ácidos grasos que entran en estas rutas degradativas lo hagan en su forma libre, y por tanto deban ser liberados previamente por la acción de enzimas acil hidrolasas (Schwab et al., 2008).

**Ruta de las lipoxigenasas:** De esta ruta derivan varios aldehídos y alcoholes de 6 y 9 carbonos (C6 y C9) tanto insaturados como saturados. Algunos de estos compuestos contribuyen al aroma de varias frutas, vegetales y hojas verdes otorgando notas descritas como "verde", "fresco" o "hierba cortada" (Derail et al., 1999). Debido a que estos compuestos además son producidos en la planta en respuesta a daño mecánico o al ataque por patógenos (Matsui, 2006), las rutas que llevan a su síntesis han recibido mucha atención. Cuatro tipos de enzimas principales están involucradas en esta ruta: lipoxigenasas (LOX), hidroperoxido liasas (HPL), 3*Z*,2*E*-Enal isomerasas (3Z,2E-EI) y alcohol dehidrogenasas (ADH). La ruta comienza con la oxigenación de ácidos grasos insaturados, como por ejemplo el ácido linolénico, por acción de la LOX (Fig. 3). Diferentes LOXs catalizan la oxigenación en el carbono 9 (9-LOX) y 13 (13-LOX), para formar 9-hidroperoxidos y 13-hydroperoxidos, respectivamente. Los productos generados por las LOXs son subsecuentemente rotos por la HPL para formar ω-oxoácidos y aldehídos volátiles de C6 y C9. Estos aldehídos pueden ser convertidos a alcoholes por ADH o sufrir isomerizaciones previamente y luego ser convertidos a los correspondientes alcoholes. Aunque estos compuestos pueden tener impacto directamente sobre el aroma, la presencia de, Alcohol Acyl Transferasas (AATs) puede llevar a la esterificacion de los alcoholes con moléculas acyl-CoA para formar diversos esteres (D´ Auria, 2006).

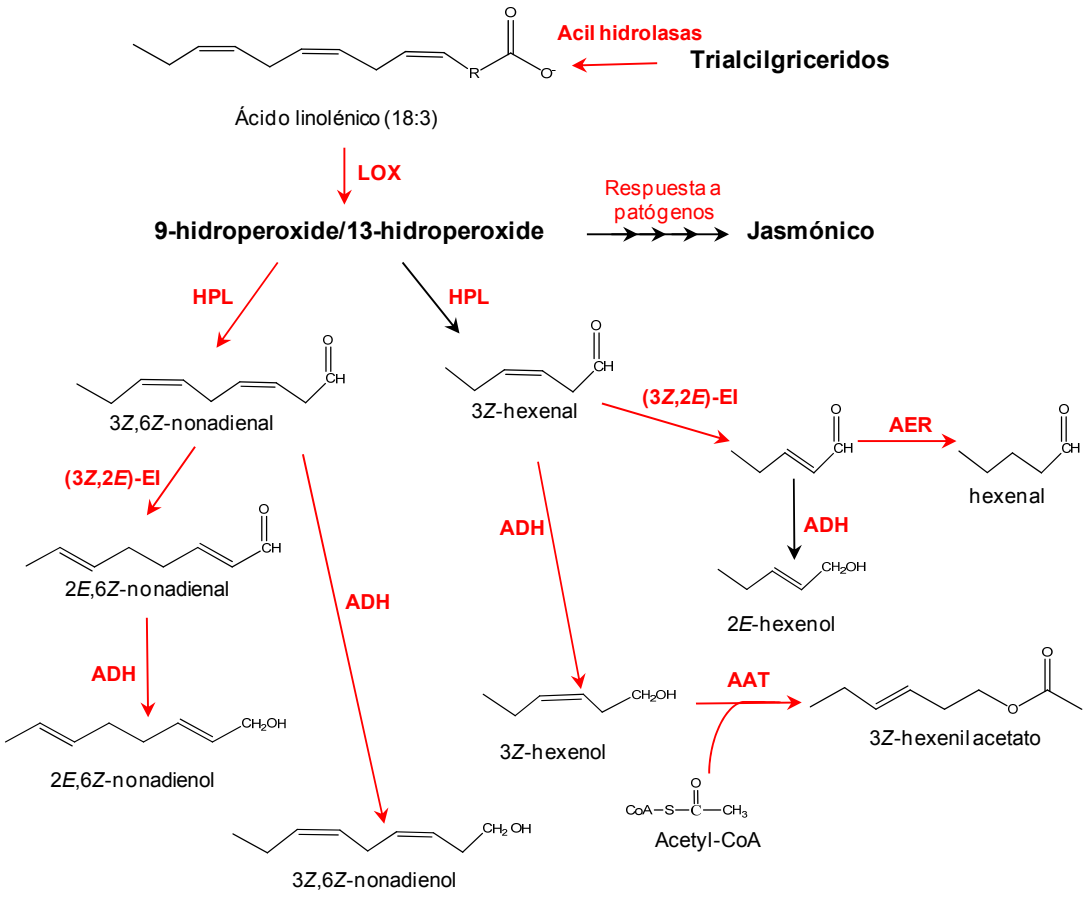



**Figura 3. Aromas derivados del catabolismo del ácido linolénico por la ruta de la lipoxigenasas.** LOX, lipoxigenasa; HPL, hidroperoxido liasa; 3Z,2E-EI, 3Z,2E-Enal isomerasa; ADH, alcohol deshidrogenasa. Modificado de Schwab et al. 2008.

***α*- y β-oxidación:** A pesar que el metabolismo de ácidos grasos mediante *α*- y β-oxidación es probablemente una vía importante para la formación de volátiles en diversos organismos, las rutas específicas en el caso de las plantas no se conocen en detalle. En la *α*-oxidación, los ácido grasos libres (C12-C18) son acortados en un carbono para formar aldehídos de cadena larga [C(N-1)] y $CO_2$ (Schwab et al., 2008). La ruta de la β-oxidación por otra parte degrada ácidos grasos activados con coenzima A (acil-CoA) mediante sucesivos pasos de remoción de unidades de dos carbonos. Como resultado se generan diversos acil-CoA de cadena corta y mediana (C6-C12) que pueden subsecuentemente ser sustratos de acil-CoA hidrolasas para formar ácidos carboxílicos libres, los

cuales alternativamente pueden ser reducidos para formar aldehídos y alcoholes (Fig. 4). Como se ha indicado anteriormente, la producción de esteres en las plantas depende del suministro de acil-CoA formados durante la β-oxidación y la presencia de diferentes alcoholes. Ambas moléculas son condensadas por la acción de AATs para formar la diversidad de esteres que muestran las plantas. Los esteres contribuyen al aroma de prácticamente todos los frutos y varias flores, por esta razón las AAT han recibido mucha atención caracterizándose molecular y bioquímicamente un gran número de ellas (D´ Auria, 2006).

En cambio, las lactonas, que químicamente son esteres cíclicos intramoleculares, parecen no estar tan ampliamente distribuidas como los esteres lineales, habiéndose identificado además de en melocotón en albaricoque, piña, coco (El Hadi et al., 2013; Guichard and Souty, 1988) y en algunos cultivares de fresa y manzano (Larsen et al., 1992; Lo Scalzo et al., 2001). Es poco lo que se conoce acerca de las rutas de biosíntesis de lactonas en plantas. Es generalmente aceptado que las lactonas provienen del catabolismo de ácidos grasos insaturados a través de la ruta de β-oxidación (Schöttler and Boland, 1996). En algún punto de la ruta de biosíntesis se introduciría un grupo hidroxilo a la molécula de acido graso, lo cual hipotéticamente podría ocurrir por: 1) reducción de oxo ácidos por NAD-reductasas, 2) hidratación del ácido graso insaturado, 3) epoxidación del ácido graso insaturado y posterior hidrólisis del grupo epóxido o 4) reducción de hidroperoxidos (Schöttler and Boland, 1996). Posteriormente, como resultado del catabolismo del hidroxi acido producirían diferentes 4- y 5-hydroxi acil-CoA que luego de un paso de ciclación, del cual no sé conoce si requiere de catálisis enzimática, formaría las γ- y δ-lactonas correspondientes (Fig. 4). El único trabajo que abordó la síntesis de lactonas en melocotón y fresa data de más de 15 años antigüedad (Schöttler and Boland,

1996). En este estudio, se infiltró en frutos un epóxido artificial marcado radiactivamente, mostrando que éste era transformado a lactonas. Mediante el análisis del marcaje de las lactonas producidas se concluyó que éstas derivan de la acción de epoxido hidrolasas endógenas y por extensión se propuso que *in vivo* la síntesis de lactonas en los frutos ocurre por esta vía (Schöttler and Boland, 1996). Sin embargo, no puede descartarse la posibilidad de que esto se deba a un artefacto experimental, siendo las rutas *in vivo* diferentes a lo que ocurre al infiltrar los frutos. Basándose en esta hipótesis varios estudios se han dirigido hacia estas enzimas, incluso se ha reportado el clonado de epóxido hidrolasas de melocotón (Vecchietti et al., 2009), sin embargo su caracterización bioquímica no se ha presentado hasta el momento.

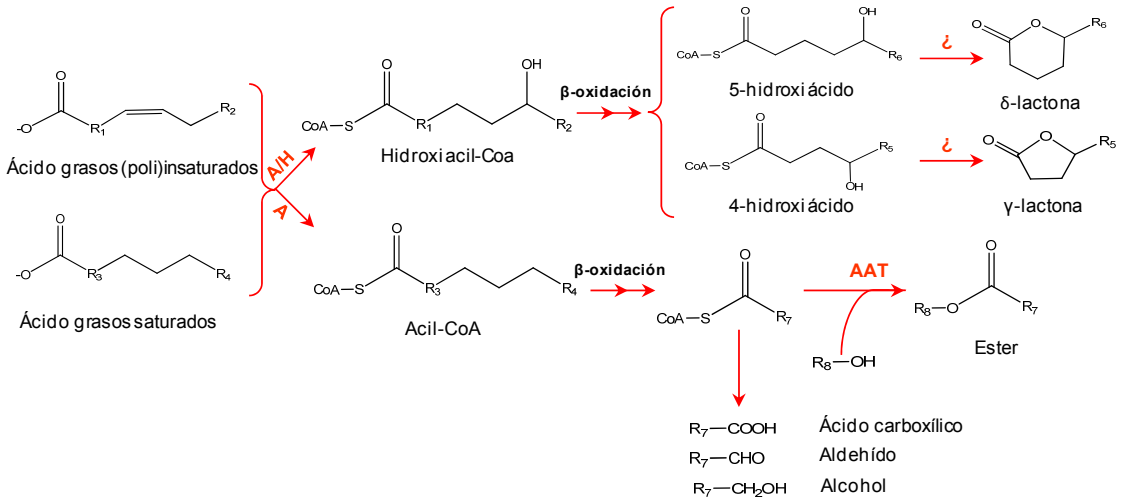



**Figura 4. Aromas derivados de la β-oxidación de ácidos grasos.** Los ácidos grasos polinsaturados, insaturados y saturados deben activarse con CoA y en el caso de los precursores de las lactonas deben además incorporar un grupo hidroxilo, lo cual se indica con A/H (A, activación con CoA; H, hidroxilación) en la figura. AAT, Alcohol Acil Transferasa. El signo de interrogación (¿) indica que no se conoce si este paso requiere actividad enzimatica u ocurre espontaneamente. Modificado de Schwab et al. 2008.

**B. Compuestos derivados de los carbohidratos**.

**Furanonas y pironas:** Estos compuestos provienen directamente del esqueleto de carbono de los carbohidratos sin ser degradados previamente y han sido

identificados en pocas especies de plantas incluidas el tomate, la fresa y piña (Schwab et al., 2008). En una cribado de los compuestos volátiles de 50 genotipos de melocotón y nectarinas, solo se ha reportado una furanona (Wang et al., 2009), sugiriendo que este tipo de compuestos también serían minoritarios en melocotón.

**Terpenoides:** Los compuestos terpenoides son sintetizado *de novo* a partir del *pool* de acetil-CoA y piruvato provenientes del catabolismo de carbohidratos. Debido a que la β-oxidación de ácidos grasos ocurre en los peroxisomas probablemente esta ruta no provea con acetil-CoA a la síntesis de los terpenoides. Los terpenoides que son volátiles posen C5 (hemiterpenos), C10 (monoterpenos), C15 (sesquiterpenos) y algunos C20 (diterpenos), pero en general, los de C10-15 son los que están involucrados en el aroma de las frutas y flores (Schwab et al., 2008). Los terpenoides se forman a partir de unidades de isopentenil bifosfato (IDP) y su isomero el dimetilalil difosfato (DMADP). Dos rutas paralelas de biosíntesis de estas unidades operan en el citosol (la ruta del mevalonato, MVA) y en plástidos (la ruta del metil eritritol fosfato, MEP). Si bien estas rutas se encuentran coordinadas por señales intracelulares, se considera que la ruta citosólica es responsable de la síntesis de sesquiterpenos entre otras moléculas, mientras que los monoterpenos y carotenoides junto con otros compuestos son producidos a partir de la plastídica (Schwab et al., 2008). Las terpenos sintasas (TS) son las enzimas involucradas en la síntesis de los hemiterpenos (C5), monoterpenos (10), sesquiterpenos (15) o diperpenos (20) a partir de 4 precursores comunes (Fig. 5). Las TS son una familia génica amplia donde varios de sus miembros tienen la capacidad de generar múltiples productos a partir de un único substrato (Dudareva et al., 2004). Mientras varios volátiles terpenoides son el producto directo de las TS muchos otros son

producidos por transformaciones posteriores como oxidaciones, deshidrogenaciones, acilaciones entre otras (Schwab et al., 2008).

Hasta la fecha, un único compuesto terpenoide, el linalool, ha sido implicado en el aroma del melocotón (Eduardo et al., 2010). El hecho que los terpenoides pueden representar hasta un 13% de la fracción volátil del fruto (Wang et al., 2009), sugiere que tendrían alguna función adicional.

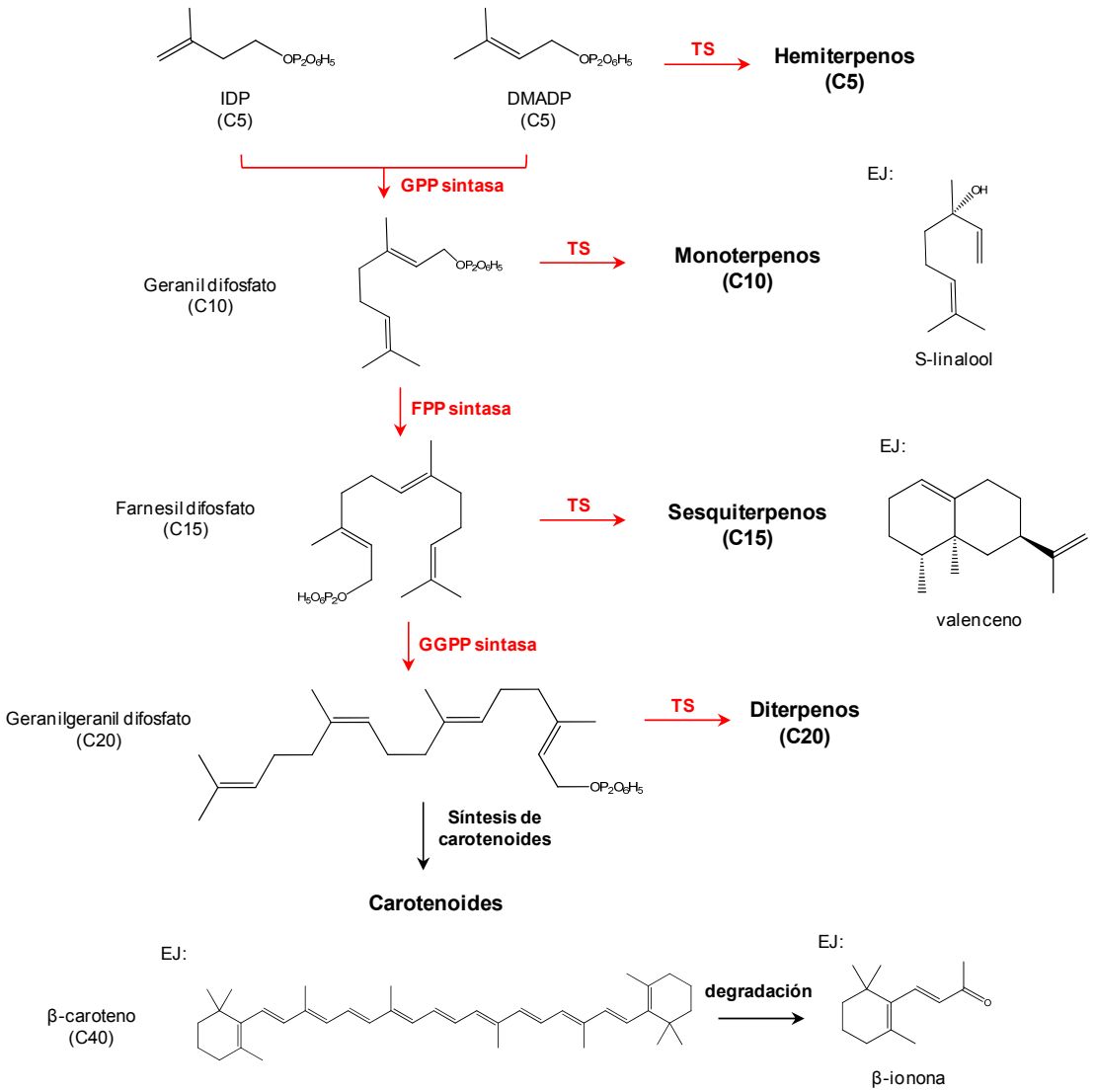



**Figura 5. Biosíntesis de aromas derivados de la ruta de los terpenoides.** Para algunos de los grupos de compuestos se muestra un compuesto volátil como ejemplo. IDP, isopentenil bifosfato; DMADP, dimetilalil difosfato; TS, Terpenoid sintasa.; GPP, Geranil difosfato sintasa; FPP, Farnesil difosfato; GGPP, Geranilgeranil difosfato sintasa. Modificado de Dubey et al. (2003).

**Apocarotenoides:** Los carotenoides son tetraterpenoides que se acumulan en plástidos de diferentes tejidos incluidos los frutos. Además de su función como pigmentos accesorios de la fotosíntesis, los carotenoides son los precursores de un amplio grupo de compuestos, los apocarotenoides (también conocidos como norisoprenoides), dentro de los cuales hay volátiles involucrados en el aroma de los frutos (Fig. 5). En el caso del melocotón, la β-ionona ha sido relacionada al aroma del fruto, estando presente en altas concentraciones en algunos genotipos pero ausentes en otros (Eduardo et al., 2010). Enzimas capaces de romper los carotenoides, como las CCDs (del inglés "*Carotenoid Cleavage Dioxygenases*"), están involucradas en la formación de los apocarotenoides. En el caso de las CCDs, se ha demostrado que exhiben cierto grado de promiscuidad de sustrato, lo cual probablemente contribuiría a la gran diversidad de apocarotenoides encontrados en la naturaleza (Schwab et al., 2008).

**C. Compuestos derivados de aminoácidos.** De la degradación de aminoácidos ramificados y aromáticos se generan alcoholes, aldehídos y ácidos carboxílicos, muchos de los cuales son abundantes en algunas especies de plantas que generan fruto como la fresa, el tomate y la uva (Aubert et al., 2005). Además, los alcoholes y ácidos carboxílicos generados pueden a su vez sufrir reacciones subsecuentes para la síntesis de otros compuestos que contribuyen en gran medida al aroma como lo son los esteres con cadenas laterales ramificados (Fig 6A). Por ejemplo, el 3-metilbutil acetato y el 3-metilbutil butanoato derivan de la degradación de la leucina y son los principales aromas del plátano (Nogueira et al., 2003). La formación de aldehídos a partir de amino ácidos requiere de la eliminación de tanto el grupo amino como el carboxilo. Se ha propuesto tres rutas alternativas para la perdida de estos grupos funcionales (Fig. 6B). En

melón, la formación de benzenacetaldehído involucra la perdida del grupo amino por acción de una transaminasa y posteriormente la decarboxilación forma el compuesto volátil (Gonda et al., 2010). En cambio, se ha demostrado que en tomate, la eliminación de los grupos funcionales sigue el orden inverso, primero se elimina el grupo carboxilo y luego el grupo amino por un segundo paso de deaminación (Tieman et al., 2006). Alternativamente, en rosa y petunia se ha identificado una aldehído sintasa capaz de catalizar una reacción combinada de decarboxilacion y deaminación generando el aldehído directamente desde el aminoácido en un único paso (Kaminaga et al., 2006).

Los volátiles benzoicos y fenilpropanoides también derivan del aminoácido fenilalanina (Phe). Los pasos iniciales de la formación del esqueleto de carbono de estos compuestos todavía no son conocidos. En general, la generación de compuestos benzoicos derivados de la Phe requiere el acortamiento del esqueleto en C2 lo cual podría ocurrir por y β-oxidación o por procesos no oxidativos (Boatright et al., 2004). Tanto los volátiles benzoicos como los fenilpropanoides sufren metilaciones posteriores por O-metil transferasas (OMTs). Por ejemplo en *Clarkia breweri*, se ha demostrado que una OMT es capaz de metilar los fenil propenos: eugenol e isoeugenol (Wang et al., 1997).

Los compuestos volátiles sulfurados se originan de la degradación de la metionina y la cisteína y algunos de ellos son los responsables del aroma característico de especies concretas de plantas como el ajo y la cebolla (Jones et al., 2004).

En melocotón se han reportado compuestos volátiles derivados de la Leu y Phe, pero no de AA sulfurados (Eduardo et al., 2010; Wang et al., 2009).

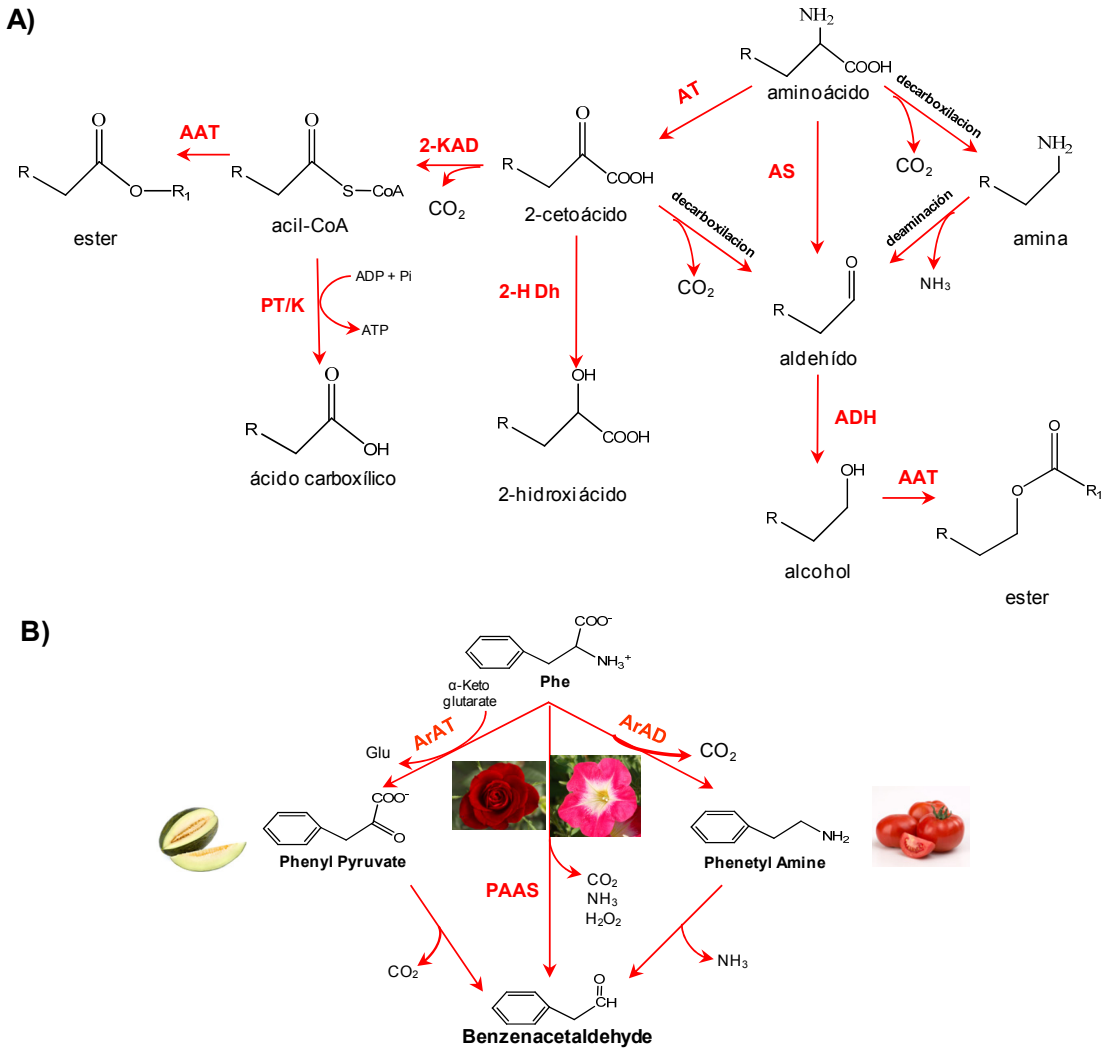



**Figura 6. Biosíntesis de aromas a partir de aminoácidos.** A) Vía general de producción de volátiles a partir de aminoácidos en plantas. B) Ruta biosintética del benzenacetaldehído en melón (izquierda), rosa y petunia (centro) y tomate (derecha). AS, Aldehído sintasa; ADH, Alcohol deshidrogenasa; AAT, Alcohol acil transferasa; AT, Aminotransferasa; 2-H Dh, 2-hidroácido deshidrogenasa; 2-KAD, 2-cetoácido deshidrogenasa; PT/K, Fosfotransferasa/quinasa. ArAAT, aminoácido aromático aminotransferasa; PAAS, Fenil acetaldehído sintasa; ArAD, aminoácido aromático decarboxilasa. Las figuras A) y B) fueron adaptado de Schwab et al. (2008) y Gonda et al. (2010), respectivamente.

## 1.6 Control genético del aroma del melocotón

La importancia capital que tienen los volátiles a la hora de definir la calidad del melocotón a través de sus aromas y sabores, convierte la identificación de genes

involucrados en la producción de estos compuestos un campo de investigación muy tentador. El descubrimiento de genes que afecten la síntesis de volátiles, y su posterior búsqueda de fuentes de variación natural, podría ser un atajo valioso para ser utilizado en programas de mejora. En este sentido se han reportado varios estudios que han seguido una aproximación experimental de genes candidatos derivados de la bibliografía. Esto frecuentemente se lleva a cabo analizando la implicación de genes que han sido caracterizados bioquímica y/o molecularmente en otras especies de plantas, en ciertos procesos fisiológicos del melocotón mediante el análisis de la expresión de genes. Cuando el conocimiento acerca del proceso estudiado es fragmentario como es el caso de la producción de aromas en melocotón, estos estudios tienden a sesgarse hacia el análisis de genes estructurales de las rutas biosintéticas, que en el caso del melocotón son incluso poco conocidas. Debido a que los compuestos con más impacto positivo en el aroma del melocotón son las lactonas y esteres mientras que los compuestos derivados de lípidos tendrían un impacto negativo por estar asociado a la percepción de inmadurez, las enzimas que llevan a la síntesis de estos compuestos son los que recibieron mayor atención. Aunque no se ha descrito la ruta de síntesis de las lactonas en plantas, se ha sugerido que estos compuestos derivan del *pool* de ácidos grasos (Schwab et al., 2008), por lo que las enzimas probablemente involucradas en la síntesis de estos posibles precursores han sido objeto de estudio. Zhang et al. (2010), clonaron genes homólogos de LOX (*PpLOX1-4*), HPL (*PpHPL1*), ADH (*PpADH1-3*), AAT (*PpAAT1*) y FAD (*PpFAD1-4*) y analizaron por PCR cuantitativa (qPCR) la expresión génica durante la maduración del fruto. Comparando los perfiles de acumulación tanto de volátiles (lactonas, esteres y compuestos derivados de lípidos) y ácidos grasos con el patrón de expresión de los genes analizados, propusieron la implicancia de alguno de los genes clonados en la síntesis de los

aromas (Zhang et al., 2010). Estos autores mostraron que una ω-6 desaturasa de ácidos grasos de retículo endoplasmático (*PpFAD1*) y una ω-3 desaturasa de ácidos grasos de plástidos (*PpFAD2*) se asocian a la acumulación de ácido linoleico (18:2) y linolenico (18:3), sugiriendo un rol en la síntesis de precursores para la ruta de la LOX (Zhang et al., 2010). Por otro lado, los mismos autores demostraron que *PpLOX1-2*, *PpHPL1* y *PpADH1-3* disminuyen durante la maduración en concordancia con los compuestos derivados de lípidos, y por tanto propusieron su posible implicancia de estos genes en relación a las notas de inmadurez que confieren estos compuestos (Zhang et al., 2010). Este mismo grupo de investigación siguió una aproximación experimental análoga donde los patrones de expresión de genes candidatos y síntesis de volátiles fue analizada en frutos sujetos a conservación por bajas temperaturas y su subsecuente maduración en condiciones poscosecha (Zhang et al., 2011). Sus resultados sugieren que la recuperación de los niveles de volátiles, principalmente esteres y lactonas, durante la maduración se debe al aumento de expresión de dos LOX: *PpLOX1,* y *2* y una AAT: *PpAAT1* (Zhang et al., 2011). Basándose en la hipótesis que postula que la síntesis de lactonas en melocotón involucra la β-oxidación de ácidos grasos (Schwab et al., 2008). Este mismo grupo se centró en analizar el rol de la primer enzima de esta ruta, la acyl-CoA oxidasa (ACX), en la producción de lactonas durante tratamientos poscosecha (Xi et al., 2012). Nuevamente, analizando el aumento de la producción de lactonas junto con la actividad ACX y la expresión de diferentes genes *PpACXs*, propusieron que la producción de las lactonas durante la maduración poscosecha está regulada mediante la expresión de *PpACX1* (Xi et al., 2012). Otros estudio han propuesto una serie de genes candidatos para la síntesis de aromas de melocotón a partir del análisis de bibliotecas de ESTs. Dado que se había demostrado que la piel y pulpa tienen diferentes concentraciones de

volátiles en el melocotón (Aubert and Milhet, 2007) y que los compuestos que más contribuyen al aroma se encuentran en la piel, la distribución de EST en estos dos tejidos fue analizada para proponer una serie de genes candidatos (Vecchietti et al., 2009). Este estudio también centró sus análisis en genes de la ruta biosintéticas propuestas anteriormente, identificando homólogos a: LOX, HPL, ADH, AAT, Epoxido hidrolasas (EPHs), Pyr decarboxilasa y O-Metil transferasas (Vecchietti et al., 2009). Este mismo grupo de investigación desarrolló un microarray basado en EST (μPeach2.0) enriquecido en genes involucrados en el metabolismo secundario para analizar la expresión de genes involucrados en la síntesis de volátiles (Pirona et al., 2012). Estudiando como varía la expresión de genes durante la maduración de dos genotipos de melocotón que contrastan en sus características aromáticas se identificó un conjunto de 12 genes relacionados al aroma que se expresan diferencialmente en ambos genotipos y por tanto fueron propuestos como posibles candidatos. Los genes identificados corresponden a enzimas relacionadas a la síntesis de esteres y lactonas (LOXs, AATs y EPHs) y a la ruta de los de los phenylpropanoides (4-coumarato CoA ligasa e hidrolasas del ácido cinámico) y apocarotenoides (CCDs, PSYs y ZDSs) (Pirona et al., 2012). La implicancia de genes en la formación de los apocarotenoides en melocotón también fue estudiada mediante comparación de un genotipo y su mutante deficiente en carotenoides (Brandi et al., 2011). Comparando la evolución durante la madurez de la acumulación de carotenoides, sus volátiles derivados (apocarotenoides) y la expresión de genes de la ruta biosintética de los carotenoides, así como también de enzimas degradativas, estos autores han propuesto que el gen *ccd4* es clave para la formación de los apocarotenoides en melocotón (Brandi et al., 2011).

Una estrategia complementaria para el análisis del control genético de la producción de volátiles, es la identificación de QTL en poblaciones segregantes. Un estudio reciente describió QTL para volátiles relacionados con el aroma del melocotón (Eduardo et al., 2013). En este estudio, analizaron el control genético para 23 compuestos volátiles, la mayoría implicados en el aroma del fruto, identificando QTL para diferentes compuestos incluidos lactonas, esteres, aromáticos y terpenenoides. La mayoría de los QTL detectados fueron menores (es decir que explican un bajo porcentaje de la variabilidad fenotípica) sin embargo un locus con efectos mayores que controla la producción de dos monoterpenos fue localizado en el brazo superior del cromosoma 4 (Eduardo et al., 2013). Analizando la variabilidad alélica, demostraron que dos terpeno sintasas localizadas en esa región cromosómica co-segregan con la variación fenotípica de estos compuestos sugiriendo una posible implicación de estos genes en la producción de monoterpenos (Eduardo et al., 2013).

Nuestra hipótesis es que si bien, se dispone de un gran cúmulo de información que asocia algunos genes estructurales con la producción de volátiles, la utilización de aproximaciones del tipo no-dirigida podría además de confirmar algunos de esos candidatos propuestos, identificar nuevos genes involucrados en la formación de aromas del melocotón.

## 1.7 El melocotonero como un nuevo modelo de fruto

La familia de las *Rosáceas* es el clado más importante entre las plantas que producen frutos. Los principales cultivos productores de fruto de las zonas templadas como el manzano (*Malus* x *domestica*), cerezo (*Prunus avium* y *Prunus cerasus*), ciruelo (*Prunus domestica* y *Prunus salicina*), albaricoque (*Prunus armeniaca*), almendro (*Prunus dulcis*), peral (*Pyrus communis*), níspero (*Eriobotrya*

*japonica*) y la fresa (*Fragaria* x *ananassa*) pertenecen al igual que el melocotón (*Prunus persica*) a las *Rosaceas*. Esta familia también incluye a especies ornamentales como la rosa (*Rosa* spp.) y hasta especies productoras de madera como el cerezo negro (*Prunus serotina*). Por tanto en esta familia están representados varios tipos de frutos (drupas, pomas, aquenios, cápsulas, etc) como hábitos de crecimiento (desde herbáceos hasta forestales) lo que justifica el uso de múltiples modelos experimentales para el estudio de las *Rosaceas* (Shulaev et al., 2008). El melocotón ha sido considerado desde hace tiempo como la especie modelo del genero *Prunus* debido a una serie de características. Entre ellas, tiene un genoma pequeño, de aproximadamente unas 230 Mb lo que representa menos doble del de *Arabidopsis thaliana* (157 Mb). Al contrario de la mayoría de las especies de este género que son auto-incompatible, el melocotón es una especie auto-compatible lo que permite la obtención de poblaciones F2. Además se dispone de plantas doble haploides homocigóticas. Otra rasgo ventajoso es que el melocotón posee un periodo juvenil relativamente corto (de 2 a 3 años) si se lo compara con otras especies que requieren entre 5-10 años para producir frutos. Por esta última razón, el melocotón es la especie mejor caracterizada desde el punto de vista genético, habiéndose descrito la herencia de aproximadamente 40 genes mayores (Arús et al., 2012) y desarrollado un gran número de marcadores moleculares lo que ha permitido disponer de un mapa consenso de alta densidad y saturado (Abbott et al., 2008). Durante la pasada década se desarrollaron diversas herramientas genómicas para melocotón, lo que afianzó aún más su condición de especie modelo para los prunoides y concretamente como modelo para fruto. Se han construido diferentes microarrays como μPeach1.0 y 2.0 (ESTree Consortium, 2005; Pirona et al., 2012) y ChillPeach (Ogundiwin et al., 2008), varias base de datos de EST (revisado en Abbot et al. 2008), y un sistema de genotipado de alta capacidad

basado en marcadores SNPs (Verde et al., 2012). Aunque no han sido usados intensivamente en el análisis funcional de genes, se han descrito un protocolo para la expresión transitoria en frutos (Spolaore et al., 2001; Tittarelli et al., 2009) y un sistema de silenciamiento génico basado en el virus TRV (Jia et al., 2011). Incluso se ha logrado la transformación estable del melocotón (Pérez-Clemente et al., 2004), aunque esta herramienta todavía es poco eficiente.

La disponibilidad de la secuencia del genoma desde el año 2010 definitivamente consolidó la posición del melocotón como un nuevo modelo de fruto. La secuenciación fue llevada a cabo por el consorcio internacional IPGI (del inglés "*International Peach Genome Initiative*") quienes utilizaron el cultivar doble haploide 'Lovell' para facilitar el ensamblado de las secuencias genómicas (Verde et al., 2013). Este hecho junto con el tipo de secuenciación utilizado (SWGS, del inglés "*Sanger Whole-Genome Shotgun*") permitió la obtención de una primera versión del genoma del melocotón (Peach v1.0) de gran calidad. La disposición de un genoma de referencia de calidad, facilitó la posterior re-secuenciación de alrededor de 70 genotipos de melocotón con tecnologías de secuenciación de fragmentos cortos como Ilumina-Solexa y Roche 454 (Ahmad et al., 2011; Verde et al., 2013; Verde et al., 2012).

Hoy en día, hay diversas herramientas genéticas y genómicas que están disponibles para el análisis funcional de genes del melocotón lo que favorece en gran medida el descubrimiento de nuevos genes. Los caracteres diferenciales del melocotón, como puede ser algunos compuestos poco distribuidos en la naturaleza, convierte a esta especie en un modelo propicio para la identificación de genes noveles.

## 1.8 Antecedentes y estructuración del presente trabajo

En este trabajo se planteó analizar el control genético de la producción de aromas en el fruto mediante aproximaciones experimentales complementarias con el objetivo de generar información útil para la mejora de la calidad del melocotón.

Disponíamos de una población de melocotón desarrollada dentro del programa de mejora del IVIA para el análisis de QTL, de modo que estructuramos nuestro trabajo en función de esta población y sus genotipos parentales (Fig. 7).

Si bien, los volátiles que contribuyen mayoritariamente al aroma del fruto son bien conocidos, nos propusimos llevar acabo un análisis del tipo no dirigido donde se identificara y cuantificara la mayor proporción posible del volatiloma del melocotón. De este modo, no solo nos proporcionaría una visión global de la fracción volátil sino también la información generada podría ser útil en otros procesos fisiológicos que involucren compuestos volátiles. En un primer paso, desarrollamos una plataforma metabolómica de alta capacidad basada en HS-SPME-CG-MS (del inglés "*HeadSpace-Solid Phase MicroExtraction-Gas Chromatography-Mass Spectroscopy*") con la que identificar y cuantificar los compuestos volátiles del fruto y que será presentada en el capitulo 2. Con el objetivo de cubrir la mayor parte del volatiloma, hemos utilizado un grupo de muestras que contienen una gran variabilidad, incluyendo cuatro genotipos relacionados con la población de mapeo, diferentes estadios de madurez, un tratamiento poscosecha y dos locaciones geográficas que aportan efecto ambiental. El análisis de esta variabilidad por métodos basados en correlación nos permitió describir los patrones de co-regulación a la que están sujetos los volátiles (capitulo 2). Sobre un subconjunto de estas muestras, que corresponde a series de maduración de los genotipos parentales ('MxR_01' y 'Granada') así como también el tratamiento poscosecha, fue analizada la expresión de genes

mediante microarray con objeto de identificar genes candidatos mediante el análisis de co-regulación entre volátiles y genes (capitulo 3). En este caso, se llevo a cabo un análisis sin hipótesis previa, es decir se investigó la relación entre volátiles y todos los genes analizados sin centrarse en aquellos que hayan sido propuestos como candidatos en la bibliografía. Por último, el análisis de la población de mapeo con la plataforma metabolómica desarrollada y una plataforma de genotipado masivo nos permitió la identificación de QTL para volátiles que se presenta en el capitulo 4.

Debido a que nuestro interés se centra en la mejora de la calidad del melocotón, los resultados de los análisis (del tipo no dirigidos y sin hipótesis previa) fueron interpretados en función de las posibles implicaciones en la producción de compuestos volátiles asociados al aroma del fruto.

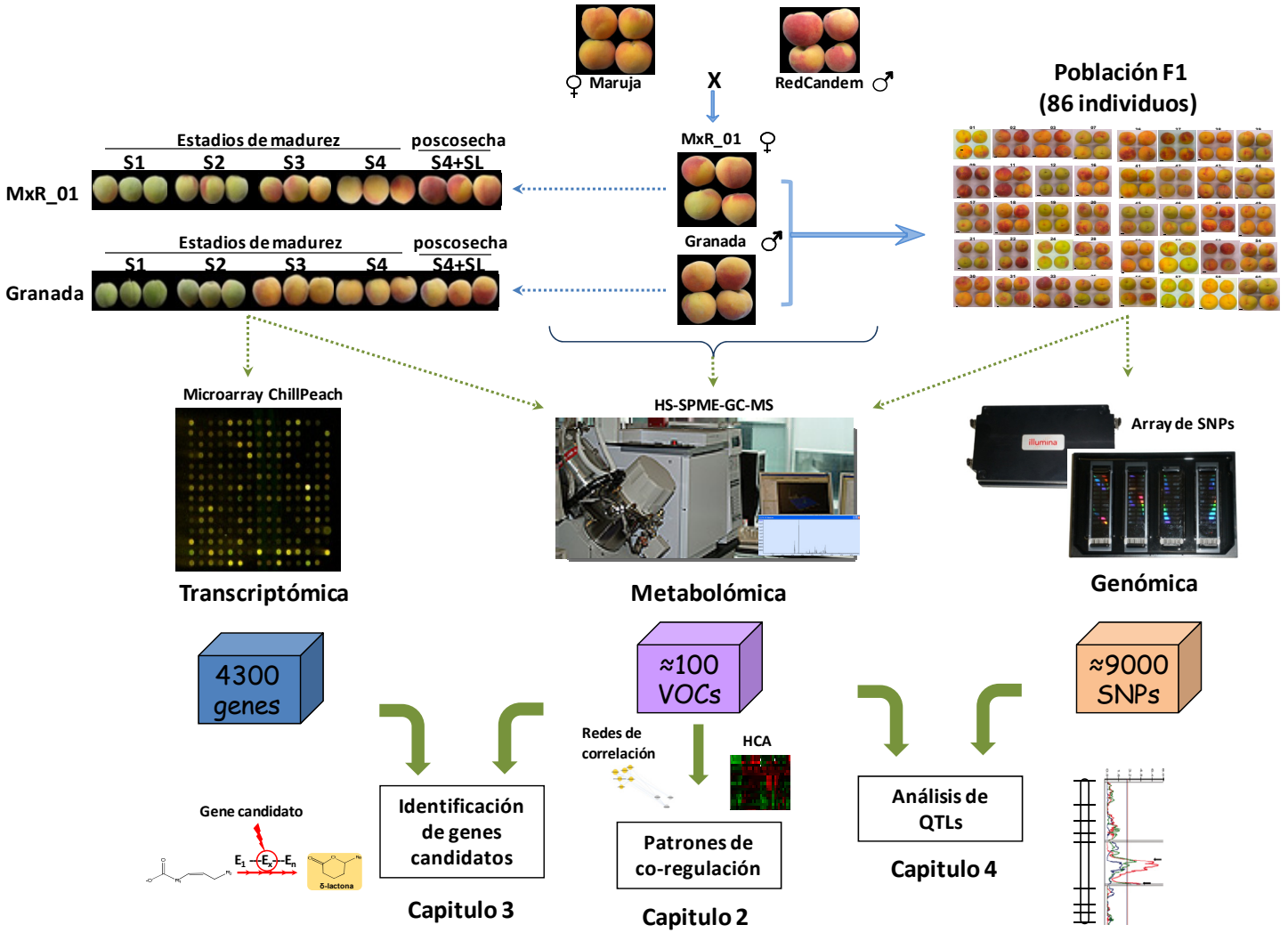



**Figura 7. Organización y diseño de la tesis.** Se muestra un resumen de las muestras analizadas, las tecnologías utilizadas así como también los set de datos obtenidos. El análisis de estos set de datos se presenta en los capítulos indicados.

## 1.9 Referencias


Abbott, A.G., Arús, P. and Scorza, R. (2008) Chapter 4: Genetic Engineering and Genomics. *The Peach: Botany, Production and Uses*, 85.

Ahmad, R., Parfitt, D., Fass, J., Ogundiwin, E., Dhingra, A., Gradziel, T., Lin, D., Joshi, N., Martinez-Garcia, P. and Crisosto, C. (2011) Whole genome sequencing of peach (*Prunus persica* L.) for SNP identification and selection. *BMC Genomics*, **12**, 569.

Arús, P., Verde, I., Sosinski, B., Zhebentyayeva, T. and Abbott, A. (2012) The peach genome. *Tree Genetics & Genomes*, **8**, 531-547.

Aubert, C., Baumann, S. and Arguel, H. (2005) Optimization of the Analysis of Flavor Volatile Compounds by Liquid-Liquid Microextraction (LLME). Application to the Aroma Analysis of Melons, Peaches, Grapes, Strawberries, and Tomatoes. *Journal of Agricultural and Food Chemistry*, **53**, 8881-8895.

Aubert, C., Gunata, Z., Ambid, C. and Baumes, R. (2003) Changes in Physicochemical Characteristics and Volatile Constituents of Yellow- and White-Fleshed Nectarines during Maturation and Artificial Ripening. *J. Agric. and Food Chem.*, **51**, 3083-3091.

Aubert, C. and Milhet, C. (2007) Distribution of the volatile compounds in the different parts of a white-fleshed peach (*Prunus persica* L. Batsch). *Food Chem.*, **102**, 375-384.

Bicchi, C. and Maffei, M. (2012) The plant volatilome: methods of analysis. In *High-Throughput Phenotyping in Plants*. Springer, pp. 289-310.

Boatright, J., Negre, F., Chen, X., Kish, C.M., Wood, B., Peel, G., Orlova, I., Gang, D., Rhodes, D. and Dudareva, N. (2004) Understanding in Vivo Benzenoid Metabolism in Petunia Petal Tissue. *Plant Physiology*, **135**, 1993-2011.

Brandi, F., Bar, E., Mourgues, F., Horvath, G., Turcsi, E., Giuliano, G., Liverani, A., Tartarini, S., Lewinsohn, E. and Rosati, C. (2011) Study of 'Redhaven' peach and its white-fleshed mutant suggests a key role of CCD4 carotenoid dioxygenase in carotenoid and norisoprenoid volatile metabolism. *BMC Plant Biology*, **11**, 24.

Bruhn, C.M. (1995) Consumer and retailer satisfaction with the quality and size of california peaches and nectarines. *Journal of Food Quality*, **18**, 241-256.

Bruhn, C.M., Feldman, N., Garlitz, C., Harwood, J., Ivans, E., Marshall, M., Riley, A., Thurber, D. and Williamson, E. (1991) Consumer Perceptions of quality: Apricots, Cantaloupes, Peaches, Pears, Strawberries, and Tomatoes. *Journal of Food Quality*, **14**, 187-195.

Buseman, C.M., Tamura, P., Sparks, A.A., Baughman, E.J., Maatta, S., Zhao, J., Roth, M.R., Esch, S.W., Shah, J. and Williams, T.D. (2006) Wounding stimulates the accumulation of glycerolipids containing oxophytodienoic acid and dinor-oxophytodienoic acid in Arabidopsis leaves. *Plant Physiology*, **142**, 28-39.

Collard, B.C.Y., Jahufer, M.Z.Z., Brouwer, J.B. and Pang, E.C.K. (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica*, **142**, 169-196.

Crisosto, C.H. (2006) Short-term approaches to increase peach fruit consumption. *Compact Fruit Tree*, **39**, 11-14.

Chapman, G.W., Horvat, R.J. and Forbus, W.R. (1991) Physical and chemical changes during the maturation of peaches (cv. Majestic). *Journal of Agricultural and Food Chemistry*, **39**, 867-870.

Christensen, C.M. (1983) Effects of color on aroma, flavor and texture judgments of foods. *Journal of Food Science*, **48**, 787-790.

D´ Auria, J.C. (2006) Acyltransferases in plants: a good time to be BAHD. *Current Opinion in Plant Biology*, **9**, 331-340.

Daorden, M.E. (2012) Comentarios generales sobre el mejoramiento genético en duraznero. In Valentini, G.H., González, J. and Gordo, M. (eds.), *Producción del duraznero en la Región Pampeana, Argentina.* INTA, San Pedro, Vol. 1, pp. 62-70.

Derail, C., Hofmann, T. and Schieberle, P. (1999) Differences in Key Odorants of Handmade Juice of Yellow-Flesh Peaches (*Prunus persica* L.) Induced by the Workup Procedure. *J. Agric. and Food Chem.*, **47**, 4742-4745.

Dudareva, N., Pichersky, E. and Gershenzon, J. (2004) Biochemistry of Plant Volatiles. *Plant Physiology*, **135**, 1893-1902.

Dufosse, L., Latrasse, A. and Spinnler, H.-E. (1994) Importance of lactones in food flavours: structure, distribution, sensory properties and biosynthesis. *Sci. Aliment.*, **14**, 17-50. (*in french*).

Eduardo, I., Chietera, G., Bassi, D., Rossini, L. and Vecchietti, A. (2010) Identification of key odor volatile compounds in the essential oil of nine peach accessions. *Journal of the Science of Food and Agriculture*, **90**, 1146-1154.

Eduardo, I., Chietera, G., Pirona, R., Pacheco, I., Troggio, M., Banchi, E., Bassi, D., Rossini, L., Vecchietti, A. and Pozzi, C. (2013) Genetic dissection of aroma volatile compounds from the essential oil of peach fruit: QTL analysis and identification of candidate genes using dense SNP maps. *Tree Genetics & Genomes*, **9**, 189-204.

El Hadi, M.A.M., Zhang, F.-J., Wu, F.-F., Zhou, C.-H. and Tao, J. (2013) Advances in Fruit Aroma Volatile Research. *Molecules*, **18**, 8200-8229.

ESTree Consortium. (2005) Development of an oligo-based microarray (uPEACH 1.0) for genomics studies in peach fruit. *Acta Hortic*, **682**, 263–268.

FAO. (2010) Accessed 25 Aug. 2013. http://faostat.fao.org.

FAO. (2011) Accessed 25 Aug. 2013. http://faostat.fao.org.

Gatfield, I.L. (1999) Biotechnological production of natural flavor materials. In. Kluwer Academic / Plenum Publishing Corporation, New York, pp. 211-227.

Gonda, I., Bar, E., Portnoy, V., Lev, S., Burger, J., Schaffer, A.A., Tadmor, Y.a., Gepstein, S., Giovannoni, J.J., Katzir, N. and Lewinsohn, E. (2010) Branched-chain and aromatic amino acid catabolism into aroma volatiles in Cucumis melo L. fruit. *J. Exp. Bot.* , **61**, 1111-1123.

Gordó, M. (2012) Origen y características botánicas. In Valentini, G.H., González, J. and Gordo, M. (eds.), *Producción del duraznero en la Región Pampeana, Argentina.* INTA, San Pedro, Vol. 1, pp. 62-70.

Greger, V. and Schieberle, P. (2007) Characterization of the Key Aroma Compounds in Apricots (Prunus armeniaca) by Application of the Molecular Sensory Science Concept. *Journal of Agricultural and Food Chemistry*, **55**, 5221-5228.

Grosch, W. (2001) Evaluation of the Key Odorants of Foods by Dilution Experiments, Aroma Models and Omission. *Chemical Senses*, **26**, 533-545.

Guichard, E. and Souty, M. (1988) Comparison of the relative quantities of aroma compounds found in fresh apricot (*Prunus armeniaca*) from six different varieties. *Zeitschrift fur Lebensmittel-Untersuchung und Forschung*, **186**, 301-307.

Guillot, S., Peytavi, L., Bureau, S., Boulanger, R., Lepoutre, J.-P., Crouzet, J. and Schorr-Galindo, S. (2006) Aroma characterization of various apricot varieties using headspace-solid phase microextraction combined with gas chromatography-mass spectrometry and gas chromatography-olfactometry. *Food Chemistry*, **96**, 147-155.

Horvat, R.J., Chapman, G.W., Robertson, J.A., Meredith, F.I., Scorza, R., Callahan, A.M. and Morgens, P. (1990) Comparison of the volatile compounds from several commercial peach cultivars. *Journal of Agricultural and Food Chemistry*, **38**, 234-237.

Iglesias, I. (2010) Producción, innovación varietal y mercado del melocotón en España. *Vida Rural*, **302**, 41-53.

Iglesias, I. and Casals, E. (2011) Producción, comercialización y consumo de melocotón en España. *Vida Rural*, **323**, 27-34.

Jia, H.-F., Chai, Y.-M., Li, C.-L., Qin, L. and Shen, Y.-Y. (2011) Cloning and Characterization of the H Subunit of a Magnesium Chelatase Gene (PpCHLH) in Peach. *Journal of Plant Growth Regulation*, **30**, 445-455.

Jia, H.-J., Araki, A. and Okamoto, G. (2005) Influence of fruit bagging on aroma volatiles and skin coloration of `Hakuho´ peach (*Prunus persica* Batsch). *Postharvest Biology and Technology*, **35**, 61-68.

Jones, M.G., Hughes, J., Tregova, A., Milne, J., Tomsett, A.B. and Collin, H.A. (2004) Biosynthesis of the flavour precursors of onion and garlic. *Journal of Experimental Botany*, **55**, 1903-1918.

Kader, A.A. (1999) Fruit maturity, ripening, and quality relationships. *Acta Hort.*, **485**, 203-208.

Kaminaga, Y., Schnepp, J., Peel, G., Kish, C.M., Ben-Nissan, G., Weiss, D., Orlova, I., Lavie, O., Rhodes, D., Wood, K., Porterfield, D.M., Cooper, A.J.L., Schloss, J.V., Pichersky, E., Vainstein, A. and Dudareva, N. (2006) Plant Phenylacetaldehyde Synthase Is a Bifunctional Homotetrameric Enzyme That Catalyzes Phenylalanine Decarboxylation and Oxidation. *Journal of Biological Chemistry*, **281**, 23357-23366.

Klee, H.J. and Giovannoni, J.J. (2011) Genetics and Control of Tomato Fruit Ripening and Quality Attributes. *Annual Review of Genetics*, **45**, 41-59.

Larsen, M., Poll, L. and Olsen, C. (1992) Evaluation of the aroma composition of some strawberry (Fragaria ananassa Duch) cultivars by use of odour threshold values. *Zeitschrift fur Lebensmittel-Untersuchung und Forschung*, **195**, 536-539.

Lo Scalzo, R., Testoni, A. and Genna, A. (2001) 'Annurca' apple fruit, a southern Italy apple cultivar: textural properties and aroma composition. *Food Chemistry*, **73**, 333-343.

Martínez-García, P., Parfitt, D., Ogundiwin, E., Fass, J., Chan, H., Ahmad, R., Lurie, S., Dandekar, A., Gradziel, T. and Crisosto, C. (2013) High density SNP mapping and QTL analysis for fruit quality characteristics in peach (*Prunus persica* L.). *Tree Genetics & Genomes*, **9**, 19-36.

Matsui, K. (2006) Green leaf volatiles: hydroperoxide lyase pathway of oxylipin metabolism. *Current opinion in plant biology*, **9**, 274-280.

Nogueira, J.M.F., Fernandes, P.J.P. and Nascimento, A.M.D. (2003) Composition of volatiles of banana cultivars from Madeira Island. *Phytochemical Analysis*, **14**, 87-90.

Ogundiwin, E.A., Marti, C., Forment, J., Pons, C., Granell, A., Gradziel, T.M., Peace, C.P. and Crisosto, C.H. (2008) Development of ChillPeach genomic tools and identification of cold-responsive genes in peach fruit. *Plant Mol Biol*, **68**, 379-397.

Olbricht, K., Grafe, C., Weiss, K. and Ulrich, D. (2008) Inheritance of aroma compounds in a model population of Fragaria × ananassa Duch. *Plant Breeding*, **127**, 87-93.

Pagliaricci, L. and Ángel, A. (2012) Evolución de la actividad frutícola en el partido de San Pedro. In Valentini, G.H., González, J. and Gordo, M. (eds.), *Producción del duraznero en la Región Pampeana, Argentina.* INTA, San Pedro, Vol. 1, pp. 19-31.

Pérez-Clemente, R.M., Pérez-Sanjuán, A., García-Férriz, L., Beltrán, J.-P. and Cañas, L.A. (2004) Transgenic peach plants (*Prunus persica* L.) produced by genetic transformation of embryo sections using the green fluorescent protein (GFP) as an in vivo marker. *Molecular Breeding*, **14**, 419-427.

Pichersky, E., Noel, J.P. and Dudareva, N. (2006) Biosynthesis of plant volatiles: nature's diversity and ingenuity. *Science*, **311**, 808-811.

Pirona, R., Vecchietti, A., Lazzari, B., Caprera, A., Malinverni, R., Consolandi, C., Severgnini, M., De Bellis, G., Chietera, G. and Rossini, L. (2012) Expression profiling of genes involved in the formation of aroma in two peach genotypes. *Plant Biology*, **15**, 443-451.

Robertson, J.A., Meredith, F.I., Horvat, R.J. and Senter, S.D. (1990) Effect of cold storage and maturity on the physical and chemical characteristics and volatile constituents of peaches (cv. Cresthaven). *Journal of Agricultural and Food Chemistry*, **38**, 620-624.

Schöttler, M. and Boland, W. (1996) Biosynthesis of Dodecano-4-lactone in Ripening Fruits: Crucial Role of an Epoxide-Hydrolase in Enantioselective Generation of Aroma Components of the Nectarine (*Prunus persica* var. nucipersica) and the Strawberry (*Fragaria ananassa*). *Helvetica Chimica Acta*, **79**, 1488-1496.

Schwab, W., Davidovich-Rikanati, R. and Lewinsohn, E. (2008) Biosynthesis of plant-derived flavor compounds. *The Plant Journal*, **54**, 712-732.

Shepherd, G.M. (2006) Smell images and the flavour system in the human brain. *Nature*, **444**, 316-321.

Shulaev, V., Korban, S.S., Sosinski, B., Abbott, A.G., Aldwinckle, H.S., Folta, K.M., Iezzoni, A., Main, D., Arús, P., Dandekar, A.M., Lewers, K., Brown, S.K., Davis, T.M., Gardiner, S.E., Potter, D. and Veilleux, R.E. (2008) Multiple Models for Rosaceae Genomics. *Plant Physiology*, **147**, 985-1003.

Spolaore, S., Trainotti, L. and Casadoro, G. (2001) A simple protocol for transient gene expression in ripe fleshy fruit mediated by *Agrobacterium*. *Journal of Experimental Botany*, **52**, 845-850.

Stommel, J., Abbott, J.A., Saftner, R.A. and Camp, M.J. (2005) Sensory and objective quality attributes of beta-carotene and lycopene-rich tomato fruit. *Journal of the American Society for Horticultural Science*, **130**, 244-251.

Sumitani, H., Suekane, S., Nakatani, A. and Tatsuka, K. (1994) Changes In Composition of Volatile Compounds in High Pressure Treated Peach. *Journal of Agricultural and Food Chemistry*, **42**, 785-790.

Takhtajan, A. (1997) *Diversity and classification of flowering plants*. Columbia University Press.

Tieman, D., Taylor, M., Schauer, N., Fernie, A.R., Hanson, A.D. and Klee, H.J. (2006) Tomato aromatic amino acid decarboxylases participate in synthesis of the flavor volatiles 2-phenylethanol and 2-phenylacetaldehyde. *Proceedings of the National Academy of Sciences*, **103**, 8287-8292.

Tittarelli, A., Santiago, M., Morales, A., Meisel, L. and Silva, H. (2009) Isolation and functional characterization of cold-regulated promoters, by digitally identifying peach fruit cold-induced genes from a large EST dataset. *BMC Plant Biology*, **9**, 121.

Vecchietti, a., Lazzari, B., Ortugno, C., Bianchi, F., Malinverni, R., Caprera, a., Mignani, I. and Pozzi, C. (2009) Comparative analysis of expressed sequence tags from tissues in ripening stages of peach (*Prunus persica* L. Batsch). *Tree Genetics & Genomes*, **5**, 377-391.

Verde, I., Abbott, A.G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., Zhebentyayeva, T., Dettori, M.T., Grimwood, J., Cattonaro, F., Zuccolo, A., Rossini, L., Jenkins, J., Vendramin, E., Meisel, L.A., Decroocq, V., Sosinski, B., Prochnik, S., Mitros, T., Policriti, A., Cipriani, G., Dondini, L., Ficklin, S., Goodstein, D.M., Xuan, P., Fabbro, C.D., Aramini, V., Copetti, D., Gonzalez, S., Horner, D.S., Falchi, R., Lucas, S., Mica, E., Maldonado, J., Lazzari, B., Bielenberg, D., Pirona, R., Miculan, M., Barakat, A., Testolin, R., Stella, A., Tartarini, S., Tonutti, P., Arus, P., Orellana, A., Wells, C., Main, D., Vizzotto, G., Silva, H., Salamini, F., Schmutz, J., Morgante, M. and Rokhsar, D.S. (2013) The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet*, **45**, 487-494.

Verde, I., Bassil, N., Scalabrin, S., Gilmore, B., Lawley, C.T., Gasic, K., Micheletti, D., Rosyara, U.R., Cattonaro, F., Vendramin, E., Main, D., Aramini, V., Blas, A.L., Mockler, T.C., Bryant, D.W., Wilhelm, L., Troggio, M., Sosinski, B., Aranzana, M.J., Arús, P., Iezzoni, A., Morgante, M. and Peace, C. (2012) Development and Evaluation of a 9K SNP Array for Peach by Internationally Coordinated SNP Detection and Validation in Breeding Germplasm. *PLoS ONE*, **7**, e35668.

Visai, C. and Vanoli, M. (1997) Volatile compound production during growth and ripening of peaches and nectarines. *Scientia Horticulturae*, **70**, 15-24.

Wang, J., Dudareva, N., Bhakta, S., Raguso, R.A. and Pichersky, E. (1997) Floral Scent Production in Clarkia breweri (Onagraceae) (II. Localization and Developmental Modulation of the Enzyme S-Adenosyl-L-Methionine:(Iso)Eugenol O-Methyltransferase and Phenylpropanoid Emission). *Plant Physiology*, **114**, 213-221.

Wang, Y., Yang, C., Li, S., Yang, L., Wang, Y., Zhao, J. and Jiang, Q. (2009) Volatile characteristics of 50 peaches and nectarines evaluated by HP-SPME with GC-MS. *Food Chemistry*, **116**, 356-364.

Xi, W.-P., Zhang, B.O., Liang, L.I., Shen, J.-Y., Wei, W.-W., Xu, C.-J., Allan, A.C., Ferguson, I.B. and Chen, K.-S. (2012) Postharvest temperature influences volatile lactone production via regulation of acyl-CoA oxidases in peach fruit. *Plant, Cell & Environment*, **35**, 534-545.

Zanor, M.I., Rambla, J.L., Chaib, J., Steppa, A., Medina, A., Granell, A., Fernie, A.R. and Causse, M. (2009) Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. *J Exp Bot*, **60**, 2139-2154.

Zhang, B., Shen, J.-y., Wei, W.-w., Xi, W.-p., Xu, C.-J., Ferguson, I. and Chen, K. (2010) Expression of Genes Associated with Aroma Formation Derived from the Fatty Acid Pathway during Peach Fruit Ripening. *Journal of Agricultural and Food Chemistry*, **58**, 6157-6165.

Zhang, B., Xi, W.-p., Wei, W.-w., Shen, J.-y., Ferguson, I. and Chen, K.-s. (2011) Changes in aroma-related volatiles and gene expression during low temperature storage and subsequent shelf-life of peach fruit. *Postharvest Biology and Technology*, **60**, 7-16.

# OBJETIVOS

# OBJETIVOS

## Objetivo general

Establecer las bases moleculares y genéticas que regulan la producción de los principales compuestos responsables del aroma del melocotón

## Objetivos específicos

1- Desarrollar una plataforma metabolómica de alto rendimiento para la identificación y cuantificación de los compuestos volátiles del melocotón.

2- Describir los patrones de co-regulación de compuestos volátiles y genes asociados en el fruto del melocotón

3- Describir las regiones del genoma de melocotón involucradas en el control genético de la producción de volátiles del fruto mediante análisis de QTL.

4- Proponer nuevos genes candidatos para el control de los volátiles del fruto a partir de los estudios de correlación (2) y de co-localización (3)

# Capítulo 2:
# A non-targeted approach unravels the volatile network in peach fruit

Los resultados presentados en este capítulo fueron publicados en un artículo científico y expuestos en dos congresos internacionales:

-**Sánchez G.**, Besada C., Badenes M.L., Monforte A.J. and Granell A. (2012) "**A Non-Targeted Approach Unravels the Volatile Network in Peach Fruit**." *PLoS ONE*, **7**(6):e38992.

**-Sánchez G.,** Besada C., Rambla J.L., Garcia J., Monforte A.J., Badenes M.L. and Granell A. (2010). **"Characterization of volatiles evolution during ripening and shelf-life of `MxR_01´ and `Granada´ peach genotypes.".** *28th International horticultural congress.* Lisboa, Portugal

**-Sánchez G**., Besada C., Monforte A.J., Badenes M.L. and Granell A. (2010). **"A non-targeted approach unravels a volatile production network on peach fruit.".** *5th International Rosaceae Genomics Conference.* Stellenbosch. Sudafrica.

# A NON-TARGETED APPROACH UNRAVELS THE VOLATILE NETWORK IN PEACH FRUIT.

Sanchez Gerardo[1,2], Besada Cristina[1], Badenes María Luisa[3], Monforte Antonio[1] and Granell Antonio[1*].

[1] Instituto de Biología Molecular y Celular de Plantas (IBMCP). Ingeniero fausto Elio s/n. 46022 Valencia. Spain.

[2] Instituto Nacional de Tecnología Agropecuaria (INTA). Ruta N°9 Km 170. 2930 San Pedro (Buenos Aires). Argentine.

[3] Instituto Valenciano de Investigaciones Agrarias (IVIA). Carretera Moncada-Náquera Km 4,5. 46113 Náquera (Valencia). Spain.

*Corresponding author:

Antonio Granell (agranell@ibmcp.upv.es)

**ABSTRACT**

Volatile compounds represent an important part of the plant metabolome and are of particular agronomic and biological interest since their contribution to the aroma and flavor of the fruit and therefore to fruit quality. By using a non-targeted approach based on HS-SPME-GC-MS, the volatile-compound complement of peach fruit was described. A total of 110 volatile compounds (including: alcohols, ketones, aldehydes, esteres, lactones, carboxylic acids, phenolics and terpenoids) were identified and quantified in peach fruit samples from different genetic background, location, maturity stage and physiological response. By using a combination of hierarchical cluster analysis and metabolomic correlation network analysis we found that previously known peach fruit volatiles are clustered according their chemical nature or known biosynthetic pathways. Moreover, novel volatiles that have not been described in peach so far were identified and assigned to co-regulated groups. In addition our analyses showed that most of the co-regulated groups showed good values of intergroup correlations and therefore consistent with the existence of a higher level of regulation which orchestrates volatiles production under different conditions and/or developmental stages. In addition, this volatile network of interactions provides the ground information for future biochemical studies as well as a useful routemap for breeding or biotechnological purposes.

## INTRODUCTION

Peach (*Prunus persica* L.) is an economically important crop with an expanding world production situated in 20 million tons in 2010 (FAO, 2010). Nevertheless, peach consumption (2 kg of fruit per capita per year) is still considered low when compared to other fresh fruit like apple (16 kg) or banana (9 kg) (Crisosto, 2006). One straight way to enhance peach consumption appear to be the improvement of quality because consumers are complaining about fruit quality since early 90s (Bruhn et al., 1991). Aroma along with fruit firmness and colour are the most important factors that contribute to peach quality according to consumers (Bruhn, 1995). Volatile organic compounds (VOCs) define fruit aroma and, in combination with sugars and organic acids, contribute to the overall peach flavor. Peach volatiles have been intensively studied and around 100 volatiles including: alcohols, aldehydes, esters, terpenoids, ketones and lactones have been described so far [(Wang et al., 2009) and reference within]. Early studies proposed γ-decalactone as the major contributor to peach aroma with smaller contributions from other volatiles such as C6 aldehydes and terpenoids (Horvat et al., 1990). Another lactone, γ-jasmolactone, with a peach like odor has been reported in handmade peach juice (Derail et al., 1999). Others lactones found in peach: γ-octalactone, γ-dodecalactone, δ-decalactone and 6-pentyl-a-pyrone (Aubert and Milhet, 2007; Wang et al., 2009; Zhang et al., 2010) also have pleasant aromas descriptions such as fruity or coconut-like (Derail et al., 1999) and contribute to the overall peach aroma (Eduardo et al., 2010).

Investigations on peach volatiles have been mainly focused on the profiling of volatiles on fruits during maturity and ripening (Aubert et al., 2003; Chapman et al., 1991; Visai and Vanoli, 1997; Zhang et al., 2010), cold storage (Robertson et al., 1990), postharvest treatments (Sumitani et al., 1994; Zhang et al., 2010), culture techniques and management (Jia et al., 2005) and germoplasm

variability analysis (Wang et al., 2009). Even, the distribution of the volatile compounds along the fruit was studied (Aubert and Milhet, 2007).

In spite of the vast amount of data regarding peach volatiles production and the organoleptic description of the main aroma-contributing compounds, nothing is known about how this complex set of volatiles is regulated. The metabolites correlation patterns, analyzed mainly by metabolomic correlation networks, is believed to provide relevant information of the underlying biological system (Morgenthal et al., 2006; Steuer et al., 2003) and could give insights into the network regulation (Camacho et al., 2005; Steuer, 2006). Metabolite-metabolite correlation was used to decipher volatile compounds co-regulated in economically important crops. For tomato, it has been found that volatile compounds derived from the same biochemical pathway are highly correlated (Tikunov et al., 2005) and the interactions between volatiles and primary metabolites was analyzed through correlations networks (Ursem et al., 2008) and hierarchical cluster analysis (Zanor et al., 2009). Using a combination of complementary metabolomic profiling platforms, the association of volatiles compound with inorganic elements and primary and secondary no-volatile metabolites was investigated in melon (Moing et al., 2011).

Currently, the gene function discovery in peach is prompted by the recent release of the whole genome sequence (http://www.rosaceae.org) and the availability of genomic tools, e.g. microarrays (ESTree Consortium, 2005; Ogundiwin et al., 2008) and transient gene expression assays in fruit (Spolaore et al., 2001). Establishing a metabolomic platform for high-throughput volatile compounds profiling for peach and describing the volatile production network is a first step that will provide the ground to aid future studies directed to the identification of genes related to aroma formation.

In this work, we have applied a non-targeted data analysis approach to describe the volatile compounds complement of peach fruit. We have analyzed metabolite-metabolite correlations to get an insight into co-regulation of peach

volatiles and we also investigated the correlations to conventional fruit quality parameters to analyze those volatiles that affect peach quality. Moreover, a correlation network analysis of the complete data set revealed the interactions between different groups of volatiles (i.e. negative interaction between lactones groups and lipid derived volatiles). In addition, several volatile compounds that have not been described in peach fruit so far could be readily assigned to co-regulated groups and/or putative metabolic pathways. These results contribute to define the peach volatile map including the regulatory and interaction patterns which we believe can be useful for breeding or biotechnological purposes.

## MATERIALS AND METHODS

### Ethics Statement

No specific permits were required for the described field studies since the authors are the responsible authority for the fields used. Our study did not imply any national park or protected area.

### Experimental design and fruit analysis.

Peach fruit were harvested in July 2009 from local commercial orchard situated in Murcia, Spain (genotypes: RedCandem, Maruja, MxR_01 and Granada) or at IVIA experimental fields in Valencia, Spain (genotype: MxR_01). RedCandem and MxR_01 are freestone melting fleshed while Maruja and Granda are clingstone non-melting peaches. Maruja is a native material from Spain. RedCandem and Granada are cultivars obtained from different breeding programs (USA and Brazilian respectively). MxR_01 is a F1 hybrid between RedCandem and Maruja genotypes. Fruit for all genotypes were harvested at commercial maturity stage. Besides for MxR_01 and Granada genotypes, three earlier stages of maturity were also evaluated. All four genotypes were analyzed at harvest and after shelf-life conditions to cover changes occurring during ripening the way normally happens during peach commercialization. Twenty fruits from each maturity stage (cultivar and location) were collected and sorted into 2 groups. Fruits in one group were immediately analyzed and those in the other were subjected to shelf-life simulation conditions (2 days at 20ºC, 85% RH). At harvest and after shelf-life simulation quality parameters: peel ground colour, flesh firmness, fruit weight, and soluble solids content (SSC) were analyzed. Peel ground colour was evaluated by a Minolta Colourimeter (Model CR-300, Ramsey, N.Y., U.S.A). L, a, b Hunter parameters were measured, and the results were expressed as colour index (IC=1000a/Lb). Flesh firmness was determined with a texturometer Instron Universal Machine

model 4301 (Instron Corp., Canton, Mass., U.S.A.), using an 8-mm plunger, after epicarp removal, at 2 equidistant locations in the equatorial region of each fruit. Texture data were expressed as the maximum force in kgf required to break the flesh. SSC values were measured twice from each fruit with a digital refractometer (model PR1, Atago, Tokyo, Japan) and the results were expressed as ºBrix. Based on quality parameter measurements, the 3-6 representing fruits per condition (genotype, maturity stage, shelf-life simulation, and location) were selected for volatile analysis as follow. A summary of the experimental design and analysis is presented in the supplementary data (Fig. S1).

**Sample preparation and HS-SPME-GC-MS conditions.**

Volatile analysis was performed essentially as described previously (Tikunov et al., 2005), with minor modifications. Immediately after firmness measurements, peach fruits were peeled and middle mesocarp tissue samples were ground to powder in liquid nitrogen. A 500 mg sample of frozen tissue powder was weighed in a 7 ml vial, which was then sealed, and incubated at 30 ºC for 10 min. Then 500 µ1 of 100 mM EDTA-NaOH (pH 7,5) solution and 1.1 g of CaCl2.2H2O were immediately added to terminate endogenous enzyme activity. The samples in closed vials were agitated and sonicated for 5 min. A 1 ml aliquot of the homogenate was then transferred into a 22 ml crimp cap vial (Perkin-Elmer), capped, and used for HS-SPME-GC-MS analysis. Volatile analysis was performed on a 6890N Agilent gas chromatograph coupled to a 5975B inert XL MSD mass spectrometer (Agilent Technologies). Volatiles were extracted and injected automatically by means of a CombiPAL autosampler (CTC Analytics). After incubating the samples for 10 min at 50°C and continuous agitation (500 rpm), volatile compounds were extracted by adsorption to a 65-µm polydimethylsiloxane-divinylbenzene fiber (Supelco) place in the head space of the vial for 10 min under the same conditions (50°C, 500 rpm). Volatiles were desorbed by direct injection into the port of the gas

chromatograph for 1 min at 250°C in splitless mode. Separation was performed on a DB-5ms column (60 m x 0.25 mm i.d., 1-μm film thickness; J&W Scientific). Helium at a flow rate of 1.2 mL min–1 was used as the carrier gas. The temperature program started at 35°C for 2 min, followed by a 5°C min–1 ramp to 250°C, with a 5 min hold at 250°C. Mass spectra were obtained at an ionization energy of 70 eV and a scan speed of 7 scans s–1, with a mass-to-charge ratio scan range of 35 to 220. In a number of cases commercial standards were used to confirm the chemical nature of the predicted compounds (see later).

**Data processing and statistical analysis.**

For non-targeted data analysis the Multivariate Mass Spectra Reconstruction (MMSR) approach developed by Tikunov et al. (2005) was used. Chromatograms were processed (automated baseline correction, mass spectra extraction, and subsequent spectral data alignment) simultaneously using the dedicated MetAlign software package (http://www.metalign.nl). A peak threshold factor of 10 was chosen for baseline correction in the MetAlign interface. The metabolic profiles aligned were subjected to Hierarchical Cluster Analysis (HCA) using Acuity software 4.0 (Axon Instruments). Pearson correlation coefficient was used as similarity metric and complete linkage as linkage method for HCA. To assign molecular fragments to compounds the HCA were manually inspected. Those molecular fragments that revealed a Pearson correlation higher than 0.8 and eluting within a 3-s retention time window (which corresponds to the maximum peak width at one-half height observed in the chromatograms we obtained) were considered as belonging to the MS spectrum of a given compound.

For compound identification, the most suitable chromatogram (eg. The one showing high abundance of the selected ion and no overlapping with other compounds at the specific position) was selected. Metabolites were putatively

identified using MSD ChemStation software (Agilent Technologies). When necessary, peak deconvolution was realized using AMDIS software (Stein, 1999). Compound name was assigned if: 1) All fragments of the cluster were present in the spectral mass of the peak and 2) the match against NIST mass spectral library (http://www.nist.gov) was higher than 800 (been 1000 a perfect match). A specific ion (m/z) was selected for each compound for quantification. The ratio of the signal relative to that of a reference sample was log 2 transformed. A reference sample consisting of a mix of all samples analyzed during the experiment and was included daily in order to correct fiber aging and the temporal variation of the system.

Acuity 4.0 software (Axon Instruments) was also used for: metabolite hierarchical cluster analysis, heatmap visualization, principal component analysis, and Pearson correlation evaluation.

For metabolomic network construction, the ExpressionCorrelation plug-in (http://www.baderlab.org/Software/ExpressionCorrelation) for Cytoscape software (Shannon et al., 2003) was used. A high and low cut-off of Pearson correlation coefficient of 0.6 and -0.6, respectively, were selected. The network was visualized with Cytoscape software v2.7.0 (www.cytoscape.org).

## RESULTS

### Identification and relative quantification of peach VOCs by a non-targeted data analysis approach.

In order to obtain an exhaustive description of the volatile compound complement of peach fruit, a set of quite diverse samples were analyzed. The sample set comprises a total of 85 samples covering different peach types, diverse genetic origins, different location, maturity stages and a physiological response as it is described in material and methods section (Fig. S1). The specific results regarding the details of this phenotypic variation (genetic background, ambient, maturity stage and shelf-life simulation) are beyond the scope of this article and will be presented elsewhere.

Multivariate Mass Spectra Reconstruction (MMSR) approach was used as non-targeted data analysis for peach volatiles identification and relative quantification. A total of 102 GC-MS chromatograms (85 samples, 12 external references and 5 blanks) were analyzed with MetAlign software. After baseline correction a total of 6.316 molecular fragments were obtained and aligned across the 102 GC-MS datasets. As first step, the 6.316 molecular fragments were subjected to HCA to identify contaminants-derived molecular fragment. Contaminant-derived molecular fragments were recognized by being present in blank samples. Fragments coming from the SPME fiber material presented polysiloxanes typical ions (e.g. m/z 207, 208, 73, etc.). These molecular fragments were excluded from de dataset

After preprocessing the dataset was re-subjected to HCA for metabolite identification. A total of 110 clusters (candidate compounds) of molecular fragments were recognized. HCA inspections permitted to identify 90 as putative VOCs while 20 remained as not-identified (Ni, Table S1). The reliability of our identification procedure was confirmed in 100% of the cases for a set of 53 volatiles for which we have commercially available standards

(Table S1). The structure and mass spectra for the 90 VOCs putatively identified (including those confirmed with standards labeled with an asterisk) are given as supplemental data (Fig. S2).

To make an assessment of the success of our pipeline, several chromatograms were inspected manually in order to obtain an estimate of the total metabolite number. A list of 124 volatiles was recognized manually, indicating that the approach was able to discover around 89% (110:124) of the total VOCs present in our experiment.

Our complex experimental set up confers our VOCs study a series of advantages: it enables the characterization of a more complete volatile compound complement than previous approaches and increases the robustness of the metabolite-metabolite correlations analysis since different variability sources are analyzed.

**Peach VOCs are clustered according their biochemical and/or structural nature**

In order to explore the metabolite-metabolite relations in peach fruit, the entire data set of 9350 elements (85 samples x 110 metabolites) was subjected to HCA. The HCA revealed that volatiles can be organized in different clusters, indicating a tight metabolite regulation (Fig. 1A). As is shown on Fig. 1B, each cluster mainly contains compounds with similar chemical structure (i.e. lactones, no-cyclic esters, carboxylic acid and long chain aldehydes) or belong to specific metabolic pathways (i.e. lipid derived metabolites and terpenoids biosynthesis).

To further explore new metabolite-metabolite correlations and metabolic group interactions we conducted a network correlation analysis using the complete data set (Fig. 2). A correlation network was constructed using the Pearson correlation coefficient. The histogram of the 5995 metabolite-metabolite correlations evaluated for all possible VOCs pairs is showed in Fig. 3. In order

to simplify the view, only strong correlations (r>0.6 in absolute value) were represented in this network.

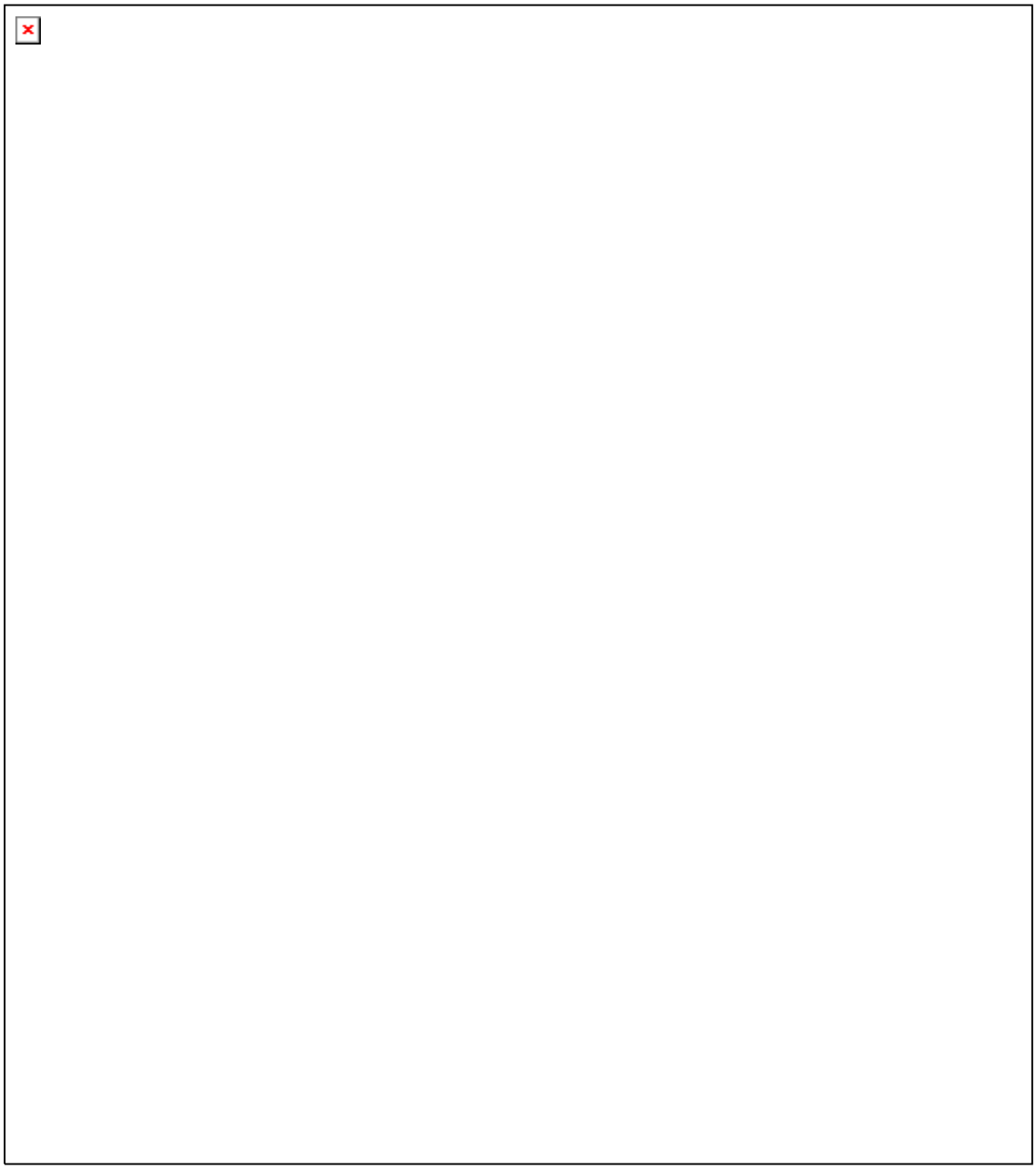


**Figure 1: Hierarchical cluster analysis and heatmap of metabolite-metabolite correlations matrix.** A) Metabolites clusters (1-11) are indicated. The heatmap was constructed with positive Pearson correlations coefficient. The darker the color, the higher the correlation coefficient. B) Detail of the main clusters showing the metabolites members. The compounds: γ-Jasmolactone, γ-Hexalactone, γ-Heptalatone, γ-Octalactone, γ-Decalactone, δ-Decalactone, δ-Limonene, β-Myrcene, and β-Ionone are indicated in the figure as: g-Jasmolactone, g-Hexalactone, g-Heptalatone, g-Octalactone, g-Decalactone, d-Decalactone, D-Limonene, b-Myrcene, and b-Ionone respectively.

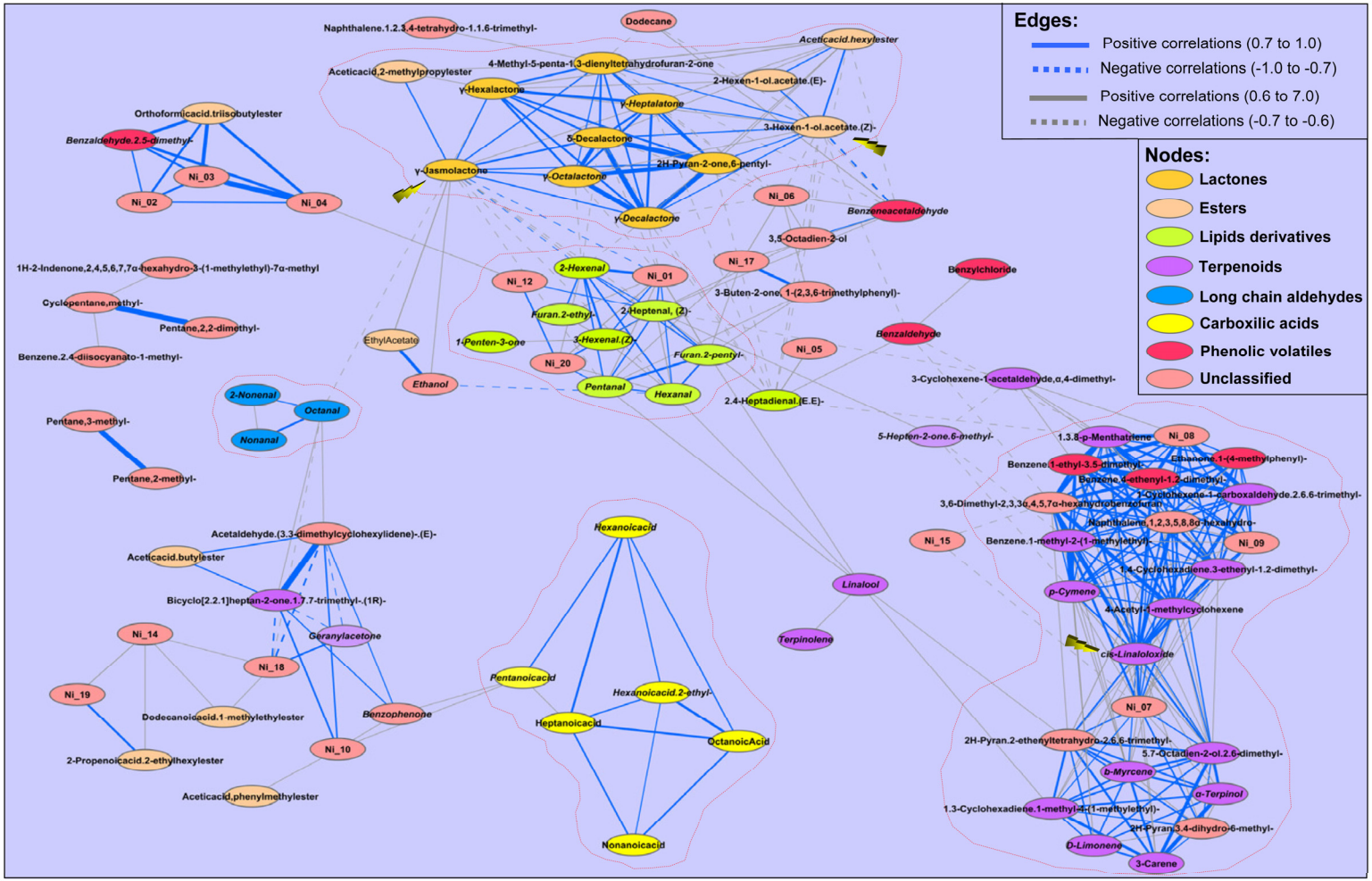



Fig. 2.

**Fig. 2 (see the legends below).**

**Figure 2: Metabolomic correlation network analysis.** The main clusters identified in Fig. 1 are highlighted with red dotted line. Nodes are codified with colors according to chemical nature or by belonging to known metabolic pathways described in others plant species. Correlations lower than 0.7 in absolute value are drawn in grey. Correlation between 0.7-1.0 (absolute value) are in blue. Line thickness indicated correlation strength, the wider the line, the stronger the correlation. Positive correlations are represented with straight lines, negatives correlations with dotted line. The yellow arrows highlight the metabolites with a central role in the metabolomic network (3-Hexen-1-ol acetate, (Z)-, γ-Jasmolactone, and cis-Linaloloxide), for a detailed description see the text. The compounds: γ-Jasmolactone, γ-Hexalactone, γ-Heptalatone, γ-Octalactone, γ-Decalactone, δ-Decalactone, δ-Limonene, and β-Myrcene are indicated in the figure as: g-Jasmolactone, g-Hexalactone, g-Heptalatone, g-Octalactone, g-Decalactone, d-Decalactone, D-Limonene, and b-Myrcene respectively.

The metabolic correlation network obtained consists of 96 nodes (VOCs) and 374 edges (correlations, 326 positives and 48 negatives). As expected, the volatile clusters observed in the HCA could be readily identified in the correlation network as groups of highly interconnected metabolites (Fig. 2). Lactones (γ-Hexalactone, γ-Heptalactone, γ-Octolactone, γ-Decalactone, δ-Decalactone, 2HPyran-2-one 6-pentyl, 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one and γ-Jasmolactona) showed high positive correlations between them. Interestingly these cyclic esters (lactones) are connected to some linear esters (Acetic acid hexyl ester, 2-Hexen-1-ol acetate (E) and 3-Hexen-1-ol acetate (Z)-) mainly through 3-Hexen-1-ol acetate (Z) suggesting an interaction between their biosynthetic pathways. Besides, γ-Jasmolactone levels were highly correlated (r=0.90) with Acetic acid 2 methyl propyl ester (described as a fruity VOC) and showed a moderate correlation (r=0.68) with ethanol and its ester (ethyl acetate) which are normally associated with over-ripening process (Pesis, 2005). Similarly, lipid derived volatiles (Pentanal, Hexanal, 2-Hexanal, 3-Hexanal (Z)-, 2-Heptenal (Z)-, Furan 2-pentyl and Furan 2-ethyl) with 3 unidentified VOCs (Ni_01, Ni_12 and Ni_20) form a

highly interconnected group. This group of volatiles is inversely correlated with the lactones/ester clusters principally through γ-Jasmolactone.

Terpenoids compounds mostly grouped in two highly correlated clusters linked through cis-linaloloxide.

The long chain aldehydes (Octanal, Nonenal and 2-Nonenal) and carboxylic acids (Pentanoic acid, Hexanoic acid, Hexanoic acid 2-ethyl, Heptanoic, Octanoic acid and Nonanoic acid) remain as groups weakly related to other volatiles groups.

In other hand, correlation network analysis revealed compounds not related structurally or by a known biochemical pathway that grouped together. This is the case for, Orthoformic acid triisobutyl ester and Benzaldehyde 2,5-dimethyl- that are highly correlated with a small group of not identified compounds (Ni_02, Ni_03, and Ni_04, Fig. 2). In addition, in our analysis two terpenoid compounds (Geranyl acetone and Bicyclo[2.2.1]heptan-2-one 1,7,7-trimethyl-(1S)-), one ester (Acetic acid butyl ester) and four not classified compounds (Ni_18, Ni_10, Benzophenone, and Acetaldehyde (3,3-dimethylcyclohexylidene)- (E)-) are interconnected by positives and negatives correlations (Fig. 2), this will be discussed later.

Finally, our correlation network analysis also identified a set of volatiles (i.e. Pentane, 2-methyl and Pentane, 3-methyl) that are apparently not related to any other compound of the volatile complement.

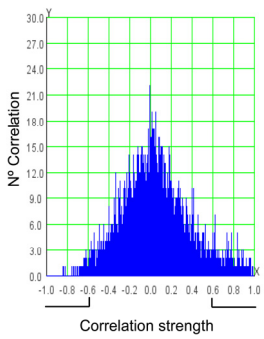



**Fig. 3.**

**Figure 3: Distribution of the correlations evaluated**. Black lines below the graph show the correlations taken to construct the network.

**Some volatile compounds can be associated to quality trait**

In addition to the description of the relationship between compounds in the peach volatile complement, correlations with quality parameters were investigated in order to identify VOCs putatively associated to a good ripen peach. As a number of quality traits such as ground colour, fruit weight and SSC increase during ripening in a genotype dependent manner, VOCs showing positive correlations to them could be associated to quality. In the same way, fruit firmness decrease during ripening and therefore VOCs associated to this fruit quality trait should be inverse correlated to this parameter. To find robust correlations, we analyzed Pearson correlations coefficient between a series of quality parameters and VOCs through the entire 85 sample set. VOCs that showed the strongest correlations with regard to ground colour (Table 1), flesh firmness (Table 2), fruit weight (Table 3) and SSC (Table 4) are shown. All lactones identified in our materials (8 in total) were highly correlated with ground colour (r=0.72 to 0.85) and 7 of them also showed strong negative correlations with firmness (r=0.75 to 0.88 in absolute value). Some lactones (γ-Hexalactone, γ-Heptalatone, and γ-Jasmolactone) also showed correlations towards weight (r=0.69, 0.57, and 0.55, respectively) but with lower values as compared to colour and firmness. A number of esters were highly correlated to quality parameters: Ethyl acetate showed strong correlation with colour (r=0.67), 3-Hexen-1-ol acetate (Z)-, and Acetic acid hexyl ester were inversely correlated to firmness (r=-0.84, and -0.73 respectively) and 2-Hexen-1-ol acetate (E)- was correlated to weight (r=0.63). In general, correlations with weight were lower (r=0.54 to 0.69) than those with colour (r=0.67 to 0.85) and firmness (r=0.73 to 0.88 in absolute value). Nevertheless, weight also showed correlations with other volatiles besides lactones and esters. Although significant, volatile compounds showed low correlations towards SSC (r=0.25 to 0.52).

**Table 1.** Correlation analysis between volatiles and ground colour.

| Ground colour | |
|---|---|
| **VOCs** | **+PCC (r)** |
| γ-Jasmolactone | 0.85 |
| δ-Decalactone | 0.79 |
| γ-Decalactone | 0.77 |
| γ-Octalactone | 0.76 |
| 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | 0.76 |
| γ-Heptalatone | 0.74 |
| 2H-Pyran-2-one, 6-pentyl- | 0.73 |
| γ-Hexalactone | 0.72 |
| Ethyl Acetate | 0.67 |
| **VOCs** | **-PCC (r)** |
| Ni_06 | -0.71 |
| Ni_01 | -0.67 |
| 2-Hexenal | -0.63 |
| Furan, 2-pentyl- | -0.61 |
| Ni_17 | -0.59 |
| 2-Heptenal, (Z)- | -0.58 |
| Pentanal | -0.58 |
| Eugenol | -0.56 |
| Benzeneacetaldehyde | -0.54 |

Volatiles showing the strongest positive and negative correlations are shown. +PCC, Positive Pearson correlation coefficient; -PCC, Negative Pearson correlation coefficient. The correlations shown are significant ($\alpha<0.05$).

**Table 2.** Correlation analysis between volatiles and flesh firmness.

| Flesh firmness | |
|---|---|
| **VOCs** | **-PCC (r)** |
| γ-Decalactone | -0.88 |
| δ-Decalactone | -0.87 |
| 2H-Pyran-2-one, 6-pentyl- | -0.85 |
| 3-Hexen-1-ol, acetate, (Z)- | -0.84 |
| 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | -0.81 |
| γ-Heptalatone | -0.81 |
| γ-Octalactone | -0.77 |
| γ-Hexalactone | -0.75 |
| Acetic acid, hexyl ester | -0.73 |
| **VOCs** | **+PCC (r)** |
| 2,4-Heptadienal, (E,E)- | 0.76 |
| Ni_06 | 0.73 |
| Benzeneacetaldehyde | 0.69 |
| Dodecane | 0.65 |
| 2-Heptenal, (Z)- | 0.64 |
| Ni_01 | 0.62 |
| 3,5-Octadien-2-ol | 0.59 |
| 2-Hexenal | 0.56 |
| Benzaldehyde | 0.56 |

Volatiles showing the strongest positive and negative correlations are shown. +PCC, Positive Pearson correlation coefficient; -PCC, Negative Pearson correlation coefficient. The correlations shown are significant ($\alpha<0.05$).

**Table 3.** Correlation analysis between volatiles and fruit weight.

| Fruit weight | |
|---|---|
| **VOCs** | **+PCC (r)** |
| γ-Hexalactone | 0.69 |
| Ethanone, 1-(4-methylphenyl)- | 0.67 |
| 2-Hexen-1-ol, acetate, (E)- | 0.63 |
| 1,3,8-p-Menthatriene | 0.60 |
| Benzene, 4-ethenyl-1,2-dimethyl- | 0.58 |
| γ-Heptalatone | 0.57 |
| 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | 0.55 |
| γ-Jasmolactone | 0.55 |
| 3-Cyclohexene-1-acetaldehyde, α,4-dimethyl- | 0.54 |
| **VOCs** | **-PCC (r)** |
| 2-Heptenal, (Z)- | -0.73 |
| Benzeneacetaldehyde | -0.60 |
| Ni_20 | -0.57 |
| Hexanal | -0.56 |
| 3-Hexenal, (Z)- | -0.55 |
| 3,5-Octadien-2-ol | -0.55 |
| Ni_01 | -0.54 |
| 2-Hexenal | -0.54 |
| Dodecane | -0.53 |

Volatiles showing the strongest positive and negative correlations are shown. +PCC, Positive Pearson correlation coefficient; -PCC, Negative Pearson correlation coefficient. The correlations shown are significant ($\alpha<0.05$).

**Table 4.** Correlation analysis between volatiles and SSC.

| SSC | |
|---|---|
| **VOCs** | **+PCC (r)** |
| 3-Hexen-1-ol, acetate, (Z)- | 0.52 |
| 3-Buten-2-ol, 2-methyl- | 0.43 |
| Acetic acid, hexyl ester | 0.36 |
| Hexanal | 0.29 |
| γ-Heptalatone | 0.28 |
| 2H-Pyran-2-one, 6-pentyl- | 0.27 |
| γ-Decalactone | 0.26 |
| Terpinolene | 0.26 |
| δ-Decalactone | 0.25 |
| **VOCs** | **-PCC (r)** |
| 1H-2-Indenone,2,4,5,6,7,77α-hexahydro .....[a] | -0.69 |
| Ni_14 | -0.57 |
| Ni_06 | -0.53 |
| Dodecanoic acid, 1-methylethyl ester | -0.45 |
| 2,4-Heptadienal, (E,E)- | -0.45 |
| Benzeneacetaldehyde | -0.44 |
| 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | -0.43 |
| 3-Buten-2-one, 1-(2,3,6-trimethylphenyl)- | -0.40 |
| Benzaldehyde | -0.39 |

Volatiles showing the strongest positive and negative correlations are shown. +PCC, Positive Pearson correlation coefficient; -PCC, Negative Pearson correlation coefficient. The correlations shown are significant ($\alpha<0.05$).

[a]The full name of the compound is: 1H-2-Indenone,2,4,5,6,7,7a-hexahydro-3-(1-methylethyl)-7a-methyl.

Volatiles showing inverse correlations with colour, weight and SSC or direct correlation to firmness could be implicated in immature fruit aroma and therefore could have a negative impact in peach quality. In general, lipid derived volatiles and a number of not-identified compounds associated to them (Fig. 1 and Fig. 2) showed moderate to high correlations to quality parameters (Tables 1, 2, 3, 4). This is the case for 2-Hexenal, Furan 2-pentyl-, 2-Heptenal (Z)-, Pentanal, and Ni_01 which showed inverse correlation with colour values (r=0.58 to 0.67 in absolute value). Regarding firmness, 2,4-Heptadienal (E,E)- showed the strongest correlation (r=0.76) and so did 2-Heptenal (Z)-, and Ni_01 although with significant but lower coefficients (r=0.64 and 0.61, respectively). In the case of weight, it is strongly but inversely correlated to 2-Heptenal (Z)-, and moderately to low correlated with Hexanal, 3-Hexenal (Z)-, and 2-Hexenal (Table 3). Other volatiles not related to lipid metabolism were also correlated to quality parameters. For example, Ni_06 showed high correlation with colour (r=-0.71) and firmness (r=0.73) and Benzeneacetaldehyde was correlated to all quality parameters ranging its value from r=0.44 to 0.69 (in absolute value). In our study, SSC was poorly correlated with volatiles compounds with the exception of 1H-2-Indenone 2,4,5,6,7,7$\alpha$-hexahydro-3-(1-methylethyl)-7$\alpha$-methyl that showed a inverse correlation of r=-0.69 (Table 4).

## DISCUSSION

### Description of the peach fruit volatile compound complement by a non-targeted approach.

A combination of analytical technique (HS-SPME-GC-MS), chromatogram aligment tools (MetAlign) and data analyses (cluster analysis and VOC identification) was used to profile the volatile compounds complement of a complex sample set of peach fruit and to reveal the network of VOC metabolic interactions. A similar approach has been previously used to analyze VOCs of tomato fruit and was estimated that approximately 80% of the total tomatoes VOCs were detected (Tikunov et al., 2005). Here, 89% of the total number of volatiles compound present in our sample set could be identified, indicating the power of the platform developed.

HS-SPME is considered an easy and inexpensive extraction methods compared to liquid–liquid extraction, solid phase extraction (SPE), vacuum distillation and dynamic headspace. Nevertheless it has been poorly used to profile peach VOCs with some exceptions. Using HS-SPME-GC-MS, Wang et al. studied fruit VOCs of a peach germoplasm collection; analyzing 50 peach and nectarine from different origin they identified 84 VOCs (Wang et al., 2009). In our study, we putatively indentified a comparable number of volatile compounds (90, Table S1). For a group of 53 VOCs we analyzed the consistency of our identification procedure using commercial standards, confirming the identity in 100% of the cases. In addition, we succeeded in the identification of several compounds that have not been described previously in peach fruit (e.g.: 34, 39, 44, 58, 67, 88, and 97 numbered according Table S1 and Figure S1). All this suggesting that the approach used here was able to profile a substantial range of the peach volatiles compound complement.

We also investigated correlations of VOCs with the main quality parameters in order to describe volatiles putatively involved in ripe peach aroma (Tables 1, 2,

3, 4). Peach quality is directly affected by maturity state of the fruit. So the conventional parameters used to characterize the fruit maturity stage (peel ground colour, flesh firmness, soluble solid content, weight and titratable acidity) are also related to fruit quality. The change in peel ground colour (from green to orange-red, increasing colour index values) is a usual criterion used to harvest the fruit at the right time in order to obtain high quality peaches. SSC and weight also rise during ripening while fruit accumulate sugars and increase it size, respectively. Contrary, flesh firmness decrease during fruit ripening as a result of the natural softening. Thus, VOCs contributing to ripe peach aroma and therefore to the peach quality should be directly correlated with colour, weight and SSC and inverse correlated with flesh firmness. In order to identify robust relationships (rather than genotype specific ones) we took advantage of our complex sample set and evaluated correlations between volatiles and fruit quality parameters throughout the 85 samples (Tables 1, 2, 3, 4). Lactones show strong correlations with colour (with positives values) and flesh firmness (in negatives values) suggesting that are associated to fruit quality assayed by these parameters irrespective of the genotype. Lactones has been previously implicated in fruit ripening (Visai and Vanoli, 1997) and have been reported as main contributors to mature fruit aroma (Derail et al., 1999; Eduardo et al., 2010; Horvat et al., 1990). The lactone putatively identified as 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one has not been described in peach so far although it appears to be the most abundant volatile in some species of chanomeles fruit (Meng et al., 2007), belonging to a different genus within the *Rosaceae* family. The fact that this compound showed high correlations with quality parameters (r=0.76 with colour and r=-0.81 with firmness) suggest that could be implicated in ripe peach aroma. Nevertheless, in order to accurately evaluate the involvement of this compound to the overall peach aroma, the odor quality as well as the odor activity and threshold values need to be analyzed.

Correlations analysis indicated also that lipid derived compounds (2-Hexenal, Furan, 2-pentyl-, Pentanal, Hexanal, 3-Hexenal, (Z)-, 2-Heptenal (Z)-, and 2,4-Heptadienal (E,E)-) could be related to immature fruit (Tables 1, 2, 3, 4). Lipid derived compound are described as ``green´´ aromas and in the case of 2-Hexenal and 3-Hexenal, (Z)- were described to confer a ``green´´ notes to handmade peach juice (Derail et al., 1999). The volatile compound 2,4-Heptadienal, (E,E)- showed the highest correlation with flesh firmness (r=0.76) suggesting that it accumulates in immature fruit. Consistently, the odor of 2,4-Heptadienal (E,E) has been described as ´´green/fatty´´ (www.thegoodscentscompany.com). This aldehyde has been found in plum and apricot cultivars (Gomez et al., 1993) but so far not in peach fruits. Our study revealed a number of compounds (e.g. Ni_01 and Ni_06) that also could be eventually implied in immature aromas of peach, but its structure remains to be elucidated.

**Volatile complement of peach fruit reveals a highly organized network.**

The ultimate goal of a metabolomics approach is to provide an unbiased identification and quantification of all metabolites present in a system but in a comprehensive manner (Hall, 2006). In this study, we described the volatile compound complement of peach metabolome and analyzed the metabolite-metabolite correlation to get an insight into co-regulated groups of volatiles compound some of which may directly be affecting fruit quality.

In order to accurately evaluate metabolite-metabolite correlations is necessary to analyze a system showing consistent variation on the variables under study. Genetic variation can be used as source for VOCs phenotypic variation and this be analyzed in an attempt to reveal the underlying metabolic networks. In a previous study, ripened fruits from a large set of commercial tomato cultivars were analyzed (Tikunov et al., 2005). Later, the same strategy was used in *Citrus* to obtain and analyze phenotypic variation (Gonzalez-Mas et al., 2011).

Recently, an alternative strategy was reported to evaluate metabolite-metabolite correlation using different developmental stages and parts of the melon fruit (Moing et al., 2011) therefore accessing both developmental and tissue variability. Here we presented a combination of these two strategies by analyzing samples from different genetic backgrounds, and maturity stages but also including different locations and a physiological response. By combining diverse origin of variation to affect the VOCs phenotypic variation our study proposes an alternative strategy to use the MMSR approach described previously to evaluate metabolite-metabolite correlations. This alternate method could be particularly useful in cases where a large number of genotypes are not available as is usually the case for fruit crops or wild species.

In our study, correlation analysis (HCA and metabolic network analysis) revealed that most of the compounds clustered according to their biochemical nature: lactones, no-cyclic esters, carboxylic acids and long chain aldehydes. In some cases they clustered according to their known metabolic pathways as described in other species: lipid derived metabolites and terpenoids biosynthesis (Fig. 1; Fig. 2). The relationship between metabolic correlation network and the underlying biochemical pathways has been studied using GC-TOF/MS derived data as a model (Camacho et al., 2005; Muller-Linow et al., 2007; Steuer et al., 2003). Currently, it is generally accepted that although correlation patterns cannot be directly extrapolated to biochemical pathways they can be very informative about the underlying system (Camacho et al., 2005). Since metabolic pathways involved in the synthesis of volatiles compounds in peach are still poorly known we have restricted our analysis in terms of pragmatic implications. Nevertheless, we provide empirical data that confirms known volatile biochemical pathways operating in peach fruit and suggest some novel ones. For example, in some fruits acetaldehyde can be interconverted to ethanol by the enzyme alcohol deshydrogenase (ADH, EC 1.1.1.1) or to acetyl coenzyme A (CoA) by the enzyme aldehyde dehydrogenase

(ALDH, EC 1.2.1.5) (Pesis, 2005). Acetyl-CoA could then undergo subsequent reactions with alcohols to generate the corresponding acetate esters. Indeed, the high correlation between ethanol and ethyl acetate (r=0.94) observed in the metabolic correlation network could represent one of these reactions (Fig. 2). Camacho et al. suggested several mechanisms to explain the correlation between metabolites in replicated experiments (Camacho et al., 2005). They stated that two metabolites in chemical equilibrium will show a high correlation across multiple experimental conditions (Camacho et al., 2005). This suggests that putative alcohol acyltransferases (AAT, EC 2.3.1.84) catalyze a reversible reaction using ethanol and acetyl-CoA as substrate in peach. Ethyl acetate has been described as fruity aroma (Guillot et al., 2006). Thus, putative AATs become an interesting target for molecular engineering or plant breeding in order to obtain enzymes with the chemical equilibrium displaced towards ethyl acetate production.

Metabolic pathways leading to lactones synthesis has not been described yet in plants. However, it is commonly accepted that lactones biosynthesis in fruit start from unsaturated fatty acids that subsequently suffer the introduction of an O atom to form hydroxy fatty acids which in turn undergo β-oxidation leading to the 4- or 5- hydroxy acids that after intramolecular esterification produce the corresponding lactones (Schwab et al., 2008). Infiltrating an artificial radiolabeled epoxy acid into fruits was demonstrated that nectarines (peach glabrous mutation) are able to produce lactones from epoxy acids (Schöttler and Boland, 1996). Based on this study, it was proposed that in nectarine (or peach) the introduction of the hydroxyl groups is achieved by fatty acid epoxidation and subsequent hydrolysis of the compound formed by epoxyde hydrolases resulted in the proper hydroxy acid (Schöttler and Boland, 1996). Here, we showed that lactones production on peach fruit is highly correlated and are central in the interactions with other volatiles groups including linear esters (Fig. 1 and Fig. 2). It have been proposed that correlation

network largely reflect biochemical regulation (Muller-Linow et al., 2007). Thus, our results indicate that the lactones biosynthesis on peach fruit is highly co-regulated supporting future efforts to identify regulatory genes that control the production of these family of compounds for biotechnological or breeding/improvement purposes. Indeed, during the preparation of this manuscript it has been proposed that lactone production during postharvest ripening of peach is regulated by the first enzyme of β-oxidation (Acyl-CoA oxidase) (Xi et al., 2012) in agreement with our hypothesis.

Ethyl acetate as well as 3-hexen1-ol acetate, 2-hexen1-ol acetate, and acetic acid 2-methyl propyl ester have been described to have fruity odour quality (Derail et al., 1999; Guillot et al., 2006). Our volatile metabolic correlation network indicated that 3-hexen1-ol acetate, 2-hexen1-ol acetate and acetic acid hexyl ester belong to a co-regulated group while ethyl acetate and acetic acid hexyl ester are regulated independently (Fig. 2). These findings suggest the existence of at least three potential ways to improve fruity notes of peach aroma. Moreover, the fact that lactones group interact with esters (3-hexen1-ol acetate, 2-hexen1-ol acetate, and acetic acid 2-methyl propyl ester) suggests that lactonization may also involved common enzymes or intermediates shared with esters. All this information may prove to be very useful for breeding.

Previously was showed that volatile compounds derived from the catabolism of linoleic acid by lipoxygenase pathway, the so-called lipid-derived volatiles, cluster together in tomato (Tikunov et al., 2005) and *Citrus* species (Gonzalez-Mas et al., 2011). Here we demonstrated that in peach lipid derived compounds (2,4-Heptadienal, (E,E)-, 1-Penten-3-one, 2-Hexanal, 3-Hexanal (Z), Furan 2-ethyl, 2-Heptenal (Z)-, and Hexanal) with other not identified compounds (Ni_01, Ni_12, and Ni_20) are grouped in the same cluster (Fig. 1) and this group negatively interact with the lactones group (Fig. 2). Since lactones biosynthesis also start from fatty acid precursors (Schwab et al., 2008), this interaction suggest that a metabolic shift of lipid metabolism during ripening

could explain the re-organization of peach aroma from ``green notes´´ conferred by hexenal, 2-hexanal, 3-hexanal towards ``fruit´´, ``floral´´ or ``peach-like´´ aromas of lactones. Targeting the genes or regulators of this metabolic shift can be an important biotechnological target.

Long chain aldehydes: octanal, nonanal and 2-nonenal could confer unpleasant aroma since their odour quality were described as ``fatty/tallowy´´ (Derail et al., 1999). The fact that, these compounds were correlated to form a group and remained as a group poorly linked to other volatiles groups (Fig. 2), could anticipated that an effort to lower these metabolites levels would result in none or little effect on the other aromas compound. Using a similar strategy, our correlation network analysis revealed cis-linaloxide as a hub for terpenoid cluster (Fig. 2) and therefore is proposed as a potential target for terpenoid manipulation in peach.

In our analysis most of the strongest correlations were positive (direct) with few exceptions of negative (inverse) correlation, e.g. between γ-Jasmolactone with 2-Hexenal or Ni_01, (Fig 2.). The correlation between volatiles and primary metabolites in tomato fruit have been investigated recently using correlation network approach to find that most of the robust correlations were direct relations (Ursem et al., 2008). It appears that this property is characteristic of metabolomic analysis approach although the reasons for that are not yet understood.

Although correlation analysis revealed that most of the compounds were clustered according to their chemical structure or their belonging to known metabolic pathways (Fig. 1 and Fig. 2), we also found groups of compounds with not obvious trend. This is the case for the group formed by: Geranyl acetone, Bicyclo[2.2.1]heptan-2-one 1,7,7-trimethyl- (1S)-, Acetic acid butyl ester, Ni_19, Ni_11, Benzophenone, and Acetaldehyde (3,3-dimethylcyclohexylidene)-(E)- (Fig. 2). Is not clear what could be the origin for the correlations found between these compounds, but is interesting to note that Acetaldehyde (3,3-

dimethylcyclohexylidene)- (E) is the active principle of commercial pheromones (PAN pesticide database: www.pesticideinfo.org) usually used in peach orchard for insect control. Although we cannot verify it we cannot exclude the possibility that foreign compounds were detected in our analysis. Since our study revealed that peach derived volatiles are highly correlated to each other (Fig. 2), a correlation network analysis could allow in the identification of contaminant volatile compounds by being poorly connected to the whole metabolome. In addition, we anticipate that the metabolic correlation network could be used to study the effect of chemical treatments (applied at pre or postharvest) on peach volatile production and therefore in the final quality of the product.

We have limited our interpretation to pragmatic application of our results to those volatiles that may affect aroma production, since its importance as reflected by the large amount of the previous work on peach volatiles that focuses in fruit aroma. Nevertheless, the metabolic correlation map established here could eventually aid in defining potential biotechnological targets to improve other characters (e.g. nutraceutical quality of the fruit or plant-insect communication) as novel functions for peach volatiles are discovered.

## CONCLUSIONS

A high resolution and sensitive platform for volatiles compound analysis of peach fruit was established. This approach allows us to identify a significant part of the volatiles complement of the peach fruit metabolome. Exploring metabolite-metabolite correlations a highly organized volatile metabolite network was discovered. In addition, several novel volatile metabolites, not described in peach fruit so far, were identified and assigned to co-regulated metabolite groups and/or putative metabolic pathways. The metabolic network established here provides a framework for future genetic, physiological and environmental analysis of peach volatile production as well as provide important information for improving volatile composition in peach breeding programs.

**ACKNOWLEDGMENTS**

We are thankful to M.Sc. José Martinez (*Instituto Valenciano de Investigaciones Agrarias*, IVIA, Spain) and Dr. Jesús Garcia Bruton (*Instituto Murciano de Investigación y Desarrollo Agrario*, IMIDA, Spain) for peach orchard handling. We also want to thank Dra. Alejandra Salvador (*Instituto Valenciano de Investigaciones Agrarias*, IVIA, Spain) for providing facilities for quality parameters determination and Dr. Christophe Aubert (*Centre technique interprofessionnel des fruits et legumes*, Ctifl, France) for his kindly supply of γ-jasmolactone standard. We are grateful to José Luís Rambla (*Instituto de Biología Molecular y Celular de Plantas*, IBMCP, Spain) for sharing his knowledge about gas-chromatography. HS-SPME-GC-MS analysis was performed at the Metabolomic lab facilities on *Instituto de Biología Molecular y Celular de Plantas* (IBMCP, CSIC) from Spain.

**AUTHORS' CONTRIBUTIONS**

Conceived and designed the experiments: GS AG. Performed the experiments: GS CB. Analyzed the data: GS. Contributed reagents/materials/analysis tools: AJM MLB AG. Wrote the paper: GS AG.

## REFERENCES


Aubert, C., Gunata, Z., Ambid, C. and Baumes, R. (2003) Changes in Physicochemical Characteristics and Volatile Constituents of Yellow- and White-Fleshed Nectarines during Maturation and Artificial Ripening. *J. Agric. and Food Chem.*, **51**, 3083-3091.

Aubert, C. and Milhet, C. (2007) Distribution of the volatile compounds in the different parts of a white-fleshed peach (*Prunus persica* L. Batsch). *Food Chem.*, **102**, 375-384.

Bruhn, C.M. (1995) Consumer and retailer satisfaction with the quality and size of california peaches and nectarines. *Journal of Food Quality*, **18**, 241-256.

Bruhn, C.M., Feldman, N., Garlitz, C., Harwood, J., Ivans, E., Marshall, M., Riley, A., Thurber, D. and Williamson, E. (1991) Consumer Perceptions of quality: Apricots, Cantaloupes, Peaches, Pears, Strawberries, and Tomatoes. *Journal of Food Quality*, **14**, 187-195.

Camacho, D., de la Fuente, A. and Mendes, P. (2005) The origin of correlations in metabolomics data. *Metabol*, **1**, 53.

Crisosto, C.H. (2006) Short-term approaches to increase peach fruit consumption. *Compact Fruit Tree*, **39**, 11-14.

Chapman, G.W., Horvat, R.J. and Forbus, W.R. (1991) Physical and chemical changes during the maturation of peaches (cv. Majestic). *Journal of Agricultural and Food Chemistry*, **39**, 867-870.

Derail, C., Hofmann, T. and Schieberle, P. (1999) Differences in Key Odorants of Handmade Juice of Yellow-Flesh Peaches (*Prunus persica* L.) Induced by the Workup Procedure. *J. Agric. and Food Chem.*, **47**, 4742-4745.

Eduardo, I., Chietera, G., Bassi, D., Rossini, L. and Vecchietti, A. (2010) Identification of key odor volatile compounds in the essential oil of nine peach accessions. *Journal of the Science of Food and Agriculture*, **90**, 1146-1154.

ESTree Consortium. (2005) Development of an oligo-based microarray (uPEACH 1.0) for genomics studies in peach fruit. *Acta Hortic*, **682**, 263–268.

FAO. (2010) Accessed 22 Feb. 2012. http://faostat.fao.org.

Gomez, E., Ledbetter, C.A. and Hartsell, P.L. (1993) Volatile compounds in apricot, plum, and their interspecific hybrids. *Journal of Agricultural and Food Chemistry*, **41**, 1669-1676.

Gonzalez-Mas, M.C., Rambla, J.L., Alamar, M.C., Gutierrez, A. and Granell, A. (2011) Comparative Analysis of the Volatile Fraction of Fruit Juice from Different *Citrus* Species. *PLoS ONE*, **6**, e22016.

Guillot, S., Peytavi, L., Bureau, S., Boulanger, R., Lepoutre, J.-P., Crouzet, J. and Schorr-Galindo, S. (2006) Aroma characterization of various apricot varieties using headspace-solid phase microextraction combined with gas chromatography-mass spectrometry and gas chromatography-olfactometry. *Food Chemistry*, **96**, 147-155.

Hall, R.D. (2006) Plant metabolomics: from holistic hope, to hype, to hot topic. *New Phytologist*, **169**, 453-468.

Horvat, R.J., Chapman, G.W., Robertson, J.A., Meredith, F.I., Scorza, R., Callahan, A.M. and Morgens, P. (1990) Comparison of the volatile compounds from several commercial peach cultivars. *Journal of Agricultural and Food Chemistry*, **38**, 234-237.

Jia, H.-J., Araki, A. and Okamoto, G. (2005) Influence of fruit bagging on aroma volatiles and skin coloration of `Hakuho´ peach (*Prunus persica* Batsch). *Postharvest Biology and Technology*, **35**, 61-68.

Meng, X., Liu, L. and Xu, H. (2007) GC-MS analysis of aroma components of different Chaenomeles fruits. *Journal of Northwest Sci-Tech University of Agriculture and Forestry*, **35**, 125-130 (*chinese*).

Moing, A., Aharoni, A., Biais, B., Rogachev, I., Meir, S., Brodsky, L., Allwood, J.W., Erban, A., Dunn, W.B., Kay, L., de Koning, S., de Vos, R.C.H., Jonker, H., Mumm, R., Deborde, C., Maucourt, M., Bernillon, S., Gibon, Y., Hansen, T.H., Husted, S., Goodacre, R., Kopka, J., Schjoerring, J.K., Rolin, D. and Hall, R.D. (2011) Extensive metabolic cross-talk in melon fruit revealed by spatial and developmental combinatorial metabolomics. *New Phytologist*, **190**, 683-696.

Morgenthal, K., Weckwerth, W. and Steuer, R. (2006) Metabolomic networks in plants: Transitions from pattern recognition to biological interpretation. *Biosystems*, **83**, 108-117.

Muller-Linow, M., Weckwerth, W. and Hutt, M.-T. (2007) Consistency analysis of metabolic correlation networks. *BMC Systems Biology*, **1**, 44.

Ogundiwin, E.A., Marti, C., Forment, J., Pons, C., Granell, A., Gradziel, T.M., Peace, C.P. and Crisosto, C.H. (2008) Development of ChillPeach genomic tools and identification of cold-responsive genes in peach fruit. *Plant Mol Biol*, **68**, 379-397.

Pesis, E. (2005) The role of the anaerobic metabolites, acetaldehyde and ethanol, in fruit ripening, enhancement of fruit quality and fruit deterioration. *Postharvest Biology and Technology*, **37**, 1-19.

Robertson, J.A., Meredith, F.I., Horvat, R.J. and Senter, S.D. (1990) Effect of cold storage and maturity on the physical and chemical characteristics and volatile constituents of peaches (cv. Cresthaven). *Journal of Agricultural and Food Chemistry*, **38**, 620-624.

Schöttler, M. and Boland, W. (1996) Biosynthesis of Dodecano-4-lactone in Ripening Fruits: Crucial Role of an Epoxide-Hydrolase in Enantioselective Generation of Aroma Components of the Nectarine (*Prunus persica* var. nucipersica) and the Strawberry (*Fragaria ananassa*). *Helvetica Chimica Acta*, **79**, 1488-1496.

Schwab, W., Davidovich-Rikanati, R. and Lewinsohn, E. (2008) Biosynthesis of plant-derived flavor compounds. *The Plant Journal*, **54**, 712-732.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, **13**, 2498-2504.

Spolaore, S., Trainotti, L. and Casadoro, G. (2001) A simple protocol for transient gene expression in ripe fleshy fruit mediated by *Agrobacterium*. *Journal of Experimental Botany*, **52**, 845-850.

Stein, S. (1999) An integrated method for spectrum extraction and compound identification from gas chromatography/mass spectrometry data. *Journal of The American Society for Mass Spectrometry*, **10**, 770-781.

Steuer, R. (2006) Review: On the analysis and interpretation of correlations in metabolomic data. *Briefings in Bioinformatics*, **7**, 151-158.

Steuer, R., Kurths, J., Fiehn, O. and Weckwerth, W. (2003) Observing and interpreting correlations in metabolomic networks. *Bioinformatics*, **19**, 1019-1026.

Sumitani, H., Suekane, S., Nakatani, A. and Tatsuka, K. (1994) Changes In Composition of Volatile Compounds in High Pressure Treated Peach. *Journal of Agricultural and Food Chemistry*, **42**, 785-790.

Tikunov, Y., Lommen, A., de Vos, C.H.R., Verhoeven, H.A., Bino, R.J., Hall, R.D. and Bovy, A.G. (2005) A Novel Approach for Nontargeted Data Analysis for Metabolomics. Large-Scale Profiling of Tomato Fruit Volatiles. *Plant Physiol.*, **139**, 1125-1137.

Ursem, R., Tikunov, Y., Bovy, A., van Berloo, R. and van Eeuwijk, F. (2008) A correlation network approach to metabolic data analysis for tomato fruits. *Euphytica*, **161**, 181-193.

Visai, C. and Vanoli, M. (1997) Volatile compound production during growth and ripening of peaches and nectarines. *Scientia Horticulturae*, **70**, 15-24.

Wang, Y., Yang, C., Li, S., Yang, L., Wang, Y., Zhao, J. and Jiang, Q. (2009) Volatile characteristics of 50 peaches and nectarines evaluated by HP-SPME with GC-MS. *Food Chemistry*, **116**, 356-364.

Xi, W.-P., Zhang, B.O., Liang, L.I., Shen, J.-Y., Wei, W.-W., Xu, C.-J., Allan, A.C., Ferguson, I.B. and Chen, K.-S. (2012) Postharvest temperature influences volatile lactone production via regulation of acyl-CoA oxidases in peach fruit. *Plant, Cell & Environment*, **35**, 534-545.

Zanor, M.I., Rambla, J.L., Chaib, J., Steppa, A., Medina, A., Granell, A., Fernie, A.R. and Causse, M. (2009) Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. *J Exp Bot*, **60**, 2139-2154.

Zhang, B., Shen, J.-y., Wei, W.-w., Xi, W.-p., Xu, C.-J., Ferguson, I. and Chen, K. (2010) Expression of Genes Associated with Aroma Formation Derived from the Fatty Acid Pathway during Peach Fruit Ripening. *Journal of Agricultural and Food Chemistry*, **58**, 6157-6165.

# Capítulo 3:
# An integrative "omics" approach identifies new candidate genes to impact aroma volatiles in peach fruit

Los resultados presentados en este capítulo fueron publicados en un artículo científico y expuestos en un congreso internacional:

**-Sánchez G.**, Venegas-Calerón M., Salas J.J., Monforte A.J., Badenes M.L. and Granell A. (2013) **"An integrative "omics" approach identifies new candidate genes to impact aroma volatiles in peach fruit".** *BMC Genomics*, **14**, 343.

**-Sánchez, G.** Besada C., Badenes M.L., Monforte A.J. and Granell A. (2012). **"An integrative "omics" approach identifies new candidate genes to impact aroma volatiles in peach fruit".** *6th Rosaceous Genomics Conference.* San Michele all'Adige, Italia. (presentación oral)

# AN INTEGRATIVE "OMICS" APPROACH IDENTIFIES NEW CANDIDATE GENES TO IMPACT AROMA VOLATILES IN PEACH FRUIT.

Sánchez Gerardo[1,2]*, Venegas-Calerón Mónica[3], Salas Joaquín Jesus[3], Monforte Antonio[1], Badenes María Luisa[4] and Granell Antonio[1]*.

[1] Instituto de Biología Molecular y Celular de Plantas (IBMCP). Ingeniero Fausto Elio s/n, 46022, Valencia, Spain

[2] Instituto Nacional de Tecnología Agropecuaria (INTA), Ruta N°9 Km 170, 2930, San Pedro, Argentine.

[3] Instituto de la Grasa (IG-CSIC), Av. Padre García Tejero, 4, 41012, Sevilla, Spain

[4] Instituto Valenciano de Investigaciones Agrarias (IVIA), Carretera Moncada-Náquera, Km 4,5, 46113, Náquera (Valencia), Spain

*Corresponding authors:

Gerardo Sánchez (gsanchez@correo.inta.gov.ar)

Antonio Granell (agranell@ibmcp.upv.es)

**ABSTRACT**

**Background**: Ever since the recent completion of the peach genome, the focus of genetic research in this area has turned to the identification of genes related to important traits, such as fruit aroma volatiles. Of the over 100 volatile compounds described in peach, lactones most likely have the strongest effect on fruit aroma, while esters, terpenoids, and aldehydes have minor, yet significant effects. The identification of key genes underlying the production of aroma compounds is of interest for any fruit-quality improvement strategy.

**Results:** Volatile (52 compounds) and gene expression (4348 genes) levels were profiled in peach fruit from a maturity time-course series belonging to two peach genotypes that showed considerable differences in maturation characteristics and postharvest ripening. This data set was analyzed by complementary correlation-based approaches to discover the genes related to the main aroma-contributing compounds: lactones, esters, and phenolic volatiles, among others. As a case study, one of the candidate genes was cloned and expressed in yeast to show specificity as an ω-6 Oleate desaturase, which may be involved in the production of a precursor of lactones/esters.

**Conclusions:** Our approach revealed a set of genes (an alcohol acyl transferase, fatty acid desaturases, transcription factors, protein kinases, cytochromes, etc.) that are highly associated with peach fruit volatiles, and which could prove useful in breeding or for biotechnological purposes.

## INTRODUCTION

Peach (*Prunus persica* L. Batsch) was definitively positioned as a new fruit model when its genome was sequenced and released in 2010 by an international initiative (Arús et al., 2012; Verde et al., 2013). This is further supported by the availability of several genetic and genomic tools, including molecular markers, genetics maps, transient fruit expression assays, microarrays, EST databases, and a 9K SNP array (Ahmad et al., 2011; Arús et al., 2012; ESTree Consortium, 2005; Ogundiwin et al., 2008; Spolaore et al., 2001; Verde et al., 2012), and by the fact that peach fruit is a drupe and therefore has a different physiology, anatomy, and metabolism from other "post-genomic" fruit crops, such as grape and tomato (berries), strawberry (an aggregate of achenes), orange (speridium), and apple (a pome-type fruit). Hence, the peach represents an excellent opportunity to isolate novel genes related to specific traits, like aroma volatiles. Nevertheless, there still exist many obstacles to discovering gene function in peach. For example, *in-silico* analyses are still quite restricted when compared to other species, like *Arabidopsis,* for which more exhaustive phenotypic and molecular database repositories are available. The use of *a priori* knowledge (e.g., based on co-expression data) to select genes for functional analyses is very limited in peach, which makes it necessary to develop in-house data in order to identify the candidate genes associated with important traits or physiological processes.

Underlying its emerging role as a fruit model is the fact that peach is an important food commodity with an estimated net worldwide production of 11 billion US$ (FAO, 2010). Aroma is one of the main attributes that affects fruit quality (Bruhn et al., 1991) and has been recognized as one of the main factors that affect peach prices in the market (Parker et al., 1991). For this reason, volatile organic compounds (VOCs), which define aroma and, in combination

with sugars and organic acids, also contribute to fruit taste, have received a great deal of attention. More than 100 VOCs have been described in peach to date [(Sánchez et al., 2012) and references therein], of which about 25 of them appear to conform the typical peach aroma. In particular, γ- and δ-decalactone play a key role in association with C6 compounds, alcohols, esters, terpenoids, and phenolic volatiles (Eduardo et al., 2010). In addition to their contribution to fruit quality, peach volatiles are also important in the food and fragrance industry, where they are used as flavoring agents. Indeed, γ-decalactone is a sought-after industrial product that confers a "peach-like" odor (Dufosse et al., 1994) with an expanding annual world demand estimated at 10,000 Kg in 1997 (Gatfield, 1999). Despite the importance of lactones, their biosynthetic pathways in peach, and in plants in general, are still poorly understood (Schwab et al., 2008). An early study suggested that epoxide hydrolases were involved in lactone production, since it was observed that nectarines (a glabrous mutation of peach) are able to produce an artificial lactone when infiltrated with a synthetic, radiolabeled epoxy acid (Schöttler and Boland, 1996). The analysis of EST libraries later showed that a homologous gene to epoxyde hydrolases was expressed in peach skin (Vecchietti et al., 2009), although this gene has not been further characterized. Indeed, no gene involved in volatile production in peach has been reported to date. Most studies on genes related to peach aroma have focused on analyzing genes whose homologs are characterized in other plant species, i.e., literature-derived candidate genes. For example, Vecchietti et al. (Vecchietti et al., 2009) analyzed an EST library to show that a set of candidate genes was expressed in peach fruit and could thus be related to the formation of different volatile compounds. Another study targeted certain members of the carotenoid cleavage dioxigenase gene family for an expression analysis of genotypes differing in carotenoid accumulation to support their involvement in the production of norisoprenoid volatiles in peach (Brandi et al., 2011). The

identification of QTL for volatile compounds of peach was recently reported (Eduardo et al., 2012). The study in question also proposed putative candidate genes for minor contributors to peach aroma, such as linalool, *p*-menth-1-al, and nonanal, based on low-resolution co-localization of candidate genes within major QTL regions for these volatiles compounds (Eduardo et al., 2012). The use of "omics" technologies is needed to increase the resolution and accuracy of candidate-gene approaches in peach. Biosynthetic pathways for the main volatile compounds in peach are still poorly understood (Schwab et al., 2008), and given the peach-specific nature of the volatiles involved, information from other model species may not be sufficient.

Integrating transcriptomic and metabolomic data by means of pair-wise correlation analyses is emerging as a promising approach for discovering novel gene functions, despite its currently being limited to the model plant *Arabidopsis thaliana* [reviewed in (Saito and Matsuda, 2010)] and tomato (Carrera et al., 2012; Mounet et al., 2009). On the other hand, integrative "omics" approaches have been widely employed for the purpose of studying diverse physiological processes of plants (Amiour et al., 2012; Carrari et al., 2006; Casati et al., 2011; Lombardo et al., 2011; Wang et al., 2009) rather than discovering candidate genes.

In order to identify the key genes underlying the production of aroma compounds in peach, we conducted an integrative analysis using in-house-developed transcriptomic and metabolomic data sets. Our reductionist approach, based on correlation analysis [Hierarchical Cluster Analysis (HCA), Pearson Correlation (PC), and Correlation Network Analysis (CNA)] between gene expression and metabolite accumulation levels, allowed us to propose candidate genes for aroma control in peach. Moreover, we cloned and expressed in yeast one of the candidate genes identified herein to analyze its enzymatic activity. Our results demonstrate that the encoded protein possesses

ω-6 Oleate desaturase activity, which might participate in the biosynthesis of esters and lactones in peach fruit by generating a common precursor, linoleic acid. In addiction, our results provide some insights into the physiology and metabolism of peach, which are discussed. To the best of our knowledge, this is the first study to use non-targeted approaches to identify candidate genes in peach. The set of genes related to aroma compounds provided herein could prove useful for marker-assisted breeding and/or biotechnology improvements of peach fruit.

## MATERIALS AND METHODS

### Fruit material and analysis

Peach fruit (genotypes 'MxR_01' and 'Granada') were harvested in July 2009 from a local commercial orchard situated in Murcia, Spain. 'MxR_01' is a freestone melting-flesh peach which was obtained through the IVIA (*Instituto Valenciano de Investigaciones Agrarias*) breeding program. 'MxR_01' is a seedling from the cross between 'RedCandem' and 'Maruja', a traditional Spanish germplasm. 'Granada'is a clingstone non-melting peach low chilling obtained from a Brazilian breeding program. Fruits corresponding to four maturity stages (S1, S2, S3, and S4) were harvested from one tree per genotype. For S1, S2, and S3, ten fruits from each maturity stage were collected and maturity parameters were immediately analyzed. For S4, "harvest ripe" (or "ready to buy") stage, 20 fruits were harvested and divided into two groups. One group (denominated S4) was immediately analyzed and the other (denominated S4+SL) was subjected to shelf-life conditions (2 days at 20°C, 85% RH) prior to the maturity analysis. S4+SL corresponded to "consumption ripe" (or "ready to eat") stage. The maturity parameters (external color, flesh firmness, weight, total soluble solids (SSC), and ethylene and $CO_2$ production) were analyzed as described in Sanchez et al. (Sánchez et al., 2012). A Principal Component Analysis was carried out with maturity parameters in order to identify the three most homogeneous fruits per maturity stage for each genotype (data not shown). These three fruits (biological replicates) were selected for volatile and microarray analyses.

### Sample preparation and HS-SPME-GC-MS conditions

Frozen samples of fruit mesocarp were ground to powder in liquid nitrogen and used for volatile and microarray analyses as follows. Volatile compounds

were analyzed from 500 mg of frozen tissue powder, as previously described (Sánchez et al., 2012). The volatile analysis was performed on an Agilent 6890N gas chromatograph coupled to a 5975B Inert XL MSD mass spectrometer (Agilent Technologies). For the chromatography and mass spectra conditions, see Sánchez et al. (Sánchez et al., 2012). A total of 36 commercial standards were used to confirm compound annotation; the VOCs confirmed are listed in Fig. S2. Volatiles were quantified relatively by means of the Multivariate Mass Spectra Reconstruction (MMSR) approach developed by Tikunov et al. (2005). A detailed description of the quantification procedure is provided in Sánchez et al. (2012).

**RNA extraction**

RNA was extracted from 3 g of frozen tissue powder as described by Meisel et al. (Meisel et al., 2005). RNA quantity and purity were determined spectrophotometrically with Nanodrop (Nanodrop Technologies Inc.; http://www.nanodrop.com/). RNA integrity was verified by agarose gel electrophoresis. After the quantity and quality checks, RNA were used for the microarray analyses and the qRT-PCR as described below.

**Microarray hybridization and scanning**

The microarray analysis was performed essentially as described in Ogundiwin et al. (Ogundiwin et al., 2008). For the microarray hybridization, the RNA from the samples and the reference pool (1μg of each sample) were amplified and aminoallyl-labeled using the MessageAmp™ II aRNA Kit (Ambion, http://www.ambion.com) and 5-(3-aminoallyl)-20-deoxyuridine-50-triphosphate (aa-dUTP, Ambion), following the manufacturer's instructions. For each RNA sample, 7.5 μg of aminoallyl-labeled aRNA were re-suspended in 0.1 M Na2CO3 (pH 9.0) and labeled with either Cy5 or Cy3 Mono NHS Ester

(CyTM Dye Postlabelling Reactive Dye Pack, Amersham), respectively. Samples were purified with MegaclearTM (Ambion) following the manufacturer's instructions. Incorporation of Cy5 and Cy3 into probes was measured with a Nanodrop spectrofluorometer (Nanodrop Technologies Inc., http://www.nanodrop.com/). Microarray hybridization of samples and references to the ChillPeach microarray slides was performed manually using Telechem Hybridization Chambers (Corning), following the manufacturer's instructions. After hybridization, slides were washed in 2x SSC, 0.1% SDS for 5 min at 42°C, 0.1x SSC, 0.1% SDS for 10 min at room temperature, 0.1x SSC for 5 min at room temperature 4 times, and 0.01x SSC for 5 min at room temperature 4 times. Arrays were drained by centrifugation at 528g for 2 min. Slides were scanned with a GenePix 4000B scanner (Axon Instruments) at 10 μm resolution, 100% laser power and with different PMT values to adjust the ratio to 1.0. Microarray images were analyzed and globally normalized using the GenePix 4.1 software (Axon Instruments). Only the spots with a background-subtracted intensity greater than 2-fold the mean background intensity in at least one channel were selected for the analysis. Data files were imported into Acuity 4.0 (Axon Instruments) and normalized by the Lowess normalization method. Finally, only the spots with valid values in at least 2 of the 3 analyzed hybridizations were considered for further analysis. The means and standard deviations of the values were calculated from each sample as log2 values and were later normalized to the median of the reference pool.

**Real-time qRT-PCR analysis**

One microgram of total RNA was used to synthesize first-strand cDNA using the SuperScript first-strand synthesis system for RT-PCR (Invitrogen). Two microliters of diluted cDNA (100 ng/μL) were used for qRT-PCR using the SYBR Green PCR master mix (Applied Biosystems), following the

manufacturer's recommendations, and an ABI Prism 7000 sequence detection system (Applied Biosystems). Each biological replicate was assayed in triplicate. Gene-specific oligonucleotide primers were designed using the Primer Express_ version 2.0 software (Applied Biosystems). Primer information is available in Supplemental Table S1. The expression levels for target genes were calculated in relation to a reference gene by the DDthreshold cycle method (DDCt, Applied Biosystems). From the microarray data, the gene (ID: PPN078E12, http://bioinfo.ibmcp.upv.es/genomics/ChillPeachDB), whose profile was corroborated as being constant throughout both time courses, was selected as a reference for the normalization of all the subsequent qRT-PCR analyses. The relative gene expression for each candidate gene was expressed by means of the DDCT method using the expression value of the reference gene to normalize expression and the sample from stage S1 in each time-course series as reference samples.

**Data analysis**

The Acuity 4.0 software (Axon Instruments) was used for: hierarchical cluster analysis, heatmap visualization, principal component analysis, Pearson correlation evaluation, and for the Student's t-test for significant differences of volatile levels.

To detect differentially expressed genes between fruit at harvest (S4) and after shelf-life simulation (S4+SL) in both genotypes, data were analyzed with the SAM (Significance Analysis of Microarray) package (Tusher et al., 2001). Statistical significance was assessed using a two-class (unpaired) SAM analysis, with a false discovery rate of 5% and a q-value of <0.05.

Correlation network analyses were conducted with the Expression Correlation (http://www.baderlab.org/Software/ExpressionCorrela) plug-in for the Cytoscape software (Shannon et al., 2003). Network topological parameters

were calculated with the NetworkAnalyzer plug-in (Assenov et al., 2008). Networks were visualized with the Cytoscape software, v2.8.2 (www.cytoscape.org).

Venn diagrams were drawn with Microsoft PowerPoint.

**Cloning and bioinformatics analysis of the peach candidate gene**

The cDNA (from 'Granada' fruit at S4+SL) synthesized by qRT-PCR was used as a template for cloning the ORF of candidate gene PP1002E07. Coding sequences were amplified by PCR (FAD_F: ATGGGTGCCGGTGGAAGAAT and FAD_R: TTATAACTTATTATTGTACC) using Taq DNA polymerase (Biotools B and M Labs SA, Spain). PCR products were cloned in pCR®8 TOPO (Invitrogen, Spain), according to the manufacturer's instructions, in order to create the pEntry-PpFAD_1B-6 vector. Cloned ORFs were verified by sequencing both DNA strands.

For the sequence analyses, different tools were used. Protein sequences were aligned by the clustalW method using the MegAlign software (DNAStar). Transmembrane domains were predicted with TMpred (http://www.ch.embnet.org) and protein localization with ProtComp v9.0 (http://linux1.softberry.com/berry.phtml).

For yeast expression vector construction, the coding sequence was PCR-amplified from the pEntry-PpFAD_1B-6 vector with forward primer FAD-Kozack_F (GGCATGGGTGCCGGTGGAAGAAT) and reverse primer FAD_R. FAD-Kozack_F includes three non-template nucleotides at the 5′ end (underlined) to improve protein synthesis in yeast. The PCR product was cloned in pYES2.1 TOPO® vectors (Invitrogen) by TA cloning following the manufacturer's instructions. After checking the insert by sequencing, the plasmid was introduced into *Saccharomyces cerevisiae* strain W303-1A MATa

[leu2-3112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15} for expression analysis as follows.

**Expression of the peach candidate gene in yeast and fatty acid analysis**

Yeast cells were grown on a rotary shaker at 200 rpm at 30°C in synthetic defined (SD) medium containing raffinose as a carbon source and mixtures of amino acids and nucleoside precursors for marker selection. The expression of the transgenes was induced by the addition of galactose to 2% (w/v) or 1% glucose as a negative control. After 3 days of induction at 20°C, cells were collected, washed with water, and dried. Fatty acids were extracted and methylated from cell pellets and the latter were silanized to obtain trimethylsilyl derivatives, as previously described (Venegas-Calerón et al., 2010). The methyl-ester derivatives were analyzed by gas chromatography (GC) with a Hewlett–Packard 6890 gas chromatograph (Palo Alto, CA, USA). The column, chromatographic, and detection conditions are described in Venegas-Calerón et al. (2010).

## RESULTS

As part of our final goal of identifying genes and sources of variability to improve peach quality, we undertook complementary genetics and genomics approaches. 'MxR_01' and 'Granada' peach genotypes vary greatly for important traits, such as melting/non-melting, freestone/clingstone, chilling requirement, aromas, and fruit flavour. To exploit this variability, an F1 population was developed to analyze quantitative trait loci, which will be presented elsewhere. Here we present together the analyses of gene expression and volatile accumulation during fruit maturity and ripening of fruit of the parental genotypes in order to find aroma-related genes.

### Physiological and shelf-life ripening of peach genotypes

In order to characterize the ripening stages of the 'Granada' and 'MxR_01' peach genotypes, typical maturity parameters were evaluated and are presented in the supplementary data (Fig. S2). The results indicate that, as expected, peel ground color and weight increased with fruit ripening, whereas flesh firmness decreased, but with differences between varieties. Soluble solids content (SSC) was not affected throughout the study period, which indicates that this parameter is not a suitable indicator of ripening for either of the genotypes analyzed. Ethylene and $CO_2$ production was also monitored, as the evolution of these compounds reflects the physiological ripening stage of fruits and could also underlie the differences observed in ripening. Both genotypes, 'Granada' and 'MxR_01', showed increased ethylene production at mature stages (S3 for 'Granada' and S4 for 'MxR_01'), which is typical of climacteric fruit with differences in degrees (Fig. S2). Shelf-life ripening was also used to increase the complexity of our data set and to evaluate the effect on volatile and gene expression. Storage at 20°C for 2 days affected peel ground color, firmness, and ethylene production in both genotypes (Fig. S2). $CO_2$ production decreased

in both genotypes with shelf-life simulation, although differences were only significant for the 'MxR_01' genotype. These results indicate that our postharvest treatment was effective in stimulating ripening off the tree in the fruit of both genotypes and that, all together, our samples represent different stages of fruit development and ripening that can be interrogated to obtain the network of interactions between transcripts and volatiles.

**Non-melting ('Granada' genotype) and melting ('MxR_01' genotype) peaches showed different volatile evolution patterns during ripening**

A total of 52 volatile compounds, most of which contribute to peach aroma, were profiled in the fruit samples of the 'Granada' and 'MxR_01' genotypes at the different ripening stages and at postharvest treatment. The odor descriptors for the analyzed volatile compounds are provided in the supplementary data (Fig. S1). A heatmap and cluster analysis of the volatiles in the two time-course series are shown (Fig. 1A). Different trends in volatile evolution can be readily discovered by a simple inspection of the heatmap: the levels of some volatiles increased during ripening (Group I), while others decreased (Group II). Moreover, some compounds, i.e., those belonging to cluster (C) 8, exhibited high levels at specific maturity stages. In addition, some compounds displayed no specific trend during either time series (C9). Several volatile evolution patterns were identified for the compounds that showed increasing or decreasing levels during ripening (Fig. 1B), and compounds often grouped according to known biochemical pathways and/or chemical structure (Fig. 1A). Clusters C1, C2, and C3 are formed by lactones and some lineal esters (Fig. 1A). They showed similar overall trends in which compound levels increased in both genotypes, but with different profiles and fold changes (Fig. 1B).

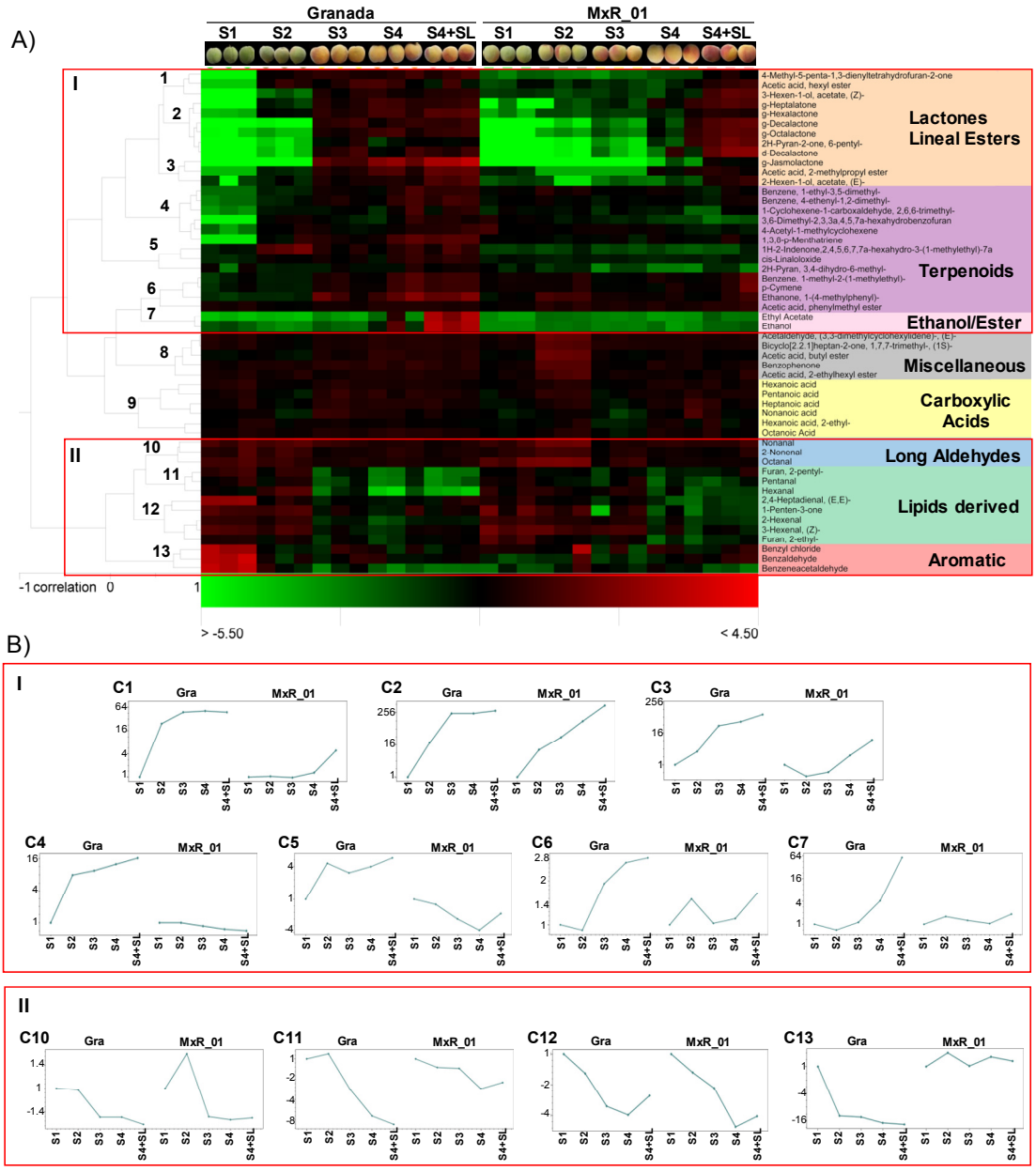



**Figure 1. Volatile evolution patterns in the 'Granada' (Gra) and 'MxR_01' time-course series**. **A)** The heatmap and cluster analysis of volatile compounds in the two time-course series (three replicates per stage are shown). The clusters observed (1-13) are indicated in the dendrogram. Data are expressed as log2 of a ratio (sample/common reference). Red boxes indicate those groups of compounds that increase (I) or decrease (II) during ripening. **B)** The evolution patterns for volatiles that change during ripening (I and II). An average profile of all the members of each cluster is shown. For each maturity stage (S1 to S4) and shelf-life response (S4+SL), the mean of the samples (n=3) is shown. Data are expressed as fold changes in relation to S1 (of each genotype) on the log2 scale. The positive region of the y axis is used for values higher than those of S1, while the negative region is for values that are lower than those of S1.

In 'Granada', the volatiles in these clusters reached high levels at S3 and remained high for the remaining time series (mainly for C1 and C2), while in 'MxR_01', the increment observed throughout the time-course series was gradual (mainly for C2 and C3). Clusters C4, C5, and C6 are formed mostly by terpenoid volatiles. They increased during ripening in 'Granada', while in 'MxR_01' they showed no changes, or even decreased (C5) during ripening.

Ethanol and Ethyl acetate (C7) levels rose at a mature stage (S4) and even reached higher levels with shelf-life conditioning (64-fold as compared to S1) for 'Granada', while in 'MxR_01', its content increased only after shelf life reached a 2-fold difference as compared to S1.

Volatiles that decreased during ripening (II) also showed different evolution profiles according to genotype (Fig. 1B). Cluster C10 is composed of aldehydes with eight and nine carbons (Octanal, Nonanal, and 2-Nonenal), which showed moderate changes during ripening (up to 1.5-fold changes in relation to S1, Fig. 1B). Clusters C11 and C12 are formed by the volatiles derived from the catabolism of linoleic and linolenic acids, the so-called "green compounds" (Fig. 1A). Both clusters displayed a similar decreasing trend in both time-course series (Fig. 1B). The phenolic volatiles Benzyl chloride, Benzaldehyde, and Benzeneacetaldehyde (C13) decreased from S1 to S2, and remained at low levels until the end of the time series for 'Granada', while they showed no significant changes in the 'MxR_01' time series.

Several aroma-related volatiles accumulated at different levels in 'Granada' and 'MxR_01' when fruits reached the commercial mature stage (S4), which means that they could display very marked differences in aroma. As shown in the supplementary data (Fig. S3), 13 volatiles showed significant differences, with most (12) exhibiting higher levels in 'Granada' (the more aromatic variety). Compound γ-jasmolactone, which has a characteristic peach-like odor, showed a 39-fold higher level in 'Granada' when compared to 'MxR_01'. Other

compounds with pleasant aroma descriptions included 2-Hexen-1-ol, acetate (E) (fruity), γ-Hexalactone (coconut), and 3-Hexen-1-ol acetate (Z) (fruity), which also had higher levels (between 3-fold and 5-fold) in 'Granada' compared to 'MxR_01'. The only compound with higher levels in 'MxR_01' at S4 was Benzeneacetaldehyde (4-fold higher than 'Granada' at S4), which has been described as confering a "green" aroma.

**The 'MxR_01' genotype showed an enhanced shelf-life response**

Due to commercial considerations (handling, long-distance transport, storage, etc.), peaches are generally harvested before their complete maturity, but the subsequent shelf life allows the fruit to ripen to the minimum quality threshold for consumers' acceptance. To analyze the shelf-life response of the 'MxR_01' and 'Granada' genotypes, we compared the changes in volatile content and gene expression.

Interestingly, the 'MxR_01' genotype showed a much enhanced volatile response as compared to the 'Granada' genotype when fruit were ripened under shelf-life conditions (Fig. 2). The principal component analysis of our volatile analysis of peach samples showed that the second principal component (explaining 19% of variance) clearly separated the fruit samples that were, or were not, subjected to shelf-life treatment in the case of 'MxR_01' (Fig 2A, left). However, the 'Granada' samples (S4 and S4+SL) were close to each other in the first two components (Fig. 2A, left), indicating a weaker postharvest response in relation to the response of 'MxR_01'. Principal component 1 explained 33% of total variance and mainly separated fruit samples according to differences in genotype (Fig. 2A, left). To graphically summarize how shelf-life simulation affects volatile production in both genotypes, a Venn diagram is shown in Figure 2B. For 'MxR_01', 10 volatiles showed significant differences (levels of 9 of the 10 volatiles increased) after shelf life, while for 'Granada', the levels of

only four volatiles changed significantly (once again, they all increased) after shelf-life simulation (Fig. 2B, left).

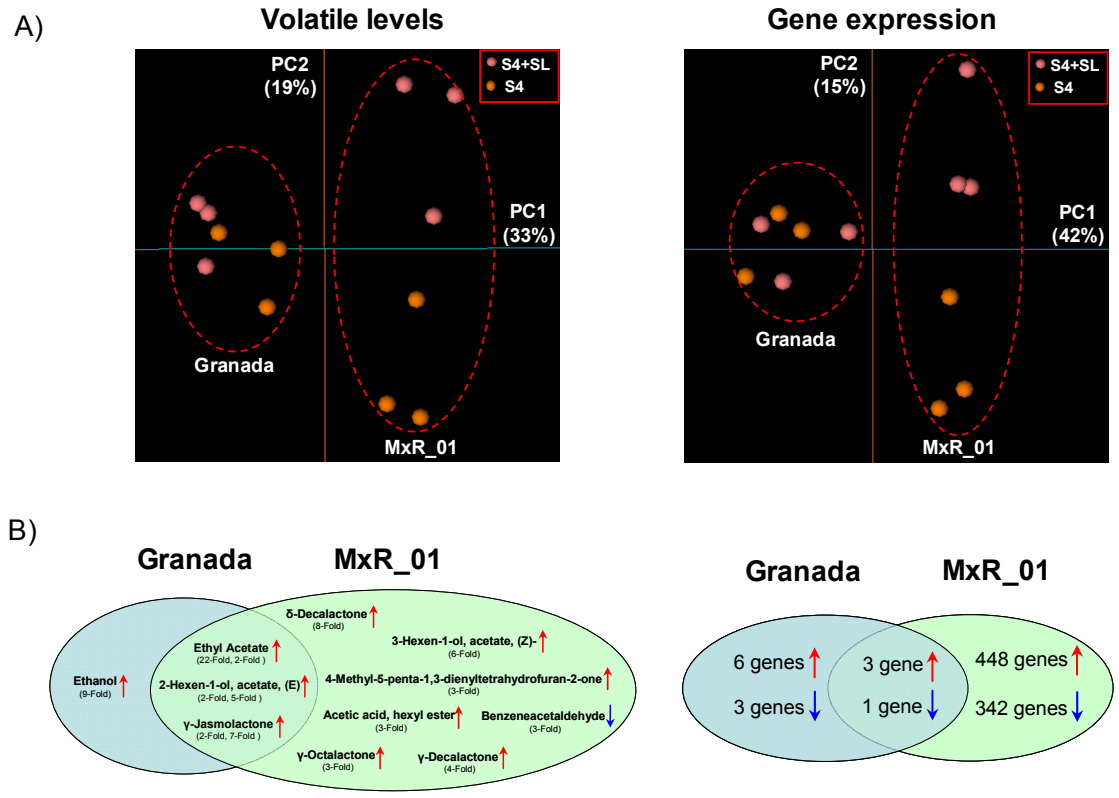



**Figure 2. Response of the 'Granada' and 'MxR_01' genotypes to shelf-life ripening. A)** A principal component analysis of the volatile compounds (left) and the gene expression (right) data sets. The first two principal components are shown (PC1 and PC2) for each analysis. Three replicates per genotype and treatment are shown. The red dotted circles indicate the genotype of the samples. **B)** Venn diagram of volatiles (left) and genes (right) that change significantly ($p<0.05$ for the volatile data, and FDR and a q-value of $<0.05$ for the gene expression data) after shelf-life ripening. Red and blue arrows indicate that the volatile or gene expression increase or decrease, respectively, after treatment. For volatiles, the fold change is indicated in parentheses below each one. For the three compounds that change in both genotypes, the fold change for 'Granada' is indicated on the left, while that for 'MxR_01' is on the right. For genes, the number of genes increasing or decreasing after treatment is indicated.

As expected, all the volatiles with increased levels after treatment in both genotypes have been described to impact the aroma of mature fruit (for a

description of the aroma of these volatiles, see Fig. S1). Accordingly, the levels of Benzeneacetaldehyde, which is associated with the aroma of immature fruit, decreased after postharvest treatment in 'MxR_01'. As a result, with shelf-life simulation (S4+SL), the initially flat-odor 'MxR_01' reached higher levels than 'Granada' (at S4+SL) for certain pleasant volatiles. This is the case of 2H-Pyran-2-one 6-pentyl, δ-Decalactone, 3-Hexen-1-ol acetate (Z)-, γ-Decalactone, γ-Octalactone, and γ-Heptalatone (Table S2). On the other hand, 'Granada' still had higher levels than 'MxR_01' for other mature fruit-related volatiles (Ethyl Acetate, γ-Jasmolactone, Acetic acid 2-methylpropyl ester, γ-Hexalactone, 2-Hexen-1-ol acetate (E), and Acetic acid butyl ester). Neither of the two genotypes showed a predominance of immature-related volatiles in the mature ripening stages, although some differences were detected between cultivars. While 'Granada' showed higher levels of 2,4-Heptadienal, (E,E)-, 1-Penten-3-one, and 2-Hexenal, 'MxR_01' displayed higher levels of Furan, 2-pentyl-, and Hexanal (Table S2). Moreover, 'Granada' at S4+SL showed higher levels for some terpenoid compounds (1,3,8-p-Menthatriene, cis-Linaloloxide, and 4-Acetyl-1-methylcyclohexene) than 'MxR_01' (Table S2).

To analyze the effect of shelf-life treatment on the transcriptome, we conducted a PCA analysis with the 4348-gene data set and studied the differentially expressed genes between S4 and S4+SL in both genotypes. The first component (PC1, explaining 42% of variance) separated samples according to genotype (Fig. 2A, right). The second component (PC2) explained 15% of variance and separated between the S4 and S4+SL replicates of the 'MxR_01' genotypes. While the 'Granada' samples remained close together in the first 2 components space, the 'MxR_01' samples separated according to treatment (Fig. 2A, right), indicating that shelf-life has a greater impact on gene expression in the 'MxR_01' genotype than in the 'Granada' one. Accordingly, the direct analysis of the differentially expressed genes between S4 and S4+SL in 'Granada' and

'MxR_01' revealed a stronger gene expression response to the shelf-life in the 'MxR_01' genotype (Fig. 2B, right). By taking a False Discovery Rate (FDR) and a q-value of < 0.05 as criteria, we found that for 'Granada', 13 genes were differentially expressed after treatment (nine up-regulated and four down-regulated), while the 'MxR_01' genotype fruits showed a drastic change in gene expression after shelf-life simulation, with 794 differentially expressed genes (451 up-regulated and 343 down-regulated). Only four genes (three up- and one down-regulated) showed the same trend in both genotypes (Fig. 2B, right, Table S3). The gene showing the largest changes after shelf-life in both genotypes (id: PPN044F04) had a calcium-binding EF hand domain. A gene related to the metabolism of amino acids (Aspartate aminotransferase) showed moderate changes in both genotypes, increasing after treatment. The third gene which increased after shelf-life (id: PPN064C01) is a homolog to a translation initiation factor, suggesting that protein synthesis could be stimulated by treatment in both genotypes. Only one gene decreased after shelf-life in both genotypes, an acid phosphatase belonging to class B. The 'MxR_01' genotype showed a higher decrease in the expression of this gene (3.2-fold) compared to 'Granada' (1.8-fold).

These results support that mostly 'MxR_01' fruits undergo significant molecular, and consequently physiological, changes during postharvest conditioning, and that very little happens to 'Granada'. It is worth mentioning that despite the larger number of genes showing significant differences in expression after shelf-life in 'MxR_01' (794 genes vs. 13 genes in 'Granada'), most involved only minor fold-change differences. Of the genes, 9% (74:794) showed a fold change of over 2 (between S4 and S4+SL) for the MxR genotype. For 'Granada', nine of the 13 differentially expressed genes showed fold differences over 2. In the supplementary data (Table S4), a complete description of differentially expressed genes with a 2-fold change cutoff is provided.

**Identification of the genes related to peach volatile compounds by hierarchical cluster analyses**

In order to obtain a global view of the interrelationships between volatile accumulation and gene expression, a Pearson correlation analysis was conducted for all the possible volatile gene pairs (Table S5). Volatiles showed wide variability in the range of the distribution of volatile-gene correlations (data not shown). For example, Hexanoic acid 2-ethyl- had correlations ranging from -0.42 to 0.52, while γ-Jasmolactone showed correlations ranging from -0.92 to 0.96. Thus, our data set contains some volatiles which have no clearly associated genes, while others show many strongly correlated genes.

In order to identify the genes related to a family of co-regulated volatiles, a Hierarchical Cluster Analysis (HCA) was conducted with the complete data set (4348 genes and 52 volatiles x 30 samples). The Pearson correlation coefficient was used as the similarity metric and a complete linkage method was analyzed in order to find the variables (i.e., VOCs or genes) correlating with all the members of each cluster. As shown in Fig. S4B, most of the volatiles remained associated with the same cluster members, as previously described in Fig. 1 (clusters C2, C5, C6, C7, C8, C9, C10, and C11), while other metabolites (those belonging to clusters C1, C3, C4, C12, and C13) correlated better with transcripts as compared to other volatiles and formed new clusters. The metabolites belonging to clusters C6, C8, C9, C10, and C11 correlated poorly to transcripts in the HCA, as the correlation values with gene clusters were below 0.6. Other volatiles (belonging to C2, C5, and C7) fall into the same cluster, and correlate well with some genes (implying higher cluster correlations than 0.6). For the volatiles in cluster C4, four genes (PP1009D02, PPN027C05, PPN063A04, and PPN069C04) were found in the same sub-cluster, indicating a good association of these genes with all of these volatiles. In contrast, the volatiles in clusters C1, C3, and C13 (according to Fig. 1) correlated better with certain

transcripts than with other volatile members of the cluster. Based on these results, a series of genes corresponding to transcripts whose levels were highly correlated to volatile levels was selected. A list and short description of the selected genes is provided in Table 1, with further details provided in the supplementary data (Table S6). Certain genes, despite their being present in the same cluster, were not selected as candidate genes for the associated volatile because their functional annotation suggested that the genes could affect general ripening. For example, all the members of cluster C2 are lactones and were associated with a cluster of five genes (Fig. 3).

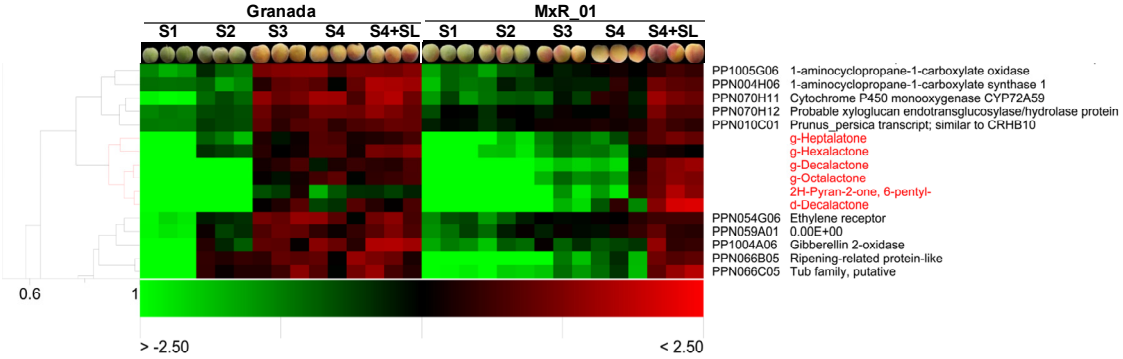



**Figure 3. Hierarchical cluster analysis for indentifying correlated genes**. Detail of the heatmap and cluster analyses where C2 lactones (in red) are found. For each gene, the id and the unigene annotation are provided. Three replicates per maturity stage are shown. Data are expressed as a log2 of a ratio (sample/common reference). The gene annotated as "0.00E+00" indicate that either no homologue was found or the homologue found has unknown function. For a detailed description of ChillPeach unigene functional annotation see Ogundiwin et al. (2008).

Two of them are genes relating to ethylene biosynthesis (PP1005G06 and PPN004H06) and another is related to cell wall softening (PPN070H12). Even though this group of volatiles and genes shared a common regulation (i.e., the ripening process), they are clearly not good candidates for the specific control of lactone accumulation. They are probably associated with a general control of other ripening aspects and were, therefore, not selected in Table 1. A

transcription factor with no homologs in *Arabidopsis* (PPN010C01) and a cytochrome P450 monooxygenase (PPN070H11), which is probably involved in lipid metabolism, were selected as putative candidates.

**Genes related to the main aroma compounds of peach**

Although aroma is a complex trait, and each volatile contributes to defining overall peach aroma, some compounds appear to be more important than others according to sensory assessments (Fig. S1). For example, the aroma of two lactones, γ-Decalactone and γ-Jasmolactone, is described as peach-like (Derail et al., 1999). We have evaluated correlations between the levels of gene transcripts and the levels of these two aroma volatiles across the different samples in order to find genes related to them that were not identified in the previous analysis. Table S7 provides the ten genes with the highest direct correlation values; in addition, the full list of genes with correlation values between the corresponding transcript and volatile levels is included in the supplementary data (Table S5). As expected, some of the genes identified in the HCA were also found in this analysis. For example, genes related to ethylene signaling proved to be highly correlated with γ-decalactone (PPN054G06 and PPN004H06) and γ-Jasmolactone (PP1005G06). Cytochrome P450 (PPN070H11) was strongly correlated with both lactones. γ-Jasmolactone also showed high correlations with a protein involved in metabolite transport (PPN032F06) and with two protein kinases (PPN069F09 and PPN010B11). Nevertheless, two other genes that were not previously identified (PPN059A01 and PPN026D01) with no homology in *Arabidopsis* showed the highest correlations with γ-decalactone (r=0.90 and r=0.87, respectively). Moreover, γ-Jasmolactone was highly correlated with the genes showing no homology with *Arabidopsis* (PPN002B03 and PPN031G12), and which were not identified in previous analyses. Interestingly, some genes involved in endoplasmatic reticulum traffic were also

associated with lactone production in peach fruit. A homolog to a Phosphatidylinositol transfer protein (PPN037A04) correlated well with γ-decalactone, while a novel plant SNARE 13 (PPN031F12) highly correlated with γ-Jasmolactone. Finally, γ-Jasmolactone showed a good correlation with the level of a transcript that encodes a protein related to the caleosin family (PP1004C02).

**Correlation network analysis for discovering hub genes**

A correlation network analysis (CNA) was conducted to further analyze the data and to identify additional genes related to aroma compounds that could have gone unnoticed in previous analyses. The CNA offers several advantages when compared to HCA or Pearson correlation analyses, which evaluate correlations one by one, since it 1) provides a complete view of the interaction between groups of variables and 2) allows the identification of variables (genes or volatiles) that play a central role in the network, the so-called hubs. Nevertheless, as the number of nodes increases, the interactions between them also increase, conceivably causing the network to become incomprehensibly complex at a certain point. To overcome this problem, a new data set comprising those genes with a correlation coefficient higher than 0.85 and with at least one VOC was created. First, a correlation network considering only the 52 VOCs was constructed. As expected, the familiar volatile clustering observed in the HCA (Fig. 1) could be readily identified as groups of highly interconnected nodes (Fig. 4A, left) in our CNA. Yet an additional interconnection between those groups emerged. Clusters C1, C2, and C3, all formed by lactones and linear esters, are interconnected by strong direct correlations ($r>0.7$). C3 is directly related to Ethanol and its ester (C7) and inversely related to lipid-derived volatiles (C11 and C12).

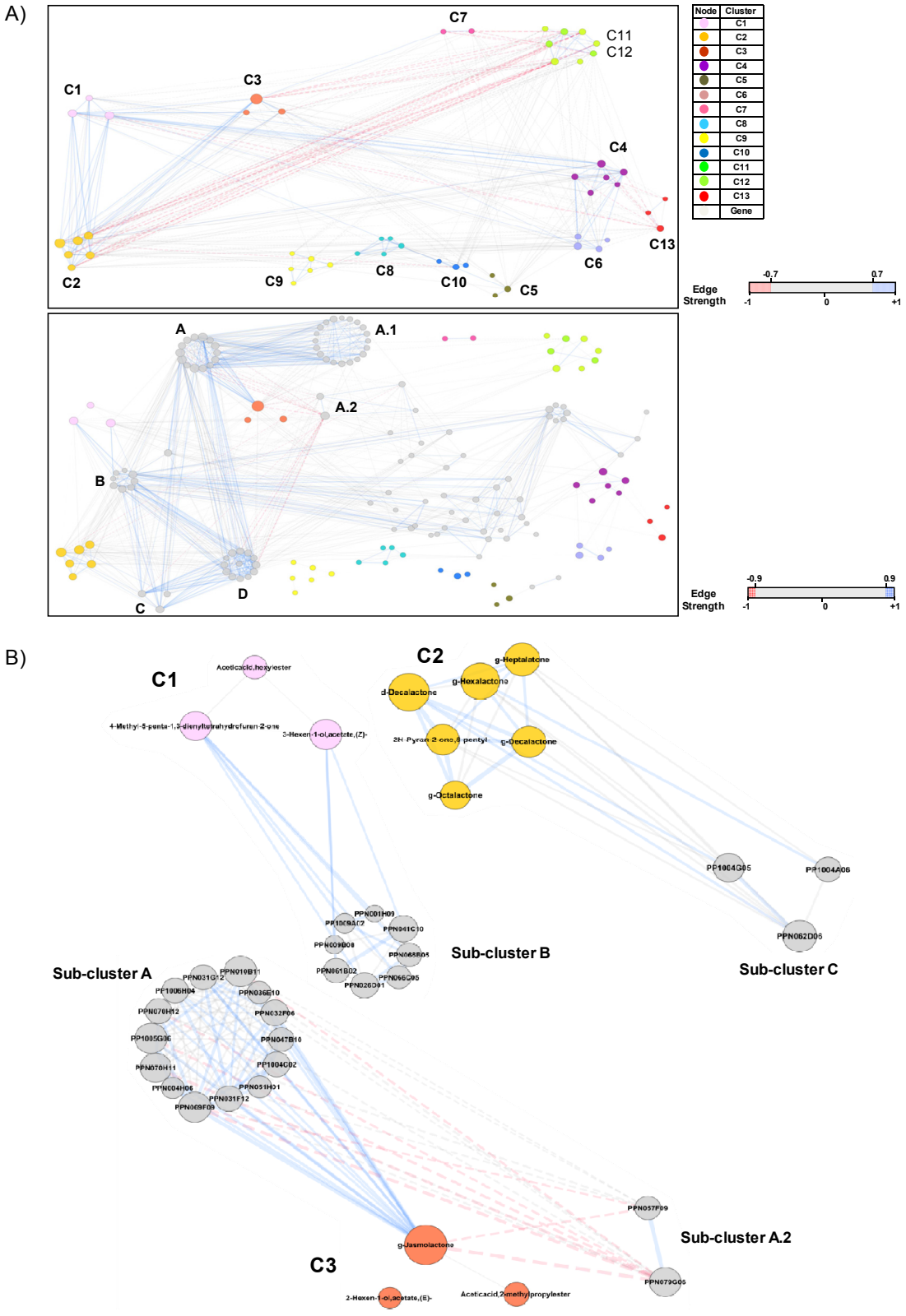



**Figure 4. Correlation network analysis for identifying candidate genes. A)** Network of volatiles (upper) and the integrated network of VOCs and genes (bottom). The nodes representing volatiles are colored according to the cluster that they belong to (C1-C13 according to Fig. 1), as indicated. Genes are

represented as gray nodes. Edges are colored according to their strength. Note that edges have different codifications for each network, which is indicated below each one. Node size indicates its connectivity; the bigger the node, the higher the connectivity measured as node degree (i.e., the number of edges connecting the node). A to E indicate the main gene sub-clusters. **B)** Magnification of volatile groups (C1, C2 and C3) showing the interactions with genes in detail. For each gene node, the id is shown. The list of genes belonging to sub-clusters A, A.1, A.2, B, C, and D, as well as some topological parameters, are provided in Table S7.

The volatiles of C2 are also inversely correlated with compounds C11 and C12. C1 metabolites are directly related to C4, which, in turn, positively correlates with C6 and negatively correlates with C13. The metabolites belonging to the other clusters (C5, C8, C9, and C10) are interconnected to various other clusters, but show weaker correlations ($r<0.70$ in absolute values).

As mentioned previously, in order to analyze interactions with genes, a correlation network was constructed by preselecting those genes with a cutoff of $r>0.85$ in absolute values. To integrate the interactions between genes and VOCs, the two networks were merged to create a new one (Fig. 4A, right). The resulting network consisted of 160 nodes (52 VOCs and 108 genes) and 1364 edges (correlations). The network clearly reflected the different distributions of the correlations between VOCs and genes observed. As the volatiles from C1, C2, and C3 showed the largest number of strong correlations with genes (both positive and negative), their nodes were highly connected to gene nodes (Fig. 4A, right). Compound γ-jasmolactone showed strong correlations ($r>0.95$) with a group of 14 genes (sub-cluster A, Fig. 4B). Some members of this sub-cluster had already been selected as candidate genes for the metabolites of C3 (PPN069F09, PPN032F06, PPN031F12, PPN036E10, PPN031G12, PPN010B11, and PP1004C02) or for lactones of C2 (PPN070H11), as described in previous sections. Nevertheless, one gene, which our previous analysis did not reveal, was selected as a candidate gene (PPN051H01, Table 1). Sub-cluster A was

related to a group of 20 genes (sub-cluster A.1) by strong positive correlations. It is interesting to note that some of the genes of sub-cluster A are involved in signal transduction processes (Table S8). For example, there are two putative protein kinases (PPN069F09 and PPN010B11) that show high connectivity (degree=39 and 40, respectively), as well as two ethylene signaling genes (PP1005G06 and PPN004H06). A gene with homology to Asparagine synthetase (PPN079G06) is strongly and inversely correlated with γ-jasmolactone and could be considered a hub given its high connectivity (degree=37, Table S8, Fig. 4B). Volatile cluster C1 was highly related to sub-cluster B by strong correlations with 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one and 3-Hexen-1-ol acetate (Z)- (Fig. 4B). Two transcription factors belong to this sub-cluster: one had already been selected (PP1009A02) and the other was a newly identified one (PPN066C05, Table 1). Furthermore, a gene that was not identified previously as showing homology to ripening-related proteins was selected (PPN066B05, Table 1). Sub-cluster B is highly interconnected with sub-cluster C, which consists of three genes with strong correlations with the lactones of C2 (Fig. 4B). The members of this sub-cluster are: a gene related to Gibberellin metabolism (Gibberellin 2-oxidase, PP1004A06), a gene that is likely related to cell wall physiology (UDP-arabinose 4-epimerase 1, PPN062D06), and a gene with no homolog in *Arabidopsis* (PP1004G05), which had already been identified (Table S8). The main sub-clusters that correlated with lactones and esters (A, B, and C) were highly interconnected to a group of 13 genes that formed sub-cluster D (Fig. 4A right, Table S8).

With the selected cutoff value (genes with a correlation higher than 0.85), the correlation network analysis failed to identify genes related to the other VOC clusters (C4 to C13). To gain insight into the genes associated with these volatile compounds, a new data set was composed by selecting genes after lowering the cutoff to > 0.8 for volatiles belonging to clusters C4 to C13, which allowed a new

correlation network to be constructed (Fig. S5). The aromatic VOCs in C13 are related to a putative tyrosine (Tyr) aminotransferase (PPN018G03) through a direct correlation with Benzeneacetaldehyde (Fig. S5B) and, therefore, was selected as a candidate gene (Table 1). The VOCs from C4 are correlated with a group of 5 genes, two of which are related to hormone signaling. One is an Auxin-responsive protein (IAA16, PP1009D02) and the other (PP1000F05) belongs to a family of proteins regulated by gibberellins. These 5 genes correlated well with a group of 42 genes (sub-cluster E), and some were also associated with auxin (PPN044D01 and PP1005H05) and gibberellin (PPN037E06) signal transduction pathways (Fig. S5A). In more detail, we see that the lipid-derived compounds Furan, 2-pentyl-, and Hexanal inversely correlated with a lipid delta 9 desaturase homolog (PPN052H12), which, in turn, was highly correlated with a gene with no homology in *Arabidopsis* (PPN023E05). A BZIP-like transcription factor (PPN019F11) strongly correlated with PPN023E05 and with some genes of the sub-cluster E. Compound cis-Linaloloxide (belonging to C5) also strongly correlated with a gene that had a short-chain dehydrogenase/reductase (SDR) domain (PPN053G07). This gene formed a sub-cluster with two other genes, one which also had a short-chain dehydrogenase/reductase (SDR) domain (PPN078H04) and the other with no homologs in *Arabidopsis thaliana* (PPN029H12). The previous selection strategy also revealed these three genes as candidates (Table 1).

| Genes for volatile cluster C1 | | | |
|---|---|---|---|
| **id** | **Unigene annotation** | **Identified by** | **Match** |
| PPN009B08 | Protein At1g45233 | HCA, CNA | ppa001502m |
| PPN078H01 | Acid phosphatase 1 precursor (EC 3.1.3.2) (Apase-1(1)) | HCA | ppa010063m |
| PP1009A02 | CAMP response element binding (CREB) protein | HCA, CNA | ppa012739m |
| PPN001H09 | YUP8H12R.39 protein | HCA, CNA | ppa005452m |
| PPN066C05 | Tub family, putative | CNA | ppa006121m |
| PPN066B05 | Ripening-related protein-like | CNA | ppa011478m |
| | | | |
| **Genes for volatile cluster C2** | | | |
| **id** | **Unigene annotation** | **Identified by** | **Match** |
| PPN010C01 | Prunus_persica transcript; similar to CRHB10 [Ceratopteris richardii (Triangle waterfern)] | HCA | ppa008984m |
| PPN070H11 | Cytochrome P450 monooxygenase CYP72A59 | HCA, PC, CNA | ppa006310m |
| PPN059A01 | 0.00E+00 | PC | ppa013582m |
| PPN026D01 | no_annotation_available | PC, CNA | ppa009231m |
| PPN037A04 | Phosphatidylinositol transfer protein/retinal degeneration b protein | HCA, PC | ppa010020m |
| | | | |
| **Genes for volatile cluster C3** | | | |
| **id** | **Unigene annotation** | **Identified by** | **Match** |
| PP1002E07 | Omega-6 fatty acid desaturase | HCA | ppa007098m |
| PP1001A05 | Pyruvate decarboxylase 1 | HCA | ppa003086m |
| PP1004F06 | Vesicle-associated membrane protein 727 | HCA | ppa010737m |
| PP1004G05 | no_annotation_available | HCA, CNA | ppa004658m |
| PP1004H08 | no_annotation_available | HCA | ppa004582m |
| PP1005B05 | AT4g00090/F6N15_8 | HCA | ppa004482m |
| PPN009C07 | AJ533535 S3II Prunus persica cDNA clone PP_S3II_T5_SP6, mRNA sequence | HCA | ppa004933m |
| PPN010B11 | Serine-threonine protein kinase, putative | HCA, PC, CNA | ppa008251m |
| PPN070B11 | Kinase-like protein | HCA | ppa003659m |
| PPN070D11 | Protein At5g25010 | HCA | ppa009131m |
| PPN075B07 | F8K7.22 protein | HCA | ppa008595m |
| PPN032F06 | PDR-like ABC-transporter | HCA, PC, CNA | ppa000267m |
| PPN036E10 | At5g03345 | HCA, CNA | ppa015891m |
| PPN069F09 | PK11-C1 | HCA, PC, CNA | ppa006108m |
| PPN002B03 | Malus_x_domestica transcript; similar to F17O7.2 [Arabidopsis thaliana] | PC | ppa007712m |
| PPN031F12 | Novel plant SNARE 13 | HCA, PC | ppa010000m |
| PPN031G12 | Os09g0509300 protein | PC, CNA | ppa002575m |
| PP1004C02 | Calcium binding protein | PC, CNA | ppa011722m |
| PPN051H01 | At1g72790/F28P22_2 | CNA | ppa002494m |
| | | | |
| **Genes for volatile cluster C4** | | | |
| **id** | **Unigene annotation** | **Identified by** | **Match** |
| PP1009D02 | IAA16 protein | HCA, CNA | ppa011570m |
| PPN027C05 | PHB2 | HCA, CNA | ppa009505m |
| PPN063A04 | T25K16.5 | HCA | ppa011444m |
| PP1000F05 | Snakin-1 | CNA | ppa014086m |
| PPN017C08 | ECA1 protein | CNA | ppa012387m |
| PPN017B08 | AT3g05350/T12H1_32 | CNA | ppa002173m |
| | | | |
| **Genes for volatile cluster C5** | | | |
| **id** | **Unigene annotation** | **Identified by** | **Match** |
| PPN029H12 | AJ827262 S3 Prunus persica cDNA clone S312E11, mRNA sequence | HCA, CNA | ppa025397m |
| PPN053G07 | Putative pod-specific dehydrogenase SAC25 | HCA, CNA | ppa008709m |
| PPN066C10 | Cytochrome P450 | HCA | ppa004664m |
| PPN078H04 | Putative pod-specific dehydrogenase SAC25 | HCA, CNA | ppa008713m |
| | | | |
| **Genes for volatile cluster C7** | | | |
| **id** | **Unigene annotation** | **Identified by** | **Match** |
| PPN030E11 | PP_LEa0004N09f Peach developing fruit mesocarp | HCA | ppa002860m |
| PPN071H07 | Enolase 1 | HCA | ppa005779m |
| PPN027H11 | Phox-like | HCA | ppa001676m |
| | | | |
| **Genes for volatile cluster C11** | | | |
| **id** | **Unigene annotation** | **Identified by** | **Match** |
| PPN052H12 | Desaturase delta 9 | CNA | ppa009359m |
| | | | |
| **Genes for volatile cluster C12** | | | |
| **id** | **Unigene annotation** | **Identified by** | **Match** |
| PP1003B11 | 0.00E+00 | HCA | ppa010788m |
| PP1009F04 | PU1_plate44_F06 PU1 Prunus persica cDNA weakly similar to putative protein (AL161573) | HCA | ppa009611m |
| PPN006F05 | GTP-binding protein SAR1B | HCA | ppa011838m |
| PPN033D11 | Putative acid cluster protein 33 | HCA | ppa006430m |
| | | | |
| **Genes for volatile cluster C13** | | | |
| **id** | **Unigene annotation** | **Identified by** | **Match** |
| PP1001G06 | Cysteine protease CP1 | HCA | ppa005199m |
| PP1004F12 | Dehydrin | HCA | ppa005514m |
| PPN018G03 | Putatative tyrosine aminotransferase | CNA | ppa006368m |

**Table 3 (see the legends below).**

**Table 1. Genes that highly correlate with volatile compounds proposed as candidate genes.** Genes were identified by either: a Hierarchical Cluster Analysis (HCA), a simple Pearson Correlation (PC) analysis and/or a Correlation Network Analysis (CNA). For each gene, microarray id (id), the unigene annotation, and the most similar gene predicted in the peach genome (Match) are shown. The genes annotated as "0.00E+00" indicate that either no homologue was found or the homologue found has unknown function. For a detailed description of ChillPeach unigene functional annotation see Ogundiwin et al. (Ogundiwin et al., 2008).

**Validation of microarray data by qRT-PCR analysis**

In order to validate the expression profile of the candidate genes identified by microarray analysis, gene-specific qRT-PCR analyses were conducted. We selected a subgroup of genes that we identified by their association with the main aroma-contributing volatiles, i.e., lactones and esters. Seven genes selected for clusters C1, C2, and C3 (PP1002E07, PPN070H11, PPN001H09, PPN032F06, PPN002B03, PPN059A01, and PPN066B05) were analyzed. Three genes that were inversely correlated with lactones/esters (PP1002D12, PPN079G06, and PPN008D02, Table S5) were also included to obtain a better validation of our results. The Pearson correlation coefficient between the expression results obtained with both RNA profiling technologies was evaluated for all the genes studied. The ten genes analyzed by transcript-specific qRT-PCR analysis corroborated the microarray analysis, showing correlation coefficients ranging from 0.710 to 0.988 (Table S1). By way of example, the comparison between the qRT-PCR and the microarray profiles of gene PPN001H09 is shown in Figure S7.

**The enzyme encoded by candidate gene PP1002E07 displays ω-6 Oleate desaturase activity**

We conducted functional studies for one of the identified candidate genes. The ORF of candidate gene PP1002E07 was amplified by RT-PCR from the RNA of 'Granada' fruits at S4+SL and was then cloned. The cloned ORF consisted of

1149 bp and was named *PpFAD_1B-6* (KC169941). The Blastp analysis on the NCBI web revealed that the deduced protein sequence contains the conserved Delta12-FADS-like domain (cd03507), typical of fatty acid desaturases, but also of other FAD-like hydroxylases and epoxidases. The presence of the typical functional elements of FAD-like enzymes was analyzed using bioinformatics tools: six transmembrane domains were identified by TMpred (data not shown) and three His motifs were identified by a sequence alignment of proteins with characterized FAD-Like enzymes (Fig. S7). ProtComp predicted that the PpFAD_1B-6 protein would be localized in the endoplasmatic reticulum (score: 7.81), as is to be expected for this kind of enzyme. To define the enzymatic activity and substrate specificity of PpFAD_1B-6, a DNA fragment encompassing the complete ORF was conditionally expressed in *S. cerevisiae*. Yeast is particularly suitable for this assay as it normally produces neither polyunsaturated nor hydroxylated fatty acids, which are the expected products. The main saturated fatty acids [palmitic acid (16:0), stearic acid (18:0)] and the corresponding unsaturated fatty acids [palmitoleic acid (16:1), oleic acid (18:1), linoleic acid (18:2), and linolenic acid (18:3)], as well as a hydroxylated fatty acid [ricinoleic acid (18:1-OH)] were analyzed. As shown in Figure 5, under inducing conditions, the yeast transformed with the *PpFAD 1B_6* construct to produce linoleic acid (18:2) at levels representing up to 6.6% of total fatty acids. Under non-inducing conditions, linoleic acid (18:2) was detected at only trace levels, representing less than 0.8% of total fatty acids, whereas in the control yeast strain (transformed with a vector where the candidate gene was replaced with the LacZ gene), linoleic acid (18:2) was not detected (less than 0.1%± 0.09), indicating that the inducible control was probably leaky. After the induction of *PpFAD 1B_6*, oleic acid (18:1) decreased from 44.2% (±0.3) to 34.5% (±0.3) of the total fatty acids content (Fig. 5), indicating that *PpFAD 1B_6* had used endogenous oleic acid as a substrate in the production of linoleic acid and was,

therefore, acting as an ω-6 Oleate desaturase. The control strain showed no differences in oleic acid (18:1) levels after induction. Ricinoleic acid (18:1-OH) was not detected in the yeast expressing *PpFAD 1B_6* (Table S9), suggesting that this enzyme has no Oleate 12-hydroxylase activity. The induction of *PpFAD 1B_6* did not affect the levels of palmitic acid (16:0), estearic acid (18:0), or palmitoleic acid (16:1) (Table S8), indicating that the encoded enzyme is not capable of introducing a double bond into unsaturated palmitic and oleic acids. Therefore, our results indicate that the candidate gene PP1002E07 encodes functional ω-6 Oleate desaturase activity.

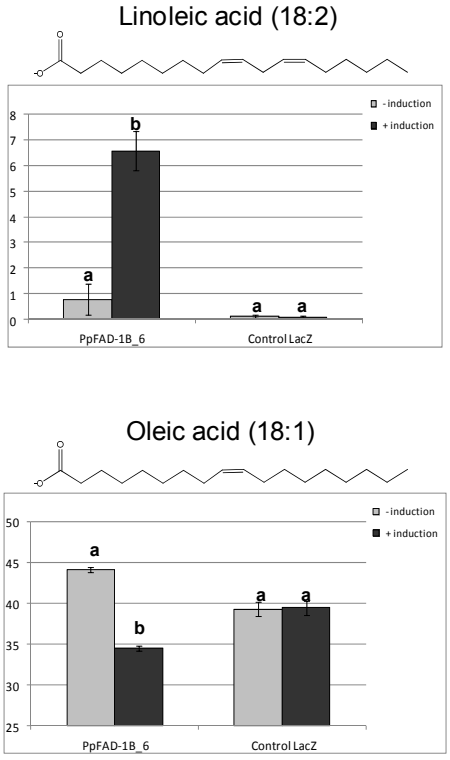



**Figure 5. Content of Linoleic acid (upper) and Oleic acid (bottom) in yeast expressing *PpFAD_1B-6*.** Fatty acid was analyzed in yeast cultures under non-inducing (-) or inducing (+) conditions. Values are provided as a % of total fatty acid content (mean of n=3). Differences between treatments (- induction vs. + induction) were stated by ANOVA ($p<0.05$) and are indicated with different letters. Error bars indicate standard deviation.

## DISCUSSION

### Non-melting ('Granada' genotype) and melting ('MxR_01' genotype) peaches show different maturity/ripening and shelf-life responses, which can be exploited to analyze co-regulation patterns

An "omics" approach based on integrative transcriptomics- and metabolomics-derived data was undertaken to identify those genes putatively involved in volatile production in peach fruit. The main aroma-contributing volatiles of peach increase during maturity and ripening following different patterns (Aubert et al., 2003; Chapman et al., 1991; Visai and Vanoli, 1997; Zhang et al., 2010). It is also known that the peach transcriptome undergoes significant re-organization during fruit ripening and maturity (Trainotti et al., 2006; Trainotti et al., 2007). Thus, our rationale was that the genes involved in aroma-related volatile production could be discovered by analyzing the co-regulation between gene expression and volatile accumulation during maturity and ripening. In addition, we hypothesized that the inclusion of different peach types could improve the robustness of metabolite-gene relationships since the gene-volatile correlation should hold across the genotype with different ripening characteristics. To this end, four maturity stages were analyzed for clingstone non-melting ('Granada') and freestone melting-flesh ('MxR_01') peaches (Fig. S2). To add further complexity to the data set and to increase the robustness of our model, one postharvest treatment was also included, since it is known that non-melting and melting peaches respond differently to shelf-life ripening (Manganaris et al., 2006). The volatile profiles analyzed in both time-course series confirmed that these genotypes show a different volatile evolution during maturity and ripening (Fig. 1), suggesting that our hypothesis was correct. Each genotype also showed a different response to shelf-life simulation. Volatile profiles were more affected after treatment in 'MxR_01' as compared to

'Granada', although the lactones and esters in both genotypes increased after treatment (Fig. 2), which is in agreement with previous studies (Aubert et al., 2003; Zhang et al., 2010). Correspondingly, 'MxR_01' also showed a more dramatic change in gene expression as compared to 'Granada' after the shelf-life simulation (Fig. 2), indicating that the restructuration of volatile content after shelf-life could be, at least in part, determined by a modification in gene expression.

Therefore, our approach combined different genetic backgrounds with a development-regulated process and an artificial post-harvest treatment in order to analyze co-regulation patterns. This strategy differs substantially from other previously used integrative approaches. In tomato, the change in gene expression and metabolite accumulation throughout fruit development was also used to identify not only candidate genes, but also variability between different fruit tissues (mesocarp and locular), which was included to improve the robustness of the study (Mounet et al., 2009). Carrera et al. (Carrera et al., 2012) applied an alternative strategy consisting in profiling (transcripts and metabolites) an RIL population to find candidate genes that affect tomato fruit aroma by modeling the "omics" data obtained. In this case, system perturbation was given by the genomic region introgressed in each line. In other study, *Arabidopsis* plants were starved in order to affect the transcriptome and the metabolome so that they could analyze the co-regulation patterns and thereby identify the genes associated with different metabolites (Hirai et al., 2005; Hirai et al., 2007). The variation in gene expression and the volatile levels obtained in our study suggest that the strategy used herein could be a suitable alternative for the analysis of co-regulation patterns.

**Lactone and ester production in peach fruit requires a tight regulation of lipid catabolism**

A number of candidate genes for peach volatiles were identified using a combination of data analysis techniques based on correlations between gene expression and volatile accumulation data (HCA, PC, and CNA). The functional annotation indicated that, in some cases, this is due to a general effect on ripening. For example, the genes related to ethylene biosynthesis and perception (PPN004H06, PP1005G06, and PPN054G06) directly correlate with lactones (Fig. 3, Fig. 4, Table S7), suggesting that they could be regulators of their biosynthesis. It is known that ethylene is the hormone that controls the ripening of climacteric fruits, such as peach (Giovannoni, 2001), which makes it an unsuitable candidate for specific aroma control since its modification could alter the whole ripening syndrome. However, the possibility of there being downstream effectors specific for aroma production, which would be much more ideal candidates, cannot be ruled out, as different aspects of ripening appear to be controlled by specific transduction elements (Klee and Giovannoni, 2011). Overall, these results are not only in agreement with Zhang et al. (2011), who indicated that ethylene plays a regulatory role in aroma formation in peach, but also provide additional support for further research to search for the specific ethylene signal transduction elements involved in the specific activation for volatile production during ripening.

Volatile clusters C1 and C3 are each formed by two lineal esters and one lactone (which, indeed, are cyclic esters), while the rest of the lactones group in C2 (Fig. 1, Fig. S1). All these volatiles accumulated in the fruit during the time-course series in both genotypes, although they trended differently (Fig. 1) and are highly interconnected in the correlation network (Fig. 3A). Esters are known to be formed by the condensation of an acyl moiety with an alcohol catalyzed by alcohol acyl transferases (AATs) (D´ Auria, 2006). The structural diversity of esters is mainly due to alcohol, which can originate from different pathways. In this study, the esters found derive from acetyl-CoA and various alcohols,

which, in the case of esters 1, 3, and 10 (numbered according to Fig. S1), derive from the lipoxygenase pathway (Fig. 6A). In fruits, acetyl-CoA is formed from Pyruvate (Pyr) by the action of two enzymes: Pyruvate decarboxylase (EC 4.1.1.1) and aldehyde dehydrogenase (EC 1.2.1.5, (Pesis, 2005)). Accordingly, a Pyruvate decarboxylase associated with ester 2-Hexen-1-ol acetate (E)- was identified by HCA (Fig. S4), and we proposed that Pyr decarboxylase is regulated at the expression level to ensure the supply of acetyl-CoA for aroma volatile ester biosynthesis (Fig. 6).

Unlike esters, the lactone biosynthesis pathway in plants is still incomplete. Nevertheless, it seems clear that lactone biosynthesis starts from fatty acids with the introduction of an O atom to form hydroxy fatty acids (Schwab et al., 2008). Hydroxy fatty acids are then shortened by β-oxidation to form 4- or 5-hydroxy acids, which in turn result in the corresponding lactones after intramolecular esterification [Fig. 6A, (Schwab et al., 2008)]. However, it is unclear how the introduction of the O atom is achieved. By infiltrating a synthetic radiolabeled epoxy acid into fruits, it has been demonstrated that nectarines (peach glabrous mutants) are able to produce an artificial lactone (Schöttler and Boland, 1996). Based on this observation, it was proposed that the introduction of the O atom is achieved by fatty acid epoxydation and that the subsequent breakdown of the epoxy group (by epoxyde hydrolase) forms hydroxy acids; unfortunately, no further evidence has been presented. Alternatively, the introduction of hydroxyl groups could be produced by the hydration of unsaturated fatty acids or the direct hydroxylation of fatty acids catalyzed by Cytochrome P450s (CYPs) or by other hydroxylases not related to CYPs (Fig. 6A). To date, neither have the proposed endogenous fatty acid modification enzymes been cloned nor has their involvement in lactone biosynthesis in peach fruit been demonstrated.

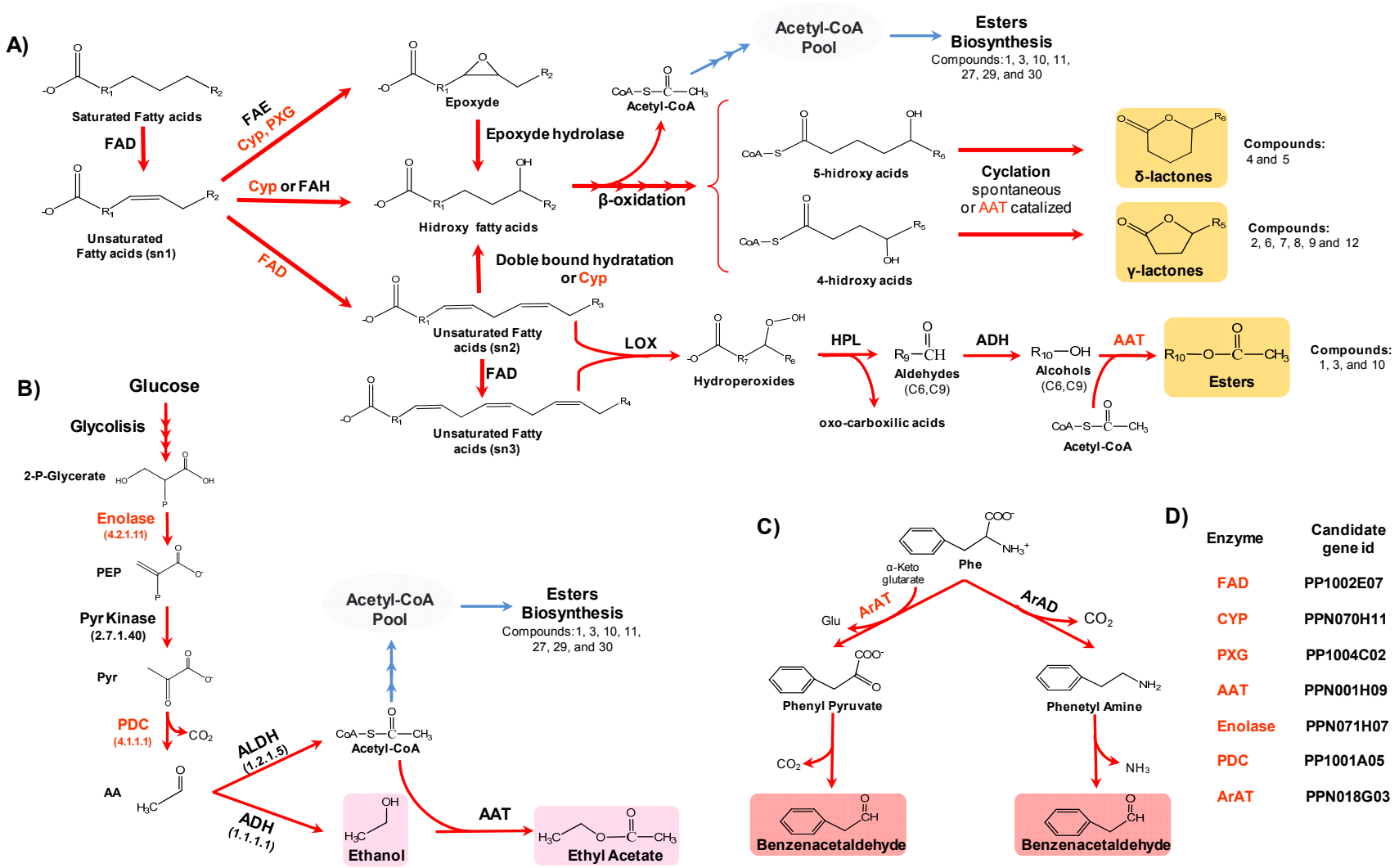



**Figure 6 (see the legends below).**

**Figure 6. Proposed models of action for the candidate genes identified**. A) The putative pathways for Lactones and Esters biosynthesis. The δ- and γ-lactones and esters produced (numbered according to Fig. S1) have different alkyl groups at R6, R5, and R10, respectively. C6 and C9 are indicated below the aldehydes and alcohols, which means that the compounds can have six or nine carbons. Please note that different enzymes use free fatty acid or are activated with different groups (CoA, Acyl carrier protein or lipidic group), but must be activated with a CoA group to enter β-oxidation. B) Ethanol and Ethyl acetate production. The Acetyl-CoA produced can also be added to the pool that is used to synthesize the numbered esters (according to Fig. S1). C) The route for Benzenacetaldehyde biosynthesis in melon (and proposed for peach) and tomato, left and right, respectively. For the three routes (A, B, and C), the enzymes codified by the candidate genes proposed are indicated in red. D) The ID of the candidate gene that is proposed for each enzyme.
Enzyme abbreviations: FAD, Fatty acid desaturase; FAE, Fatty acid Epoxydase; CYP, Cytochrome P450; PXG, Peroxygenase; FAH, Fatty acid Hydroxylase; LOX, Lipoxygenase; HPL, hydroperoxide lyase; ADH, alcohol dehydrogenase; AAT, Alcohol Acyl Transferase; PDC, Pyr decarboxylase; ALDH, Aldehyde dehydrogenase; ArAT, Aromatic aminoacid Aminotransferase; ArAD, Aromatic amino acid decarboxylase. Compound abbreviations: 2-P-Glycerate, 2-Phospho Glycerate; PEP, Phosphoenol Pyr; Pyr, Pyruvate; AA, Acetaldehyde; Phe, Phenylalanine; Glu, Glutamate.

For volatile cluster C1, an AAT homolog (PPN001H09) was selected as a putative candidate (Table 1, Table S6) because it is highly associated with the lactone of C1 in the correlation network analysis (Fig. 4). Moreover, the expression of this gene is moderately to highly correlated to the other lactones (from r=0.64 to 0.86, Table S5). It is still unclear whether the last cyclation step leading to lactone biosynthesis requires an enzyme or not (Schwab et al., 2008), but if this were the case, this gene could represent a novel type of AAT that produces cyclic esters (lactones). This AAT (PPN001H09) also correlated with lineal esters (r=0.69 to 0.75, Table S5), making it therefore possible that this gene encodes a standard AAT. Nonetheless, a phylogenetic analysis has revealed that the AAT found here (PPN001H09) is not related to previously characterized fruit AATs (data not shown). As it is unlikely that this gene can be identified by candidate gene approaches based solely on sequence similarity

to the genes reported in the bibliography, these results emphasize the advantage of using non-targeted analyses based on correlation rather than more targeted gene approaches based solely on homology with genes previously described as being associated with a given process. To date, no AAT from peach has been characterized, although an EST analysis has revealed that some AATs are expressed in fruit mesocarp (Vecchietti et al., 2009). Moreover, a targeted approach has demonstrated that the expression of another AAT (ppAAT1) did not correlate with ester and lactone production during postharvest peach ripening (Zhang et al., 2010). We can confirm this result as ppAAT1 is represented in our microarray (PPN018D08) and it poorly or moderately correlates with lactones and lineal esters (r between 0.27 and 0.60, Table S5).

Correlation analyses revealed a CYP, PPN070H11, which is highly associated with the lactones in C2 (Table 1). CYPs represent the largest family of genes involved in plant metabolism, which include enzymes that are capable of epoxidizing or directly hydroxylating fatty acids. A CYP77A4 (belonging to the CYP77 subfamily) from *Arabidopsis* catalyzes the epoxidation of free oleic acid (18:1) to form 9,10 epoxyestearic acid (Sauveplane et al., 2009), which has been proposed to be the precursor of γ-dodecalactone in peach (Schöttler and Boland, 1996). Other CYPs (belonging to the CYP71 family), which are capable of epoxydating unsaturated C18 fatty acids (18:2 and 18:3) as phosphatidylcholine esters, have been characterized in a *Euphorbia* species (Cahoon et al., 2002). Moreover, the position of the introduction of the hydroxyl group is crucial to producing a long hydroxy fatty acid, which, after β-oxidation, will produce the proper hydroxy acid molecule. The CYP703A2 (CYP703 subfamily) from *Arabidopsis* can catalyze the hydroxylation of saturated medium chain fatty acids (C10-C14) with a preference for hydroxylate at position C7 (Morant et al., 2007), and can therefore produce hydroxy fatty acids, which could be precursors for lactone generation after β-oxidation and

cyclation. According to phylogenetic analyses (data not shown), the CYP identified here belongs to the CYP72A subfamily. CYP72A subfamily members are quite diverse, and most of them have unknown functions (Nelson and Werck-Reichhart, 2011). The CYP72 clan is also associated with the metabolism of other fairly hydrophobic compounds besides fatty acids, such as isoprenoids. Other members have been associated with the catabolism of hormones (brassinosteroids and gibberellins) or with the biosynthesis of cytokinins (Nelson and Werck-Reichhart, 2011). Thus, we cannot exclude the possibility that this candidate gene (CYP, PPN070H11) could be involved in lactone production by, for example, controlling a hormone metabolism instead of participating directly in lactone biosynthesis.

The HCA also identified a gene, PP1002E07, with a homology to fatty acid desaturase (FAD), as being highly associated with 2-Hexen-1-ol acetate (E)- (Fig. S4, Table 1). Nevertheless, this gene also correlated well with lactones (r= 0.65 to 0.89, Table S5) and, in some cases, has correlation coefficients higher than those of 2-Hexen-1-ol, acetate (E)- (r=0.73, Table S5). The *Arabidopsis* homolog of this gene is an ω-6 fatty acid desaturase, which catalyzes the reduction of oleic acid (18:1) esterified to the sn-2 position of the membrane lipid phosphatidylcholine (PC) to linoleic acid (18:2). Yet desaturase genes also share a good homology with other fatty acid modification enzymes, like hydroxylases and epoxidases (Shanklin and Cahoon, 1998). For example, a hydroxylase that can hydroxylate oleic acid (18:1) and is linked to PC to form ricinoleic acid (18:1-OH) has been described in castor plants (van de Loo et al., 1995). In addition, a bifunctional enzyme that can both catalyze hydroxylation and desaturate the same fatty acid has been reported in a *Brassicacae* species (Broun et al., 1998a). Moreover, it has been demonstrated that a plant desaturase can be converted into a hydroxylase, and *vice versa,* by swapping a few specific amino acids by targeted mutagenesis (Broun et al., 1998b). Since

hydroxylase, desaturase, and epoxidase activity could lead to the production of hydroxy acids (Fig. 6A), we expressed the functional protein encoded by PP1002E07 in yeast to clarify the reaction catalyzed by the putative encoded enzyme. Yeasts expressing the peach ORF gene PP1002E07 accumulated linoleic acid (18:2) when oleic acid (18:1) levels were reduced (Fig. 5), indicating that the candidate gene has ω-6 Oleate desaturase activity. No ricinoleic acid was produced in the yeast expressing the candidate gene, suggesting that the encoded protein is not a bifunctional enzyme. Since no enzymes in plants have been shown to simultaneously desaturate and epoxidate fatty acid (Shanklin and Cahoon, 1998), our results suggest that the candidate gene identified herein is a monofunctional ω-6 Oleate desaturase.

Linoleic acid (18:2) may be further desaturated to linolenic acid (18:3), and both compounds could enter the Lipoxygenase (LOX) pathway to be catabolized into C6 and C9 alcohols, which are the substrates of AAT enzymes in the synthesis of a variety of esters [Fig. 6A, (Schwab et al., 2008)]. Moreover, double bonds of unsaturated fatty acids could undergo the introduction of oxygen by hydratation to form hydroxy acids, which could be biosynthetic precursors of lactones (Fig. 6A). In a previous report, we suggested that lipid-derived compounds and lactones are inversely regulated in peach fruit (Sánchez et al., 2012). An analysis of the current data set also revealed lactones, specifically those in C2 and C3, showing strong inverse correlations with lipid-derived compounds belonging to C11 and C12 (Fig. 4A). This is in agreement with previous results indicating that lactone content increases during maturity and ripening, while lipid-derived compounds decrease (Aubert et al., 2003; Chapman et al., 1991; Visai and Vanoli, 1997; Zhang et al., 2010). The fact that ω-6 Oleate desaturase positively correlates with lactones, i.e., its levels increase during ripening, while lipid-derived compounds decrease, seems to suggest that the formation of lipid-derived compounds downstream of linoleic acid

(18:2) is regulated during peach ripening, and that this pathway can feed into the precursors for lactone synthesis.

Beside CYPs and desaturase-like epoxidases, a third kind of enzyme, Peroxygenase (PXG), which can catalyze the epoxygenation of fatty acids, has been reported in plants. PXGs share no structural similarity with either peroxidase or cytochrome P450, but instead do so with caleosin, a small oil body-associated protein with both heme- and calcium-binding motifs (Hanano et al., 2006). Here, a caleosin-related protein (PP1004C02), showing a strong correlation with g-jasmolactone, was indentified (Table 1, Table S6). Recently, a PXG capable of epoxidating free oleic acid (18:1) was described in oat seeds (Meesapyodsuk and Qiu, 2011). It is interesting to note that PXG has potential kinase phosphorylation sites (Murphy et al., 2001), and our approach also identified kinases that correlate well with C3 (PPN010B11, PPN069F09).

PXG, like FAD and CYPs, is found in plant microsomes, which implies that it could be associated with the endoplasmatic reticulum *in vivo* (Cahoon et al., 2002; Murphy et al., 2001). Interestingly, two SNARE genes associated with the endomembrane trafficking system (PP1004F06 and PPN031F12) also correlated with the C3 volatile cluster (Table 1, Table S6), indicating the co-activation of both structural and metabolic enzymes.

It has been proposed that highly correlated metabolites in replicated experiments reflect the existence of a common regulatory mechanism in the system under study (Camacho et al., 2005). In this sense, our approach identifies several transcription factors relating to volatile clusters C1 (PP1009A02, PPN066C05) and C2 (PPN010C01), which are tentatively involved in the regulation of volatile production during peach fruit ripening.

The highly stringent criteria used in our approach indicate that our candidate genes are unlikely to be the result of random correlations. In any case, further

functional studies are required to unequivocally prove the implication of the candidate gene in controlling the proposed volatile accumulation.

**Selection of candidate genes for other aroma-related compounds**

Candidate genes for other minor aroma volatiles in peach have also been discovered. Ethyl acetate has been described to have a "fruity" aroma (Guillot et al., 2006) and has also been proposed to be regulated independently of other esters in peach (Sánchez et al., 2012). Here we demonstrate that the HCA identifies an enolase (phosphopyruvate hydratase, PPN071H07), which is highly related to the accumulation of both Ethanol and its acetate ester, Ethyl acetate (Fig. S4, Table 1). Enolase catalyzes one of the final glycolysis steps (EC. 4.2.1.11), which generates Pyr as an end product (Fig. 6B). As previously mentioned, Pyr is also converted into acetaldehyde by the action of Pyr decarboxylase (Pesis, 2005). From acetaldehyde, the pathway splits into two branches, one producing ethanol and the other acetyl-CoA [Fig. 6B, (Pesis, 2005)]. Ethyl acetate can be formed by the condensation of ethanol with Acetyl-CoA by a putative AAT, which is yet to be described. The correlation between ethanol and ethyl acetate with the enolase suggests that this gene is a good candidate for driving the production of glycolysis end products, i.e., Pyr, into aroma-contributing compounds.

The aroma of Benzeneacetaldehyde is described as "green" (Fig. S1) and could, therefore, confer immature fruit notes. Accordingly, the levels of this compound decreased during the ripening of the 'Granada' genotype (Fig. 1) and after the shelf-life ripening in 'MxR_01' (Fig. 2), which is in accordance with the shift of aroma from immature to ripe fruit. The correlation network analysis reveals that putative tyrosine (Tyr) aminotransferase (PPN018G03) correlates well (r=0.85) with this volatile, which is, in turn, related to other phenolic volatiles (Benzaldehyde and Benzyl chloride, Fig. 1, Fig. 4). Two pathways for

the biosynthesis of aromatic volatiles have been described in different fruits (Fig. 6C). In tomato, Phenylalanine (Phe) undergoes two sequential enzymatic steps: a decarboxylation step followed by a deamination step to form Benzeneacetaldehyde (Baldwin et al., 2000). In contrast, in melon, Phe first loses the amine groups by transamination reactions to then undergo a decarboxylation reaction (Gonda et al., 2010). Aromatic acid transaminase (CmArAT1), which is involved in Benzeneacetaldehyde biosynthesis in melon, is capable of catalyzing the transamination of either Tyr or Phe (Gonda et al., 2010). The gene identified herein (PPN018G03) is phylogenetically related to CmArAT1 (data not shown), suggesting that it may have both transaminase activities. Therefore, we hypothesize that Benzeneacetaldehyde in peach is produced by a similar transaminase pathway to that in melon (Fig. 6C).

As previously mentioned, the catabolism of linoleic and linolenic acid by the LOX/HPL pathway produces the so-called lipid-derived compounds (Schwab et al., 2008). As some of these volatiles are described to confer ”green” aromas (Fig. S1), which is the typical aroma of freshly-cut grass, these volatiles have been traditionally associated with unripe fruits. Accordingly, we previously found that lipid-derived volatiles correlated highly among themselves and also with parameters that measure the peach ripening stage, indicating that high levels are linked to unripe fruit (Sánchez et al., 2012). In addition, some of these compounds have been described to have other aromas, which could even be unpleasant when present in fruit, these being ”fermented”, ”spicy”, ”fatty”, or even ”chemical” (Fig. S1). In the data set used herein, these compounds are grouped into two highly interrelated clusters (C11 and C12, Fig. 1, Fig. 4A). The correlation network analysis (Fig. S5) reveals that a transcription factor (PPN019F11) strongly correlates with the expression levels of a gene with no homolog in *Arabidopsis* (PPN023E05) which, in turn, is highly correlated to a lipid δ 9 desaturase (PPN052H12). This lipid desaturase strongly and inversely

correlates with lipid-derived compounds Furan 2-pentyl- and Hexanal. It is possible that the lipid desaturase catalyzes the formation of unsaturated lipids, which are not substrates for the LOX/HPL pathway, thus driving the carbon allocation toward other products. If this were the case, this gene would be a candidate for reducing lipid-derived compound production. Four candidate genes for C12 volatiles were selected because they correlated with 2,4-Heptadienal (E,E)- in the HCA (Fig. S4). Fatty acids are stored as triacylglycerides, so acyl hydrolases should free them to enter the LOX/HPL pathway (Schwab et al., 2008). One candidate gene is an esterase homolog (PPN033D11) that may be involved in the supply of fatty acids to the LOX/HPL pathway that produces lipid-derived volatiles, which makes it an interesting target for decreasing undesired aromas in each fruit.

Cluster C4 is formed by two volatiles deriving from the terpenoid biosynthesis pathway and by four volatiles of unknown origin (Fig. 1, Fig. S1). Our approach was able to identify the genes associated with C4 volatiles, although their functional annotations do not suggest a mechanism (Table 1, Table S6). Nonetheless, it is quite likely that hormones are involved in the regulation of these compounds since gibberellin- and auxin-responsive proteins were identified (PP1000F05 and PP1009D02, respectively), as was aminopeptidase P (PPN017B08; Table 1, Table S6). In *Arabidopsis*, it has been demonstrated that auxin signaling can be blocked by auxin-transport inhibitors, and that the specific Aminopeptidase P regulates hormone perception by degrading these inhibitors (Murphy et al., 2002).

Our "omics" approach could be improved by including additional samples, i.e., fruits from other genotypes or those subjected to other treatments. For example, it has been demonstrated that low temperature storage and hot water treatments lower the volatile content of peach, while other quality parameters, e.g., flesh firmness, remain unaffected (Raffo et al., 2008; Zhang et al., 2011).

Thus, if the volatile reduction is due to changes in gene expression, these treatments may be used to unlink both processes and therefore improve gene identification by co-regulation patterns. Although further experiments are required to unequivocally prove some of the links between the genes and volatile compounds found herein, we provide empirical data, which may prove useful to boost gene-function discovery studies in peach fruit.

## CONCLUSIONS

In the present study, a hypothesis-free approach based on correlation analyses was applied to discover genes that are highly associated with the main aroma-contributing compounds of peach fruit. Consequently, a set of candidate genes for lactones, esters, terpenoids, phenolics and lipid-derived volatiles was identified. Some have functional annotations, which clearly suggest a possible role in controlling the production of related volatiles, which qualifies them as strong candidates to be selected for functional studies. In addition, for one of the candidate genes, we demonstrate that the encoded protein possesses enzymatic activity that produces a volatile peach precursor. Although our main interest was related to the practical implications of our results, the data provided herein not only support a previous hypothesis related to the physiology and metabolism of peach fruit, but also suggest novel ones. Although more studies are required to unequivocally prove the associations found herein, we provide experimental evidence and data analyses that support the role of a number of candidate genes in the control of volatile compounds, some of which may prove useful as targets for peach improvement and/or biotechnology industry interests.

**ACKNOWLEDGMENTS**

We are grateful to Cristina Besada, PhD (Instituto Valenciano de Investigaciones Agrarias, IVIA, Spain) for her help with the fruit quality parameter analyses. We are also thankful to Cristina Martí and Clara Pons (*Instituto de Biología Molecular y Celular de Plantas*, IBMCP, Spain) for their advice on microarray analyses. Jesús Garcia Brunton, PhD for providing the fruits used in this study (*Instituto Murciano de Investigación y Desarrollo Agrario*, IMIDA, Spain). HS-SPME-GC-MS analyses were performed at the Metabolomic lab facilities at the IBMCP (CSIC) in Spain. GS has financial support from INTA (*Instituto Nacional de Tecnología Agropecuaria*, Argentine). This project has been funded by the Ministry of Economy and Competitivity grant AGL2010-20595.

**COMPETING INTERESTS**

The authors declare that they have no competing interests.

**AUTHORS' CONTRIBUTIONS**

GS conceived and designed the work, performed the metabolomics and microarray experiments, analyzed the data, and wrote the manuscript. MV-C and JJS performed the yeast expression experiment and the fatty acid analysis. AM and MLB contributed reagent material and analysis tools. AG conceived, designed, and supervised the work.

## REFERENCES


Ahmad, R., Parfitt, D., Fass, J., Ogundiwin, E., Dhingra, A., Gradziel, T., Lin, D., Joshi, N., Martinez-Garcia, P. and Crisosto, C. (2011) Whole genome sequencing of peach (*Prunus persica* L.) for SNP identification and selection. *BMC Genomics*, **12**, 569.

Amiour, N., Imbaud, S., Clément, G., Agier, N., Zivy, M., Valot, B., Balliau, T., Armengaud, P., Quilleré, I., Cañas, R., Tercet-Laforgue, T. and Hirel, B. (2012) The use of metabolomics integrated with transcriptomic and proteomic studies for identifying key steps involved in the control of nitrogen metabolism in crops such as maize. *Journal of Experimental Botany*, **63**, 5017-5033.

Arús, P., Verde, I., Sosinski, B., Zhebentyayeva, T. and Abbott, A. (2012) The peach genome. *Tree Genetics & Genomes*, **8**, 531-547.

Assenov, Y., Ramírez, F., Schelhorn, S.-E., Lengauer, T. and Albrecht, M. (2008) Computing topological parameters of biological networks. *Bioinformatics*, **24**, 282-284.

Aubert, C., Gunata, Z., Ambid, C. and Baumes, R. (2003) Changes in Physicochemical Characteristics and Volatile Constituents of Yellow- and White-Fleshed Nectarines during Maturation and Artificial Ripening. *J. Agric. and Food Chem.*, **51**, 3083-3091.

Baldwin, E.A., Scott, J.W., Shewmaker, C.K. and Schuch, W. (2000) Flavor trivia and tomato aroma: biochemistry and possible mechanisms for control of important aroma components. *HortScience*, **35**, 1013-1022.

Brandi, F., Bar, E., Mourgues, F., Horvath, G., Turcsi, E., Giuliano, G., Liverani, A., Tartarini, S., Lewinsohn, E. and Rosati, C. (2011) Study of 'Redhaven' peach and its white-fleshed mutant suggests a key role of CCD4 carotenoid dioxygenase in carotenoid and norisoprenoid volatile metabolism. *BMC Plant Biology*, **11**, 24.

Broun, P., Boddupalli, S. and Somerville, C. (1998a) A bifunctional oleate 12-hydroxylase: desaturase from Lesquerella fendleri. *The Plant Journal*, **13**, 201-210.

Broun, P., Shanklin, J., Whittle, E. and Somerville, C. (1998b) Catalytic Plasticity of Fatty Acid Modification Enzymes Underlying Chemical Diversity of Plant Lipids. *Science*, **282**, 1315-1317.

Bruhn, C.M., Feldman, N., Garlitz, C., Harwood, J., Ivans, E., Marshall, M., Riley, A., Thurber, D. and Williamson, E. (1991) Consumer Perceptions of quality: Apricots, Cantaloupes, Peaches, Pears, Strawberries, and Tomatoes. *Journal of Food Quality*, **14**, 187-195.

Cahoon, E.B., Ripp, K.G., Hall, S.E. and McGonigle, B. (2002) Transgenic Production of Epoxy Fatty Acids by Expression of a Cytochrome P450 Enzyme from Euphorbia lagascaeSeed. *Plant Physiology*, **128**, 615-624.

Camacho, D., de la Fuente, A. and Mendes, P. (2005) The origin of correlations in metabolomics data. *Metabol*, **1**, 53.

Carrari, F., Baxter, C., Usadel, B., Urbanczyk-Wochniak, E., Zanor, M.I., Nunes-Nesi, A., Nikiforova, V., Centero, D., Ratzka, A. and Pauly, M. (2006) Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiology*, **142**, 1380-1396.

Carrera, J., Fernández del Carmen, A., Fernández-Muñoz, R., Rambla, J.L., Pons, C., Jaramillo, A., Elena, S.F. and Granell, A. (2012) Fine-Tuning Tomato Agronomic Properties by Computational Genome Redesign. *PLoS Comput Biol*, **8**, e1002528.

Casati, P., Campi, M., Morrow, D., Fernandes, J. and Walbot, V. (2011) Transcriptomic, proteomic and metabolomic analysis of UV-B signaling in maize. *BMC Genomics*, **12**, 321.

Chapman, G.W., Horvat, R.J. and Forbus, W.R. (1991) Physical and chemical changes during the maturation of peaches (cv. Majestic). *Journal of Agricultural and Food Chemistry*, **39**, 867-870.

D´ Auria, J.C. (2006) Acyltransferases in plants: a good time to be BAHD. *Current Opinion in Plant Biology*, **9**, 331-340.

Derail, C., Hofmann, T. and Schieberle, P. (1999) Differences in Key Odorants of Handmade Juice of Yellow-Flesh Peaches (*Prunus persica* L.) Induced by the Workup Procedure. *J. Agric. and Food Chem.*, **47**, 4742-4745.

Dufosse, L., Latrasse, A. and Spinnler, H.-E. (1994) Importance of lactones in food flavours: structure, distribution, sensory properties and biosynthesis. *Sci. Aliment.*, **14**, 17-50. (*in french*).

Eduardo, I., Chietera, G., Bassi, D., Rossini, L. and Vecchietti, A. (2010) Identification of key odor volatile compounds in the essential oil of nine peach accessions. *Journal of the Science of Food and Agriculture*, **90**, 1146-1154.

Eduardo, I., Chietera, G., Pirona, R., Pacheco, I., Troggio, M., Banchi, E., Bassi, D., Rossini, L., Vecchietti, A. and Pozzi, C. (2012) Genetic dissection of aroma volatile compounds from the essential oil of peach fruit: QTL analysis and identification of candidate genes using dense SNP maps. *Tree Genetics & Genomes*, 1-16.

ESTree Consortium. (2005) Development of an oligo-based microarray (uPEACH 1.0) for genomics studies in peach fruit. *Acta Hortic*, **682**, 263–268.

FAO. (2010) Accessed 25 Aug. 2013. http://faostat.fao.org.

Gatfield, I.L. (1999) Biotechnological production of natural flavor materials. In. Kluwer Academic / Plenum Publishing Corporation, New York, pp. 211-227.

Giovannoni, J. (2001) Molecular biology of fruit maturation and ripening. *Annual review of plant biology*, **52**, 725-749.

Gonda, I., Bar, E., Portnoy, V., Lev, S., Burger, J., Schaffer, A.A., Tadmor, Y.a., Gepstein, S., Giovannoni, J.J., Katzir, N. and Lewinsohn, E. (2010) Branched-chain and aromatic amino acid catabolism into aroma volatiles in Cucumis melo L. fruit. *J. Exp. Bot.* , **61**, 1111-1123.

Guillot, S., Peytavi, L., Bureau, S., Boulanger, R., Lepoutre, J.-P., Crouzet, J. and Schorr-Galindo, S. (2006) Aroma characterization of various apricot varieties using headspace-solid phase microextraction combined with gas chromatography-mass spectrometry and gas chromatography-olfactometry. *Food Chemistry*, **96**, 147-155.

Hanano, A., Burcklen, M., Flenet, M., Ivancich, A., Louwagie, M., Garin, J.r.m. and Blée, E. (2006) Plant Seed Peroxygenase Is an Original Heme-oxygenase with an EF-hand Calcium Binding Motif. *Journal of Biological Chemistry*, **281**, 33140-33151.

Hirai, M.Y., Klein, M., Fujikawa, Y., Yano, M., Goodenowe, D.B., Yamazaki, Y., Kanaya, S., Nakamura, Y., Kitayama, M., Suzuki, H., Sakurai, N., Shibata, D., Tokuhisa, J., Reichelt, M., Gershenzon, J., Papenbrock, J. and Saito, K. (2005) Elucidation of Gene-to-Gene and Metabolite-to-Gene Networks in Arabidopsis

by Integration of Metabolomics and Transcriptomics. *Journal of Biological Chemistry*, **280**, 25590-25595.

Hirai, M.Y., Sugiyama, K., Sawada, Y., Tohge, T., Obayashi, T., Suzuki, A., Araki, R., Sakurai, N., Suzuki, H., Aoki, K., Goda, H., Nishizawa, O.I., Shibata, D. and Saito, K. (2007) Omics-based identification of Arabidopsis Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proceedings of the National Academy of Sciences*, **104**, 6478-6483.

Klee, H.J. and Giovannoni, J.J. (2011) Genetics and Control of Tomato Fruit Ripening and Quality Attributes. *Annual Review of Genetics*, **45**, 41-59.

Lombardo, V.A., Osorio, S., Borsani, J., Lauxmann, M.A., Bustamante, C.A., Budde, C.O., Andreo, C.S., Lara, M.V., Fernie, A.R. and Drincovich, M.F. (2011) Metabolic Profiling during Peach Fruit Development and Ripening Reveals the Metabolic Networks That Underpin Each Developmental Stage. *Plant Physiology*, **157**, 1696-1710.

Manganaris, G.A., Vasilakakis, M., Diamantidis, G. and Mignani, I. (2006) Diverse metabolism of cell wall components of melting and non-melting peach genotypes during ripening after harvest or cold storage. *Journal of the Science of Food and Agriculture*, **86**, 243-250.

Meesapyodsuk, D. and Qiu, X. (2011) A Peroxygenase Pathway Involved in the Biosynthesis of Epoxy Fatty Acids in Oat. *Plant Physiology*, **157**, 454-463.

Meisel, L., Fonseca, B., Gonzalez, S., Baeza-Yates, R., Cambiazo, V., Campos, R., Gonzalez, M., Orellana, A., Retamales, J. and Silva, H. (2005) A rapid and efficient method for purifying high quality total RNA from peaches (Prunus persica) for functional genomics analyses. *Biol Res*, **38**, 83-88.

Morant, M., Jørgensen, K., Schaller, H., Pinot, F., Møller, B.L., Werck-Reichhart, D.l. and Bak, S. (2007) CYP703 Is an Ancient Cytochrome P450 in Land Plants Catalyzing in-Chain Hydroxylation of Lauric Acid to Provide Building Blocks for Sporopollenin Synthesis in Pollen. *The Plant Cell Online*, **19**, 1473-1487.

Mounet, F., Moing, A., Garcia, V., Petit, J., Maucourt, M., Deborde, C., Bernillon, S., Le Gall, G., Colquhoun, I. and Defernez, M. (2009) Gene and metabolite regulatory network analysis of early developing fruit tissues highlights new candidate genes for the control of tomato fruit composition and development. *Plant Physiology*, **149**, 1505-1528.

Murphy, A.S., Hoogner, K.R., Peer, W.A. and Taiz, L. (2002) Identification, Purification, and Molecular Cloning of N-1-Naphthylphthalmic Acid-Binding Plasma Membrane-Associated Aminopeptidases from *Arabidopsis*. *Plant Physiology*, **128**, 935-950.

Murphy, D.J., Hernández-Pinzón, I. and Patel, K. (2001) Role of lipid bodies and lipid-body proteins in seeds and other tissues. *Journal of Plant Physiology*, **158**, 471-478.

Nelson, D. and Werck-Reichhart, D. (2011) A P450-centric view of plant evolution. *The Plant Journal*, **66**, 194-211.

Ogundiwin, E.A., Marti, C., Forment, J., Pons, C., Granell, A., Gradziel, T.M., Peace, C.P. and Crisosto, C.H. (2008) Development of ChillPeach genomic tools and identification of cold-responsive genes in peach fruit. *Plant Mol Biol*, **68**, 379-397.

Parker, D.D., Zilberman, D. and Moulton, K. (1991) How quality relates to price in California fresh peaches. *Calif. Agr.*, **45**, 14-16.

Pesis, E. (2005) The role of the anaerobic metabolites, acetaldehyde and ethanol, in fruit ripening, enhancement of fruit quality and fruit deterioration. *Postharvest Biology and Technology*, **37**, 1-19.

Raffo, A., Nardo, N., Tabilio, M. and Paoletti, F. (2008) Effects of cold storage on aroma compounds of white- and yellow-fleshed peaches. *European Food Research and Technology*, **226**, 1503-1512.

Saito, K. and Matsuda, F. (2010) Metabolomics for Functional Genomics, Systems Biology, and Biotechnology. *Annual Review of Plant Biology*, **61**, 463-489.

Sánchez, G., Besada, C., Badenes, M.L., Monforte, A.J. and Granell, A. (2012) A Non-Targeted Approach Unravels the Volatile Network in Peach Fruit. *PLoS ONE*, **7**, e38992.

Sauveplane, V., Kandel, S., Kastner, P.-E., Ehlting, J., Compagnon, V., Werck-Reichhart, D. and Pinot, F. (2009) *Arabidopsis thaliana* CYP77A4 is the first cytochrome P450 able to catalyze the epoxidation of free fatty acids in plants. *FEBS Journal*, **276**, 719-735.

Schöttler, M. and Boland, W. (1996) Biosynthesis of Dodecano-4-lactone in Ripening Fruits: Crucial Role of an Epoxide-Hydrolase in Enantioselective Generation of Aroma Components of the Nectarine (*Prunus persica* var. nucipersica) and the Strawberry (*Fragaria ananassa*). *Helvetica Chimica Acta*, **79**, 1488-1496.

Schwab, W., Davidovich-Rikanati, R. and Lewinsohn, E. (2008) Biosynthesis of plant-derived flavor compounds. *The Plant Journal*, **54**, 712-732.

Shanklin, J. and Cahoon, E.B. (1998) Desaturation and related modifications of fatty acids. *Annual Review of Plant Physiology and Plant Molecular Biology*, **49**, 611-641.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, **13**, 2498-2504.

Spolaore, S., Trainotti, L. and Casadoro, G. (2001) A simple protocol for transient gene expression in ripe fleshy fruit mediated by *Agrobacterium*. *Journal of Experimental Botany*, **52**, 845-850.

Tikunov, Y., Lommen, A., de Vos, C.H.R., Verhoeven, H.A., Bino, R.J., Hall, R.D. and Bovy, A.G. (2005) A Novel Approach for Nontargeted Data Analysis for Metabolomics. Large-Scale Profiling of Tomato Fruit Volatiles. *Plant Physiol.*, **139**, 1125-1137.

Trainotti, L., Bonghi, C., Ziliotto, F., Zanin, D., Rasori, A., Casadoro, G., Ramina, A. and Tonutti, P. (2006) The use of microarray μPEACH1.0 to investigate transcriptome changes during transition from pre-climacteric to climacteric phase in peach fruit. *Plant Science*, **170**, 606-613.

Trainotti, L., Tadiello, A. and Casadoro, G. (2007) The involvement of auxin in the ripening of climacteric fruits comes of age: the hormone plays a role of its own and has an intense interplay with ethylene in ripening peaches. *Journal of Experimental Botany*, **58**, 3299-3308.

Tusher, V.G., Tibshirani, R. and Chu, G. (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences*, **98**, 5116-5121.

van de Loo, F.J., Broun, P., Turner, S. and Somerville, C. (1995) An oleate 12-hydroxylase from Ricinus communis L. is a fatty acyl desaturase homolog. *Proceedings of the National Academy of Sciences*, **92**, 6743-6747.

Vecchietti, a., Lazzari, B., Ortugno, C., Bianchi, F., Malinverni, R., Caprera, a., Mignani, I. and Pozzi, C. (2009) Comparative analysis of expressed sequence tags from tissues in ripening stages of peach (*Prunus persica* L. Batsch). *Tree Genetics & Genomes*, **5**, 377-391.

Venegas-Calerón, M., Beaudoin, F., Garcés, R., Napier, J.A. and Martínez-Force, E. (2010) The sunflower plastidial ω3-fatty acid desaturase (HaFAD7) contains the signalling determinants required for targeting to, and retention in, the endoplasmic reticulum membrane in yeast but requires co-expressed ferredoxin for activity. *Phytochemistry*, **71**, 1050-1058.

Verde, I., Abbott, A.G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., Zhebentyayeva, T., Dettori, M.T., Grimwood, J., Cattonaro, F., Zuccolo, A., Rossini, L., Jenkins, J., Vendramin, E., Meisel, L.A., Decroocq, V., Sosinski, B., Prochnik, S., Mitros, T., Policriti, A., Cipriani, G., Dondini, L., Ficklin, S., Goodstein, D.M., Xuan, P., Fabbro, C.D., Aramini, V., Copetti, D., Gonzalez, S., Horner, D.S., Falchi, R., Lucas, S., Mica, E., Maldonado, J., Lazzari, B., Bielenberg, D., Pirona, R., Miculan, M., Barakat, A., Testolin, R., Stella, A., Tartarini, S., Tonutti, P., Arus, P., Orellana, A., Wells, C., Main, D., Vizzotto, G., Silva, H., Salamini, F., Schmutz, J., Morgante, M. and Rokhsar, D.S. (2013) The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet*, **45**, 487-494.

Verde, I., Bassil, N., Scalabrin, S., Gilmore, B., Lawley, C.T., Gasic, K., Micheletti, D., Rosyara, U.R., Cattonaro, F., Vendramin, E., Main, D., Aramini, V., Blas, A.L., Mockler, T.C., Bryant, D.W., Wilhelm, L., Troggio, M., Sosinski, B., Aranzana, M.J., Arús, P., Iezzoni, A., Morgante, M. and Peace, C. (2012) Development and Evaluation of a 9K SNP Array for Peach by Internationally Coordinated SNP Detection and Validation in Breeding Germplasm. *PLoS ONE*, **7**, e35668.

Visai, C. and Vanoli, M. (1997) Volatile compound production during growth and ripening of peaches and nectarines. *Scientia Horticulturae*, **70**, 15-24.

Wang, Y., Yang, C., Li, S., Yang, L., Wang, Y., Zhao, J. and Jiang, Q. (2009) Volatile characteristics of 50 peaches and nectarines evaluated by HP-SPME with GC-MS. *Food Chemistry*, **116**, 356-364.

Zhang, B., Shen, J.-y., Wei, W.-w., Xi, W.-p., Xu, C.-J., Ferguson, I. and Chen, K. (2010) Expression of Genes Associated with Aroma Formation Derived from the Fatty Acid Pathway during Peach Fruit Ripening. *Journal of Agricultural and Food Chemistry*, **58**, 6157-6165.

Zhang, B., Xi, W.-p., Wei, W.-w., Shen, J.-y., Ferguson, I. and Chen, K.-s. (2011) Changes in aroma-related volatiles and gene expression during low temperature storage and subsequent shelf-life of peach fruit. *Postharvest Biology and Technology*, **60**, 7-16.

# Capítulo 4:
# The peach volatilome modularity is reflected at the genetic and environmental response levels

Los resultados presentados en este capítulo fueron enviados para considerar su publicación en una revista científica:

-**Sánchez** et al. (2013) "**The peach volatiloma modularity is reflected at the genetic and environmental response levels".** *Molecular Breeding* (submitted).

# THE PEACH VOLATILOME MODULARITY IS REFLECTED AT THE GENETIC AND ENVIRONMENTAL REPONSE LEVELS

## ABSTRACT

The improvement of peach fruit aroma is currently a desirable aim in breeding programs. To better characterize and asses the genetic potential of peach genotypes for increasing the aroma of the fruit by breeding, a quantitative trait locus (QTL) analysis was carried out in a F1 population segregating largely for fruit traits. A set of 9000 SNPs were evaluated in order to find polymorphic markers for linkage maps construction. For QTL analysis, a metabolomic approach based on GC-MS was used to profile 81 volatile compounds across the population in two locations. The peach volatilome in the segregating population presented a wide composition and dynamic rage and a correlation-based analysis of this volatile data set revealed that the peach volatilome is organized in modules formed by compounds from the same biosynthetic origin or sharing similar chemical structures. QTL clustering of members of those volatile modules indicated that some of them are subjected to joint genetic control. The monoterpene module is controlled by a unique locus at the top of LG4, a locus previously showed to affect the levels of two terpenoid compounds. At the bottom of LG4 a locus controlling several volatiles but also Melting/noMelting and maturity-related traits was found, suggesting putative pleiotropic effects. In addition, two novel locus controlling lactones and esters in linkage group 5 and 6 were discovered. The results presented here confirmed previously identified loci controlling the aroma of peach but also identified novel ones.

## INTRODUCTION

Traditionally, peach (*Prunus persica* (L.) Batsch) breeding programs were focused on obtaining elite genotypes with high productivity, resistant to pathogen and plagues, and producing fruit of large size and overall good appearance covering most of the season (early and late cultivars). As result, many cultivars with excellent agronomic performance have been developed. Nevertheless, breeding for agronomic traits often occurs in detriment of the organoleptic quality of the fruit as was demonstrated in the case of "greek basil", strawberry and tomato where most of the typical aroma was lost during recent breeding process (Klee and Giovannoni, 2011; Koutsos et al., 2009; Olbricht et al., 2008). In peach, the decrease of organoleptic fruit quality was perceived by consumers as the principal cause of dissatisfaction (Bruhn et al., 1991). A likely consequence of that is the low consumption of peaches when compared with other fruits like apple and banana (Crisosto, 2006). Early studies established that fruit aroma, along with flesh firmness and color, is the main attribute that consumers used to judge peach quality (Bruhn, 1995), and one of the main factors affecting peach prices at the market (Parker et al., 1991). Therefore, genetic improvement of organoleptic fruit quality could lead not only to an increased consumption but also would add value to this food commodity.

Peach breeding is hindered by the reduced genetic variability in the available germplasm and by certain aspects of the tree physiology like a short blossoming time and a juvenile phase of 2 to 3 years (Abbott et al., 2008). Related to the latter is the fact that peach breeding not only requires an investment of time but also results in high operative costs associated to the maintenance of the trees in the field until fruit are evaluated. Consequently, the implementation of marker assisted selection (MAS) becomes, almost exclusively, the only feasible option for saving space and time and therefore reduce costs in efficient breeding

programs. However, improvement of fruit flavor is not a trivial task since the aroma is formed by the qualitative and quantitative combination of a large number of Volatile Organic Compounds (VOCs) released by the fruit. To add complexity, VOCs also contribute to the taste of the fruit acting in combination with sugars and organic acids. In the case of peach, around 100 compounds have been described so far (Sánchez et al., 2012 and references within), but few of them seems to contribute to the aroma of the fruit (Eduardo et al., 2010). Among this volatiles, lactones appear to be the main contributors to peach aroma (Eduardo et al., 2010; Horvat et al., 1990) and particularly γ-decalactone, an intramolecular ester with aroma described as "peach-like" (Derail et al., 1999). Esters such as (Z)-3-Hexen-1-ol acetate, (E)-2-Hexen-1-ol acetate, and Ethyl Acetate could contribute with "fruity" notes to the overall fruit aroma (Derail et al., 1999; Eduardo et al., 2010; Guillot et al., 2006), while terpenoid compounds like Linalool and β-Ionone could provide "floral" notes (Eduardo et al., 2010; Greger and Schieberle, 2007; Guillot et al., 2006). On the other hand, the aroma of the lipid derived compounds such as (Z)-3-Hexenal and (E)-2-Hexenal have been described as "green" notes (Derail et al., 1999), and are usually associated to unripe fruit. Several investigations demonstrated that aroma formation on peach is a dynamic process since volatile change dramatically during maturity and ripening (Aubert et al., 2003; Chapman et al., 1991; Visai and Vanoli, 1997; Zhang et al., 2010), cold storage (Robertson et al., 1990) , by application of postharvest treatments (Sumitani et al., 1994; Zhang et al., 2010) and are also influenced by culture techniques and management of the trees in the field (Jia et al., 2005).

The large impact VOCs have on peach acceptability and marketability encouraged several groups to find genes and loci controlling aroma production. Recently, Eduardo et al. (2013) performed the QTL analysis for 23 volatile most of them contributing to peach fruit aroma. Among the QTL identified, a locus

with major effect on the production of two monoterpene compounds was described in LG4 and moreover the co-localization with terpene synthase genes was shown (Eduardo et al., 2013). Earlier the same group performed a microarray-based RNA profiling during ripening to describe changes on aroma-related gene expression during ripening (Pirona et al., 2012). In addition, an EST library was analyzed to find a set of candidate genes expressed in peach fruit as related to the formation of different volatile compounds (Vecchietti et al., 2009). Additional studies targeted literature-derived candidate genes to analyze its involvement in the production of lactones and esters (Xi et al., 2012; Zhang et al., 2010; Zhang et al., 2011) and carotenoid-derived volatiles (Brandi et al., 2011). More recently, novel candidate genes for controlling diverse groups of volatiles were proposed by using a non-targeted genomic approach which analyzed the correlation between transcript and compound levels (Sánchez et al., 2013). Nowadays, a high-quality genome of peach is available (Verde et al., 2013) and it is envisaged that next-generation sequencing technologies such as RNA-seq will be soon applied to discover more genes related to the aroma of peach. In this context, additional studies delimiting the chromosome regions linked to the aroma formation will help to interconnect this emerging wealth of information to elucidate aroma-associated gene function in peach.

The recent release of a high-quality genome of peach (Verde et al., 2013) and the development of a 9K SNP array by The International Peach SNP Consortium (IPSC) (Verde et al., 2012) facilitates the development of linkage maps which were in some extend hampered by the low genetic variability of intraspecific populations (Abbott et al., 2008). Complementarily, the recent advances in high-throughput technologies based on GC-MS for volatile profiling (Bicchi and Maffei, 2012) enabled to describe the peach volatilome at a more exhaustive level (Sánchez et al., 2012). Similar profiling platforms combined with natural variability and mapping information have been applied recently to a large scale

analyses of volatiles QTL in strawberry (Zorrilla-Fontanesi et al., 2012) and tomato (Zanor et al., 2009).

In this study we take advantage of a high-throughput SNPs genotyping array coupled to GC-MS based-metabolomic approach to discover QTL for volatile compounds on peach fruit. The data presented here confirmed a locus controlling linalool and p-mentha-1-en-9-al described previously (Eduardo et al., 2013) but also showed that this locus controls the content of additional monoterpene compounds. Moreover, novel sources of variability on LG5 and LG6 were identified for the most important peach aroma-related compounds: lactones and esters which are independent of major pleiotropic loci and therefore could be used for improvement of peach flavor. The results presented here strengthen the actual knowledge regarding the genetic control of aroma and confirmed the genetic potential for improving the peach flavor by marker-assisted breeding.

## MATERIALS AND METHODS

### Plant material

The peach progeny investigated in this study was an F1 population derived from 86 seedlings which were obtained from a cross between the genotypes 'MxR_01' and 'Granada'. 'MxR_01' is a freestone melting-flesh peach which was obtained within the IVIA (*Instituto Valenciano de Investigaciones Agrarias*) breeding program and selected from the cross between the melting peach 'RedCandem' (obtained by a U.S. breeding program) and the non-melting peach 'Maruja' (a traditional Spanish variety). 'Granada' is a clingstone non-melting peach with low chilling requirement obtained from a Brazilian breeding program (Bassols Raseira and Bonifacio, 2006). Female parent of 'Granada' is Conserva 471 while the male parent is unknown. Replicate clones derived from each seedling in the collection were cultivated in three experimental orchards: "El Jimeneo" (EJ) and "Aguas Amargas" (AA), both of them situated in the Murcia region in Spain and at the *Insituto Valenciano de Investigaciones Agrarias* (IVIA) in Valencia, Spain. EJ is located at 80m of altitude in latitude: 37º 45' 31,5 N; longitude: 1º 01' 35,1 O. AA is located at 344m of altitude in latitude: 38° 31' N; longitude: 1° 31' O. IVIA is located at 55m altitude in latitude: 39° 34' N, longitude 0° 24' W. A total of 86 genotypes were grown at EJ, 74 at AA and 71 at IVIA. Horticultural practices such as pruning, irrigation, fertilization and control of weeds, insects and diseases were performed according to standard procedures in all orchards.

Peach fruit from F1 hybrids and parental genotypes were harvested from June to August 2012. The harvest date (HD) for each genotype analyzed was expressed as the difference in days with the date of the earliest genotype. Fruits harvested at EJ and AA were used for fruit and volatile analyses as is described

later. Fruits from IVIA location were analyzed only for the other fruit traits as indicated.

**Population genotyping and map construction**

DNA was extracted from 50 mg of young leaves following the method of Doyle & Doyle (Doyle, 1987). The concentration of DNA was checked by comparison with standard DNA labels in agarose geles and with Quant-iT™ PicoGreen H Assay (Life Technologies, Grand Island, NY, USA) Samples were genotyped using the IPSC peach 9 K Infinium ® II array which includes around 9000 SNPs peach markers (Verde et al., 2012) at the Genotyping and Genetic Diagnosis Unit (Health Research Institute, INCLIVA, Valencia, Spain).

Polymorphic markers were codified as cross-pollinator (CP) for linkage map construction with JoinMap® V4 (Kyazma B.V, Netherland) (Van Ooijen, 2006). Monomorphic SNPs and SNPs with more than 5% of missing data were removed. For genetic map construction, we followed the two-way pseudo-test cross approach (Grattapaglia and Sederoff, 1994). Homozygous SNPs in one parent and heterozygous in the other parent (and, therefore segregating 1:1 through the progeny) were selected to generate a genetic map for each parent, discarding SNPs heterozygous for both parents. Linkage groups with LOD of 6.0 to 8.0 were selected. Map construction was performed using the regression mapping algorithm (Stam, 1993) and the default JoinMap® parameters (Rec = 0.40, LOD = 1, Jump = 5.0, and ripple = 1). The order of markers in each linkage map was double-checked with MAPMAKER/EXP version 3.0b (Lander et al., 1987).The Kosambi mapping function was used to convert recombination frequencies into map distances. Maps were drawn with MapChart 2.2 (Voorrips, 2002)

**Fruit and volatile analyses**

Ten fruits corresponding to "harvest ripe" (also know as "ready to buy") stage were harvested from each tree in each of the EJ, AA and IVIA locations. Fruits were transported to the IBMCP laboratories in Valencia, Spain. Transportation was carried out at room temperature (20-28 °C) within the next 16 to 20 h. This period allowed fruit to ripe to "consumption ripe" (or "ready to eat") stage. Fruit with no-evident defects (disease, damaged, etc) and most homogeneous were selected (four to ten fruit) for further analysis. The maturity parameters (peel ground color, flesh firmness, weight, and total soluble solids, SSC) were analyzed as described previously (Sánchez et al., 2012) for fruit of EJ, AA and IVIA. Fruit were weighted and peel ground color parameters (L, lightness, C, chroma and H, color measured in Hue degree) were recorded. For fruits of EJ and AA, immediately after firmness measured, the fruit mesocarp was frozen in liquid nitrogen. For each genotype, the frozen samples were pooled to obtain a composite samples (biological replicates) that were assessed thrice for volatile analyses (technical replicates). Frozen samples of fruit mesocarp were ground to powder in liquid nitrogen. Volatile compounds were analyzed from 500 mg of frozen tissue powder, following the method described previously (Sánchez et al., 2012). The volatile analysis was performed on an Agilent 6890N gas chromatograph coupled to a 5975B Inert XL MSD mass spectrometer (Agilent Technologies). For the chromatography and mass spectra conditions, see Sánchez et al. (2012). A total of 47 commercial standards were used to confirm compound annotation. Volatiles were quantified relatively by means of the Multivariate Mass Spectra Reconstruction (MMSR) approach developed by Tikunov et al. (Tikunov et al., 2005). A detailed description of the quantification procedure is provided in Sánchez et al. (2012). The data was expressed as log2 of a ratio (sample/common reference) and the mean of the three replicates (per genotype, per location) was used for all the analyses performed.

**Data and QTL analysis**

The Acuity 4.0 software (Axon Instruments) was used for: hierarchical cluster analysis, heatmap visualization, principal component analysis, and ANOVA analyses.

Correlation network analysis was conducted with the Expression Correlation (www.baderlab.org/Software/ExpressionCorrelation) plug-in for the Cytoscape software (Shannon et al., 2003). Networks were visualized with the Cytoscape software, v2.8.2 (www.cytoscape.org).

Genetic linkage maps were condensed, eliminating co-segregating or very close linked SNPs, in order to reduce computer requirements for QTL analysis without loosing map resolution. Maps for each parental were analyzed independently and coded as two independent backcross populations. QTL analysis was performed by single marker analysis and composite interval mapping (CIM) methods with Windows QTL Cartographer v2.5 (Wang et al., 2012). A QTL was considered statistically significant if its LOD was higher than the threshold value score after 1000 permutation tests (at $\alpha$=0.05). Maps and QTL were plotted using Mapchart 2.2 software (Voorrips, 2002), taking one and two LOD intervals for QTL localization. Epistatic effect was assayed with QTLNetwork v2.1 (Yang et al., 2008) using default parameters.

## RESULTS

### SNP genotyping and map construction

The IPSC 9 K Infinium ® II array (Verde et al., 2012), which interrogates 8144 marker positions, was used to genotype at deep coverage our mapping population. To analyze only high-quality SNP data, markers with four or more missing data (around 300 SNPs in total) were eliminated from the data set. Non informative SNPs, i.e. those monomorphic and therefore not segregating were also eliminated, resulting finally in 3630 polymorphic markers. The marker segregation was tested against normal Mendelian expectation ratio (1:1) in order to analyze segregation distortion, and those markers showing segregation distortion (stated at $\alpha<0.05$) were eliminated to avoid map artefacts. Thus, a total of 2864 polymorphic SNPs (40% of the total) were identified (Table 1), from which 1970 segregated (1:1) for the 'MxR_01' parent and 895 for 'Granada' and were selected for their respective map construction

An example of the way we proceeded is shown in figure S2 for scaffold 1. A total of 282 polymorphic SNPs were located in the scaffold 1 of the peach genome assembly v1.0 segregating for 'MxR_01' parental. From them 265 markers could be grouped and ordered in a single linkage group with several markers co-segregating in the same position (Fig. S2). One SNP for each position was selected (26 in total) to obtain a condensed map. Similarly, map corresponding to the others scaffolds (3, 4, 5, 6, 7, and 8) were obtained except for Sc2 where a map inconsistent with genome position and with large gaps (bigger than 30 cM) was obtained and was discarded for being not suitable for QTL analyses. A total 178 SNPs were located in the 'MxR_01' condensed map, representing a total distance of 480 cM (Table 1). The marker density varies between 1.98 cM/marker (for LG8) to 4.08 cM/marker (for LG6). On average, one marker per 2.94 cM was found in the 'MxR_01' map (Table 1).

For 'Granada', a lower number of polymorphic markers were obtained as compared to 'MxR_01' (Table. 1). Following the same strategy as described for 'MxR_01', the maps for Sc: 2, 4, 5, 6, 7, and 8 were obtained for 'Granada'. Whereas for Sc 1 only two linked markers were found and no map was obtained for Sc3. Only linkage groups of Sc6 and Sc7 showed evenly distributed markers with good coverage (as is presented later). The map obtained covered less distance compared to 'MxR_01' (276 vs 480 cM) with less marker density (3.87 vs 2.94 cM/marker in average).

Marker orders were double-checked with MAPMAKER 3.0. No important differences were observed with Joinmap 4.0, so the order obtained with the latter was used for our analysis.

**Evaluation of volatile variability of the mapping population**

Volatile compounds were analysed from the populations grown in the different agro-ecological zones EJ and AA. As an example of the variability of the mapping population, a picture of some representative fruits (grown at EJ) are shown in figure S2. Genotypes growing at EJ ripened on average 7.9 days earlier as compared to AA (stated by ANOVA at $\alpha$<0.01), probably due to the warmer weather in AA compared with EJ. This results, suggested that both locations represent significant different environments.

A total of 81 volatiles were profiled (Table 2). To asses the environmental effect, the Pearson correlation of volatiles levels between EJ and AA locations was analyzed. Around half of the metabolites (41) showed significant correlation but only 18 of them showed correlation higher than 0.40 (Table 2) indicating that a large proportion of the volatiles are influenced by the environment.

| | | Polymorphic SNPs | | | SNPs selected | | Map distance (cM) | | Marker density (cM/marker) | |
|---|---|---|---|---|---|---|---|---|---|---|
| **Scafold** | **Total SNPs** | **SNPs (% of Total)** | **MxR_01'** | **Granada'** | **MxR_01'** | **Granada'** | **MxR_01'** | **Granada'** | **MxR_01'** | **Granada'** |
| Sc1 | 959 | 319 (33%) | 282 | 37 | 26 | 2 | 75.01 | 12.00 | 2.89 | 6.00 |
| Sc2 | 1226 | 461 (38%) | 273 | 188 | 0 | 13 | 0 | 59.08 | X | 4.54 |
| Sc3 | 700 | 336 (48%) | 325 | 11 | 40 | 0 | 87.28 | 0 | 2.18 | X |
| Sc4 | 1439 | 496 (34%) | 269 | 227 | 29 | 10 | 69.95 | 22.46 | 2.41 | 2.25 |
| Sc5 | 476 | 243 (51%) | 196 | 47 | 14 | 8 | 50.8 | 39.61 | 3.63 | 4.95 |
| Sc6 | 827 | 364 (44%) | 188 | 176 | 15 | 20 | 61.18 | 75.75 | 4.08 | 3.79 |
| Sc7 | 686 | 318 (46%) | 168 | 150 | 21 | 16 | 70.45 | 50.87 | 3.35 | 3.18 |
| Sc8 | 804 | 328 (41%) | 269 | 59 | 33 | 7 | 65.37 | 16.70 | 1.98 | 2.39 |
| **TOTAL** | **7117** | **2865 (40%)** | **1970** | **895** | **178** | **76** | **480** | **276** | | |

**Table 1. Summary of the SNPs analyzed for scaffold 1-8.** For each scaffold the total number of SNPs present in the array (Total SNPs) and the number of polymorphic marker with the percentage of the total (in parenthesis) are indicated. Also for each parental map ('MxR_01' and 'Granada'), the total number of polymorphic SNPs found at each scaffold and the number of SNPs selected for map construction are indicated. Map distance (in cM) indicates the length of linkage group corresponding to each chromosome and the total map distance covered for both parental maps. Marker density indicates the distance between contiguous markers (in average) for each linkage group.

| Nº | id | Name | Cluster | Family | corr_EJ-AA |
|---|---|---|---|---|---|
| 1 | 43_7136 | Acetic acid, butyl ester | C1 | Ester | 0.34** |
| 2 | 73_9306 | 3,4-Dimethyl-3-hexanol | C1 | Alcohol | 0.25* |
| 3 | 41_6638 | Propanoic acid, 2,2-dimethyl- | C1 | Carboxilic acid | 0.10 |
| 4 | 73_11030 | **Heptanoic acid** | C1 | Carboxilic acid | 0.21 |
| 5 | 165_16433 | 8,8,9-Trimethyl-deca-3,5-diene-2,7-dione | C2 | Long Ketone/Ester | 0.02 |
| 6 | 74_14613 | Decanoic acid, methyl ester | C2 | Long Ketone/Ester | 0.17 |
| 7 | 58_12978 | 2-Decanone | C2 | Long Ketone/Ester | 0.11 |
| 8 | 72_12911 | 3-Decanone | C2 | Long Ketone/Ester | 0.11 |
| 9 | 74_11980 | Octanoic acid, methyl ester | C2 | Long Ketone/Ester | 0.18 |
| 10 | 70_12284 | Acetic acid, 2-ethylhexyl ester | C2 | Long Ketone/Ester | 0.08 |
| 11 | 74_13339 | Nonanoic acid, methyl ester | C2 | Long Ketone/Ester | 0.09 |
| 12 | 74_16928 | Dodecanoic acid, methyl ester | C2 | Long Ketone/Ester | 0.07 |
| 13 | 70_13492 | 2-Propenoic acid, 2-ethylhexyl ester | C2 | Long Ketone/Ester | -0.03 |
| 14 | 55_13398 | 4-(Prop-2-enoyloxy)octane | C2 | Carboxilic acid | -0.03 |
| 15 | 91_11117 | **Benzeneacetaldehyde** | C2 | Aromatic | 0.25* |
| 16 | 73_11681 | **Hexanoic acid, 2-ethyl-** | C3 | Carboxilic acid | -0.04 |
| 17 | 73_12448 | **Octanoic Acid** | C3 | Carboxilic acid | -0.05 |
| 18 | 60_9627 | **Hexanoic acid** | C3 | Carboxilic acid | 0.07 |
| 19 | 73_13753 | **Nonanoic acid** | C3 | Carboxilic acid | 0.14 |
| 20 | 73_14964 | Decanoic acid | C3 | Carboxilic acid | -0.02 |
| 21 | 172_17688 | 2,7-Ethanonaphth[2,3-b]oxirene, 1a,2,7,7a-tetrahydro-, (1aa,2a,7a,7aa)- | C4 | un. | 0.06 |
| 22 | 178_15640 | Isoeugenol methyl ether | C4 | Aromatic Ether | 0.17 |
| 23 | 105_18590 | **Benzophenone** | C4 | Aromatic Ketone | 0.10 |
| 24 | 196_20012 | Benzophenone, 3-methyl- | C4 | Aromatic Ketone | -0.06 |
| 25 | 108_12704 | **Acetic acid, phenylmethyl ester** | C4 | Aromatic Ester | 0.13 |
| 26 | 95_12787 | Bicyclo[2.2.1]heptan-2-one, 1,7,7-trimethyl-, (1R)- | C4 | Ketone | 0.20 |
| 27 | 105_14972 | Benzene, (1,3-dimethylbutyl)- | C4 | Aromatic | 0.21 |
| 28 | 120_14256 | Salicylic acid, ethyl ester | C4 | Aromatic Ester | 0.23 |
| 29 | 121_11693 | **Terpinolene** | C5a | Monoterpene | 0.31* |
| 30 | 121_13280 | **α-Terpinol** | C5a | Monoterpene | 0.31* |
| 31 | 139_9925 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | C5a | Monoterpene | 0.39** |
| 32 | 93_11733 | **Linalool** | C5a | Monoterpene | 0.39** |
| 33 | 93_10867 | **δ-Limonene** | C5b | Monoterpene | 0.33* |
| 34 | 59_11450 | ***cis* -Linaloloxide** | C5b | Monoterpene | 0.73* |
| 35 | 59_11676 | α-Methyl-α-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | 0.67* |
| 36 | 119_13188 | Ethanone, 1-(4-methylphenyl)- | C5b | un. | 0.81* |
| 37 | 138_12371 | 4-Acetyl-1-methylcyclohexene | C5b | Monoterpene | 0.87* |
| 38 | 117_11758 | Benzene, 4-ethenyl-1,2-dimethyl- | C5b | un | 0.77* |
| 39 | 43_13124 | Benzenemethanol, a,a,4-trimethyl- | C5b | un. | 0.75* |
| 40 | 123_10728 | 1-Cyclohexene-1-carboxaldehyde,2,6,6-trimethyl | C5b | Monoterpene | 0.86* |
| 41 | 134_12449 | 1,3,8-p-Menthatriene | C5b | Monoterpene | 0.72* |
| 42 | 137_13215 | 3,6-Dimethyl-2,3,3a,4,5,7a-hexahydrobenzofuran | C5b | un. | 0.84* |
| 43 | 94_13611 | 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl | C5b | Monoterpene | 0.77* |
| 44 | 69_13073 | 2-Cyclohexen-1-ol, 2-methyl-5-(1-methylethenyl)-, cis | C5b | un. | 0.66* |
| 45 | 71_11820 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5b | un. | 0.68* |
| 46 | 119_10786 | ***p* -Cymene** | C5c | Monoterpene | 0.76* |
| 47 | 43_3992 | **Ethyl Acetate** | C6 | Ester | 0.35* |
| 48 | 45_2418 | **Ethanol** | C6 | Alcohol | 0.22* |
| 49 | 68_15955 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | C6 | Lactone | 0.48* |
| 50 | 43_10318 | **Acetic acid, hexyl ester** | C7 | Ester | 0.27** |
| 51 | 67_10239 | **3-Hexen-1-ol, acetate, (Z)-** | C7 | Ester | 0.18 |
| 52 | 43_10351 | **2-Hexen-1-ol, acetate, (E)-** | C7 | Ester | 0.40** |
| 53 | 80_10583 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | 0.06 |
| 54 | 43_6475 | **Acetic acid, 2-methylpropyl ester** | C8 | Ester | 0.36** |
| 55 | 85_11181 | **γ-Hexalactone** | C8 | Lactone | 0.50** |
| 56 | 95_16515 | **Pyran-2-one, 6-pentyl-** | C8 | Lactone | 0.27* |
| 57 | 99_16910 | **δ-Decalactone** | C8 | Lactone | 0.36** |
| 58 | 85_16556 | **γ-Decalactone** | C8 | Lactone | 0.14* |
| 59 | 85_14019 | **γ-Octalactone** | C8 | Lactone | 0.33* |
| 60 | 85_15326 | γ-Nonalactone | C8 | Lactone | 0.33* |
| 61 | 68_16405 | **γ-Jasmolactone** | C8 | Lactone | 0.39** |
| 62 | 85_12613 | **γ-Heptalactone** | C8 | Lactone | 0.28* |
| 63 | 177_16736 | **β-Ionone** | C9 | Carotenoid-derived | 0.14 |
| 64 | 69_15595 | **β-Damascenone** | C9 | Carotenoid-derived | 0.30* |
| 65 | 108_9983 | **5-Hepten-2-one, 6-methyl-** | C9 | Carotenoid-derived | 0.27* |
| 66 | 43_16192 | **Geranyl acetone** | C9 | Carotenoid-derived | 0.20 |
| 67 | 56_12707 | **Nonalol** | C9 | Alcohol | 0.12 |
| 68 | 56_9747 | **Heptanol** | C9 | Alcohol | -0.01 |
| 69 | 45_3430 | Acetic acid | C10 | Carboxilic acid | 0.33* |
| 70 | 57_10633 | 1-Hexanol, 2-ethyl- | C10 | Alcohol | 0.09 |
| 71 | 56_11249 | **1-Octanol** | C10 | Alcohol | 0.01 |
| 72 | 57_5029 | **1-Penten-3-ol** | C11 | Lipid-derived | -0.14 |
| 73 | 69_6953 | **3-Hexenal, (Z)-** | C11 | Lipid-derived | 0.22 |
| 74 | 81_5323 | **Furan, 2-ethyl-** | C11 | Lipid-derived | 0.18 |
| 75 | 55_5080 | **1-Penten-3-one** | C11 | Lipid-derived | 0.19 |
| 76 | 81_10494 | **2,4-Heptadienal, (E,E)-** | C11 | Lipid-derived | 0.22 |
| 77 | 44_5291 | **Pentanal** | C12 | Lipid-derived | 0.06 |
| 78 | 81_10119 | **Furan, 2-pentyl-** | C12 | Lipid-derived | 0.57** |
| 79 | 56_6998 | **Hexanal** | C12 | Lipid-derived | 0.25* |
| 80 | 83_7941 | 2-Hexenal, (E)- | C12 | Lipid-derived | 0.24 |
| 81 | 164_15243 | **Eugenol** | C12 | Aromatic alcohol | 0.46** |

**Table 2 (see the legends below).**

**Table 2. Volatile analyzed in this study.** For each volatile, the cluster (C1-C12) where the compound was found in the HCA of Fig. 2 is shown. Cluster 5 is divided in three sub-clusters indicated with letters: a, b and c. The volatile number (Nº) indicated the compound position on the HCA. For each compound an identification code (id) was given that is formed by the ion use for quantification and the retention time (given in scan number) where the peak was found. Compounds identified by comparing of its retention time with authentic standards are highlighted in bold letters. Family indicates the biosynthetic origin or chemical nature of the volatile. un., unknown. The Pearson correlation coefficients of volatile levels between EJ and AA locations are indicated (corr_EJ-AA). * and ** indicate that the correlation is significant at $\alpha$=0.05 and $\alpha$=0.01, respectively. The correlation coefficient shaded indicates that stable QTL for that volatiles were found.

To get a deeper understanding of the structure of the volatile data set, a PCA were conducted. Genotypes were distributed in the first two components (PC1 and PC2 explaining 22% and 20% of the variance respectively) without forming clear groups (Fig. 1A). Genotypes located in EJ and AA were no clearly separated, although at extreme PC2 values, samples tends to group according location, pointing an environmental effect. Loading score plot (Figure 1B) indicated that lipid-derived compounds (73-80, numbered according Table 2), long chained esters (6, 9, and 11), ketones (5, 7 and 8) along with Acetic acid, 2-ethylhexyl ester (10) would be the VOCs more influenced by location (Fig. 1B). According with this analysis, fruits harvested at EJ are expected to have higher levels lipid-derived compounds while long chained esters and ketones and Acetic acid 2-ethylhexyl ester would accumulate to higher levels in fruits harvested in AA. This result indicated that these compounds are likely the most influenced by the local environment conditions. On the other hand, PC1 separated the lines mainly on the basis of the concentration of lactones (49 and 56-62), lineal esters (47, 50, 51, 53, and 54) and monoterpenes and other related compounds from unknown origin (29-46), so it is expected that those VOCs are under stronger genetic control.

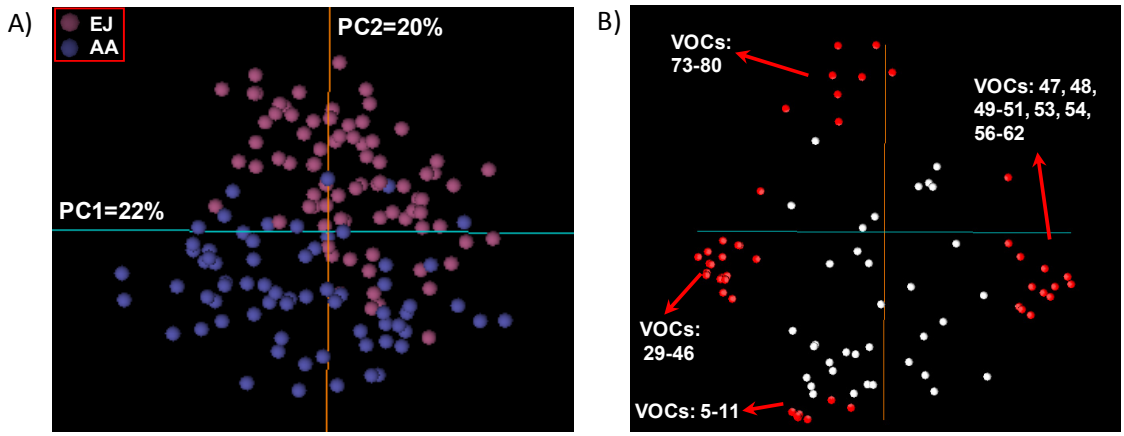



**Figure 1. Principal component analysis of the volatile data set. A)** Principal component analysis of the population mapping. Hybrids harvested at locations: EJ and AA are indicated with different colors. **B)** Loading plots of PC1 and PC2. In red are pointed the volatiles (numbered according Table 2) that accounted most for the variability of the aroma profiles across PC1 and PC2, for more detail see the text.

To analyze the relationship between metabolites, a HCA was conducted for volatile data recorded in both locations. This analysis revealed that volatile compounds grouped in 12 main clusters; most of them having members of known metabolic pathways or similar chemical nature (Fig. 2, Table 2). Cluster 2 is enriched with methyl esters of long carboxylic acids, i.e. 8-12 carbons (6, 9, 11, and 12), other esters (10 and 13) and ketones of 10 carbons (5, 7, and 8). Similarly, carboxylic acids of 6-10 carbons are grouped in cluster 3 (16-20). Cluster 4 is mainly formed by volatiles having aromatic rings. In turns, monoterpenes (29-34, 37, 40, 41, 43, and 46) are grouped in cluster 5 with other compounds with ten carbons from yet unknown origin. Ethanol and its acetate ester (47) clustered together in C6. Esters derived of acetyl-CoA and six carbons alcohols (50-53) grouped in cluster 7. All detected lactones, with the exception of number 49, were grouped in cluster in C8. Four carotenoid-derived volatile (63-66) are found in C9, while lipid-derived compounds are grouped in C11 and C12. These results suggested that volatiles are co-regulated according specific modules in this F1 population represented by fruits normalized by ripening stage. The heat map revealed that the genotypes contain different combinations

of these volatile modules. For example, the clusters of genotypes S7-S9 have high levels of volatiles belonging to C5 (which is rich in monoterpenes), while clusters S5 and S6 have low levels of those compounds (Fig. 2). Even there are genotypes, those of S1-S4, with different concentration of volatiles in the sub-cluster of C5.

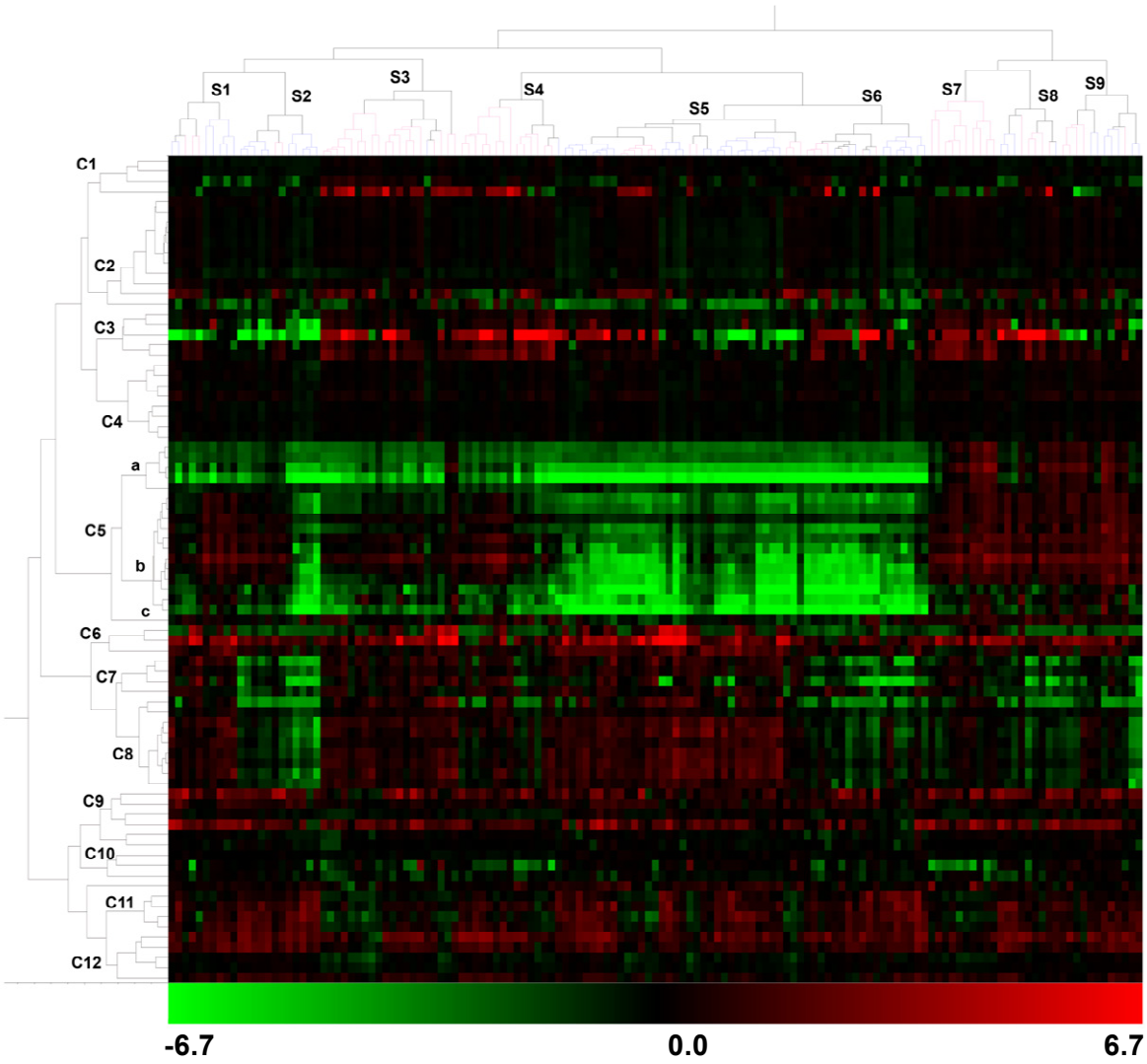



**Figure 2. Hierarchical cluster analysis and heatmap.** On the dendrogram of volatiles (at left) are indicated the clusters obtained: C1-C12. The order of the volatile in the dendrogram is in correspondence with the one indicated on Table 2. The upper dendogram corresponds to genotypes where the samples clusters are indicated with: S1-S9. Data are expressed as a log2 of a ratio (sample/common reference). Below the heatmap is indicated the scale used.

A correlation network analysis (CNA) was conducted to further investigate the association between metabolites as well as the interrelationship between volatile modules. As expected, the volatiles that clustered together on the HCA were interconnected by positive interaction represented with blue lines in CNA (Fig. 3). As previously reported (Sánchez et al., 2012), lactones and lipid derived compounds showed negative interactions mainly through 2-Hexenal (E)-. Also lactones showed high correlation with lineal esters of C7 (50-53), Ethyl Acetate and Acetic acid butyl ester, the only ester on C1. Volatile of C2 and C4 are interconnected with high positives correlations. And these two modules showed positive correlation with C1 volatiles through the interaction with 3,4-Dimethyl-3-hexanol. In turn, volatiles from C2 interact negative with lipid-derived compounds of C11. On the other side, compounds of C5 are highly correlated to each other but remains quite isolated from the rest of the compounds.

Taken together, these results suggested that, within our population, volatiles are co-regulated according to their specific groups and that the genotypes have different combination of volatile modules what should condition their aroma profiles.

**Genetic control analysis of volatile compound synthesis and fruit quality traits.**

Peach volatile biosynthesis is highly dependent on the ripening stage of fruit (Aubert et al., 2003; Chapman et al., 1991; Sánchez et al., 2012; Sánchez et al., 2013; Visai and Vanoli, 1997; Zhang et al., 2010). For that reason we also analyzed QTL for the main characteristics that have been traditionally used to asses maturity stage of the peach fruit (and therefore quality): flesh firmness, weight, SSC and peel color-related variables in order to compare the locus controlled by maturity with those controlling volatile production.

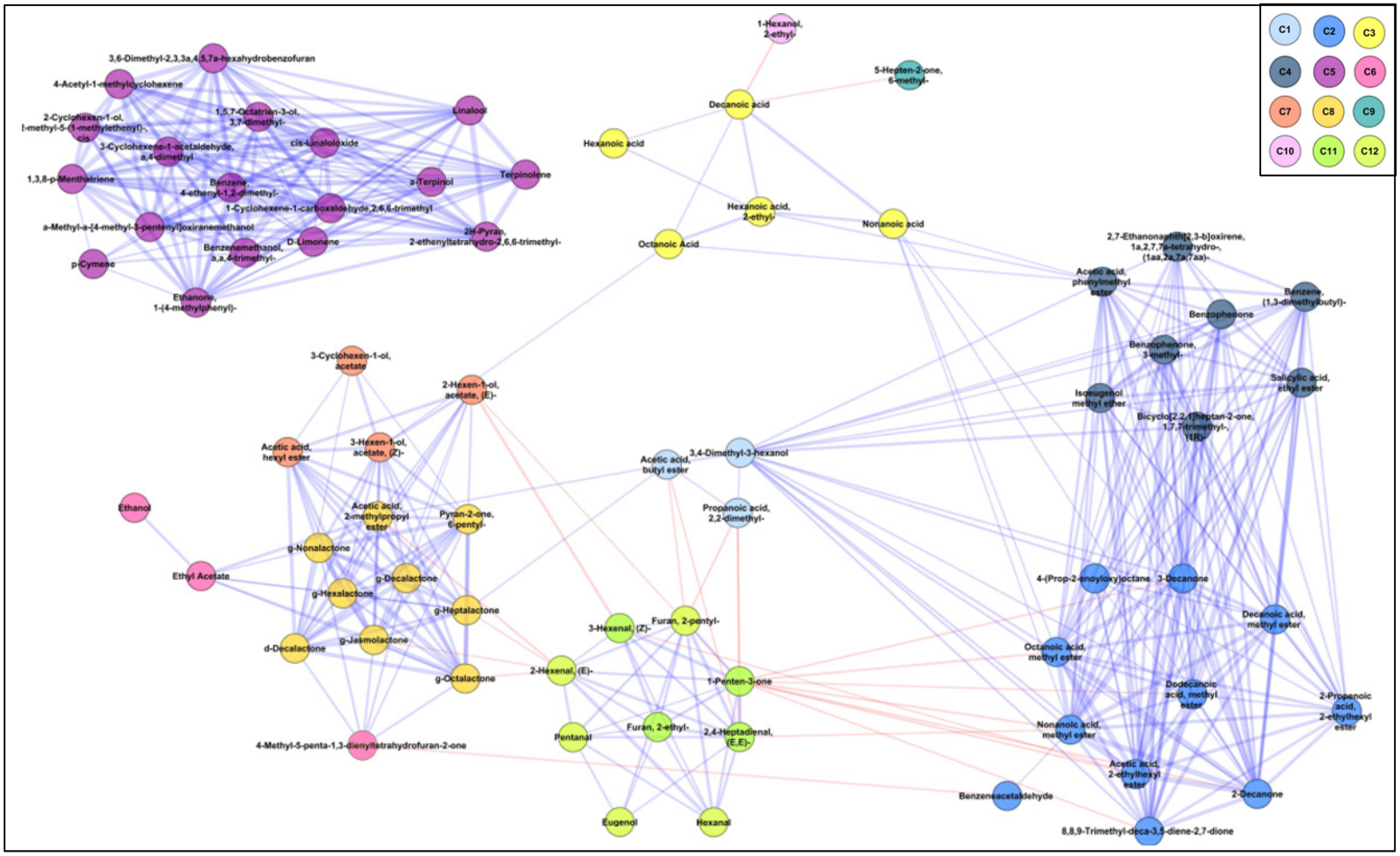



**Figure 3. Correlation network analysis of the data set.** The nodes representing volatiles are colored according to the cluster that they were found (C1-C12) according to Fig. 2, as indicated on the top-right corner. Positives and negatives correlation are indicated with blue and red edges respectively. Line thickness indicated correlation strength, the wider the line, the stronger the correlation.

In a similar way, the Harvest Date (HD) (analyzed also by Romeu et al., submitted) was also included in our analysis since it has been proposed that a major HD QTL at the south end of LG4 has a pleiotropic effect on volatile production on peach (Eduardo et al., 2013). Additionally as our mapping population segregated for melting/non-melting flesh (MnM) this trait was also included to analyze if there is a possible pleiotropic effect of the locus controlling flesh type on volatile production.

A large number of QTL were detected for both fruit quality traits and volatile production (Tables S1, S2 and S3). Most of them were detected only in the 'MxR_01' map, probably due to the higher genetic diversity among the progenitors of 'MxR_01' than among the progenitors of 'Granada'. To graphically summarize the genetic control of volatile, the likelihood of association between markers and compounds are presented as heatmaps in supplementary data (Fig. S3, Fig. S4). A proportion of the QTL identified (in general, between 20-40% depending on the trait) were consistently detected in at least two locations. These consistent QTL are represented in figures 4 and 5.

In general, volatile modules showed similar LOD profiles in defined regions of chromosomes, suggesting the presence of loci that increase the production of some volatile modules. For example in 'MxR_01', volatiles bellowing to the monoterpene-enriched cluster C5 showed similar LOD profiles on LG1, LG4, and LG5 in both locations (Fig. S3). Additionally, this analysis showed that LG8 of 'MxR_01' map exerted a very little control of the peach volatilome. On the contrary variability of compounds belonging to C3 and C10 clusters (all formed by carboxylic acids and alcohols) were not associated to any genomic region, indicating absence of allelic variability in the control of those compounds in the sources of variability analyzed (Fig. S3).

For 'MxR_01' map, most of the consistent QTL were found in LG4 forming two clusters (Fig. 4).

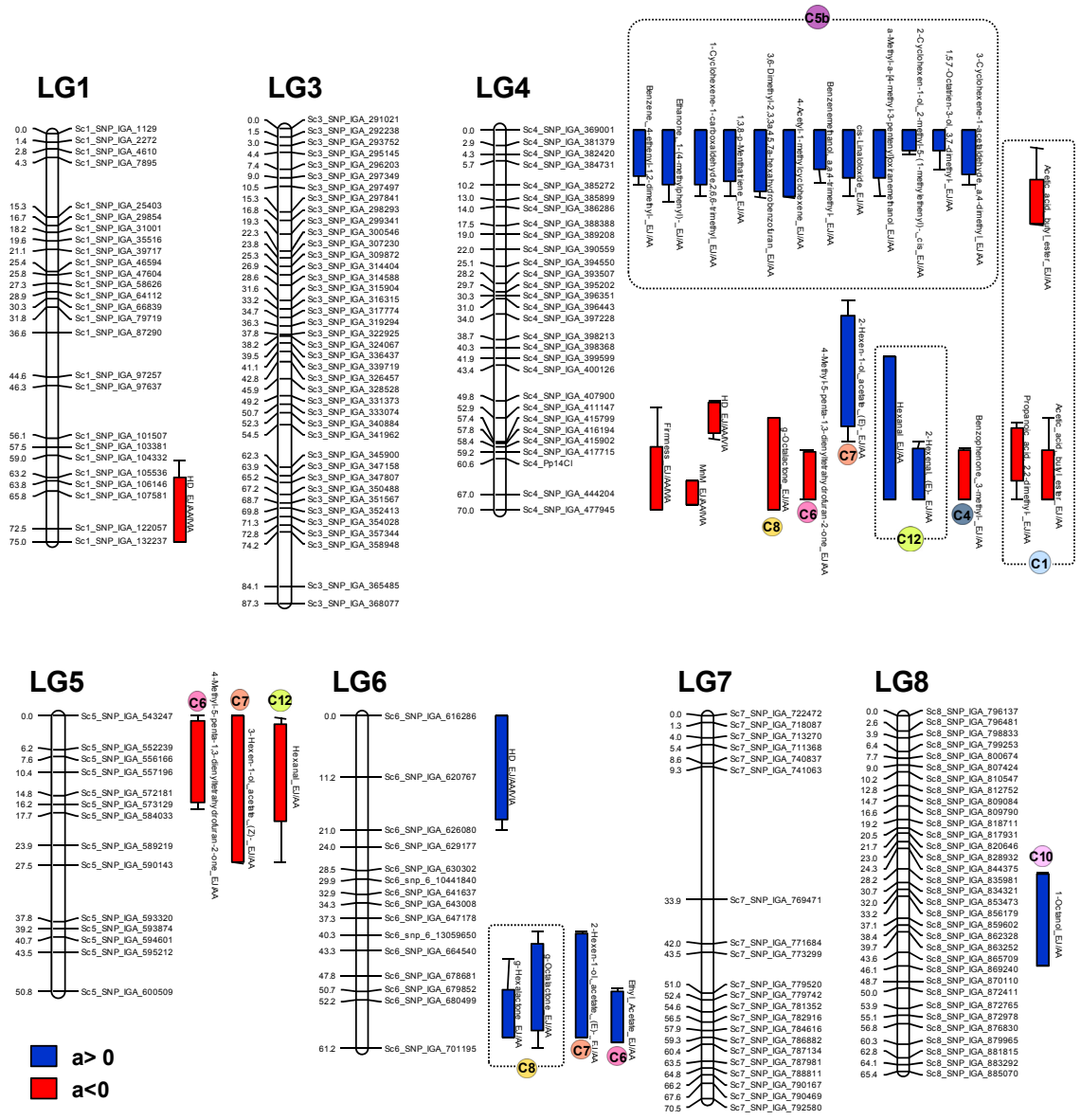



**Figure 4. Location of volatile QTL stable across location for 'MxR_01' map.** Consistent QLTs found at the location: EJ, AA (for volatiles) and IVIA (for traits: MD, Firmness and MnM) are shown. The QTL are colored according the additive effect (a) that exerts. Bars and lines represent the 1-LOD and 2-LOD support intervals for each.

At the upper end of LG4, QTL for 12 (out of 13) volatiles of cluster C5b were identified. At the southern end of the LG4, QTL for lactones, esters, lipid derived compounds and others volatiles co-localizing with the loci controlling HD, MnM and firmness were found. In the later QTL cluster, QTL controlling the production of lactones 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one

and γ-Octalactone showed negative additive effects, whereas those affecting two lipid derived compounds (Hexanal and 2-Hexenal (E)-), and a lineal esters (2-Hexen-1-ol acetate (E)-) showed positive additive effect. Another cluster of QTL controlling the production of a lactone, an ester and a lipid derived compound was also found at the top of LG5. In addition a cluster of QTL was fond at the southern end of the LG6 thus defining a locus controlling the content of two lactones (γ-Hexalactone and γ-Octalactone) and two esters (Ethyl Acetate and 2-Hexen-1-ol acetate (E)-) with the same direction of additive effect.

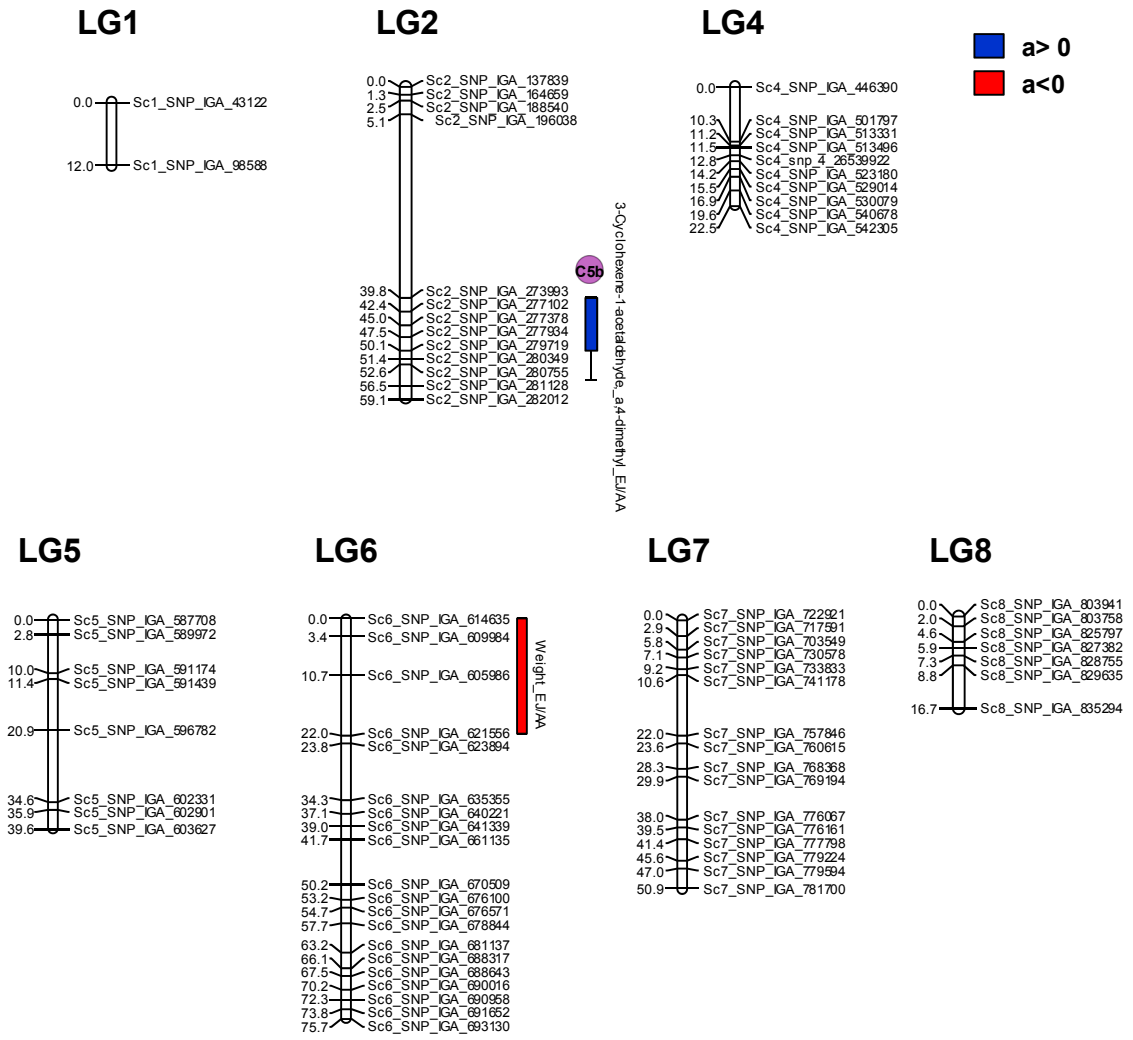



**Figure 5. Location of volatile QTL stable across location for 'Granada' map.** The two consistent QLTs found are shown. The QTL are colored according the additive effect (a) that exerts. For volatiles QTL, the circles with different colors

(according Fig. 2) indicate the cluster that the controlled-volatile belongs to. QTL for volatiles of the same cluster at the same linkage group are indicated with dashed-line rectangles. Bars and lines represent 1-LOD and 2-LOD support intervals.

To further analyze the potential of this materials and information for volatile improvement, the epistatic effects between QTL were analyzed for all traits, but no significant effects were detected for the stable QTL indicated on figure 4 (data no shown).

For 'Granada' map, fewer QTL were found compared to 'MxR_01' (Table S2) and only for the compound 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl a QTL stable along locations was found (Fig. 5). Also a stable QTL for fruit weight explaining between 14-16% of the variance was identified on LG6 (Fig. 5).

**Assessment of the breeding population potential for improvement**

Since QTL analysis showed that MnM locus co-localized with a cluster of volatile QTL (Fig. 4), we compared the volatile profile of melting and no-melting genotypes within our population. As expected, melting and no-melting peaches showed different levels of volatiles with QTL co-localizing in that region (Table S4). According to the direction of additive effects observed, melting peaches showed lower levels of γ-Octalactone and 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one but also of other six lactones (Table S4). In the same way, Acetic acid butyl ester and Propanoic acid 2,2-dimethyl- levels were lower in melting peaches compared to no-melting ones. On the contrary, melting genotypes showed high levels of Hexanal and 2-Hexenal, (E)- along with other lipid derived compound (Pentanal).

The genotypes showed a similar trend of ripening in EJ, AA and IVIA, being the HD highly correlated between locations (r=0.94 to 0.97). According to the mean of HD across locations, the genotypes were divided in early, medium and late season. In our population, around half of the peaches were melting and the

other half no-melting (54% and 46%, respectively). Since the QTL for HD with major effects was found near the MnM locus, the effect of this linkage was analyzed. As expected according the direction of additive effects, early genotypes tend to be melting type (83%) while among the late genotypes most of the peaches are no-melting (79%, Table S5). The potential of prediction of fruit type was assessed. The genotypes were divided according to the ideotype of the two markers closest to the locus MnM (Sc4_SNP_IGA_444204, Sc4_SNP_IGA_477945). In the group with ideotype corresponding to melting peaches, 96% of the genotypes were actually phenotyped as melting type. In the group predicted to be no-melting according the ideotype, 83% were actually phenotyped as such.

To evaluate the potential for volatile improvement, the breeding population was divided according the ideotype at the different locus controlling the aroma production. For the locus controlling most of the monoterpenes of C5b (Fig. 4), the population was divided according the ideotype of the region expanding the QTL at LG4 (Sc4_SNP_IGA_369001 to Sc4_SNP_IGA_386286). The levels of all volatiles were compared between the group expected to have high levels of these compound with other group formed by the rest of the genotypes (i.e. having the contrary ideotype or recombinants in that region). The expected rich-monoterpene ideotype group showed high levels for all the compounds of the C5b but also for the rest of monoterpenes of C5 (Table S6). As a side effect, the monoterpene-rich group showed lower levels of Acetic acid butyl ester as a QTL with opposite effect was located near the locus tagged (Fig. 4).

Similarly the genotypes were divided according the ideotype at the three loci that showed QLTs for lactones at LG4 (Sc4_SNP_IGA_411147 to Sc4_SNP_IGA_415902), LG5 (Sc5_SNP_IGA_543247 to Sc5_SNP_IGA_584033) and LG6 (Sc6_snp_6_13059650 to Sc6_SNP_IGA_701195). Only four genotypes have rich-lactone ideotype, all are no-melting peaches for medium (3 genotype)

or late (1 genotype) season. This group has in mean higher levels of five lactones compared to the rest of the genotypes (Table S7).

## DISCUSSION

As part of our ongoing efforts dedicated to the identification of genes and locus controlling important fruit-quality traits in peach, we investigated the genetic control of aroma production and its relationship with other fruit quality characteristics. In this work, we took advantage of high-throughput genotyping and metabolite-profiling technologies in order to perform a large scale QTL analysis.

### Maps construction by high-throughput SNP genotyping

Despite the wide genome coverage represented in the IPSC peach 9 K SNPs array (Verde et al., 2012), the chromosomes 2 in 'MxR_01' map and 3 in 'Granada' did not have enough polymorphic SNPs markers to obtain a minimum genetic map for those chromosomes (Table 1, Fig. 4). In the case of 'Granada', only linkage maps covering the whole chromosomes 6 and 7 were obtained, whereas only partial coverage linkage groups were obtained for the rest of chromosomes. The most likely explanation for the extensive homozygosity detected for chromosome 2 of 'MxR_01' is the identity by descent, i.e. 'Maruja' and 'RedCandem' shares the same ideotype of chromosome 2. Since 'Maruja' is a traditional variety which pedigree is unknown and therefore is not possible to check this hypothesis. The male parental of 'Granada' is unknown (Bassols Raseira and Bonifacio, 2006) so its possible that this genotype is self-pollinated which could explained the extensive homozygosity found.

The linkage map quality depends on the characteristics of the population used (population type, number of individuals genotyped, the genetic origin of the parentals, etc) but also on related to the power of the genotyping platform utilized. The F1 population analyzed by Eduardo et al. (2013) was also genotyped with the IPSC 9K SNP array and also showed a low level of

polymorphic SNPs (1748 in total vs the 2864 SNPs found here, Table 1) but the total genetic distance are comparable (405 cM and 228 cM for the two parental map vs 480 cM and 276 cM in the maps obtained here, Table 1). An alternative high-density SNP genotyping approach based on parent sequencing for SNPs discovery was used for the detection of QTL for peach quality traits (Martínez-García et al., 2013). In that case, the number of polymorphic marker (1775 SNPs) and the map coverage (422 cM and 369 cM) reported were comparable to ours results although the map was denser (0.81 cM/markers in average vs 3.87 and 2.94 cM/marker for each map). SNP genotyping chips are an inflexible assay, subject to assortment bias, i. e., they may be suitable for a certain sample of germplasm but not appropriate for other samples. In our case, we can not discard wheter the lack of polymorphic SNPs in certain chromosomes is caused by actual homozygosis or by design bias of the chip. Currently, Genotype-by-Sequence technologies (Elshire et al., 2011) could allow to overcome assortment bias.

**Monoterpenes module is controlled by a unique locus while lactones and other lineal esters showed several QTL.**

To get a first insight of the structure of the data set, a series of correlation-based analyses (HCA and CNA) and a data reduction method (PCA) were conducted (Fig. 1-3). Previously, we analyzed the correlation patterns of volatiles in a complex sample set (formed by four genotypes analyzed in different locations, at different maturity stages and after a post-harvest treatment) to define groups of co-regulated compounds (Sánchez et al., 2012). Here, the correlation-based analyses also showed that the volatile complement in genetically diverse siblings of our population represented by fruits at the same ripening stage is highly organized in modules (Fig. 2, Fig. 3) and the co-regulations patterns found are markedly similar to the formerly described. However, the novel

results presented here revealed that several of the co-regulated groups are not necessarily genetically controlled or at least are strongly affected by the environment. Related to the environmental control, the PCA suggested a group of compounds that account for a separation among locations (Fig. 1) and therefore expected to reflect the influence of environment on volatile production in our population. To further emphasize the importance of environment, only 50% of the volatile analyzed showed significant correlation between locations (Table 1). Conversely, PCA showed that lactones, esters and monoterpenes accounted for the separation among genotypes independent of the location and therefore points to an important genetic control for those volatiles in our current population (Fig. 1). Accordingly, most of the stable QTL found were for these compounds: lactones, esters and monoterpenes (Fig. 4, Fig. S3). Eduardo et al. (2013) also found a strong environmental effect with less than 9% of the volatiles analyzed in that case showing significant correlation between the years of evaluation.

Previously, we proposed that lipid derived-compounds and lactones are inversely regulated during ripening and speculated that this could be due to a shift in fatty acid metabolism (Sánchez et al., 2012). Here we indentified a locus that controls the levels of some of the members of these two groups of volatiles in antagonist manner (i.e. with opposite additive effects). Accordingly, this locus, located at the end of LG4, co-localized with a major QTL that control the harvest date (Fig. 4). Recently, a cluster of QTL for some esters, lactones and other volatiles was identified at the lower half of linkage group LG4 (Eduardo et al., 2013) and the authors interpreted that a locus with pleiotropic effect is the responsible since at the southern end of that chromosome a locus controlling maturity related trait (including HD) had been identified earlier by the same research group (Eduardo et al., 2011). QTL for HD had been detected in different mapping populations at LG1, LG2, LG3, LG4 and LG6, with those

located on LG4 and LG6 having the most important effect (Dirlewanger et al., 2012; Eduardo et al., 2011; Quilot et al., 2004; Verde et al., 2002). Here we detected three QTL controlling HD on LG1, LG4 and LG6 of the 'MxR_01' map that coincide with the positions reported (Fig. 4). Among them the one on LG4 explained the larger proportion of the variance (50% on average across locations: EJ, AA and IVIA) and has the larger additive effect (-23.4 days on average). Early ripening cultivars are often a desirable objective of breeding programs since its fruits reach better market prices because of the "novelty" phenomenon. Since the QTL located on LG4 partially overlaps with a locus controlling the production of the important fruit aromas (γ-Octalactone, 2-Hexen-1-ol acetate (E)- and Hexanal), the use of this QTL to reduce the harvest time would affect the aroma profile and *vice versa*. On the contrary, the QTL for HD in LG1 and LG6 (with lower effects than the previous one, 18% and 9%, respectively) did not co-localized with aroma QTL, being therefore more suitable for breeding for earliness without affecting quality.

Our analysis found a locus controlling MnM trait coinciding with the localization previously reported (Dirlewanger et al., 2004). The Melting locus (M) co-localized with flesh firmness and volatile QTL (Fig. 4), what may be likely due to pleiotropic effects of this locus. The additive effect of these QTL suggests that selecting for no-melting flesh type in our current program would increase the levels of two lactones (γ-Octalactone, 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one) and a ester (Acetic acid, butyl ester) and decrease the levels of 2-Hexenal, (E)-, which would probably improve the overall fruit aroma. Accordingly, no-melting and melting genotypes showed differences in these volatiles as well as in other important aroma-related compounds (Table S4) and resulting in that all the four genotypes with a lactone-rich ideotype were no-melting peaches.

The co-localization of QTL that control HD and MnM (and also firmness) with those affecting volatile production could be due to pleiotropic effects or independent linked loci. In the case of the later scenario, increasing the number of individuals of the population mapping could improve the resolution of QTL localization and probably unlink some of the QTL in this region and clarify if these fruit traits and volatile levels could be improved independently. Most of the market peaches for fresh consumption are melting type with the exception of those from countries as Spain, Italy and Mexico where no-melting peaches are preferred (Okie et al., 2008). Our data (Table S5) point that, if the ideotype pursued involves early no-melting peach, a high number of hybrids should be developed in order to generate enough variability for cultivar selection.

Although, the most likely explanation for the cluster of QTL identified at the bottom of LG4 is two loci with pleiotropic effect, is also interesting to note that a delta 9 fatty acid desaturase (ppa009359m) which we identified as a putative candidate gene for being inversely correlated to Hexanal (Sánchez et al., 2013) co-localized with its QTL (Fig. S5A). In the same way, the QTL controlling 2-Hexen-1-ol acetate (E)- falls in the same region of Cytochrome P450 homologs (ppa006310m) which we identified for being highly correlated to this compound (Sánchez et al., 2013).

We identified three QTL that control the production of several volatiles but do not affect the other fruit traits analyzed. A locus controlling the synthesis of 12 volatiles from C5, formed mostly by monoterpenes, was identified at the top of LG4 (Fig. 4). Previously, Eduardo et al. (2013) mapped in the same region a major QTL for the monoterpenes: linalool and p-menth-1-en-9-al. By analyzing the allelic variation, they also showed that two terpene synthases co-segregate with the QTL. In the current report, we studied both compounds but only a stable QTL for p-menth-1-en-9-al (we reported this compound with the synonym: 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl) was detected (Figure 4).

Regarding accumulation of linalool the correlation between locations was significant, but not high(r=0.39, Table 2), indicating environmental factors also affect the variability of this volatile and probably causing that a significant QTL was only detected in EJ location (Table S1). In fact, all the compound of cluster C5 showed high likelihood (LOD > 3) for association with markers at the top of LG4 in both locations (Fig. S3) but after permutation test only members of C5b (with the exception of 33) were significant in both locations (Table S1). Concomitantly, compounds of C5a showed weak correlation between locations (r=0.31 to r=0.39, Table 2), whereas QTL for C5b where detected in both locations and also those traits showed higher correlation among locations (r=0.66 to r=0.86, Table 2). In addition, the group of monoterpene-rich ideotype showed high levels of all the compounds of C5 compared to the rest of the genotypes (Table S7). Therefore, it is possible that this locus control the whole monoterpene module but in our experiment only detected stable QTL for some of them, a plausible explanation of that is a sampling effects due to the limited experiment size. In summary, our data confirmed the presence of QTL for p-menth-1-en-9-al in the upper end of LG4 but also showed that this locus controls other members of the monoterpenes family in peach. This locus explained between 10-40% of the volatile variance and, the volatile content could be increased between 2 to 11-Fold (a=1.0-3.5) by selecting this locus (Table S3). By analyzing the homology to 90 biochemical characterized monoterpene synthases genes described in (Degenhardt et al., 2009) we found a monoterpene synthase (ppa003423m), additional to the two terpenoid synthase reported by Eduardo et al. (2013) in the LG 4 QTL genome region (data no shown). Further research would be necessary to assess whether these three structural genes could account for the variation of the 12 compounds controlled by this locus (and likely all the monoterpenes) or there are others regulatory genes (e.g. a transcription factor) that control the whole biochemical pathway. In any case,

our data support the exploitation of this locus to modify the concentration of monoterpenes in fruit and also encourage further functional studies on the candidate genes located in this locus.

The volatiles γ-Hexalactone and γ-Octalactone have a coconut-like odor while the esters 2-Hexen-1-ol, acetate, (E)- Ethyl Acetate confer a "fruity" note to fruit aroma (Derail et al., 1999; Guillot et al., 2006). QTL controlling these four aroma-related volatiles were discovered at the same locus in the bottom of LG6 (Fig. 4). The QTL explain between 14% and 31% of the volatile variance and has additive effects of the same sign (Table S1) indicating that the levels of these compounds could be improved (between 1.7 to 3.5 fold according the additive effect) in conjunction. This source of volatile variability was not indentified previously and resulted appealing for aroma improvement. Several genes previously associated with different volatiles by a combined genomic approach (Sánchez et al., 2013) are localized in this region (Fig. S5). Among them one protein kinase (ppa008251m) and two genes with unknown function (ppa004582m and ppa003086m) highly correlated to lactones (Fig. S5B). A Pyruvate decarboxylase (ppa003086m) associated with ester 2-Hexen-1-ol acetate (E)- that we proposed that is regulated at the expression level to ensure the supply of acetyl-CoA for ester biosynthesis (Sánchez et al., 2013), co-localized with a stable QTL for this ester that explains 14% of the variance in mean and has an additive effect that suggested a potential for increasing this volatile of around 3-fold (Table S1, Fig. S5). In addition, a gene with no-homolog in *Arabidopsis* (ppa002860m) that was associated to the levels of Ethyl acetate (Sánchez et al., 2013) is also co-localized in this locus (Fig. S5).

Similarly, QTL with the same additive effect for a lactone (4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one) an ester (3-Hexen-1-ol, acetate (Z)-) and a lipid derived compound (Hexanal) were identified at the top of the LG5 (Fig. 4). In the case of the ester and Hexanal, the QTL detected in EJ and AA locations

partially overlap and spanned a region of near 25 cM, so is unclear if this three QTL are controlled by the same locus or by linked loci. Since the levels of volatiles in the group of lipid-derived compounds are inverse correlated with lactones and linear esters (Fig. 3), we would expect opposite effects if the same locus control their production. Therefore is likely that these two QTL are controlled by independent linked loci. Accordingly with this scenario, the genome position of a protein kinase (ppa006108m) associated a lactones and ester (Sánchez et al., 2013) overlap with the position of the QTL detected in EJ for the lactone and the lineal ester detected at AA. Fatty acids are stored as triacylglycerides, and acyl hydrolases should free them to enter the LOX/HPL pathway in order to produce the so called lipid-derived compounds (Schwab et al., 2008). As we identified an esterase homolog (ppa006430m) we proposed that this gene could be involved in supply of fatty acids to the LOX/HPL pathway (Sánchez et al., 2013). Here this gene co-localized with the QTL controlling the levels of Hexanal in LG5 (Fig S5).

In any case, the co-localization of QTL with the position of the candidate genes previously identified by a genomic approach proves in any way a cause-effect relationship. QTL positions estimated by a low-resolution map span over several hundreds and even thousands of genes in addition to those candidate (not to mention other regulatory elements like microRNAs that could explained the phenotypic variance). Moreover, several of the candidate genes indentified previously for being associated to a given volatile, here failed to co-localize with the QTL controlling this compounds. In addition, evidence for allelic variation within the genes involved must be first presented to become a true candidate. Our results provide additional genetic evidence for linking genes to traits that could be used as starting point for these studies.

Probably as a result of the high level of homozygosity revealed by the SNP genotyping, the genetic map of 'Granada' had a low coverage (e.g. for

chromosome 1, 2, 3, 4, 5 and 8), and consequently a small number of QTL detected (Fig. 5, Table S2, Table S3). Only two QTL stable among locations, one for a monoterpene (43) and the other for fruit weight, were identified in LG2 and LG6, respectively (Fig. 5). Our previous genomic approach indentified four candidate genes in LG2 (Sánchez et al., 2013) but only the one that was correlated to terpenoid leveles (ppa011444m) co-localized with the QTL (Fig. S5). A minor QTL for peach weight had been identified in another locus in LG6 (Eduardo et al., 2011), indicating that the one found here represent a novel source of variability. The QTL for fruit weigh identified here also has minor effect (r=0.15 in mean) and the additive effect is 22 g, but since its localization do not overlap with QTL for volatiles, it should be possible to use it to increase fruit size to some extend without modifying the aroma profile of the fruit.

In conclusion, our data confirmed loci identified previously but also discovered novel ones for important aroma-related volatiles of peach. Furthermore our results confirms the modularity of the genetic control of volatile production in peach, suggesting that groups of related volatiles rather than single volatiles could be the target of aroma improvement . The source of variability described here could be used in the improvement of quality of peach and also could aid in the discovery of genes controlling the aroma of peach fruit.

## REFERENCES


Abbott, A.G., Arús, P. and Scorza, R. (2008) Chapter 4: Genetic Engineering and Genomics. *The Peach: Botany, Production and Uses*, 85.

Aubert, C., Gunata, Z., Ambid, C. and Baumes, R. (2003) Changes in Physicochemical Characteristics and Volatile Constituents of Yellow- and White-Fleshed Nectarines during Maturation and Artificial Ripening. *J. Agric. and Food Chem.*, **51**, 3083-3091.

Bassols Raseira, M.C. and Bonifacio, H. (2006) Peach breeding program in souther brazil. *Acta Hort.*, **713**, 93-98.

Bicchi, C. and Maffei, M. (2012) The plant volatilome: methods of analysis. In *High-Throughput Phenotyping in Plants*. Springer, pp. 289-310.

Brandi, F., Bar, E., Mourgues, F., Horvath, G., Turcsi, E., Giuliano, G., Liverani, A., Tartarini, S., Lewinsohn, E. and Rosati, C. (2011) Study of 'Redhaven' peach and its white-fleshed mutant suggests a key role of CCD4 carotenoid dioxygenase in carotenoid and norisoprenoid volatile metabolism. *BMC Plant Biology*, **11**, 24.

Bruhn, C.M. (1995) Consumer and retailer satisfaction with the quality and size of california peaches and nectarines. *Journal of Food Quality*, **18**, 241-256.

Bruhn, C.M., Feldman, N., Garlitz, C., Harwood, J., Ivans, E., Marshall, M., Riley, A., Thurber, D. and Williamson, E. (1991) Consumer Perceptions of quality: Apricots, Cantaloupes, Peaches, Pears, Strawberries, and Tomatoes. *Journal of Food Quality*, **14**, 187-195.

Crisosto, C.H. (2006) Short-term approaches to increase peach fruit consumption. *Compact Fruit Tree*, **39**, 11-14.

Chapman, G.W., Horvat, R.J. and Forbus, W.R. (1991) Physical and chemical changes during the maturation of peaches (cv. Majestic). *Journal of Agricultural and Food Chemistry*, **39**, 867-870.

Degenhardt, J., Köllner, T.G. and Gershenzon, J. (2009) Monoterpene and sesquiterpene synthases and the origin of terpene skeletal diversity in plants. *Phytochemistry*, **70**, 1621-1637.

Derail, C., Hofmann, T. and Schieberle, P. (1999) Differences in Key Odorants of Handmade Juice of Yellow-Flesh Peaches (*Prunus persica* L.) Induced by the Workup Procedure. *J. Agric. and Food Chem.*, **47**, 4742-4745.

Dirlewanger, E., Graziano, E., Joobeur, T., Garriga-Calderé, F., Cosson, P., Howad, W. and Arús, P. (2004) Comparative mapping and marker-assisted selection in Rosaceae fruit crops. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 9891-9896.

Dirlewanger, E., Quero-Garcia, J., Le Dantec, L., Lambert, P., Ruiz, D., Dondini, L., Illa, E., Quilot-Turion, B., Audergon, J.M., Tartarini, S., Letourmy, P. and Arus, P. (2012) Comparison of the genetic determinism of two key phenological traits, flowering and maturity dates, in three Prunus species: peach, apricot and sweet cherry. *Heredity*, **109**, 280-292.

Doyle, J.J. (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem bull*, **19**, 11-15.

Eduardo, I., Chietera, G., Bassi, D., Rossini, L. and Vecchietti, A. (2010) Identification of key odor volatile compounds in the essential oil of nine peach accessions. *Journal of the Science of Food and Agriculture*, **90**, 1146-1154.

Eduardo, I., Chietera, G., Pirona, R., Pacheco, I., Troggio, M., Banchi, E., Bassi, D., Rossini, L., Vecchietti, A. and Pozzi, C. (2013) Genetic dissection of aroma volatile compounds from the essential oil of peach fruit: QTL analysis and identification of candidate genes using dense SNP maps. *Tree Genetics & Genomes*, **9**, 189-204.

Eduardo, I., Pacheco, I., Chietera, G., Bassi, D., Pozzi, C., Vecchietti, A. and Rossini, L. (2011) QTL analysis of fruit quality traits in two peach intraspecific populations and importance of maturity date pleiotropic effect. *Tree Genetics & Genomes*, **7**, 323-335.

Elshire, R.J., Glaubitz, J.C., Sun, Q., Poland, J.A., Kawamoto, K., Buckler, E.S. and Mitchell, S.E. (2011) A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS one*, **6**, e19379.

Grattapaglia, D. and Sederoff, R. (1994) Genetic linkage maps of Eucalyptus grandis and Eucalyptus urophylla using a pseudo-testcross: mapping strategy and RAPD markers. *Genetics*, **137**, 1121-1137.

Greger, V. and Schieberle, P. (2007) Characterization of the Key Aroma Compounds in Apricots (Prunus armeniaca) by Application of the Molecular Sensory Science Concept. *Journal of Agricultural and Food Chemistry*, **55**, 5221-5228.

Guillot, S., Peytavi, L., Bureau, S., Boulanger, R., Lepoutre, J.-P., Crouzet, J. and Schorr-Galindo, S. (2006) Aroma characterization of various apricot varieties using headspace-solid phase microextraction combined with gas chromatography-mass spectrometry and gas chromatography-olfactometry. *Food Chemistry*, **96**, 147-155.

Horvat, R.J., Chapman, G.W., Robertson, J.A., Meredith, F.I., Scorza, R., Callahan, A.M. and Morgens, P. (1990) Comparison of the volatile compounds from several commercial peach cultivars. *Journal of Agricultural and Food Chemistry*, **38**, 234-237.

Jia, H.-J., Araki, A. and Okamoto, G. (2005) Influence of fruit bagging on aroma volatiles and skin coloration of `Hakuho´ peach (*Prunus persica* Batsch). *Postharvest Biology and Technology*, **35**, 61-68.

Klee, H.J. and Giovannoni, J.J. (2011) Genetics and Control of Tomato Fruit Ripening and Quality Attributes. *Annual Review of Genetics*, **45**, 41-59.

Koutsos, T.V., Chatzopoulou, P.S. and Katsiotis, S.T. (2009) Effects of individual selection on agronomical and morphological traits and essential oil of a "Greek basil" population. *Euphytica*, **170**, 365-370.

Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M.J., Lincoln, S.E. and Newburg, L. (1987) MAPMAKER: An interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. *Genomics*, **1**, 174-181.

Martínez-García, P., Parfitt, D., Ogundiwin, E., Fass, J., Chan, H., Ahmad, R., Lurie, S., Dandekar, A., Gradziel, T. and Crisosto, C. (2013) High density SNP mapping and QTL analysis for fruit quality characteristics in peach (Prunus persica L.). *Tree Genetics & Genomes*, **9**, 19-36.

Okie, W.R., Bacon, T. and Bassi, D. (2008) 6 Fresh Market Cultivar Development. *The Peach: Botany, Production and Uses*, 139.

Olbricht, K., Grafe, C., Weiss, K. and Ulrich, D. (2008) Inheritance of aroma compounds in a model population of Fragaria × ananassa Duch. *Plant Breeding*, **127**, 87-93.

Parker, D.D., Zilberman, D. and Moulton, K. (1991) How quality relates to price in California fresh peaches. *Calif. Agr.*, **45**, 14-16.

Pirona, R., Vecchietti, A., Lazzari, B., Caprera, A., Malinverni, R., Consolandi, C., Severgnini, M., De Bellis, G., Chietera, G. and Rossini, L. (2012) Expression profiling of genes involved in the formation of aroma in two peach genotypes. *Plant Biology*, **15**, 443-451.

Quilot, B., Wu, B.H., Kervella, J., Genard, M., Foulongne, M. and Moreau, K. (2004) QTL analysis of quality traits in an advanced backcross between Prunus persica cultivars and the wild relative species P. davidiana. *Theoretical and Applied Genetics*, **109**, 884-897.

Robertson, J.A., Meredith, F.I., Horvat, R.J. and Senter, S.D. (1990) Effect of cold storage and maturity on the physical and chemical characteristics and volatile constituents of peaches (cv. Cresthaven). *Journal of Agricultural and Food Chemistry*, **38**, 620-624.

Sánchez, G., Besada, C., Badenes, M.L., Monforte, A.J. and Granell, A. (2012) A Non-Targeted Approach Unravels the Volatile Network in Peach Fruit. *PLoS ONE*, **7**, e38992.

Sánchez, G., Venegas-Caleron, M., Salas, J., Monforte, A., Badenes, M. and Granell, A. (2013) An integrative "omics" approach identifies new candidate genes to impact aroma volatiles in peach fruit. *BMC Genomics*, **14**, 343.

Schwab, W., Davidovich-Rikanati, R. and Lewinsohn, E. (2008) Biosynthesis of plant-derived flavor compounds. *The Plant Journal*, **54**, 712-732.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Research*, **13**, 2498-2504.

Stam, P. (1993) Construction of integrated genetic linkage maps by means of a new computer package: Join Map. *The Plant Journal*, **3**, 739-744.

Sumitani, H., Suekane, S., Nakatani, A. and Tatsuka, K. (1994) Changes In Composition of Volatile Compounds in High Pressure Treated Peach. *Journal of Agricultural and Food Chemistry*, **42**, 785-790.

Tikunov, Y., Lommen, A., de Vos, C.H.R., Verhoeven, H.A., Bino, R.J., Hall, R.D. and Bovy, A.G. (2005) A Novel Approach for Nontargeted Data Analysis for Metabolomics. Large-Scale Profiling of Tomato Fruit Volatiles. *Plant Physiol.*, **139**, 1125-1137.

Van Ooijen, J.W. (2006) JoinMap 4. *Software for the calculation of genetic linkage maps in experimental populations. Kyazma BV, Wageningen, Netherlands*.

Vecchietti, a., Lazzari, B., Ortugno, C., Bianchi, F., Malinverni, R., Caprera, a., Mignani, I. and Pozzi, C. (2009) Comparative analysis of expressed sequence tags from tissues in ripening stages of peach (Prunus persica L. Batsch). *Tree Genetics & Genomes*, **5**, 377-391.

Verde, I., Abbott, A.G., Scalabrin, S., Jung, S., Shu, S., Marroni, F., Zhebentyayeva, T., Dettori, M.T., Grimwood, J., Cattonaro, F., Zuccolo, A., Rossini, L., Jenkins, J., Vendramin, E., Meisel, L.A., Decroocq, V., Sosinski, B., Prochnik, S., Mitros, T., Policriti, A., Cipriani, G., Dondini, L., Ficklin, S., Goodstein, D.M., Xuan, P., Fabbro, C.D., Aramini, V., Copetti, D., Gonzalez, S., Horner, D.S., Falchi, R., Lucas, S., Mica, E., Maldonado, J., Lazzari, B., Bielenberg, D., Pirona, R., Miculan, M., Barakat, A., Testolin, R., Stella, A., Tartarini, S., Tonutti, P., Arus, P., Orellana, A., Wells, C., Main, D., Vizzotto, G., Silva, H., Salamini, F., Schmutz, J., Morgante, M. and Rokhsar, D.S. (2013) The high-quality draft genome of peach (Prunus persica) identifies unique patterns of genetic diversity, domestication and genome evolution. *Nat Genet*, **45**, 487-494.

Verde, I., Bassil, N., Scalabrin, S., Gilmore, B., Lawley, C.T., Gasic, K., Micheletti, D., Rosyara, U.R., Cattonaro, F., Vendramin, E., Main, D., Aramini, V., Blas, A.L., Mockler, T.C., Bryant, D.W., Wilhelm, L., Troggio, M., Sosinski, B., Aranzana, M.J., Arús, P., Iezzoni, A., Morgante, M. and Peace, C. (2012) Development and Evaluation of a 9K SNP Array for Peach by Internationally Coordinated SNP Detection and Validation in Breeding Germplasm. *PLoS ONE*, **7**, e35668.

Verde, I., Quarta, R., Cedrola, C. and Dettori, M.T. (2002) QTL analysis of agronomic traits in a BC1 peach population. *Acta Hort.* , **592**, 291-297.

Visai, C. and Vanoli, M. (1997) Volatile compound production during growth and ripening of peaches and nectarines. *Scientia Horticulturae*, **70**, 15-24.

Voorrips, R.E. (2002) MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *Journal of Heredity*, **93**, 77-78.

Wang, S., Basten, C.J. and Zeng, Z.-B. (2012) Windows QTL Cartographer 2.5. In Department of Statistics, N.C.S.U., Raleigh, NC. (ed.), *(http://statgen.ncsu.edu/qtlcart/WQTLCart.htm)*

Xi, W.-P., Zhang, B.O., Liang, L.I., Shen, J.-Y., Wei, W.-W., Xu, C.-J., Allan, A.C., Ferguson, I.B. and Chen, K.-S. (2012) Postharvest temperature influences volatile lactone production via regulation of acyl-CoA oxidases in peach fruit. *Plant, Cell & Environment*, **35**, 534-545.

Yang, J., Hu, C., Hu, H., Yu, R., Xia, Z., Ye, X. and Zhu, J. (2008) QTLNetwork: mapping and visualizing genetic architecture of complex traits in experimental populations. *Bioinformatics*, **24**, 721-723.

Zanor, M.I., Rambla, J.L., Chaib, J., Steppa, A., Medina, A., Granell, A., Fernie, A.R. and Causse, M. (2009) Metabolic characterization of loci affecting sensory attributes in tomato allows an assessment of the influence of the levels of primary metabolites and volatile organic contents. *J Exp Bot*, **60**, 2139-2154.

Zhang, B., Shen, J.-y., Wei, W.-w., Xi, W.-p., Xu, C.-J., Ferguson, I. and Chen, K. (2010) Expression of Genes Associated with Aroma Formation Derived from the Fatty Acid Pathway during Peach Fruit Ripening. *Journal of Agricultural and Food Chemistry*, **58**, 6157-6165.

Zhang, B., Xi, W.-p., Wei, W.-w., Shen, J.-y., Ferguson, I. and Chen, K.-s. (2011) Changes in aroma-related volatiles and gene expression during low temperature storage and subsequent shelf-life of peach fruit. *Postharvest Biology and Technology*, **60**, 7-16.

Zorrilla-Fontanesi, Y., Rambla, J.-L., Cabeza, A., Medina, J.J., Sánchez-Sevilla, J.F., Valpuesta, V., Botella, M.A., Granell, A. and Amaya, I. (2012) Genetic Analysis of Strawberry Fruit Aroma and Identification of O-Methyltransferase FaOMT as the Locus Controlling Natural Variation in Mesifurane Content. *Plant Physiology*, **159**, 851-870.

# Capítulo 5:
# DISCUSIÓN GENERAL

## DISCUSIÓN GENERAL

El objetivo último de este trabajo es la generación tanto de conocimiento como de herramientas para la mejora de la calidad del melocotón. Debido a que desde el punto de vista de los consumidores el aroma y el sabor son los atributos que más afectan a la calidad del melocotón, este trabajo se centró en los compuestos volátiles emitidos por el fruto. Como objetivos concretos se planteó la identificación de genes candidatos y el cartografiado de loci que controlan la producción de volátiles en el fruto. Si bien, se ha demostrado que al menos una veintena de compuestos claves están involucrados en la percepción del aroma del melocotón (Eduardo et al., 2010), se decidió abordar este trabajo con una visión holística, debido a que una comprensión global de la producción de volátiles permitiría una mejor interpretación de los resultados. Para llevar a cabo esta estrategia se utilizaron tecnologías de alto rendimiento que permiten la adquisición de un gran volumen de datos de expresión génica y de niveles de volátiles que luego fueron interpretados con una combinación de herramientas de análisis multivariante. En primer lugar se infirió la red reguladora de la síntesis de volátiles en el fruto para posteriormente explotar este conocimiento para la identificación de genes involucrados en el proceso y por último elucidar el control génico subyacente. Dado que en cada capítulo de esta tesis ya se ha presentado una discusión detallada de los resultados obtenidos, en este apartado se realizará una discusión integradora de los tres capítulos intentando no repetir argumentos ya expuestos.

Desde que el termino "metabolómica" fue acuñado a fines del siglo pasado (Oliver et al., 1998), los estudios basados en esta disciplina han progresado paulatinamente desde ser meros análisis descriptivos hasta la actualidad donde

presentan estudios cada vez más informativos. En la presente tesis, a pesar de utilizar tecnologías de alto rendimiento en combinación con aproximaciones del tipo no-dirigidas, que son capaces de generar gran cantidad de datos, se intentó no perder de vista el fin último que es generar conocimiento y no datos.

Otro cambio que ha experimentado la utilización de las tecnologías de alta capacidad está relacionado con el abaratamiento de los costes de los análisis. De esta forma, en un principio los altos costes permitían evaluar solo un número reducido de muestras y por tanto muchos estudios se centraron en comparar el efecto de un tratamiento dado sobre el metaboloma de la planta (siendo el ejemplo típico la comparación entre plantas transgénicas vs no-transgénicas). Con la evolución de las tecnologías los costes se han ido reduciendo haciendo factible el análisis de una gran cantidad de individuos/muestras permitiendo el llevar a acabo estudios a gran escala como es el caso de su utilización en poblaciones de mapeo para el análisis de QTL. Por lo tanto, mientras que inicialmente el factor limitante resultaba ser el número de muestras a analizar (o mejor dicho el coste asociado a ello), en la actualidad la tendencia parece revertirse siendo limitante el número de genotipos/individuos disponibles para ser estudiados. En algunos casos como lo es el del melocotón, este factor limitante puede tener aún más peso. Esto se debe a que el coste asociado al mantenimiento de una población de melocotoneros (durante un cierto período de tiempo) en los campos experimentales puede resultar incluso mayor que los costes asociados al análisis. En estos casos el material vegetal resulta el factor limitante para realizar los estudios deseados, cobrando especial importancia su generación.

Una de las ventajas que proporciona el análisis global del metaboloma es el permitir describir como se relacionan estos compuestos entre sí. El análisis de la correlación metabolito-metabolito en un conjunto complejo de muestras reveló que el volatiloma del melocotón se encuentra altamente estructurado en

módulos formados por compuestos con identidad química similar o bien que pertenecen a rutas metabólicas conocidas (capítulo 2). Los resultados obtenidos en esta tesis están en concordancia con el concepto de modularidad jerárquica que establece que para optimizar la funcionalidad celular de los organismos, el metabolismo está organizado en un conjunto de módulos funcionales que a su vez se organizan en grupos mayores (Jeong et al., 2000; Ravasz et al., 2002). Esta organización en módulos jerárquicos de la red metabólica permitiría un funcionamiento energéticamente eficiente, tolerante a errores y robusto. La alta estructuración del volatiloma también ha sido descrita en tomate, melón, fresa y en especies de *Citrus* (Gonzalez-Mas et al., 2011; Moing et al., 2011; Tikunov et al., 2005; Zorrilla-Fontanesi et al., 2012). Estos estudios utilizaron estrategias diferentes a la presentada aquí para generar la variabilidad fenotípica en el que basan el análisis de correlaciones. Es posible que la selección del conjunto de muestras de una especie dada influya sobre la red metabolómica obtenida. Los resultados presentados en esta tesis permiten analizar esta cuestión ya que las diferentes redes de correlación fueron calculadas a partir de muestras cuya variabilidad fenotípica a nivel de volátiles tiene diferentes orígenes (capítulos 2-4). En el caso de la red del capítulo 2 se utilizó una combinación de diferentes fuentes de variabilidad conformada por: cuatro genotipos, cuatro estados de maduración en el árbol y uno de poscosecha, además de dos ambientes diferentes. En cambio en el capítulo 3, se analizó un sub-grupo de estas muestras que corresponde a series temporales de maduración de dos genotipos. Y finalmente en la red del capítulo 4 la variabilidad utilizada fue la correspondiente a la combinación genotipo/ambiente ya que se estudió el perfil de volátiles de la población de mejora analizada en dos ubicaciones. Las redes de correlación obtenidas mostraron una consistencia muy elevada, observándose en todas ellas los siguientes módulos: 1) lactonas y algunos esteres, 2) los monoterpenos en dos grupos altamente correlacionados entre si, 3) los compuestos derivados de lípidos, 4) los ácidos carboxílicos y 5) el etanol

junto a su ester, etil acetato. Si bien previamente no se reportó ningún análisis a gran escala del volatiloma del melocotón, algunos estudio ya habían señalado que las lactonas se encuentran correlacionadas entre si (Eduardo et al., 2013; Xi et al., 2012; Zhang et al., 2010), apoyando la hipótesis de que este grupo de compuestos está sujeto a una regulación robusta. En términos generales, los compuestos volátiles producidos y liberados por las plantas al medio que las circunda tienen como finalidad la comunicación con otras plantas (u otras partes de las misma planta que los emite), o repeler organismos antagonistas o atraer organismo mutualistas (Baldwin et al., 2006; Bednarek et al., 2009; Degenhardt et al., 2009; Kessler and Baldwin, 2001; Kessler et al., 2008; Vickers et al., 2009). Resulta lógico pensar que en el caso de los volátiles que contribuyen al aroma del fruto con aromas placenteros (e.j., las lactonas y esteres) se regulen conjuntamente para atraer a los frugívoros de forma más eficiente. De forma similar tiene sentido que lo hagan aquellos con descripciones no placenteras (e.j.: derivados de lípidos) y a su vez que estos grupos se encuentren coordinados entre si durante el desarrollo del fruto. De este modo se aseguraría que la atracción de frugívoros ocurra en el momento adecuado, es decir cuando el fruto se encuentra maduro y la semilla haya completado su programa de desarrollo, con lo que se podría lograr una exitosa dispersión de las semillas. Los resultados presentados aquí se ajustan a este esquema e indican que los módulos funcionales involucrados en la formación de aromas derivan del *pool* de ácidos grasos. En el capitulo 2 se ha demostrado que el módulo de las lactonas/esteres está inversamente regulado al de los compuestos derivados de lípidos. En concordancia con la implicación de estos compuestos en la formación de aromas se ha mostrado que volátiles perteneciente a estos dos módulos mostraron correlaciones fuertes con las variables que usualmente se utilizan para medir la calidad del fruto y por lo tanto la maduración (firmeza y color principalmente) pero en sentidos opuestos para cada módulo. Estos resultados reafirman el papel central que se ha

propuesto para estos compuestos en definir el aroma del melocotón (Derail et al., 1999; Eduardo et al., 2010; Horvat et al., 1990). Más aún, en el capítulo 3 se ha mostrado que a medida que avanza la maduración del fruto las lactonas y esteres se acumulan mientras que los compuestos derivados de lípidos disminuyen, confirmando resultados anteriores (Visai and Vanoli, 1997). Además, el análisis de los patrones de co-regulación de estos módulos identificó genes relacionados con la señalización del proceso de maduración (capítulo 3). Por último en el capitulo 4, se ha demostrado que un locus situado hacia el final del LG4 controla la síntesis de algunos de los compuestos de estos dos módulos de forma antagónica y que genes en este locus a su vez afectan la fecha de maduración del fruto.

El módulo de los monoterpenos también parece estar sometido a una regulación robusta según se desprende de nuestros resultados (capítulos 2-4) y además su síntesis está sujeta a un fuerte control genético como fue demostrado en el capítulo 4 y simultáneamente, aunque de forma parcial, por otros (Eduardo et al., 2013). La relevancia de este módulo implicado en la producción de monoterpenos en la formación del aroma del melocotón está menos clara debido a que solo uno de sus miembros, el linalool, posee una descripción sensorial placentera ("floral") (Derail et al., 1999; Greger and Schieberle, 2007) y ha sido implicado en el aroma de acuerdo a su valor activo de aroma (Eduardo et al., 2010). La principal función de los volátiles terpenoides emitidos por los tejidos vegetativos de las plantas está relacionada a la resistencia hacia plagas (Mithofer and Boland, 2012). Las formas de acción de los volátiles terpenoides son variadas desde la intoxicación directa, repeler o disuadir insectos (Aharoni et al., 2003; Seybold et al., 2006) así como también atraer a predadores naturales (Kessler and Baldwin, 2001). Incluso, se ha reportado que ciertos monoterpenos tiene un efecto alelopático sobre plantas vecinas (Maffei et al., 2011). Por tanto no puede descartarse que este módulo de compuestos posea funciones

diferentes a la atracción de frugívoros. Por lo tanto el locus que hemos identificado en la parte superior del LG4 y que controla la síntesis de los compuestos de este módulo podría tener utilidades adicionales a la modificación del aroma, las cuales se desvelarán cuando la función de estos compuestos se esclarezca.

La descripción del control genético subyacente a la síntesis de los compuestos volátiles sólo es el comienzo de la mejora del sabor y el aroma (conocido conjuntamente como *flavor*) del melocotón. Aún se requiere de una gran cantidad de información para llevar a cabo una mejora eficiente de este carácter. Por ejemplo, se ha reportado que existen dos grupos de personas con preferencias diferentes según el carácter dulce o ácido del melocotón (Bonany, 2010). De igual forma puede ocurrir con las diferentes notas gustativas y el aroma del melocotón y por tanto debería conocerse previamente las preferencias del mercado a satisfacer e incorporar esta información a los objetivos de mejora. Por otro lado, el *flavor* del fruto puede depender en gran medida de su apariencia y textura. Se ha propuesto que un tomate con una coloración más rojiza al igual que una manzana con un textura más crujiente serían percibidos como más sabrosos por parte de los consumidores (Baldwin et al., 2000). Es probable que en el caso del melocotón ocurran alteraciones similares de la percepción subjetiva del *flavor*, las cuales una vez descriptas deberían tenerse en cuenta a la hora de definir los ideotipos buscados. A pesar de todas las dificultades mencionadas, en el caso del melocotón, el mejoramiento del *flavor* parece prometedor debido a que se han identificado dos compuestos involucrados en el aroma, la γ-decalactona y la γ -jasmolactona (Eduardo et al., 2010), cuyo descripción sensorial es "aroma típico de melocotón" (Derail et al., 1999). Este hecho es poco común ya que no ocurre frecuentemente en otras especies de interés agronómico como por ejemplo el tomate donde ningunos de los más de 400 volátiles identificados tienen una

descripción sensorial que se lo pueda vincular a su aroma o sabor del fruto. En este sentido, parecería lógico intentar incrementar los niveles de estas lactonas. Los resultados obtenidos en esta tesis sugieren que el aumento de estos dos compuestos traería acarreado el aumento de las lactonas en conjunto y algunos esteres. Debido a que todos estos volátiles poseen descripciones sensoriales "placenteras", es posible especular que tal aumento en conjunto resultaría beneficioso a la hora de mejorar el *flavor* general. No obstante, en algunos casos podría resultar ventajoso modificar el contenido de volátiles individuales para lograr frutos con notas aromáticas diferenciales. El esclarecimiento de las rutas metabólicas de las lactonas en melocotón permitirá plantear hipótesis acerca de esta posibilidad.

## REFERENCIAS


Aharoni, A., Giri, A.P., Deuerlein, S., Griepink, F., de Kogel, W.-J., Verstappen, F.W.A., Verhoeven, H.A., Jongsma, M.A., Schwab, W. and Bouwmeester, H.J. (2003) Terpenoid metabolism in wild-type and transgenic Arabidopsis plants. *The Plant Cell Online*, **15**, 2866-2884.

Baldwin, E.A., Scott, J.W., Shewmaker, C.K. and Schuch, W. (2000) Flavor trivia and tomato aroma: biochemistry and possible mechanisms for control of important aroma components. *HortScience*, **35**, 1013-1022.

Baldwin, I.T., Halitschke, R., Paschold, A., Von Dahl, C.C. and Preston, C.A. (2006) Volatile signaling in plant-plant interactions:" talking trees" in the genomics era. *Science*, **311**, 812-815.

Bednarek, P., Pislewska-Bednarek, M., Svatos, A., Schneider, B., Doubsky, J., Mansurova, M., Humphry, M., Consonni, C., Panstruga, R. and Sanchez-Vallet, A. (2009) A glucosinolate metabolism pathway in living plant cells mediates broad-spectrum antifungal defense. *Science*, **323**, 101-106.

Bonany, J. (2010) Choosing the right apple and peach variety for each consumer. In Brunton, R.G.N. (ed.), *Fruit, Health and the Consumer: Outcomes from ISAFRUIT Fruit Forum Conference*. Teagasc Food Research Ashtown, Dublin, Ireland.

Degenhardt, J., Hiltpold, I., Köner, T.G., Frey, M., Gierl, A., Gershenzon, J., Hibbard, B.E., Ellersieck, M.R. and Turlings, T.C.J. (2009) Restoring a maize root signal that attracts insect-killing nematodes to control a major pest. *Proceedings of the National Academy of Sciences*, **106**, 13213-13218.

Derail, C., Hofmann, T. and Schieberle, P. (1999) Differences in Key Odorants of Handmade Juice of Yellow-Flesh Peaches (*Prunus persica* L.) Induced by the Workup Procedure. *J. Agric. and Food Chem.*, **47**, 4742-4745.

Eduardo, I., Chietera, G., Bassi, D., Rossini, L. and Vecchietti, A. (2010) Identification of key odor volatile compounds in the essential oil of nine peach accessions. *Journal of the Science of Food and Agriculture*, **90**, 1146-1154.

Eduardo, I., Chietera, G., Pirona, R., Pacheco, I., Troggio, M., Banchi, E., Bassi, D., Rossini, L., Vecchietti, A. and Pozzi, C. (2013) Genetic dissection of aroma volatile compounds from the essential oil of peach fruit: QTL analysis and identification of candidate genes using dense SNP maps. *Tree Genetics & Genomes*, **9**, 189-204.

Gonzalez-Mas, M.C., Rambla, J.L., Alamar, M.C., Gutierrez, A. and Granell, A. (2011) Comparative Analysis of the Volatile Fraction of Fruit Juice from Different *Citrus* Species. *PLoS ONE*, **6**, e22016.

Greger, V. and Schieberle, P. (2007) Characterization of the Key Aroma Compounds in Apricots (Prunus armeniaca) by Application of the Molecular Sensory Science Concept. *Journal of Agricultural and Food Chemistry*, **55**, 5221-5228.

Horvat, R.J., Chapman, G.W., Robertson, J.A., Meredith, F.I., Scorza, R., Callahan, A.M. and Morgens, P. (1990) Comparison of the volatile compounds from several commercial peach cultivars. *Journal of Agricultural and Food Chemistry*, **38**, 234-237.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N. and Barabási, A.L. (2000) The large-scale organization of metabolic networks. *Nature*, **407**, 651-654.

Kessler, A. and Baldwin, I.T. (2001) Defensive function of herbivore-induced plant volatile emissions in nature. *Science*, **291**, 2141-2144.

Kessler, D., Gase, K. and Baldwin, I.T. (2008) Field experiments with transformed plants reveal the sense of floral scents. *Science*, **321**, 1200-1202.

Maffei, M.E., Gertsch, J.r. and Appendino, G. (2011) Plant volatiles: production, function and pharmacology. *Natural Product Reports*, **28**, 1359-1380.

Mithofer, A. and Boland, W. (2012) Plant Defense Against Herbivores: Chemical Aspects. *Annual Review of Plant Biology*, **63**, 431-450.

Moing, A., Aharoni, A., Biais, B., Rogachev, I., Meir, S., Brodsky, L., Allwood, J.W., Erban, A., Dunn, W.B., Kay, L., de Koning, S., de Vos, R.C.H., Jonker, H., Mumm, R., Deborde, C., Maucourt, M., Bernillon, S., Gibon, Y., Hansen, T.H., Husted, S., Goodacre, R., Kopka, J., Schjoerring, J.K., Rolin, D. and Hall, R.D. (2011) Extensive metabolic cross-talk in melon fruit revealed by spatial and developmental combinatorial metabolomics. *New Phytologist*, **190**, 683-696.

Oliver, S.G., Winson, M.K., Kell, D.B. and Baganz, F. (1998) Systematic functional analysis of the yeast genome. *Trends in biotechnology*, **16**, 373-378.

Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N. and Barabási, A.L. (2002) Hierarchical organization of modularity in metabolic networks. *Science*, **297**, 1551-1555.

Seybold, S.J., Huber, D.P.W., Lee, J.C., Graves, A.D. and Bohlmann, J.r. (2006) Pine monoterpenes and pine bark beetles: a marriage of convenience for defense and chemical communication. *Phytochemistry Reviews*, **5**, 143-178.

Tikunov, Y., Lommen, A., de Vos, C.H.R., Verhoeven, H.A., Bino, R.J., Hall, R.D. and Bovy, A.G. (2005) A Novel Approach for Nontargeted Data Analysis for Metabolomics. Large-Scale Profiling of Tomato Fruit Volatiles. *Plant Physiol.*, **139**, 1125-1137.

Vickers, C.E., Gershenzon, J., Lerdau, M.T. and Loreto, F. (2009) A unified mechanism of action for volatile isoprenoids in plant abiotic stress. *Nature Chemical Biology*, **5**, 283-291.

Visai, C. and Vanoli, M. (1997) Volatile compound production during growth and ripening of peaches and nectarines. *Scientia Horticulturae*, **70**, 15-24.

Xi, W.-P., Zhang, B.O., Liang, L.I., Shen, J.-Y., Wei, W.-W., Xu, C.-J., Allan, A.C., Ferguson, I.B. and Chen, K.-S. (2012) Postharvest temperature influences volatile lactone production via regulation of acyl-CoA oxidases in peach fruit. *Plant, Cell & Environment*, **35**, 534-545.

Zhang, B., Shen, J.-y., Wei, W.-w., Xi, W.-p., Xu, C.-J., Ferguson, I. and Chen, K. (2010) Expression of Genes Associated with Aroma Formation Derived from the Fatty Acid Pathway during Peach Fruit Ripening. *Journal of Agricultural and Food Chemistry*, **58**, 6157-6165.

Zorrilla-Fontanesi, Y., Rambla, J.-L., Cabeza, A., Medina, J.J., Sánchez-Sevilla, J.F., Valpuesta, V., Botella, M.A., Granell, A. and Amaya, I. (2012) Genetic Analysis of Strawberry Fruit Aroma and Identification of O-Methyltransferase FaOMT as the Locus Controlling Natural Variation in Mesifurane Content. *Plant Physiology*, **159**, 851-870.

# CONCLUSIONES

## CONCLUSIONES

1. Se ha puesto a punto una plataforma de alto rendimiento para el perfilado de compuestos volátiles de melocotón. La plataforma desarrollada es capaz de detectar, identificar y cuantificar relativamente una parte significativa de la fracción volátil del metaboloma del fruto incluidos una serie de metabolitos no descritos previamente en melocotón y que podrían contribuir al aroma del fruto.

2. Mediante una combinación de estrategias de análisis de datos basada en correlaciones (Análisis de cluster jerárquico y Análisis de redes de correlación) se ha demostrado que la producción de volátiles en melocotón está altamente organizada y ello ha permitido identificar grupos de compuestos que comparten una regulación común.

3. El análisis en conjunto de los patrones de co-regulación entre la síntesis de volátiles y la expresión de genes permitió la identificación de genes candidatos para controlar la acumulación de los principales aromas del melocotón

4. Uno de los genes candidatos identificado fue clonado y mediante la expresión en un sistema heterólogo se demostró que posee una actividad ω-6 desaturasa de oleico que es compatible con la formación de un posible precursor de lactonas y esteres del melocotón.

5. Se describió el control genético de la producción de volátiles en el fruto. Se identificaron loci que afectan los niveles de los monoterpenos, lactonas y esteres entre otros.

6. Los resultados presentados indican que es posible mejorar la calidad del melocotón mediante la selección de materiales que lleven alelos adecuados para aumentar el contenido de compuestos asociados al aroma y sabor del fruto.

# ANEXOS

**Información suplementaria de Sánchez et al., 2012 (Capitulo 2)***

*La información presentada en esta sección puede ser encontrada en:
http://dx.plos.org/10.1371/journal.pone.0038992

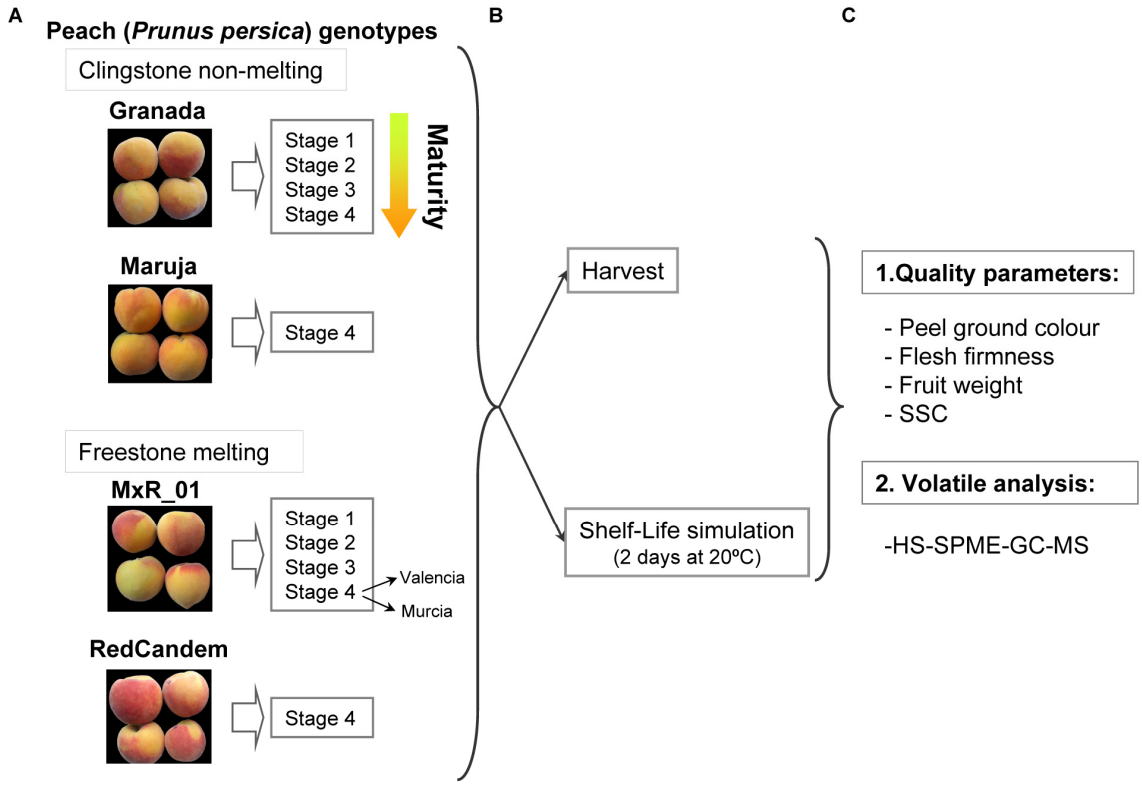



**Figure S1: Experimental design and analysis. A)** Genotypes and maturity stages analyzed (S1-S4). For each genotype, four representative fruits at commercial maturity stage (S4) are shown. Valencia and Murcia indicate that fruits at S4 of MxR_01 genotype were analyzed in two locations. **B)** The post-treatment applied. **C)** Fruit and volatile organic compound analysis. HS-SPME-GC-MS: Head Space-Solid Phase Microextraction-Gas Chromatography-Mass Spectroscopy.

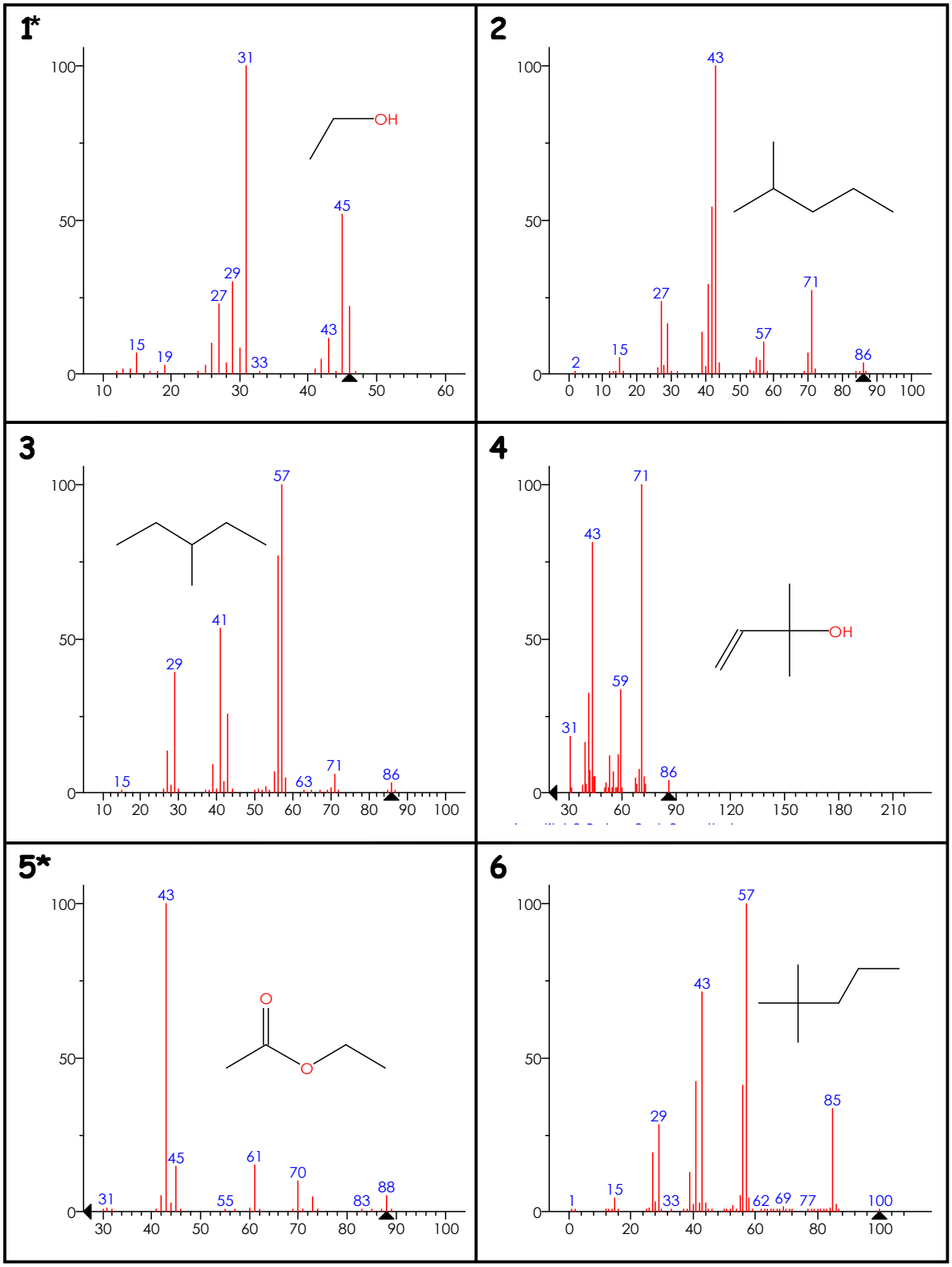
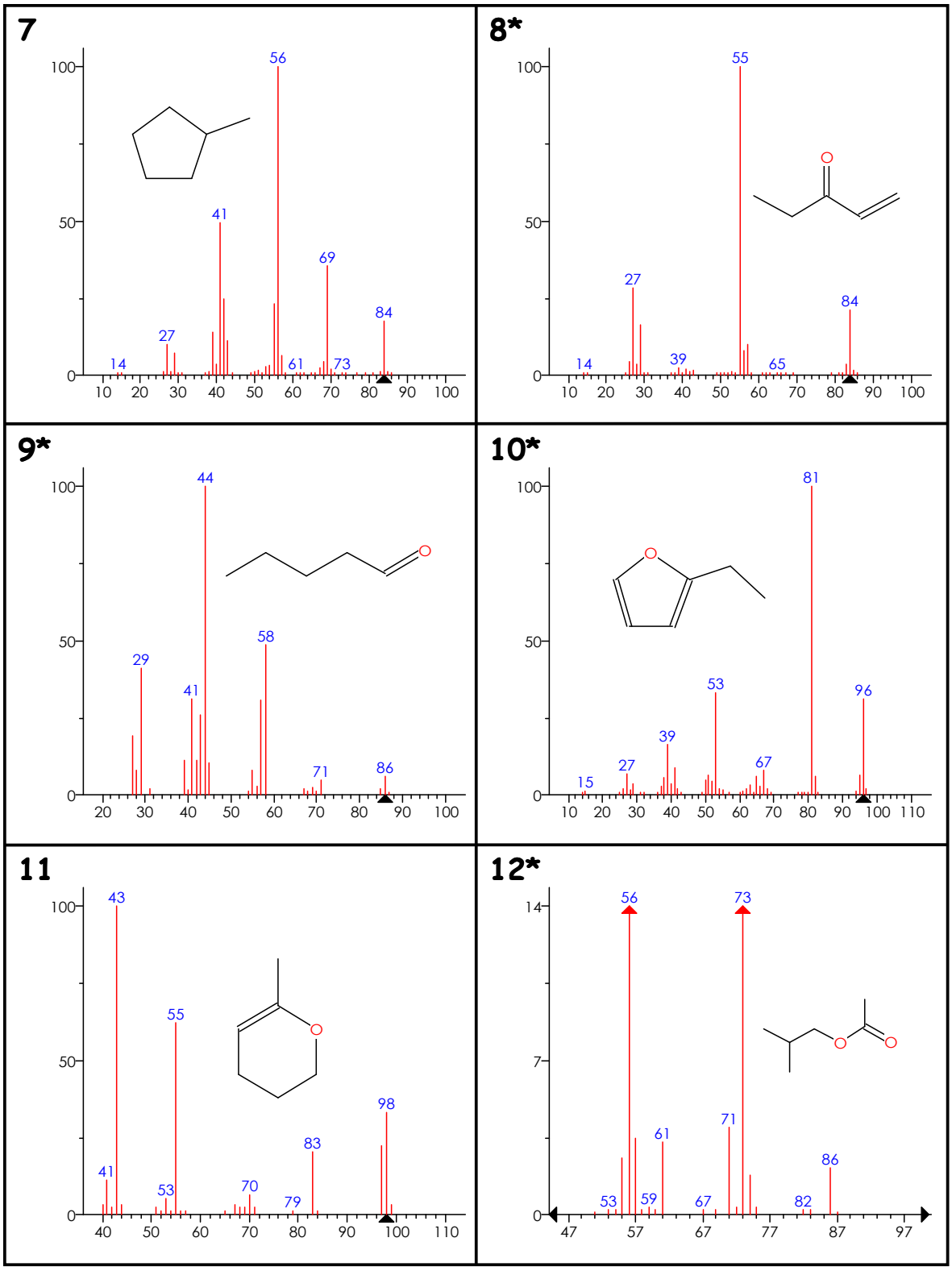



**Fig. S2 (see the legends below).**




**S2-continued**

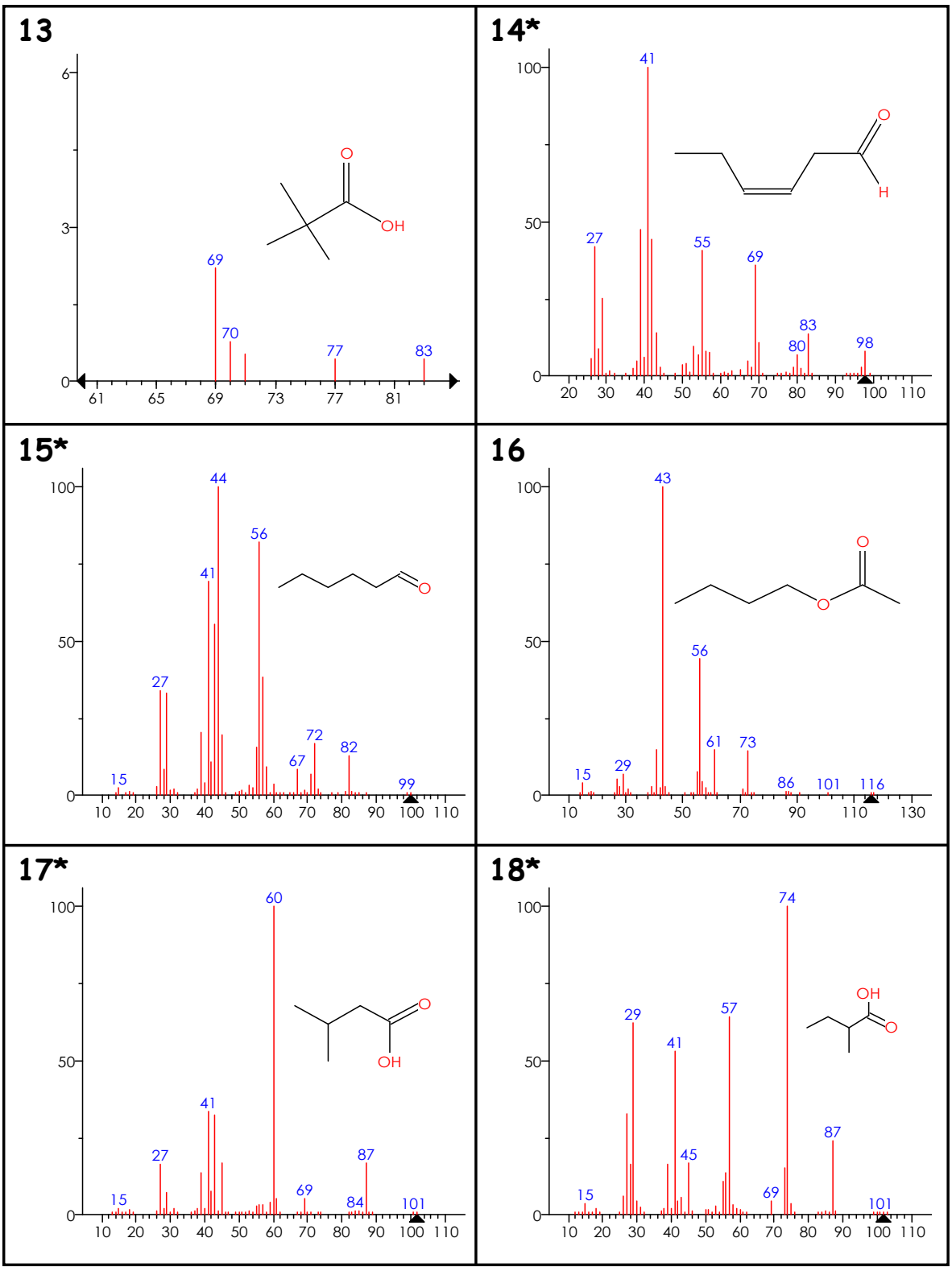



**Fig. S2-continued**

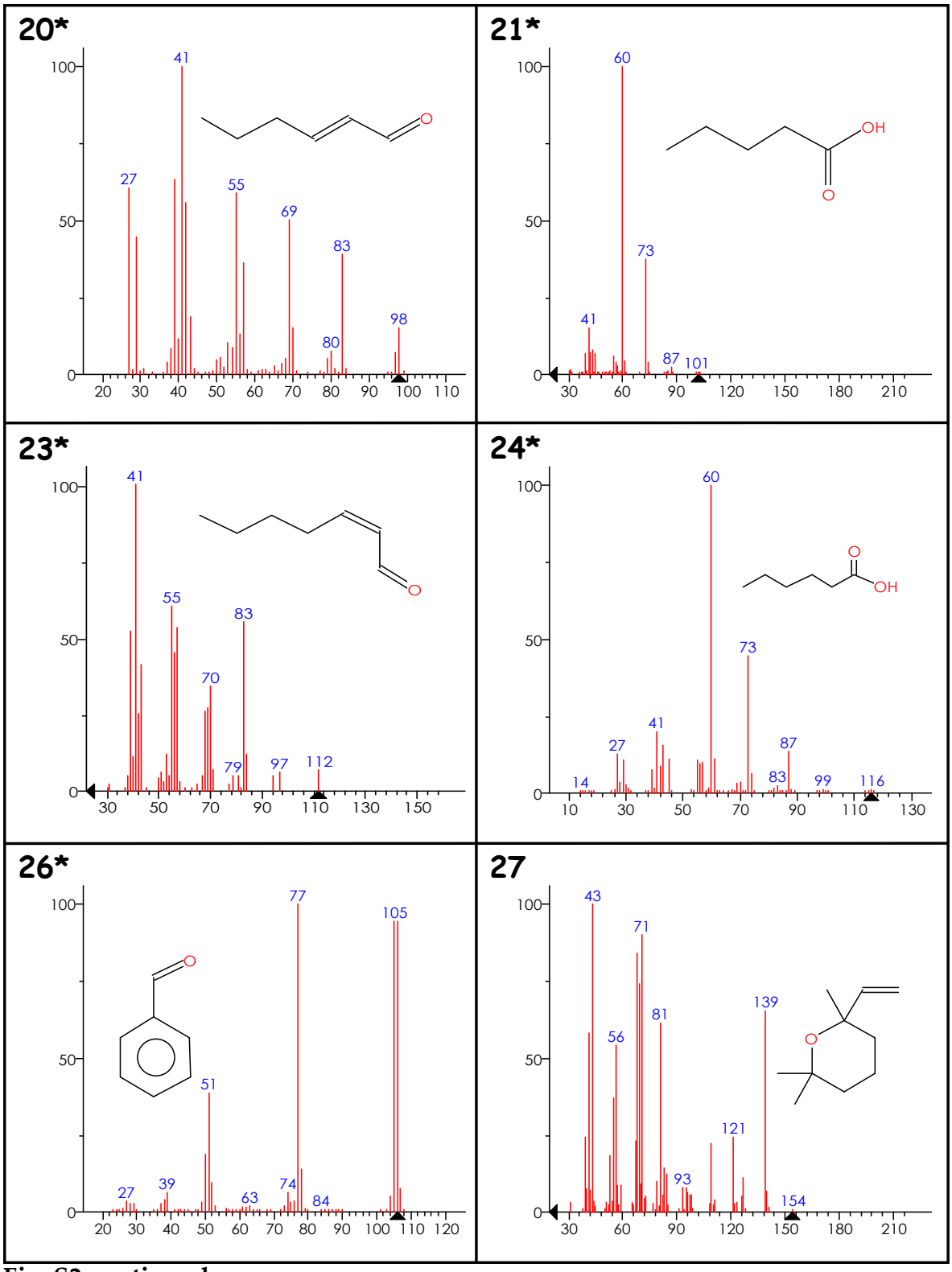



**Fig. S2-continued**

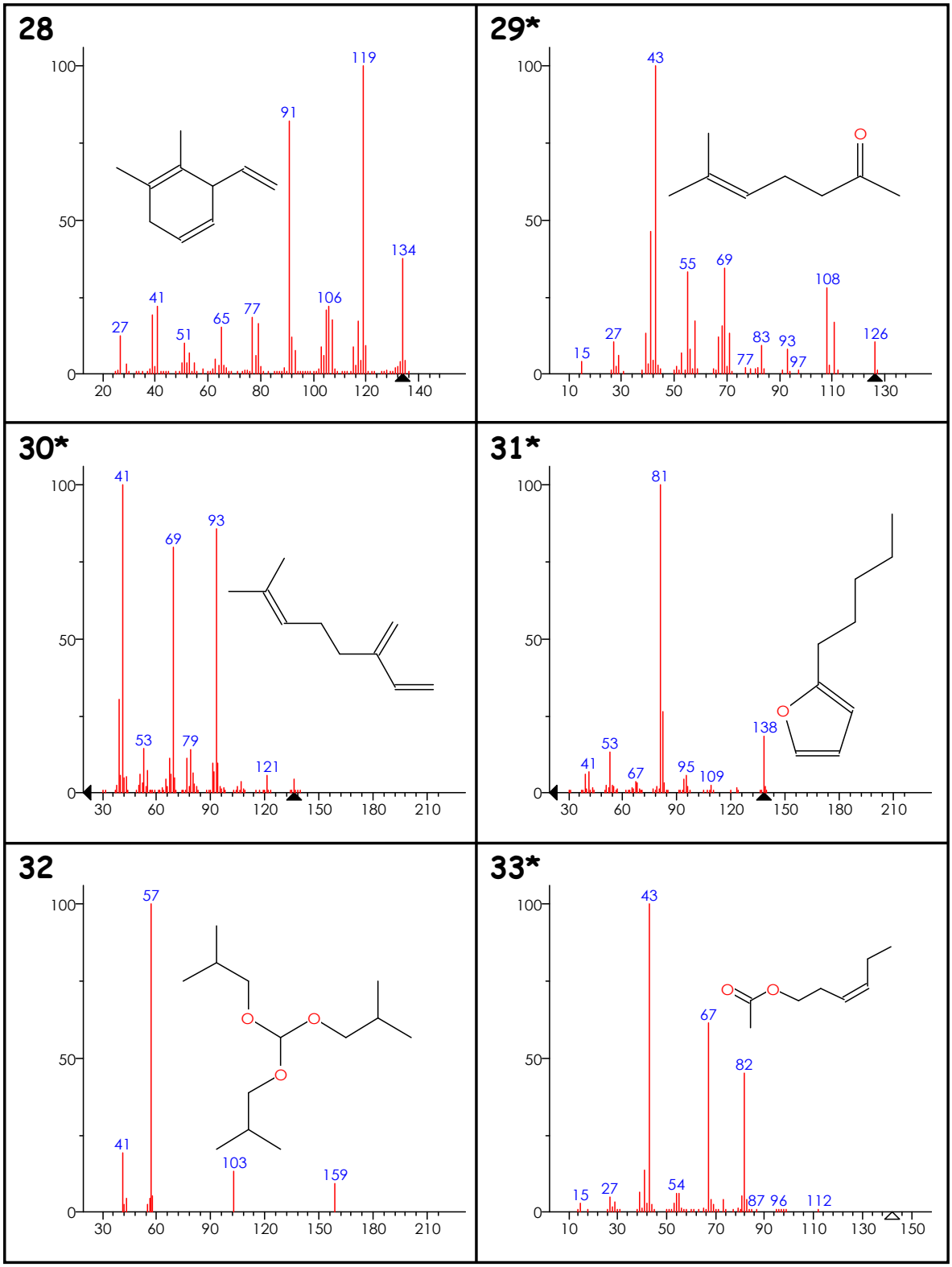



**Fig. S2-continued**

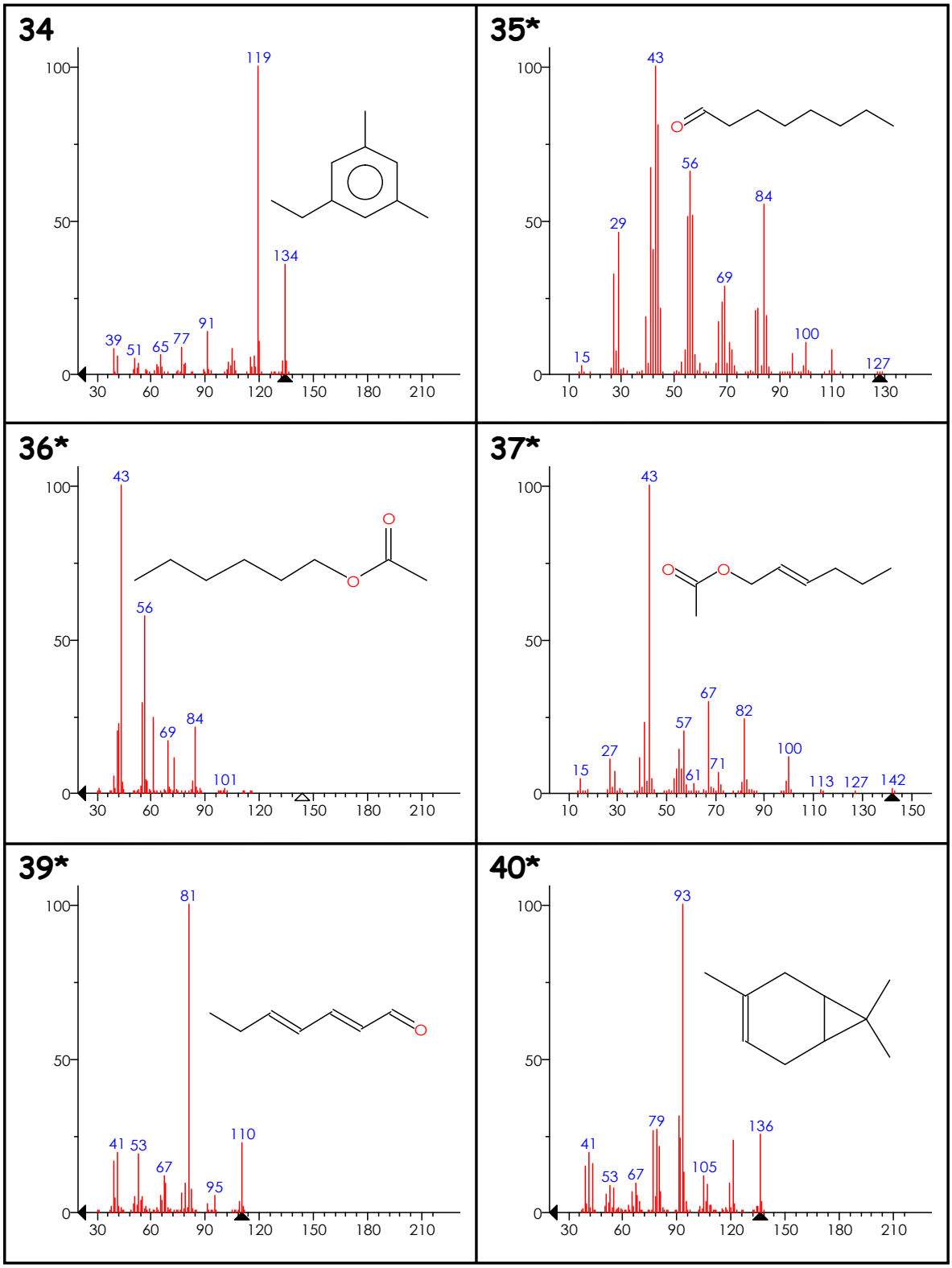



**Fig. S2-continued**

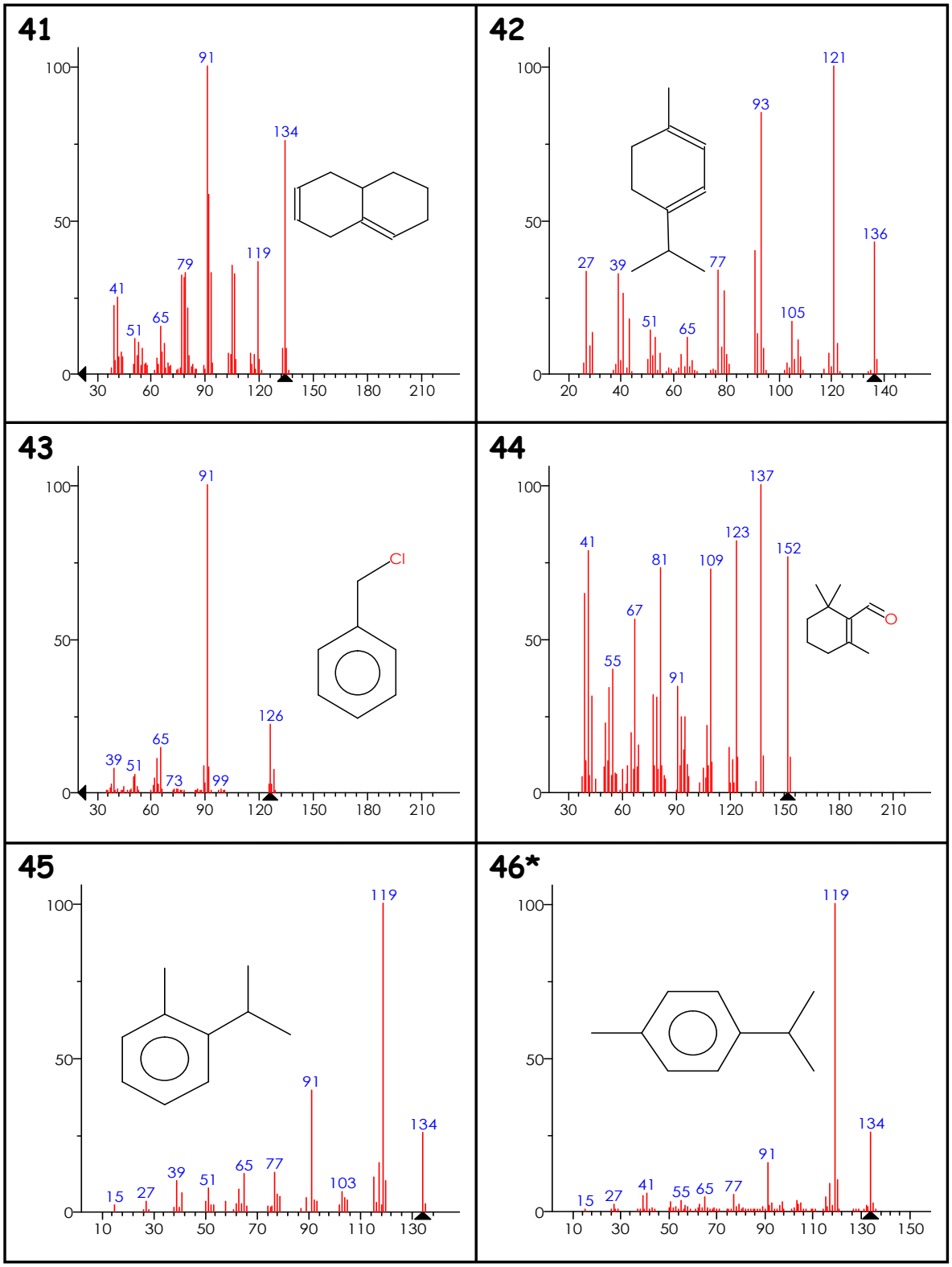



**Fig. S2-continued**

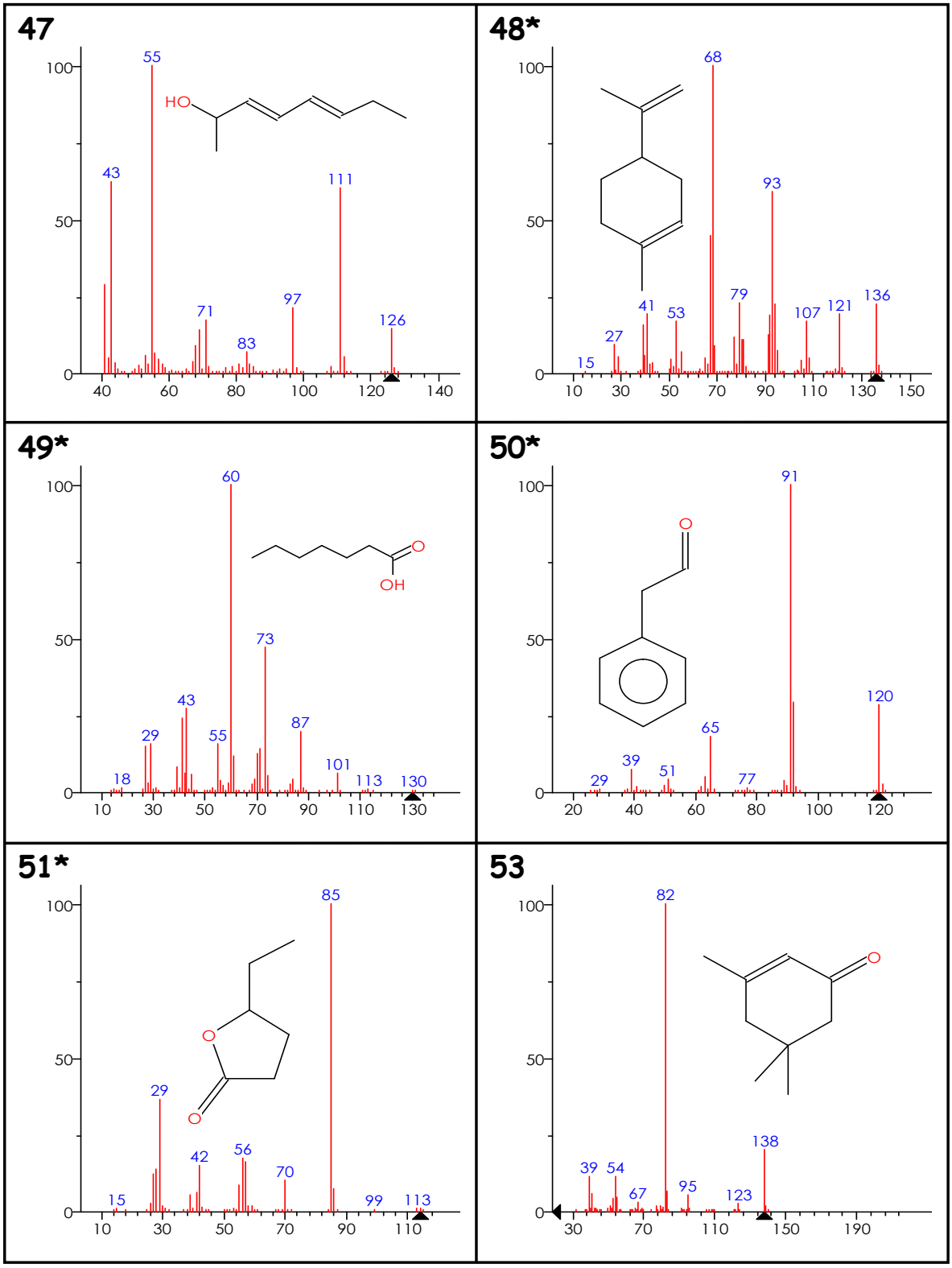



**Fig. S2-continued**

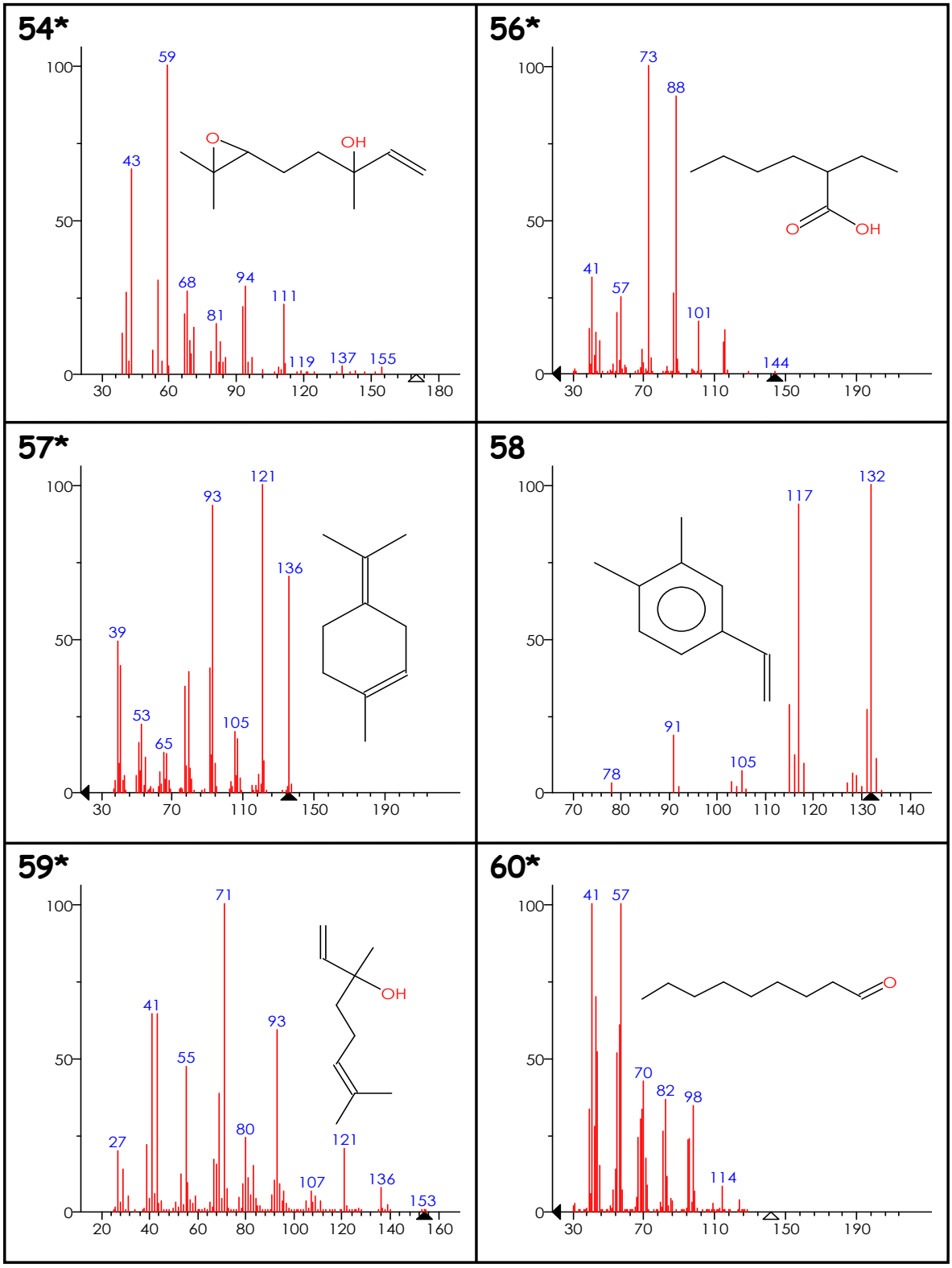



**Fig. S2-continued**

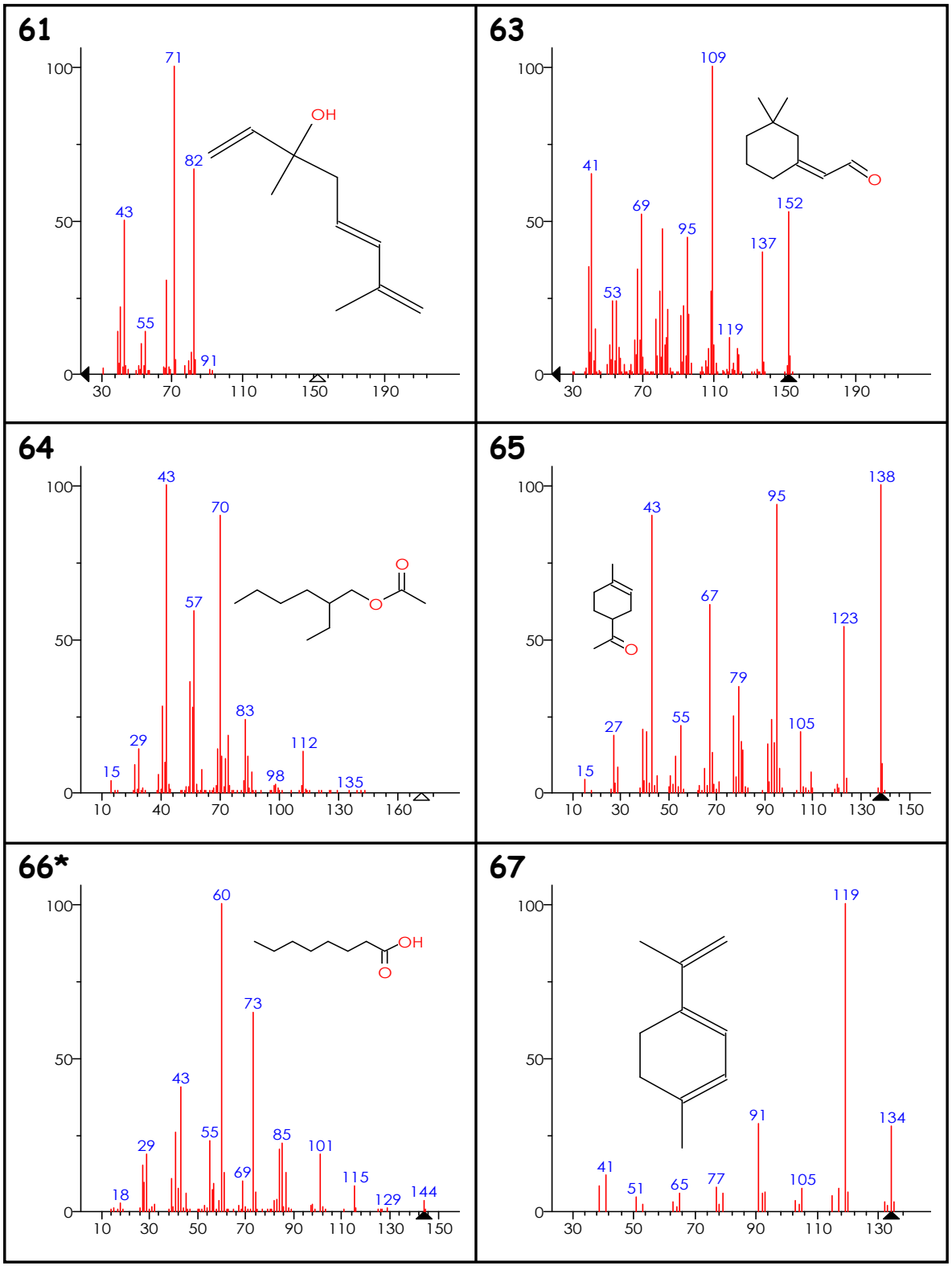



**Fig. S2-continued**

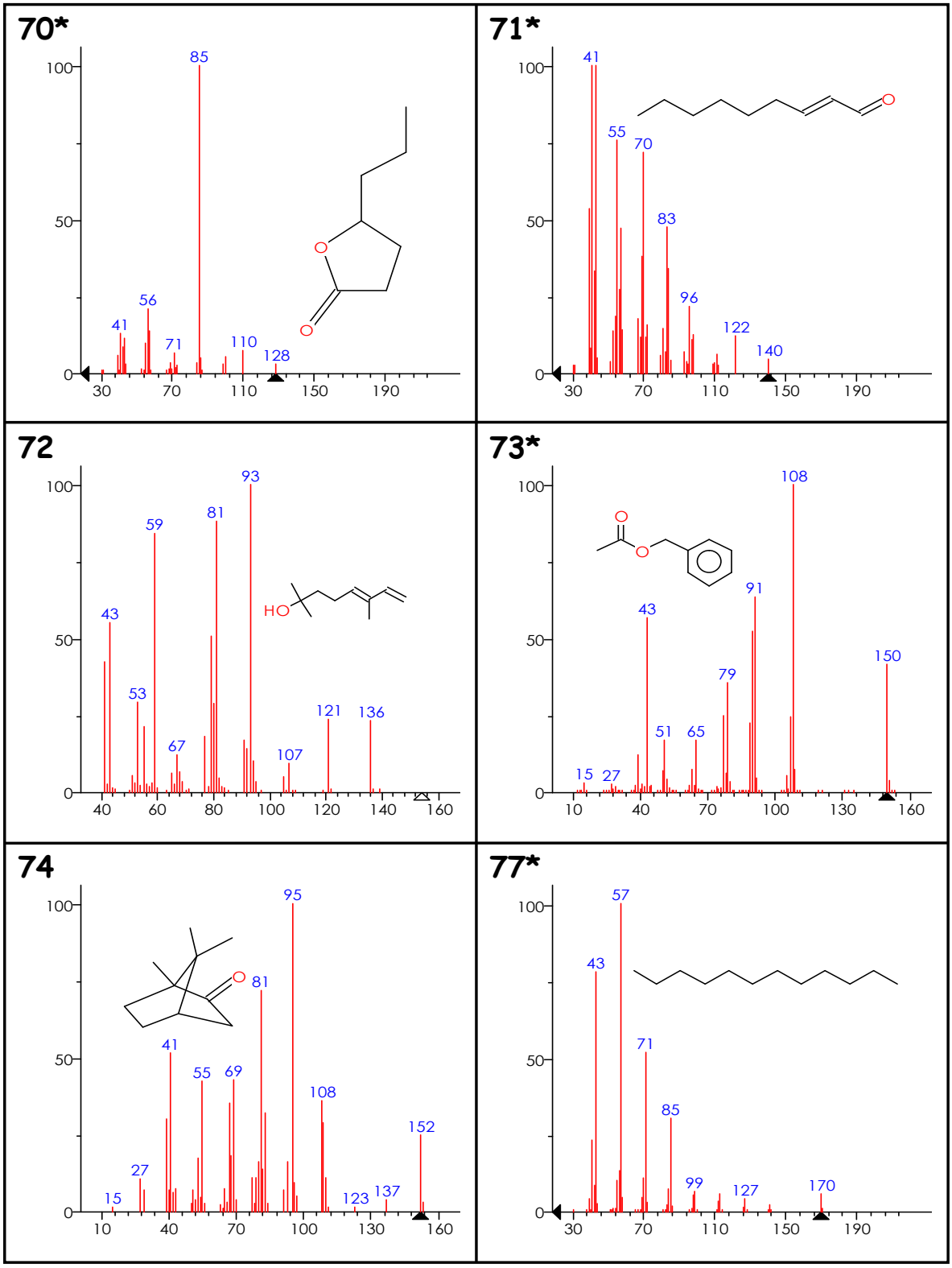



**Fig. S2-continued**

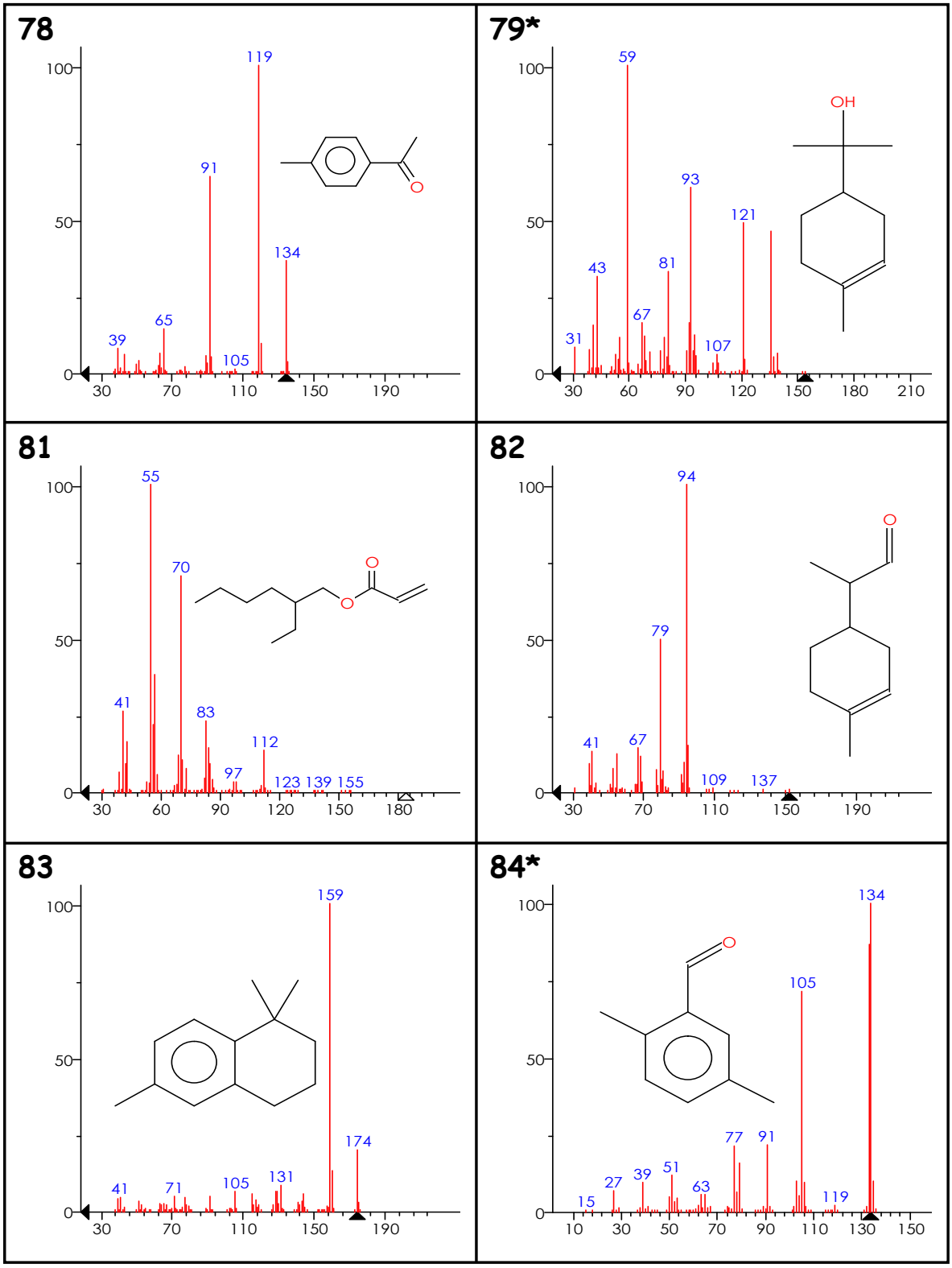



**Fig. S2-continued**

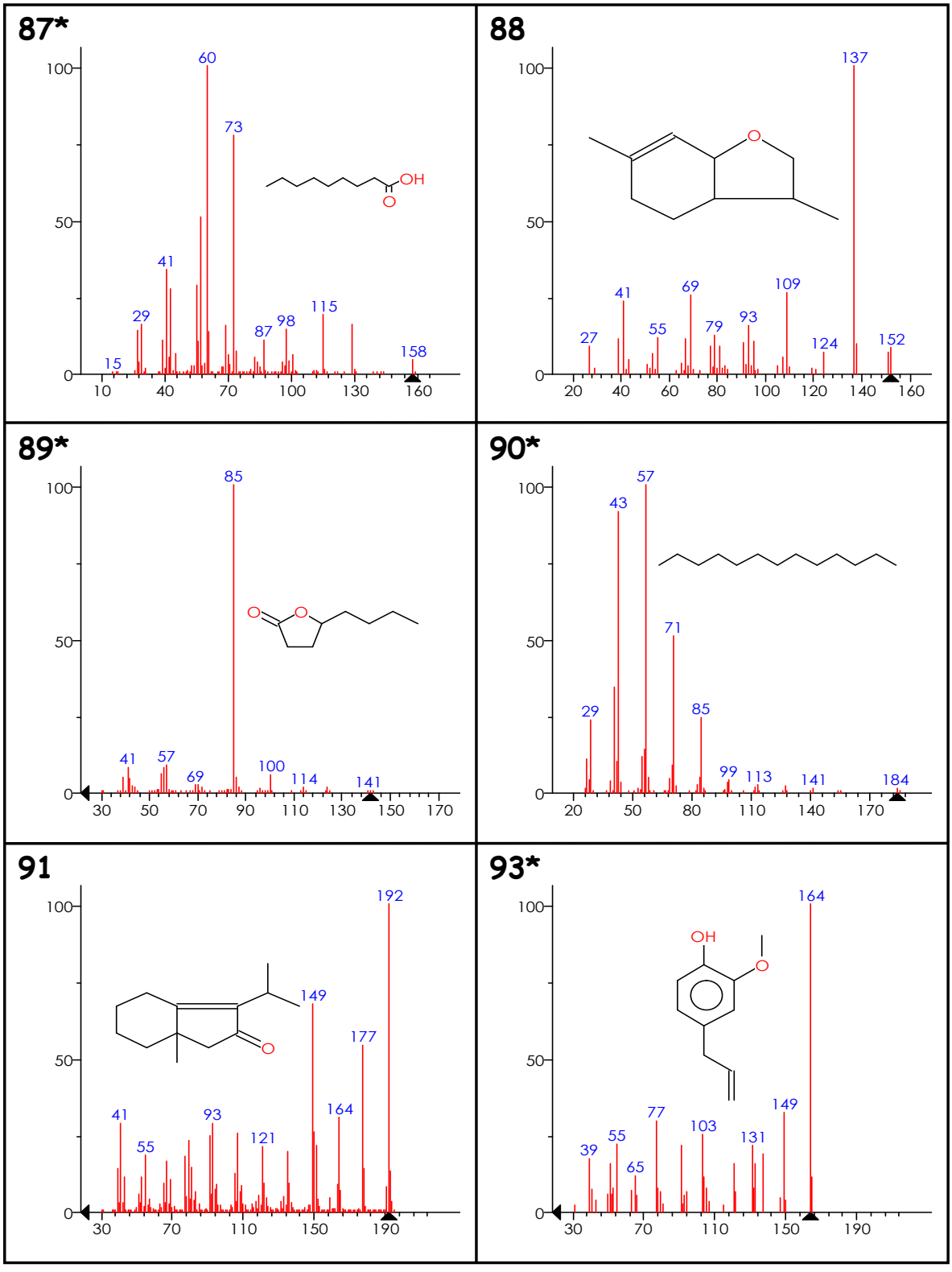



**Fig. S2-continued**

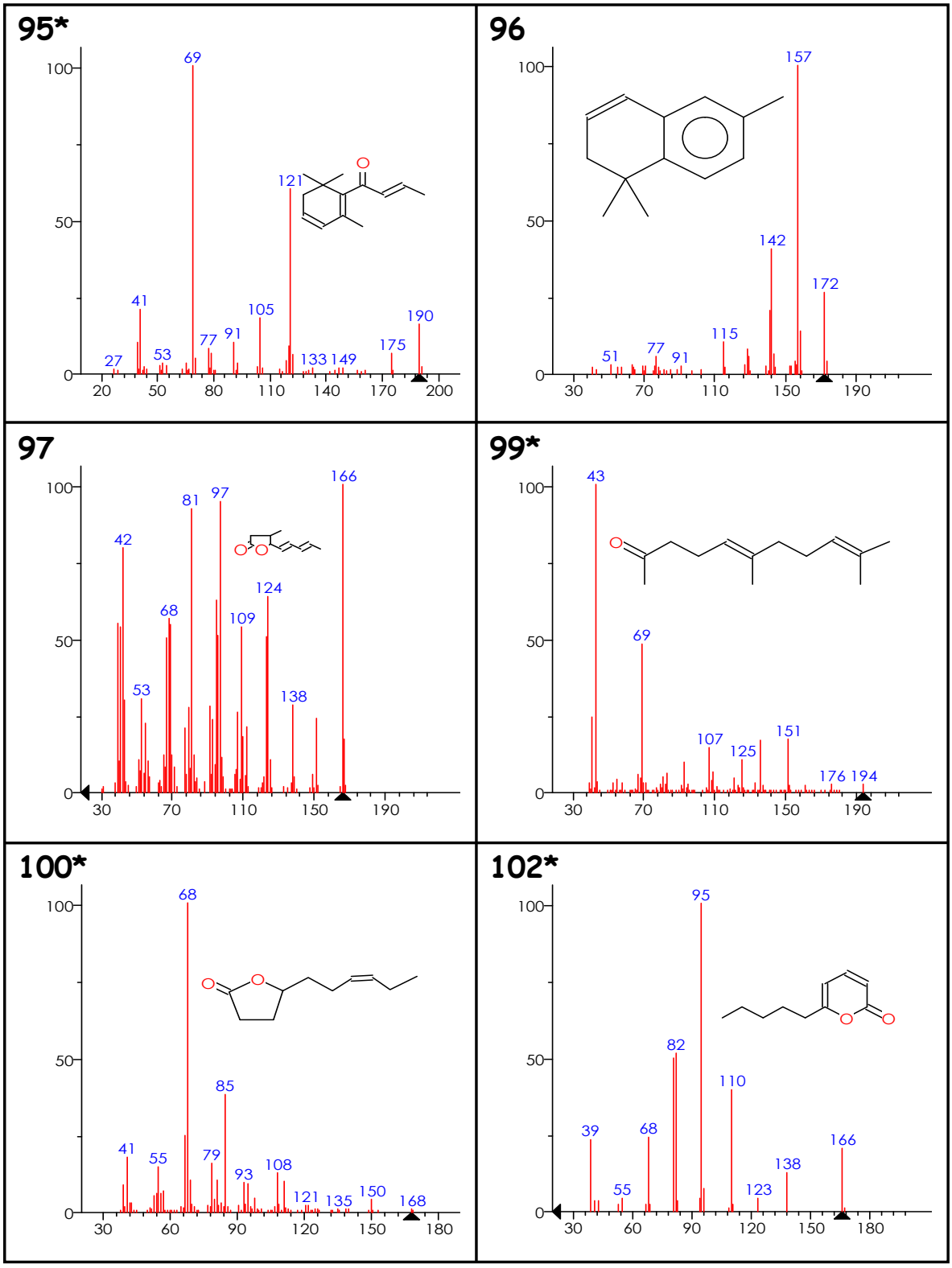



**Fig. S2-continued**

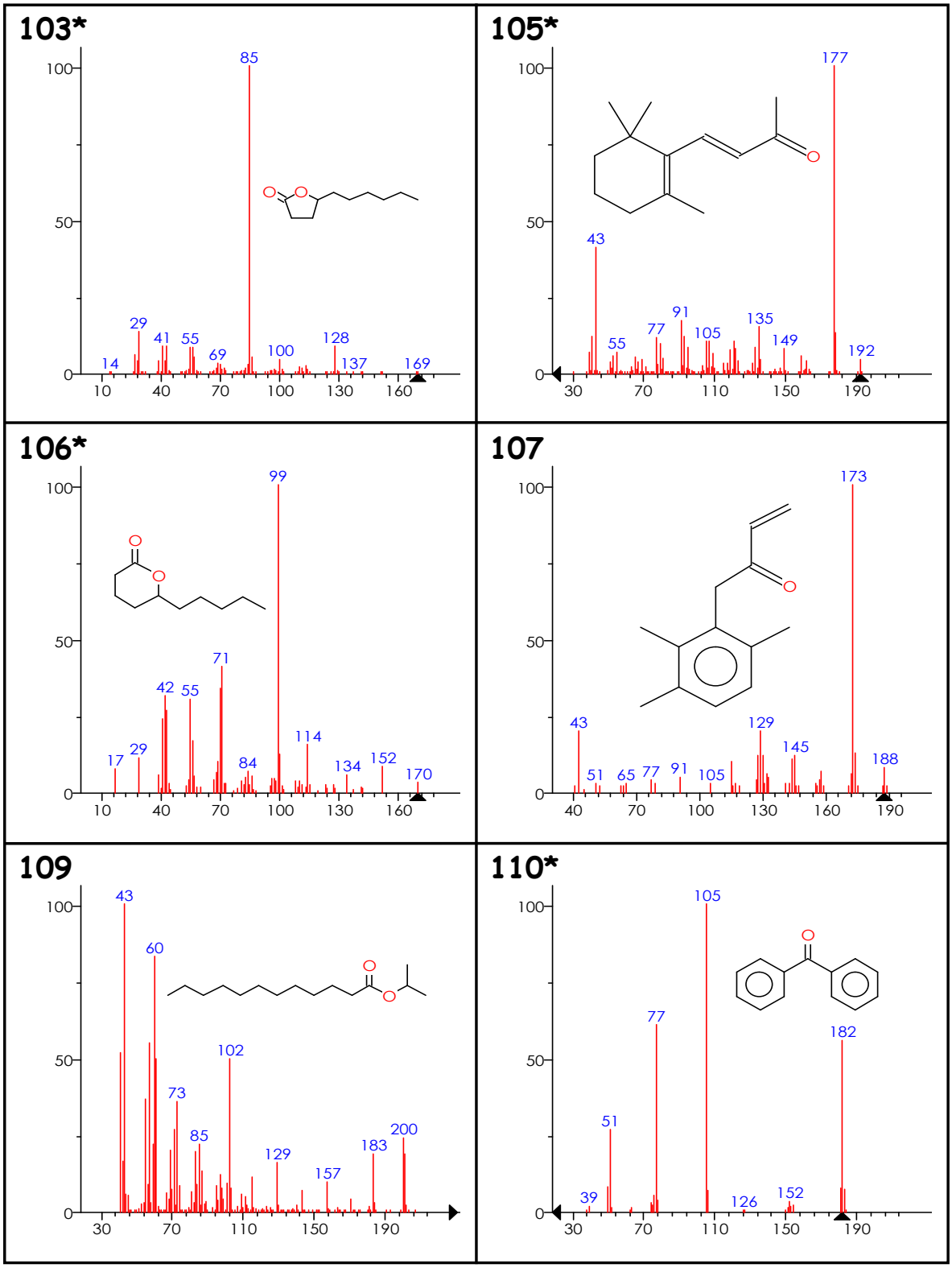



**Figure S2: Chemical structure and mass spectra of volatiles identified in peach.** All spectra were taken from NIST/EPA/NIH mass spectral library version 2.0 with the exception of the spectra of γ-Jasmolactone (Nº 100) that was obtained from the authentic standard. Compounds are numbered according

Table 1. The numbers marked with an asterisk indicate that the compound identity was confirmed with authentic standard.

| Nº | RT (min) | CI | VOCs* | Ion | Nist Match (F) | Nist Match (R) | CAS Nº |
|---|---|---|---|---|---|---|---|
| 1 | 5.659 | 6 | **Ethanol** | 45 | 842 | 856 | 64-17-5 |
| 2 | 7.785 | 5 | Pentane, 2-methyl- | 71 | 880 | 905 | 107-83-5 |
| 3 | 8.201 | 5 | Pentane, 3-methyl- | 57 | 877 | 887 | 96-14-0 |
| 4 | 9.051 | 8 | 3-Buten-2-ol, 2-methyl- | 71 | 834 | 883 | 115-18-4 |
| 5 | 9.108 | 8 | **Ethyl Acetate** | 43 | 871 | 876 | 141-78-6 |
| 6 | 9.410 | 4 | Pentane, 2,2-dimethyl- | 57 | 806 | 905 | 590-35-2 |
| 7 | 9.731 | 4 | Cyclopentane, methyl- | 56 | 848 | 876 | 96-37-7 |
| 8 | 11.545 | 4 | **1-Penten-3-one** | 55 | 816 | 866 | 1629-58-9 |
| 9 | 12.023 | 3 | **Pentanal** | 44 | 932 | 942 | 110-62-3 |
| 10 | 12.110 | 4 | **Furan, 2-ethyl-** | 81 | 821 | 875 | 3208-16-0 |
| 11 | 14.297 | 1 | 2H-Pyran, 3,4-dihydro-6-methyl- | 83 | 815 | 847 | 16015-11-5 |
| 12 | 14.798 | 7 | **Acetic acid, 2-methylpropyl ester** | 56 | 849 | 875 | 110-19-0 |
| 13 | 15.244 | 11 | Propanoic acid, 2,2-dimethyl- | 57 | 875 | 906 | 75-98-9 |
| 14 | 15.925 | 4 | **3-Hexenal, (*Z*)-** | 69 | 831 | 835 | 6789-80-6 |
| 15 | 16.028 | 4 | **Hexanal** | 72 | 922 | 922 | 66-25-1 |
| 16 | 16.413 | 11 | Acetic acid, butyl ester | 43 | 913 | 922 | 123-86-4 |
| 17 | 16.862 | 3 | **Butanoic acid, 3-methyl-** | 60 | 862 | 875 | 503-74-2 |
| 18 | 17.216 | 6 | **Butanoic acid, 2-methyl-** | 74 | 807 | 848 | 116-53-0 |
| 19 | 17.828 | 4 | Ni_01 | 83 | <800 | <800 | na |
| 20 | 18.211 | 4 | **2-Hexenal** | 83 | 927 | 928 | 6728-26-3 |
| 21 | 18.315 | 11 | **Pentanoic acid** | 60 | 801 | 890 | 109-52-4 |
| 22 | 21.350 | 9 | Ni_02 | 69 | <800 | <800 | na |
| 23 | 22.134 | 4 | **2-Heptenal, (Z)-** | 83 | 915 | 943 | 57266-86-1 |
| 24 | 22.137 | 10 | **Hexanoic acid** | 60 | 833 | 866 | 142-62-1 |
| 25 | 22.373 | 9 | Ni_03 | 71 | <800 | <800 | na |
| 26 | 22.753 | 6 | **Benzaldehyde** | 77 | 832 | 869 | 100-52-7 |
| 27 | 22.864 | 1 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | 139 | 812 | 833 | 7392-19-0 |
| 28 | 22.952 | 2 | 1,4-Cyclohexadiene, 3-ethenyl-1,2-dimethyl- | 119 | 840 | 863 | 62338-57-2 |
| 29 | 23.031 | 4 | **5-Hepten-2-one, 6-methyl-** | 108 | 847 | 893 | 110-93-0 |
| 30 | 23.279 | 1 | **β-Myrcene** | 93 | 820 | 842 | 123-35-3 |
| 31 | 23.350 | 3 | **Furan, 2-pentyl-** | 81 | 805 | 852 | 3777-69-3 |
| 32 | 23.365 | 9 | Orthoformic acid, triisobutyl ester | 57 | 803 | 905 | 16754-49-7 |
| 33 | 23.614 | 7 | **3-Hexen-1-ol, acetate, (*Z*)-** | 67 | 908 | 908 | 3681-71-8 |
| 34 | 23.731 | 2 | Benzene, 1-ethyl-3,5-dimethyl- | 119 | 802 | 835 | 934-74-7 |
| 35 | 23.783 | 5 | **Octanal** | 84 | 860 | 880 | 124-13-0 |
| 36 | 23.837 | 7 | **Acetic acid, hexyl ester** | 69 | 853 | 895 | 142-92-7 |
| 37 | 23.906 | 7 | **2-Hexen-1-ol, acetate, (*E*)-** | 67 | 805 | 831 | 2497-18-9 |
| 38 | 23.965 | 9 | Ni_04 | 89 | <800 | <800 | na |
| 39 | 24.210 | 4 | **2,4-Heptadienal, (*E*,*E*)-** | 81 | 889 | 892 | 881395 |
| 40 | 24.424 | 1 | **3-Carene** | 93 | 819 | 854 | 13466-78-9 |
| 41 | 24.534 | 2 | Naphthalene, 1,2,3,5,8,8α-hexahydro- | 91 | 801 | 839 | 62690-65-7 |
| 42 | 24.631 | 1 | 1,3-Cyclohexadiene, 1-methyl-4-(1-methylethyl)- | 121 | 832 | 870 | 99-86-5 |
| 43 | 24.645 | 3 | Benzyl chloride | 91 | 852 | 886 | 100-44-7 |
| 44 | 24.775 | 2 | 1-Cyclohexene-1-carboxaldehyde,2,6,6-trimethyl | 109 | 802 | 823 | 432-25-7 |
| 45 | 24.893 | 2 | Benzene, 1-methyl-2-(1-methylethyl)- | 119 | 861 | 884 | 535-77-3 |
| 46 | 24.912 | 2 | **p-Cymene** | 119 | 803 | 860 | 99-87-6 |
| 47 | 24.994 | 5 | 3,5-Octadien-2-ol | 111 | 834 | 857 | 69668-82-2 |
| 48 | 25.112 | 1 | **δ-Limonene** | 68 | 847 | 856 | 138-86-3 |
| 49 | 25.504 | 10 | **Heptanoic acid** | 60 | 852 | 881 | 111-14-8 |
| 50 | 25.632 | 5 | **Benzeneacetaldehyde** | 91 | 829 | 884 | 122-78-1 |
| 51 | 25.774 | 7 | **γ-Hexalactone** | 85 | 843 | 888 | 695-06-7 |
| 52 | 25.779 | 6 | Ni_05 | 70 | <800 | <800 | na |
| 53 | 26.187 | 5 | 2-Cyclohexen-1-one, 3,5,5-trimethyl- | 82 | 808 | 850 | 78-59-1 |
| 54 | 26.477 | 2 | **cis-Linaloloxide** | 59 | 840 | 916 | na |
| 55 | 26.998 | 3 | Ni_06 | 95 | <800 | <800 | na |
| 56 | 27.047 | 10 | **Hexanoic acid, 2-ethyl-** | 73 | 838 | 851 | 149-57-5 |
| 57 | 27.049 | 1 | **Terpinolene** | 121 | 801 | 854 | 586-62-9 |
| 58 | 27.175 | 2 | Benzene, 4-ethenyl-1,2-dimethyl- | 132 | 865 | 924 | 27831-13-6 |
| 59 | 27.184 | 1 | **Linalool** | 93 | 830 | 847 | 78-70-6 |
| 60 | 27.295 | 5 | **Nonanal** | 57 | 817 | 851 | 124-19-6 |
| 61 | 27.319 | 2 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | 71 | 804 | 868 | 29957-43-5 |

**Table S1 (see the legends below).**

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 62 | 27.491 | 2 | Ni_07 | 67 | <800 | <800 | na |
| 63 | 27.994 | 11 | Acetaldehyde, (3,3-dimethylcyclohexylidene)-,(E) | 108 | 801 | 805 | 26532-25-1 |
| 64 | 28.464 | 9 | Acetic acid, 2-ethylhexyl ester | 70 | 822 | 877 | 103-09-3 |
| 65 | 28.610 | 2 | 4-Acetyl-1-methylcyclohexene | 138 | 905 | 917 | 70286-20-3 |
| 66 | 28.795 | 10 | **Octanoic Acid** | 73 | 844 | 908 | 124-07-2 |
| 67 | 28.812 | 2 | 1,3,8-p-Menthatriene | 119 | 838 | 886 | 21195-59-5 |
| 68 | 29.005 | 2 | Ni_08 | 84 | <800 | <800 | na |
| 69 | 29.012 | 2 | Ni_09 | 85 | <800 | <800 | na |
| 70 | 29.133 | 7 | **γ-Heptalatone** | 85 | 805 | 884 | 105-21-5 |
| 71 | 29.220 | 5 | **2-Nonenal** | 83 | 801 | 876 | 2463-53-8 |
| 72 | 29.333 | 1 | 5,7-Octadien-2-ol, 2,6-dimethyl- | 93 | 872 | 882 | 5986-38-9 |
| 73 | 29.383 | 11 | **Acetic acid, phenylmethyl ester** | 108 | 804 | 905 | 140-11-4 |
| 74 | 29.553 | 11 | Bicyclo[2.2.1]heptan-2-one, 1,7,7-trimethyl-, (1*S*)- | 95 | 913 | 933 | 464-48-2 |
| 75 | 29.556 | 11 | Ni_10 | 112 | <800 | <800 | na |
| 76 | 30.245 | 1 | Ni_11 | 69 | <800 | <800 | na |
| 77 | 30.276 | 5 | **Dodecane** | 57 | 838 | 886 | 112-40-3 |
| 78 | 30.498 | 2 | Ethanone, 1-(4-methylphenyl)- | 119 | 832 | 851 | 122-00-9 |
| 79 | 30.753 | 1 | **α-Terpinol** | 121 | 815 | 827 | 98-55-5 |
| 80 | 31.025 | 4 | Ni_12 | 88 | <800 | <800 | na |
| 81 | 31.067 | 9 | 2-Propenoic acid, 2-ethylhexyl ester | 70 | 894 | 906 | 103-11-7 |
| 82 | 31.516 | 11 | 3-Cyclohexene-1-acetaldehyde, α,4-dimethyl- | 94 | 815 | 837 | 29548-14-9 |
| 83 | 31.516 | 8 | Naphthalene, 1,2,3,4-tetrahydro-1,1,6-trimethyl- | 159 | 801 | 845 | 475-03-6 |
| 84 | 31.584 | 9 | **Benzaldehyde, 2,5-dimethyl-** | 133 | 833 | 844 | 5779-94-2 |
| 85 | 31.620 | 4 | Ni_13 | 137 | <800 | <800 | na |
| 86 | 31.729 | 9 | Ni_14 | 140 | <800 | <800 | na |
| 87 | 31.917 | 10 | **Nonanoic acid** | 73 | 889 | 895 | 112-05-0 |
| 88 | 32.097 | 2 | 3,6-Dimethyl-2,3,3α,4,5,7α-hexahydrobenzofuran | 137 | 904 | 934 | 70786-44-6 |
| 89 | 32.451 | 7 | **γ-Octalactone** | 85 | 920 | 923 | 104-50-7 |
| 90 | 33.313 | 5 | **Tridecane** | 57 | 871 | 891 | 629-50-5 |
| 91 | 33.677 | 6 | 1H-2-Indenone,2,4,5,6,7,77α-hexahydro .....[a] | 192 | 806 | 806 | na |
| 92 | 34.313 | 6 | Ni_15 | 138 | <800 | <800 | na |
| 93 | 35.340 | 3 | **Eugenol** | 164 | 885 | 952 | 97-53-0 |
| 94 | 35.378 | 4 | Ni_16 | 174 | na | na | na |
| 95 | 36.206 | 8 | **β-Damascenone** | 69 | 846 | 904 | 23726-93-4 |
| 96 | 36.934 | 8 | Naphthalene, 1,2-dihydro-1,1,6-trimethyl- | 157 | 805 | 889 | 30364-38-6 |
| 97 | 37.000 | 7 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | 68 | 823 | 835 | 185211 |
| 98 | 37.350 | 3 | Ni_17 | 173 | <800 | <800 | na |
| 99 | 37.621 | 6 | **Geranyl acetone** | 69 | 803 | 816 | 3796-70-1 |
| 100 | 38.071 | 7 | **γ-Jasmolactone** | 68 | 948 | 948 | 93787-95-2 |
| 101 | 38.165 | 9 | Ni_18 | 165 | <800 | <800 | na |
| 102 | 38.314 | 7 | **2H-Pyran-2-one, 6-pentyl-** | 95 | 857 | 906 | 27593-23-3 |
| 103 | 38.427 | 7 | **γ-Decalactone** | 85 | 924 | 924 | 706-14-9 |
| 104 | 38.486 | 9 | Ni_19 | 177 | <800 | <800 | na |
| 105 | 38.883 | 4 | **β-Ionone** | 177 | 844 | 851 | 14901-07-6 |
| 106 | 39.237 | 7 | **δ-Decalactone** | 99 | 855 | 902 | 705-86-2 |
| 107 | 39.736 | 3 | 3-Buten-2-one, 1-(2,3,6-trimethylphenyl)- | 173 | 825 | 874 | 54789-45-6 |
| 108 | 39.882 | 4 | Ni_20 | 124 | <800 | <800 | na |
| 109 | 41.883 | 9 | Dodecanoic acid, 1-methylethyl ester | 102 | 805 | 807 | 10233-13-3 |
| 110 | 43.170 | 11 | **Benzophenone** | 182 | 806 | 917 | 119-61-9 |

**Table S1. Volatile organic compounds detected in the sample set.** For each volatile, the retention time (RT) in min, the cluster (Cl) that it belong according Fig. 1, the specific ion (Ion, m/z) used for quantification, the forward (F) and reverse (R) match against Nist library (with the exception for γ-Jasmolactone where the F and R match against its authentic standard is shown) and the CAS number are indicated. *compound identified by comparing of its retention time with authentic standard are highlighted in bold. [a]The full name of compound Nº91 is: 1H-2-Indenone,2,4,5,6,7,7*α*-hexahydro-3-(1-methylethyl)-7*α*-methyl. na, not available.

**Información suplementaria de Sánchez et al., 2013 (Capitulo 3)***

*La información presentada en esta sección puede ser encontrada en:
http://www.biomedcentral.com/1471-2164/14/343/additional

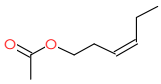
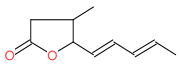
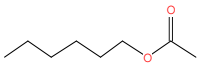
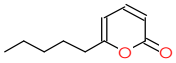
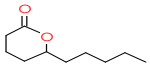
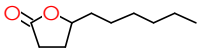
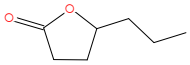
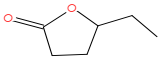
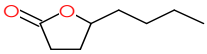
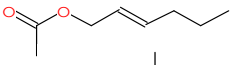
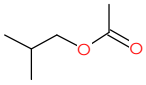
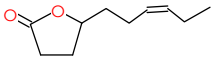
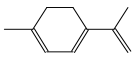
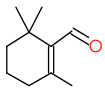
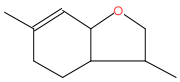
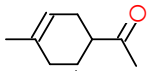
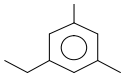
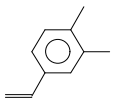
| Nº | VOCs* | Structure | CAS Nº | Group | Odor description** |
|---|---|---|---|---|---|
| **C1** | | | | | |
| 1 | **3-Hexen-1-ol, acetate, (Z)-** |  | 3681-71-8 | Ester | fruity[1] |
| 2 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one |  | 185211 | Lactone | |
| 3 | **Acetic acid, hexyl ester** |  | 142-92-7 | Ester | fruity[W] |
| **C2** | | | | | |
| 4 | **2H-Pyran-2-one, 6-pentyl-** |  | 27593-23-3 | Lactone | coconut-like[1] |
| 5 | **δ-Decalactone** |  | 705-86-2 | Lactone | coconut-like[1] |
| 6 | **γ-Decalactone** |  | 706-14-9 | Lactone | peach-like[1] |
| 7 | **γ-Heptalatone** |  | 105-21-5 | Lactone | coconut-like[w] |
| 8 | **γ-Hexalactone** |  | 695-06-7 | Lactone | coconut-like, fruity[1] |
| 9 | **γ-Octalactone** |  | 104-50-7 | Lactone | coconut-like[1] |
| **C3** | | | | | |
| 10 | **2-Hexen-1-ol, acetate, (E)-** |  | 2497-18-9 | Ester | fruity[w] |
| 11 | **Acetic acid, 2-methylpropyl ester** |  | 110-19-0 | Ester | fruity[w] |
| 12 | **γ-Jasmolactone** |  | 93787-95-2 | Lactone | peach-like[1] |
| **C4** | | | | | |
| 13 | 1,3,8-p-Menthatriene |  | 21195-59-5 | Terpenoid | |
| 14 | 1-Cyclohexene-1-carboxaldehyde -2,6,6-trimethyl- |  | 432-25-7 | | |
| 15 | 3,6-Dimethyl-2,3,3α,4,5,7 α-hexahydrobenzofuran |  | 70786-44-6 | | |
| 16 | 4-Acetyl-1-methylcyclohexene |  | 70286-20-3 | Terpenoid | |
| 17 | Benzene, 1-ethyl-3,5-dimethyl- |  | 934-74-7 | | |
| 18 | Benzene, 4-ethenyl-1,2-dimethyl- |  | 27831-13-6 | | |

**Fig. S1 (see the legends below).**

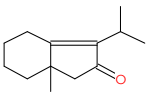
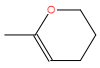
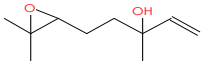
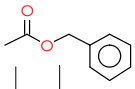
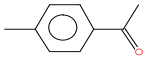
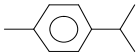
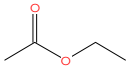
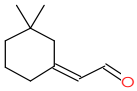
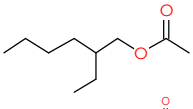
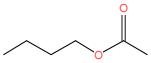
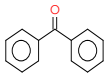
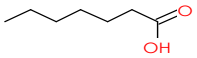
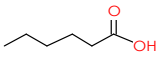
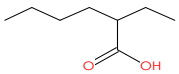
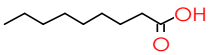
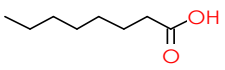
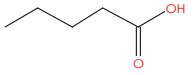
| | | | | | |
|---|---|---|---|---|---|
| **C5** | | | | | |
| 19 | 1H-2-Indenone,2,4,5,6,7,7 α-hexahydro-3-(1-methylethyl)-7α-methyl |  | n.a. | | |
| 20 | 2H-Pyran, 3,4-dihydro-6-methyl- |  | 16015-11-5 | | |
| 21 | **cis-Linaloloxide** |  | n.a . | Terpenoid | |
| **C6** | | | | | |
| 22 | **Acetic acid, phenylmethyl ester** |  | 140-11-4 | Ester | floral [w] |
| 23 | Benzene, 1-methyl-2-(1-methylethyl)- |  | 535-77-3 | | |
| 24 | Ethanone, 1-(4-methylphenyl)- |  | 122-00-9 | | |
| 25 | **p-Cymene** |  | 99-87-6 | Terpenoid | |
| **C7** | | | | | |
| 26 | **Ethanol** |  | 64-17-5 | | |
| 27 | **Ethyl Acetate** |  | 141-78-6 | Ester | fruity[2] |
| **C8** | | | | | |
| 28 | Acetaldehyde, (3,3-dimethylcyclohexylidene)-, (E)- |  | 26532-25-1 | | |
| 29 | Acetic acid, 2-ethylhexyl ester |  | 103-09-3 | Ester | |
| 30 | Acetic acid, butyl ester |  | 123-86-4 | Ester | |
| 31 | **Benzophenone** |  | 119-61-9 | | balsamic[w] |
| 32 | Bicyclo[2.2.1]heptan-2-one 1,7,7-trimethyl-, (1S)- |  | 464-48-2 | | |
| **C9** | | | | | |
| 33 | **Heptanoic acid** |  | 111-14-8 | Carboxilic acid | |
| 34 | **Hexanoic acid** |  | 142-62-1 | Carboxilic acid | |
| 35 | **Hexanoic acid, 2-ethyl-** |  | 149-57-5 | Carboxilic acid | |
| 36 | **Nonanoic acid** |  | 112-05-0 | Carboxilic acid | |
| 37 | **Octanoic Acid** |  | 124-07-2 | Carboxilic acid | fatty[w] |
| 38 | **Pentanoic acid** |  | 109-52-4 | Carboxilic acid | |

**Fig. S1-continued**

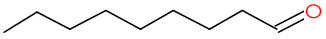
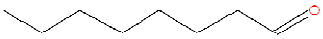
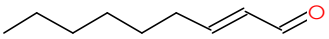
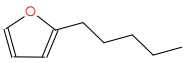
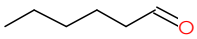
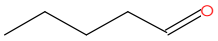
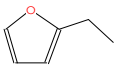
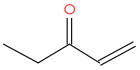
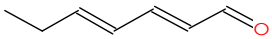
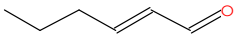
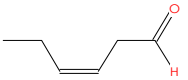
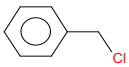
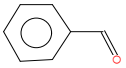
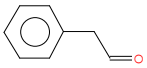
| | | | | | |
|---|---|---|---|---|---|
| **C10** | | | | | |
| 39 | **Nonanal** |  | 124-19-6 | Long chain aldheyde | aldehydic[w] |
| 40 | **Octanal** |  | 124-13-0 | Long chain aldheyde | aldehydic[w] |
| 41 | **2-Nonenal** |  | 2463-53-8 | Long chain aldheyde | fatty[w] |
| **C11** | | | | | |
| 42 | **Furan, 2-pentyl-** |  | 3777-69-3 | Lipid derivative | fruity[w] |
| 43 | **Hexanal** |  | 66-25-1 | Lipid derivative | green[w] |
| 44 | **Pentanal** |  | 110-62-3 | Lipid derivative | fermented[w] |
| **C12** | | | | | |
| 45 | **Furan, 2-ethyl-** |  | 3208-16-0 | Lipid derivative | chemical[w] |
| 46 | **1-Penten-3-one** |  | 1629-58-9 | Lipid derivative | spicy[w] |
| 47 | **2,4-Heptadienal, (E,E)-** |  | 881395 | Lipid derivative | fatty[w] |
| 48 | **2-Hexenal** |  | 6728-26-3 | Lipid derivative | green, banana-like [1] |
| 49 | **3-Hexenal, (Z)-** |  | 6789-80-6 | Lipid derivative | green, apple-like [w] |
| **C13** | | | | | |
| 50 | Benzyl chloride |  | 100-44-7 | Phenolic | |
| 51 | **Benzaldehyde** |  | 100-52-7 | Phenolic | |
| 52 | **Benzeneacetaldehyde** |  | 122-78-1 | Phenolic | green[w] |

**Figure S1**. **Volatile compounds analyzed in this study**. For each volatile, the cluster that they belongs to according to Fig. 1 (upper left corner), the chemical structure, the CAS number, the group and the odor description are shown. n.a., not available. *Bold indicates those volatiles whose retention time was verified by an authentic standard. **References for odor descriptions are: 1, Derail et al., 1999 [34]; 2, Guillot et al., 2006 [61]; and w, www.thegoodscentscompany.com.

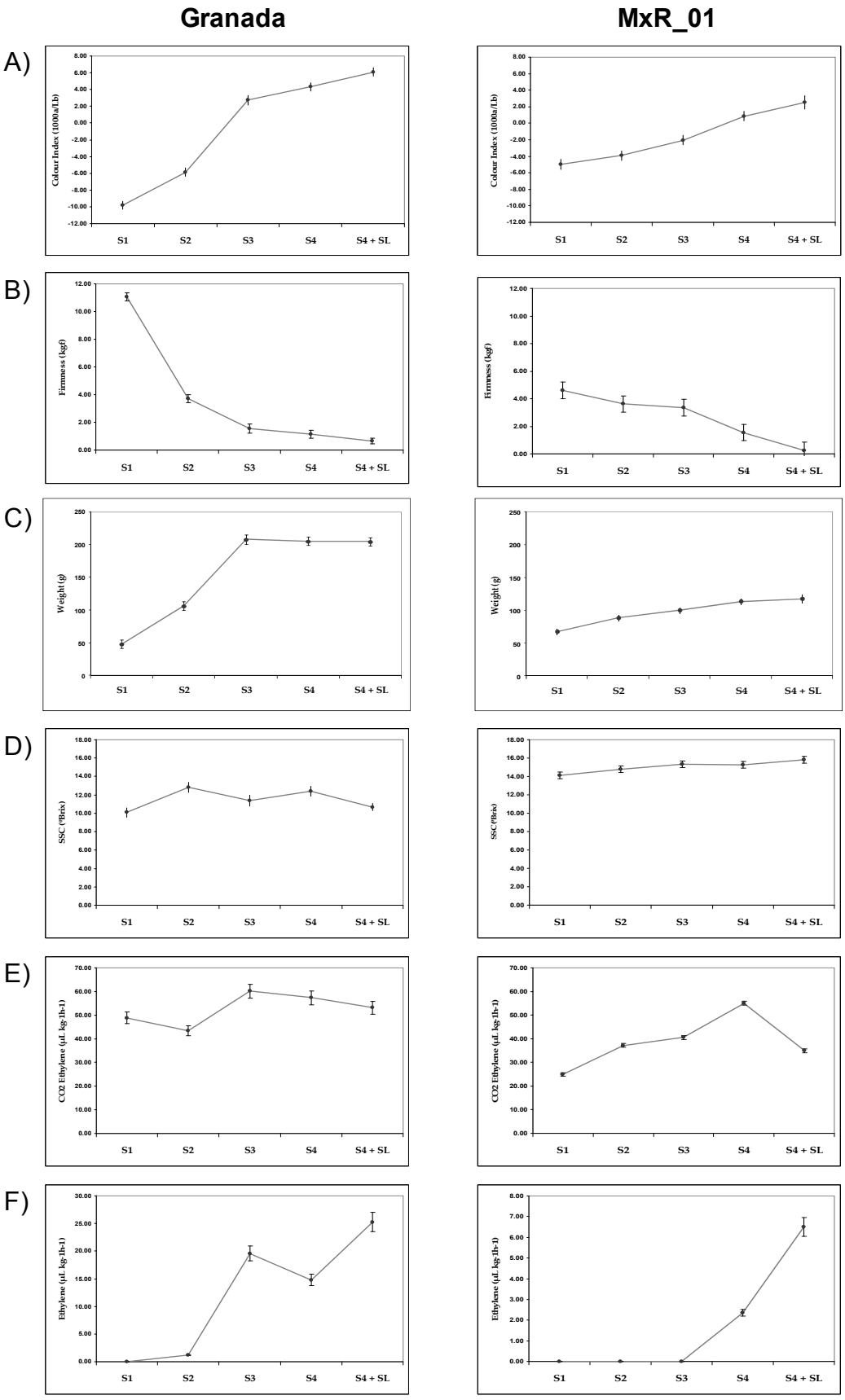



**Fig. S2 (see the legends below).**

**Figure S2. Maturity time-course series of the 'Granada' and 'MxR_01' genotypes. A)** Color index **B)** Firmness **C)** Weight **D)** Soluble Solids Content (SSC) **E)** CO2 consumption **F)** Ethylene production. Bars represent the LSD range.

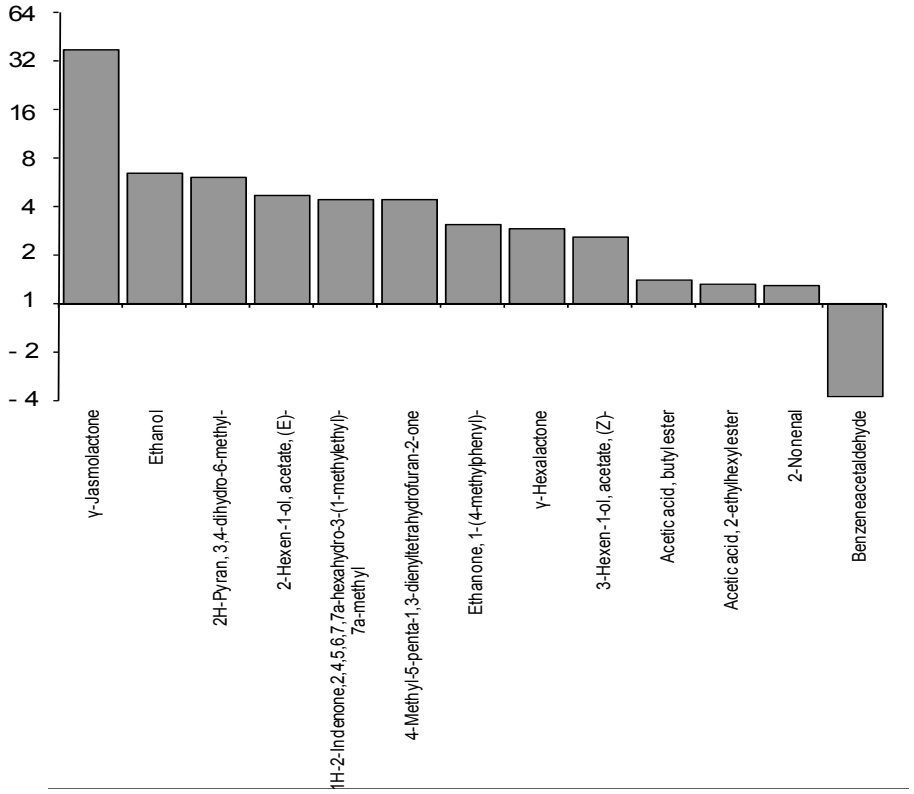



**Figure S3. Comparison of volatile contents in 'Granada 'and 'MxR_01' at commercial maturity stage (S4).** The values are expressed as fold changes on the Log2 scale. The positive region of the y-axis was used for values higher in 'Granada' as compared to 'MxR_01', and the indicated fold change is 'Granada'/'MxR_01', while the negative region is used for values that are higher in 'MxR_01' as compared to 'Granada', and the indicated fold change is 'MxR_01'/'Granada'. All the differences are significant ($p<0.05$).

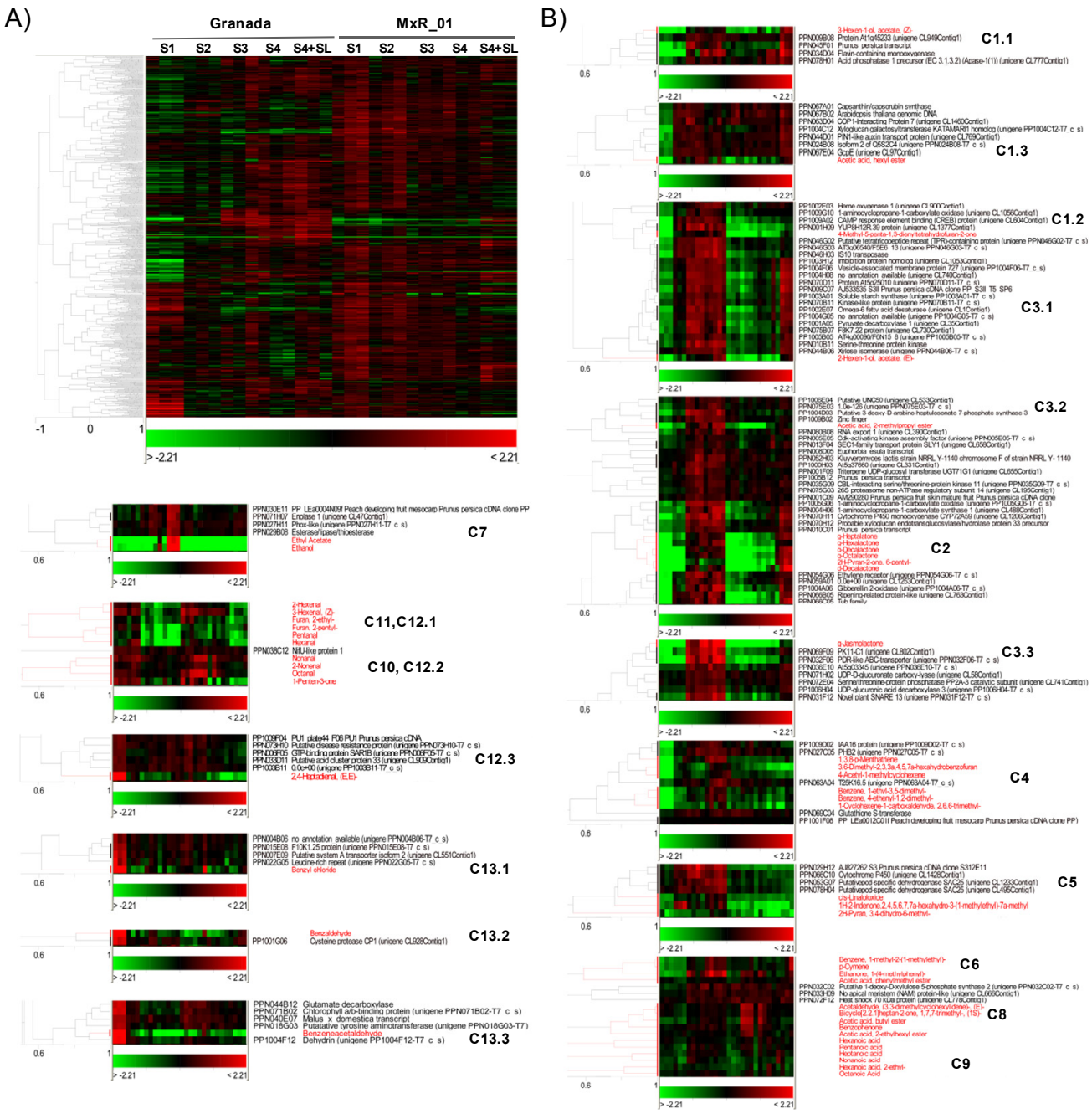



**Figure S4. Hierarchical cluster analysis for identifying the genes correlating with the 52 VOCs. A)** The heatmap and cluster analyses of the gene-volatile data set (4348 genes and 52 volatiles). Three replicates per stage are shown. Data are expressed as the log2 of a ratio (sample/common reference). **B)** Details of the HCA where volatile compounds are present with genes. Volatiles are indicated with a red letter. For each gene, the id and unigene annotation and identifier are provided (in parentheses) when available. Three replicates per stage are shown. Sub-clusters are named according to the volatile members that they have. For example, the volatiles of C1, according to Figure 1, appear in three sub-clusters named C1.1, C1.2, and C1.3. (para mejor resolución ver: http://www.biomedcentral.com/content/supplementary/1471-2164-14-343-s6.xlsx).

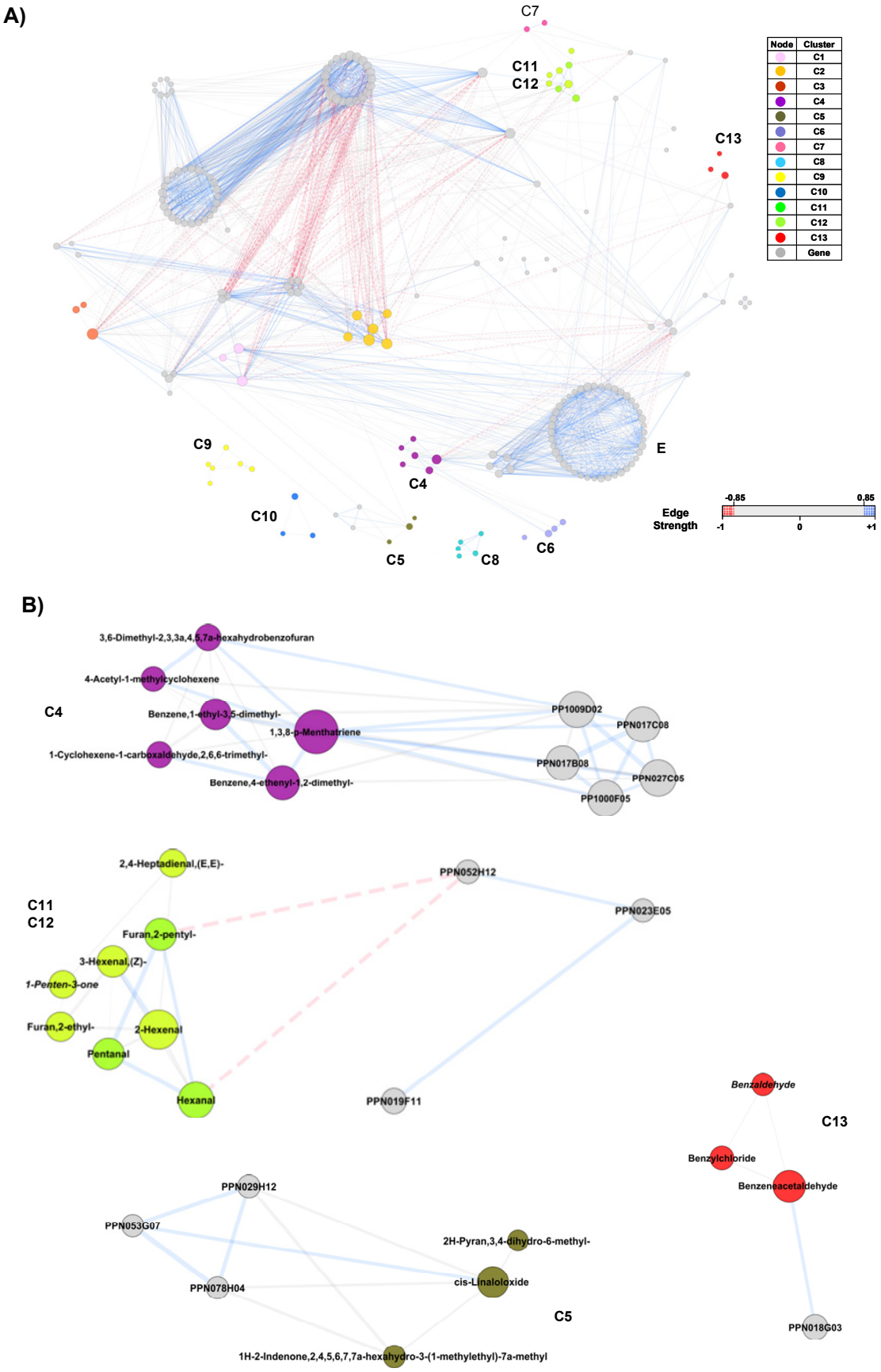



**Fig. S5 (see the legends below).**

**Figure S5. Correlation network of VOCs with genes correlating with the compounds of clusters C4 to C13. A)** Network of VOCs and genes. The nodes representing volatiles are colored according to the cluster that they belong to (according to Fig. 1). Genes are represented as gray nodes. Edges are colored according to their strength. The edge codification is indicated to the right of the network. Node size indicates its connectivity. The bigger the node, the higher the connectivity measured with the node degree (i.e., the number of edges connecting the node). A sub-cluster of genes is indicated with E. **B)** Magnification of volatile groups (C4, C11, C12, C5, and C13) showing the interactions with genes in detail. The genes annotated as "0.00E+00" indicate that either no homologue was found or the homologue found has unknown function. For a detailed description of ChillPeach unigene functional annotation see Ogundiwin et al. [5].

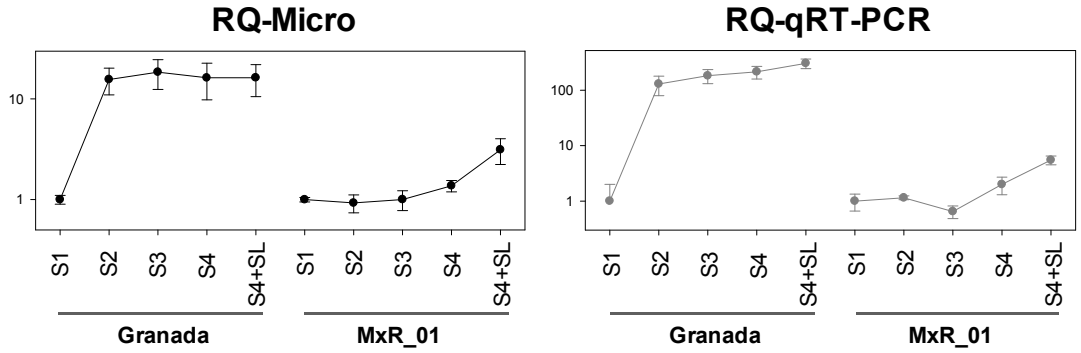



**Figure S6. Profile of the candidate gene (PPN001H09) expression assayed by microarray (left) and qRT-PCR (right) analysis.** For both analyses, the Relative Quantitation (RQ) in arbitrary units is shown. The y-axis is on the log10 scale.

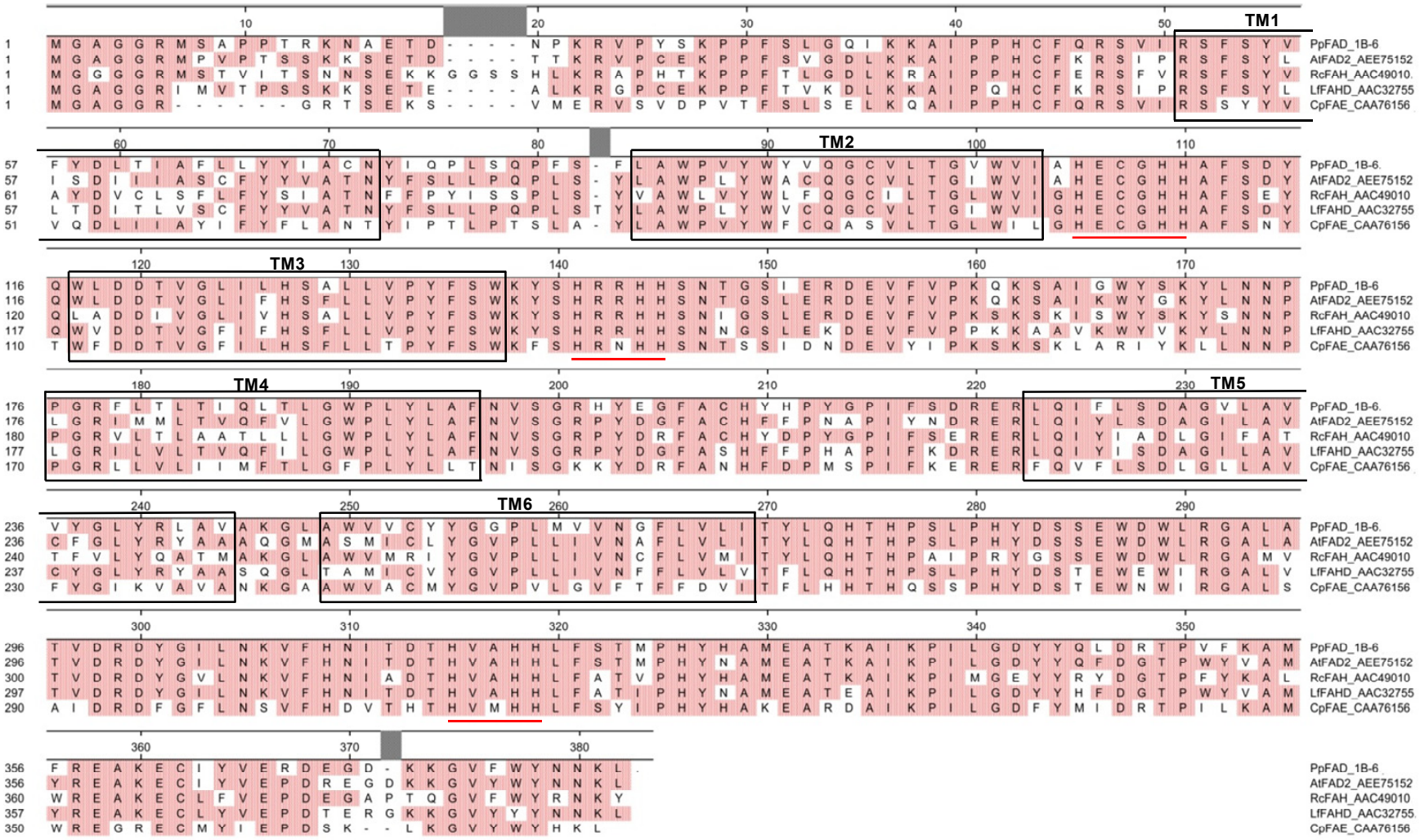



**Fig. S7 (see the legends below).**

**Figure S7. Alignment of PpFAD 1B_6 with FAD-type enzymes.** Conserved amino acids are shaded in red. Open boxes indicate the predicted transmembrane domains (TM). The three His motifs (one HXXXHH and two HXXHH) are underlined. Sequence alignment was performed with MegAlign (DNAStar). PpFAD 1B-6, *Prunus persica* Fatty Acid Desaturase allele 1B-6; AtFAD2, *Arabidopsis thaliana*, Fatty Acid Desaturase type 2; RcFAH, *Ricinus communis* Fatty Acid Hydroxylase; LfFAHD, *Lesquerella fendleri* Fatty Acid Hydroxylase/Desaturase; CpFAE, *Crepis palaestina* Fatty Acid Epoxygenase. For AtFAD2, RcFAH, LfFAHD, and CpFAE, the NCBI accession numbers are provided following each name.

| id | Unigene id | Unigene functional annotation | Match | Forward primer (5' - 3') | Reverse primer (5' - 3') | Amplicon lenght (bp) | Validation (PCC) |
|---|---|---|---|---|---|---|---|
| PPN078E12 | PPN078E12-T7_c_s | Esterase/lipase/thioesterase; Phospholipid/glycerol acyltransferase | **ppa003154m** | ATGCCTCTCGTCTTTGGCTG | TCTTTGCCACTGGCCAACA | 141 | Norm. |
| PP1002E07 | CL1Contig1 | Omega-6 fatty acid desaturase | **ppa007098m** | CCTGACTGGTGTTTGGGTGAT | CGGCGATGGCTATACTTCCA | 141 | 0.710 |
| PPN070H11 | CL1206Contig1 | Cytochrome P450 monooxygenase CYP72A59 | **ppa006310m** | GCCGCACGAAGTGTATACATTC | CTTCATCCTCTTGTTATGCTTGGTT | 66 | 0.789 |
| PPN001H09 | CL1377Contig1 | YUP8H12R.39 protein | **ppa005452m** | GTCCAGGCAACTGAGCTCAAG | GAGCCATCTCAGCCCATGA | 115 | 0.921 |
| PPN032F06 | PPN032F06-T7_c_s | PDR-like ABC-transporter | **ppa000267m** | ACTAATCATAACAAGAAAAGAGGAATGGT | CTTCATTGACTGTGGCATATCCA | 102 | 0.927 |
| PPN002B03 | PPN002B03-T7_c_s | Malus_x_domestica transcript; similar to F17O7.2 [Arabidopsis thaliana (Mouse-ear cress)] | **ppa007712m** | CTGCGAAAGTGTCGTAGATTCG | CGACCAACCCGTCATTGAAC | 68 | 0.863 |
| PPN059A01 | CL1253Contig1 | 0.00E+00 | **ppa013582m** | ACGATCTTGCAAATCAATAAAACCT | TCAGTTGGGTGCTGTGTGTGA | 81 | 0.923 |
| PPN066B05 | CL763Contig1 | Ripening-related protein-like | **ppa011478m** | CCCCAACCTCAAATTCACACA | GCATGTTTGGCATGAAGATCA | 75 | 0.998 |
| PPN079G06 | CL283Contig1 | Asparagine synthetase | **ppa003265m** | GCGTCTGCCGAATCACAAGT | AAGCTGTTGTCGCGGGTATC | 141 | 0.921 |
| PP1002D12 | CL29Contig1 | Leucine zipper, homeobox-associated; Homeodomain-related | **ppa010647m** | GAAGGCACACACTCCCAAAAC | GCATAAGTACTAATTGGGTCTGCAAA | 75 | 0.976 |
| PPN008D02 | CL985Contig1 | Flavonoid 3'-hydroxylase | **ppa004433m** | GCGTCCCAGTTCCTGAAGAC | AGATGCTTGGCGCCAGAGT | 71 | 0.711 |

**Table S1. Genes analyzed by qRT-PCR analysis.** For each gene, the microarray id (id), the identifier of the ChillPeach database (unigene id), the functional annotation, and the most similar gene from the peach genome sequence (Match) are shown. The forward and reverse primers and amplicon length are indicated for each gene tested. Validation indicates the Pearson correlation coefficient (PCC) between the microarray and qRT-PCR data for each gene. The gene profile was considered validated if PCC was higher than 0.7. Norm: indicates the gene used as normalizator in the qRT-PCR analysis. The gene annotated as "0.00E+00" indicate that either no homologue was found or the homologue found has unknown function. For a detailed description of ChillPeach unigene functional annotation see Ogundiwin et al. [5].

| VOCs with higher content in Granada | Fold | p value |
|---|---|---|
| Ethyl Acetate | 38.4 | <0.01 |
| Ethanol | 35.5 | <0.01 |
| γ-Jasmolactone | 11.1 | <0.01 |
| 3,6-Dimethyl-2,3,3a,4,5,7a-hexahydrobenzofuran | 5.2 | <0.01 |
| 1H-2-Indenone,2,4,5,6,7,7a-hexahydro-3-(1-methylethyl)-7a-methyl | 4.4 | 0.01 |
| 2H-Pyran, 3,4-dihydro-6-methyl- | 3.1 | 0.04 |
| 2,4-Heptadienal, (E,E)- | 2.8 | 0.03 |
| 1,3,8-p-Menthatriene | 2.7 | 0.01 |
| Ethanone, 1-(4-methylphenyl)- | 2.7 | 0.03 |
| cis-Linaloloxide | 2.6 | <0.01 |
| Acetic acid, 2-methylpropyl ester | 2.5 | <0.01 |
| 4-Acetyl-1-methylcyclohexene | 2.4 | <0.01 |
| 1-Penten-3-one | 2.3 | 0.01 |
| γ-Hexalactone | 2.1 | <0.01 |
| Benzene, 1-ethyl-3,5-dimethyl- | 2.0 | 0.04 |
| 2-Hexen-1-ol, acetate, (E)- | 1.7 | 0.02 |
| Acetic acid, butyl ester | 1.5 | <0.01 |
| 2-Hexenal | 1.4 | 0.02 |
| Bicyclo[2.2.1]heptan-2-one, 1,7,7-trimethyl-, (1S)- | 1.2 | <0.01 |
| **VOCs with higher content in MxR_01** | **Fold** | **p value** |
| Hexanal | 5.4 | <0.01 |
| Furan, 2-pentyl- | 5.2 | 0.01 |
| 2H-Pyran-2-one, 6-pentyl- | 3.5 | <0.01 |
| δ-Decalactone | 2.8 | 0.01 |
| 3-Hexen-1-ol, acetate, (Z)- | 1.8 | 0.01 |
| γ-Decalactone | 1.7 | 0.01 |
| γ-Octalactone | 1.5 | 0.01 |
| γ-Heptalatone | 1.4 | 0.02 |

**Table S2. Comparison of volatile content in 'Granada' and 'MxR_01' after shelf-life simulation (S4+SL).** The fold change between genotypes is shown. The ANOVA p value for each volatile compound is indicated.

| id | Unigene annotation | Fold Change Granada | Fold Change MxR_01 | Match | iana_protein | GO Molecular Function | Arabidopsis_thaliana_proteins most similar description | Arabidopsis |
|---|---|---|---|---|---|---|---|---|
| | **Up-regulated** | | | | | | | |
| PPN044F04 | Calcium-binding EF hand family protein (Gb\|AAC24386.1) | **2.2** | **1.5** | ppa005459m | AT5G54130 | GO:0005488, binding | calcium-binding EF hand family protein, contains INTERPRO:IPR002048 calcium-b | 4E-26 |
| PPN067B11 | Aspartate aminotransferase, chloroplast precursor | **1.7** | **1.2** | ppa005315m | AT5G11520 | GO:0003824, catalytic activity; GO:0016740, transferase activity | aspartate aminotransferase, chloroplast / transaminase A (ASP3) (YLS4), identical | 4E-78 |
| PPN064C01 | EIF4-gamma/eIF5/eIF2-epsilon | **1.7** | **1.4** | ppa005197m | AT5G36230 | GO:0003676, nucleic acid binding; GO:0045182, translation regulator activity | eIF4-gamma/eIF5/eIF2-epsilon domain-containing protein, low similarity to SP:Q13 | 5E-71 |
| | **Down-Regulated** | | | | | | | |
| PPN041F05 | Acid phosphatase class B family protein (F21M11.2 protein) | **1.8** | **3.2** | ppa010275m | AT1G04040 | GO:0016787, hydrolase activity | acid phosphatase class B family protein, similar to SP:P15490 STEM 28 kDa glyc | 1E-53 |

**Table S3. Genes showing the same trends after the shelf-life condition in both genotypes.** The fold change (up-regulated: S4+SL/S4; down-regulated: S4/S4+SL) for the 'Granada' and 'MxR_01' genotypes is indicated. For each gene, the microarray id (id), the unigene id and annotation and the most similar gene predicted in the peach genome (Match) are shown. The most similar *Arabidopsis* gene with the GO molecular function, the description, and the E-value are also provided.

**Table S4. Genes affected by the shelf-life response with a 2-fold cut-off.** The fold change for up-regulated (S4+SL/S4) and down-regulated (S4/S4+SL) genes in both genotypes is shown. The number of genes for each list are indicated in parentheses. For each gene, the microarray id (id), the unigene id and annotation and the most similar gene predicted in the peach genome (Match) are shown. The most similar *Arabidopsis* gene with the GO molecular function, the description and the E-value are also provided. n.f.: not found. (Información incompatible con la presentación en formato físico, disponible en: http://www.biomedcentral.com/content/supplementary/1471-2164-14-343-s7.xls)

**Table S5. Correlation matrix of the complete dataset.** The Pearson correlation coefficient is shown. (Información incompatible con la presentación en formato físico, disponible en: http://www.biomedcentral.com/content/supplementary/1471-2164-14-343-s8.xlsx)

**Table S6. Extended description for the candidate genes identified.** Besides the information provided in Table 1, in this table, the ChillPeach database identifier (Unigene ID) and the localization of the gene in the peach genome is shown. The most similar *Arabidopsis* gene with its description and E-value are also provided. n.f.: not found. The genes annotated as "0.00E+00" indicate that either no homologue was found or the homologue found has unknown function. For a detailed description of ChillPeach unigene functional annotation see Ogundiwin et al. [5]. (Información incompatible con la presentación en formato físico, disponible en: http://www.biomedcentral.com/content/supplementary/1471-2164-14-343-s10.xlsx

**Table S7. The ten genes that best correlated with γ-decalactone (upper) and γ-jasmodecalactone (bottom).** For each gene, the microarray id (id), the unigene id and annotation, the Pearson Correlation coefficient (PC), and the most similar gene predicted in the peach genome are shown (Match). The most similar *Arabidopsis* gene with the GO molecular function, the description, and the E-value are also provided. n.f.: not found. The gene annotated as "0.00E+00" indicate that either no homologue was found or the homologue found has unknown function. For a detailed description of ChillPeach unigene functional annotation see Ogundiwin et al. [5]. (Información incompatible con la presentación en formato físico, disponible en: http://www.biomedcentral.com/content/supplementary/1471-2164-14-343-s8.xlsx)

| id | Unigene ID | Unigene annotation | Degree | Average Shortest Path Length | Betweenness Centrality | Closeness Centrality | Clustering Coefficient | Neighborhood Connectivity | Radiality | Stress |
|---|---|---|---|---|---|---|---|---|---|---|
| PP1003H12 | CL1053Contig1 | Imbibition protein homolog | 31 | 2.211 | 9.30E-03 | 0.452 | 0.436 | 33.81 | 0.798 | 19676 |
| PP1002E07 | CL1Contig1 | Omega-6 fatty acid desaturase | 29 | 2.313 | 4.00E-03 | 0.432 | 0.591 | 34.07 | 0.781 | 7732 |
| PPN009C07 | CL761Contig1 | AJ533535 S3II Prunus persica cDNA clone PP_S3II_T5_SP6, mRNA sequence | 27 | 2.453 | 1.69E-03 | 0.408 | 0.655 | 32.63 | 0.758 | 4212 |
| PPN070D11 | PPN070D11-T7_c_s | Protein At5g25010 | 27 | 2.260 | 7.85E-03 | 0.442 | 0.457 | 34.20 | 0.790 | 17286 |
| PP1004F06 | PP1004F06-T7_c_s | Vesicle-associated membrane protein 727 | 26 | 2.400 | 1.76E-03 | 0.417 | 0.643 | 34.19 | 0.767 | 3886 |
| PPN054G06 | PPN054G06-T7_c_s | Ethylene receptor | 25 | 2.358 | 4.07E-03 | 0.424 | 0.523 | 32.80 | 0.774 | 8324 |
| PP1004H08 | CL740Contig1 | no_annotation_available | 24 | 2.272 | 6.18E-03 | 0.440 | 0.491 | 34.95 | 0.788 | 14062 |
| PPN059A01 | CL1253Contig1 | 0.00E+00 | 23 | 2.411 | 9.89E-03 | 0.415 | 0.458 | 31.00 | 0.765 | 18578 |
| PP1001A05 | CL35Contig1 | Pyruvate decarboxylase 1 | 22 | 2.664 | 9.66E-04 | 0.375 | 0.706 | 31.18 | 0.723 | 2590 |
| PP1003A01 | PP1003A01-T7_c_s | Soluble starch synthase | 19 | 2.551 | 4.26E-04 | 0.392 | 0.842 | 33.32 | 0.742 | 1408 |
| PPN070B11 | PPN070B11-T7_c_s | Kinase-like protein | 16 | 2.728 | 9.36E-05 | 0.367 | 0.900 | 30.88 | 0.712 | 348 |
| PP1001C04 | PP1001C04-T7_c_s | Ring zinc finger protein | 15 | 2.853 | 1.81E-03 | 0.351 | 0.648 | 26.53 | 0.691 | 2672 |
| PPN032A08 | PPN032A08-T7_c_s | F19G10.11 protein | 9 | 2.894 | 2.37E-05 | 0.346 | 0.917 | 30.56 | 0.684 | 60 |

**Table S8. Sub-cluster of genes obtained by CNA.** For each gene, the microarray id (id), the unigene id and annotation, and a set of topological network parameters of its node are shown. The gene annotated as "0.00E+00" indicate that either no homologue was found or the homologue found has unknown function. For a detailed description of ChillPeach unigene functional annotation see Ogundiwin et al. [5].

| | | Fatty Acid content (% of Total) | | | |
|---|---|---|---|---|---|
| Construction | Induction | C16:0 | C16:1 | C18:0 | C18:1-OH |
| PpFAD-1B_6 | + | 17.28a | 31.22a | 7.72a | n.d. |
| | - | 19.71a | 30.89a | 8.64a | n.d. |
| Control LacZ | + | 19.86a | 32.86a | 7.86a | n.d. |
| | - | 20.76a | 31.33a | 8.42a | n.d. |

**Table S9. Composition of fatty acid in yeast expressing PpFAD_1B-6.** The % of total fatty acid content is shown for palmitic acid (16:0), palmitoleic acid (16:1), estearic acid (18:0), and ricinoleic acid (18:1-OH). No significant differences in fatty acid levels between treatments were detected by ANOVA ($p<0.05$), indicated by the same letter. The inducing and non-inducing conditions are indicated by + and -, respectively. The mean of three determinations (n=3) are shown. n.d.: not detected.

## Información suplementaria del Capitulo 4

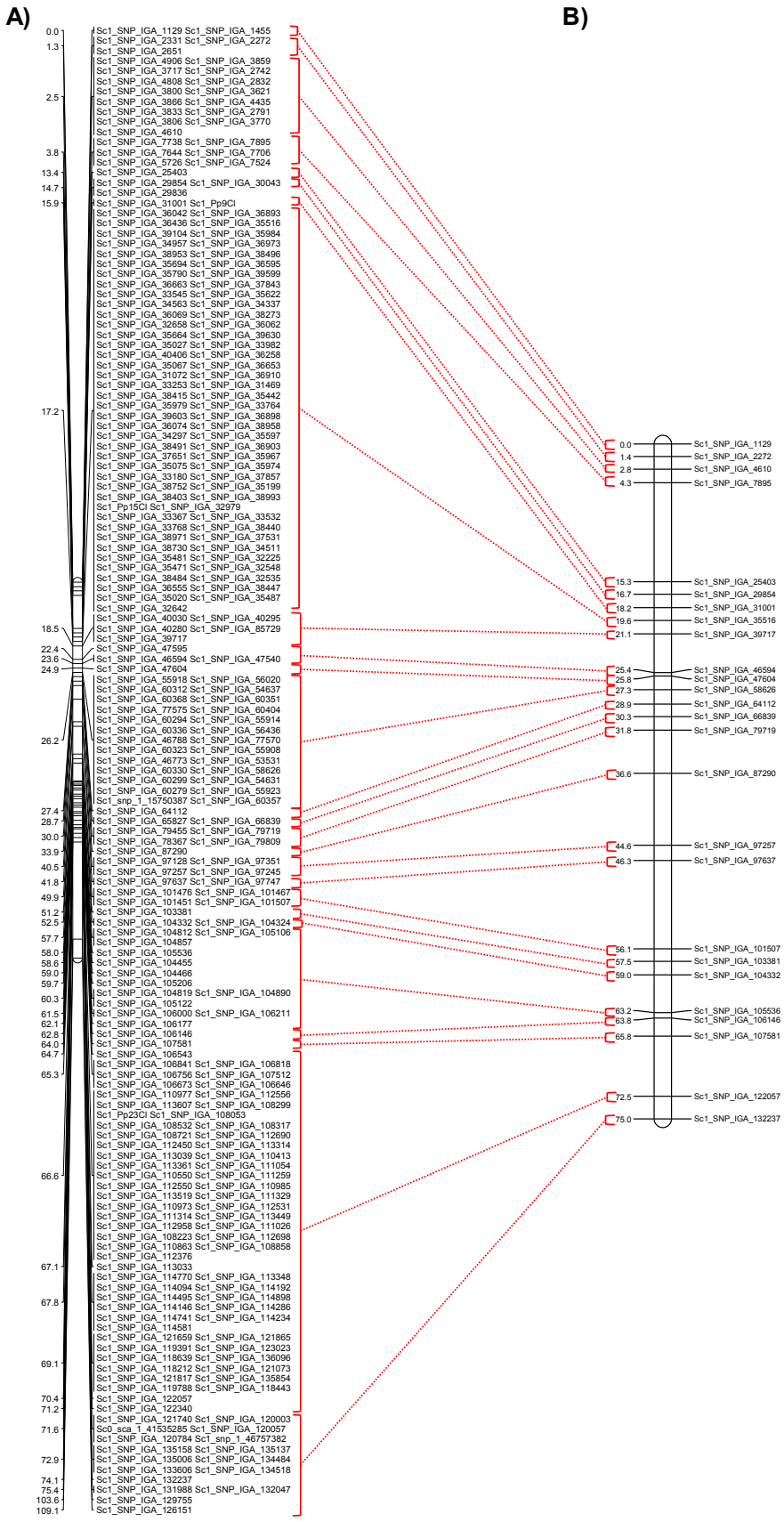



**Figure S1. SNPs selected for Sc1 of 'MxR_01'. A)** Linkage group obtained with all the polymorphic SNPs mapped to scaffold 1 for 'MxR_01' (265 markers). **B)** The map obtained after selecting informative SNPs for each locus (26 markers).

For each map, the SNP positions in cM are given at the left of each one. SNP name indicate with the first 3 characters the scaffold that the marker was mapped to (e.g. Sc1 indicate Scaffold 1). The relative position in the genome of each SNP is indicated with the last number (e.g. 1129 for Sc1_SNP_IGA_1129). The exact genome position could be found searching at the genome browser (http://www.rosaceae.org/gb/gbrowse/prunus_persica/) with the last part of the SNP name (e.g. SNP_IGA_1129 for Sc1_SNP_IGA_1129).

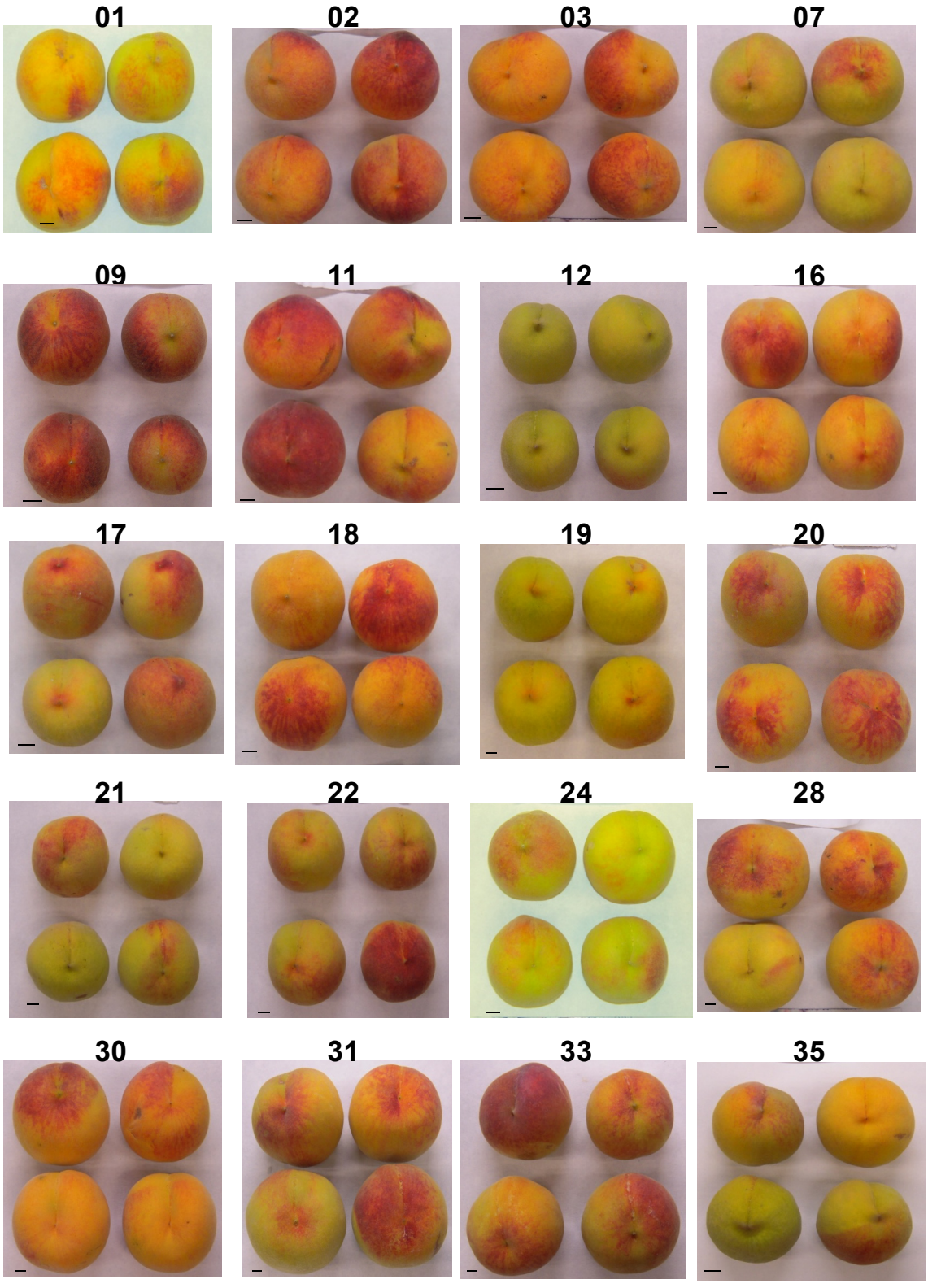



**Fig. S2 (see the legends below).**

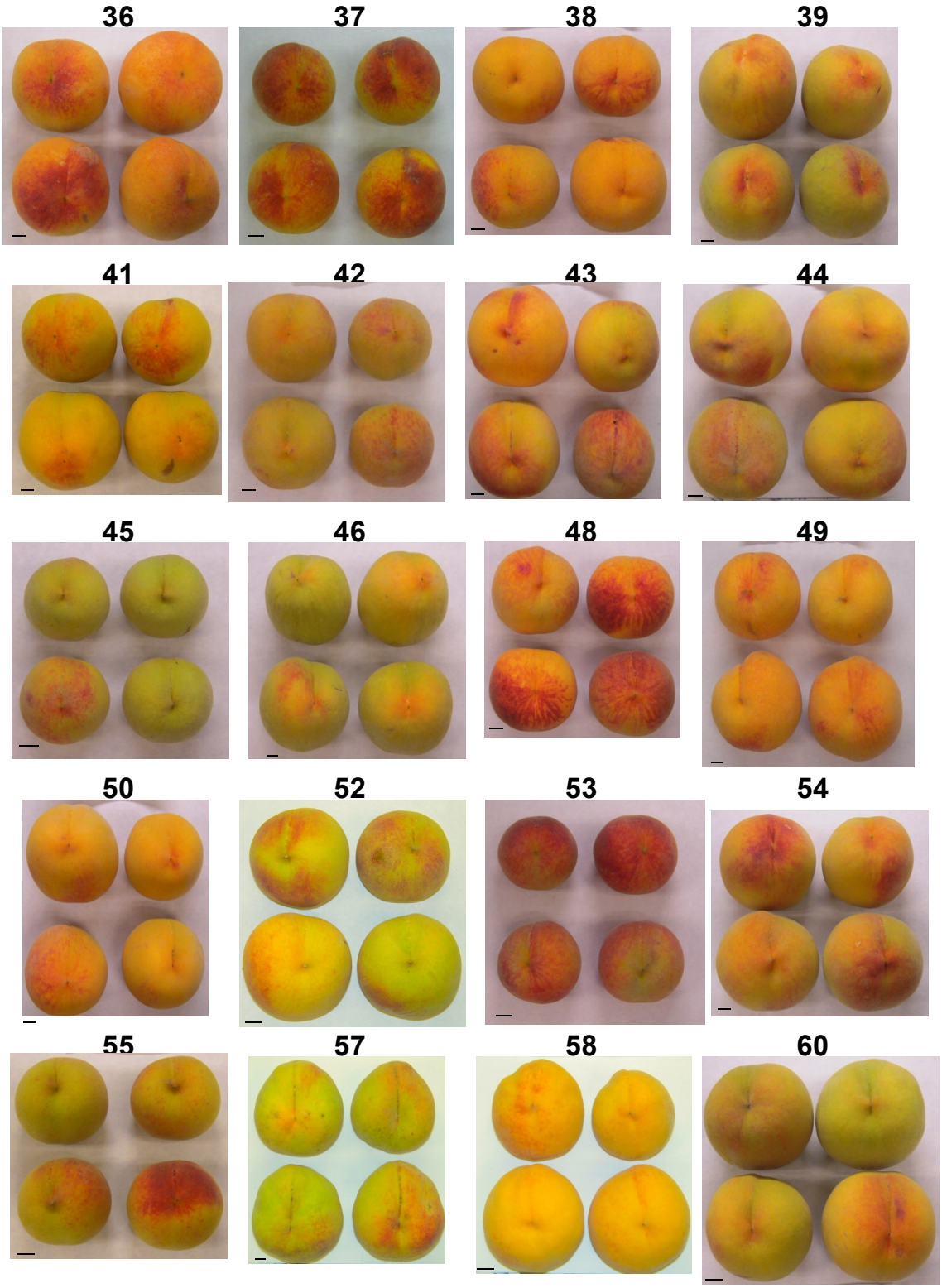



**Fig. S2-continued**

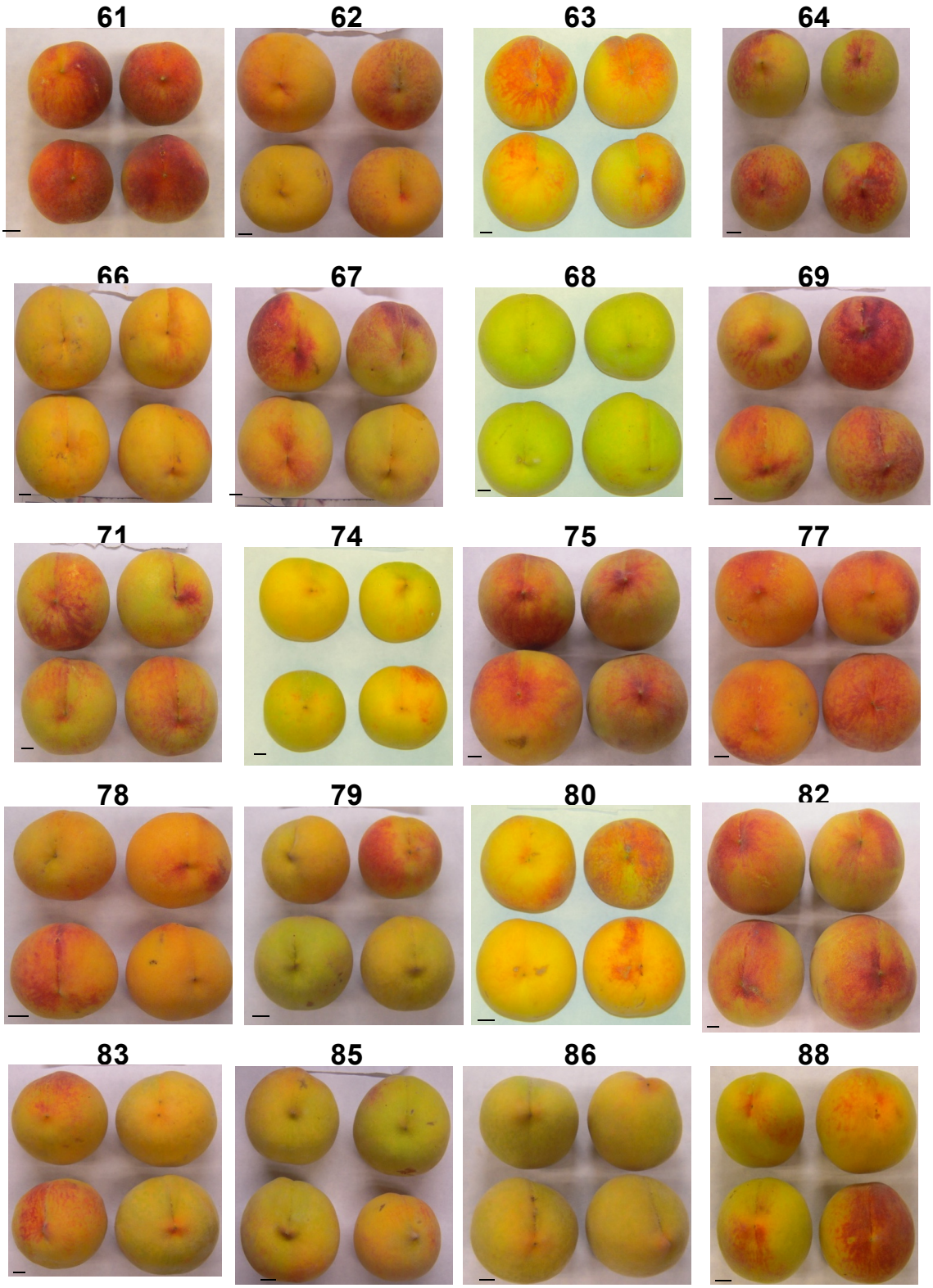



**Fig. S2-continued**

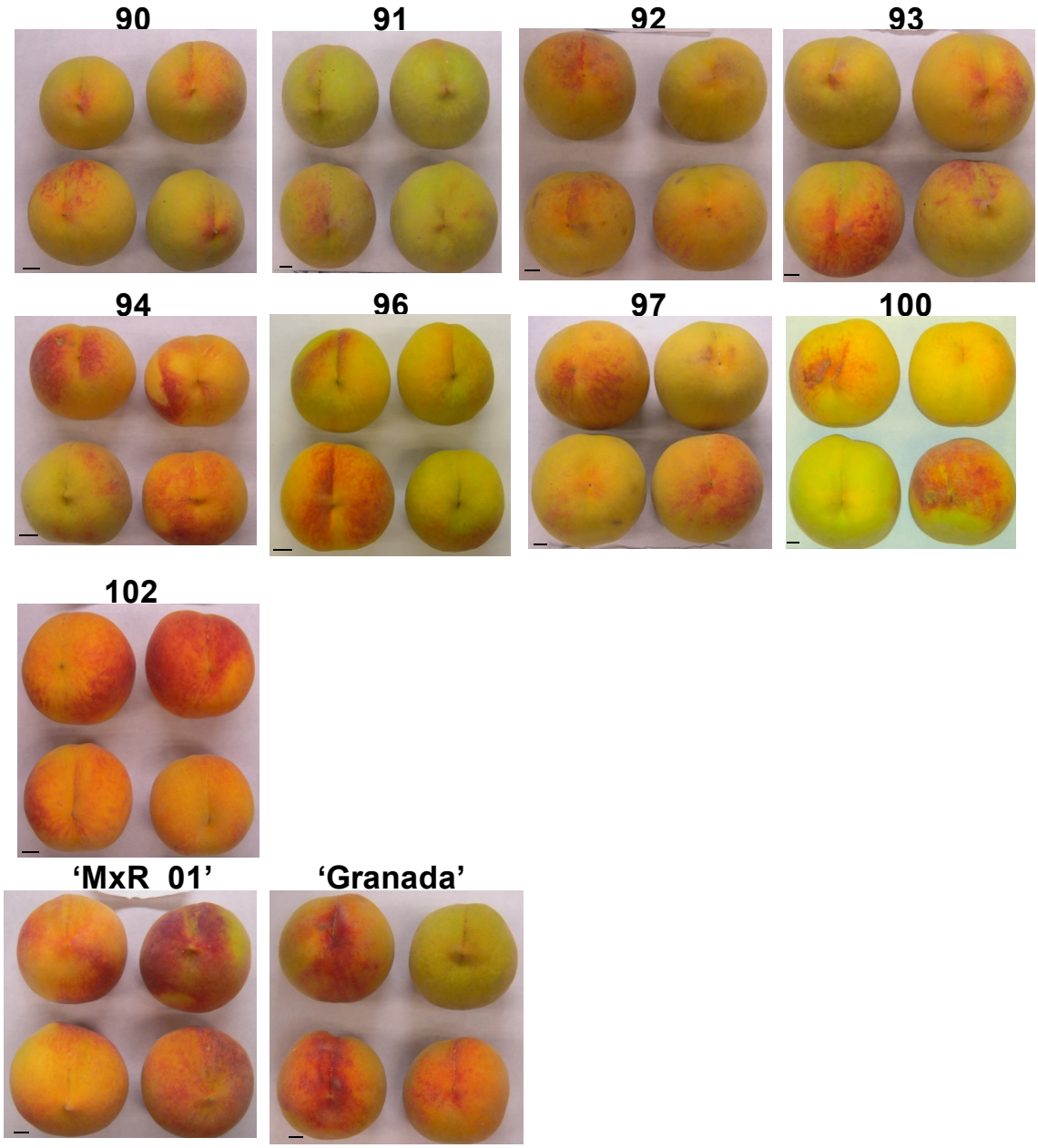



**Figure S2. Fruit variability within the population mapping from "El Jimeno" trial .** Four representative fruits for each F1 hybrid and parental genotypes are shown. Upon each photo the number (for hybrid) or name (for parental) of the genotype is indicated. The bar at the left bottom corner, indicate a 1 cm scale.

‘MxR_01‘

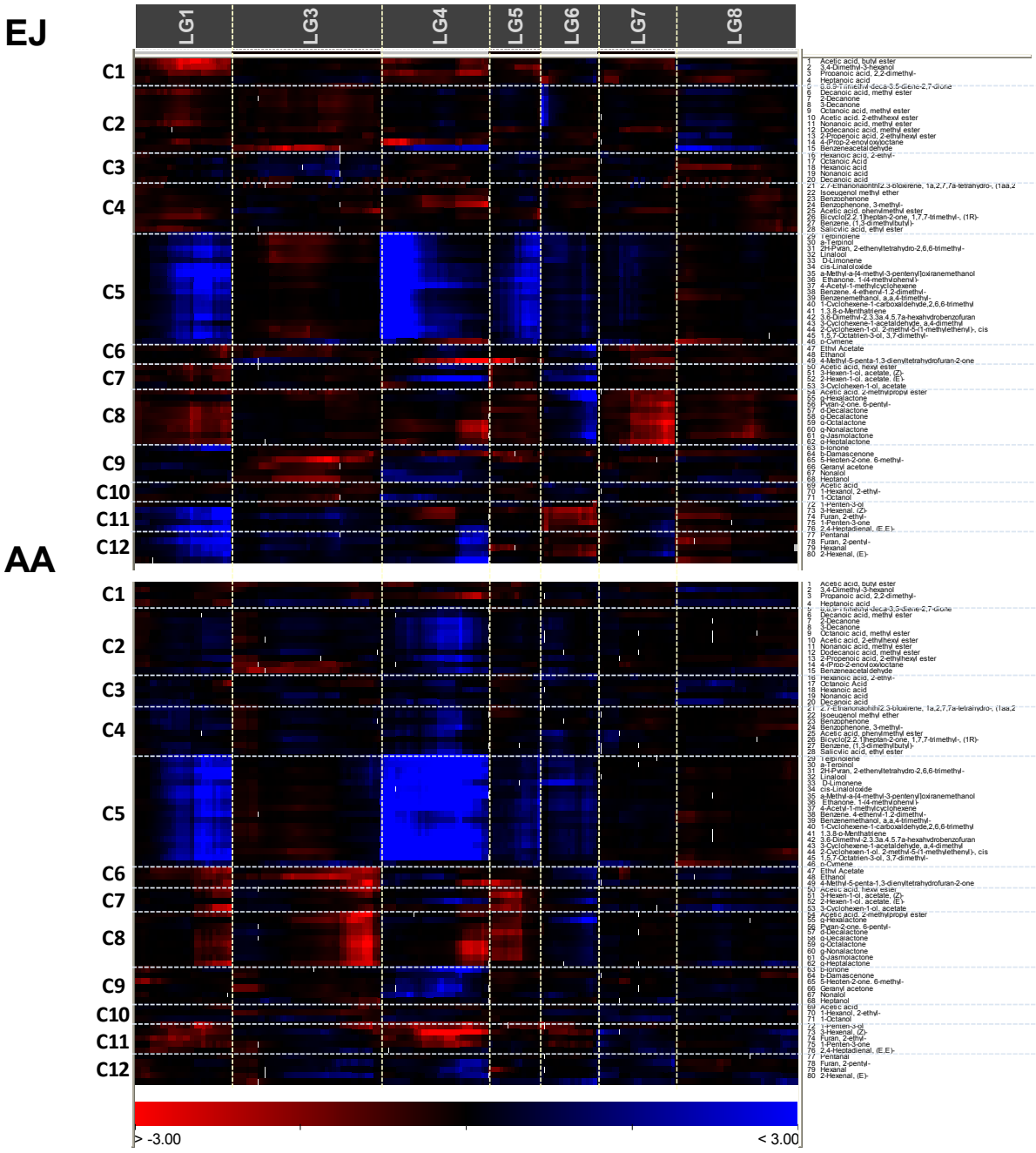



**Figure S3. Heatmap of likelihood for ‘MxR_01’ at EJ (top) and AA (bottom) locations.** The LOD score (computed by single correlation analysis) for each marker/volatile pair is presented in different color according to their additive effects (a), red for negatives a, and blue for positive a. The color intensity is according the LOD value, the higher the intensity the higher the LOD score. For each linkage group (LG1, LG2-LG8) the markers are ordered from left to right according the position at the peach genome. The volatiles are ordered according the position on the HCA of Fig. 2. C1-C12 indicates the volatile clusters. Vertical and horizontal lines divide the linkage groups and the volatile clusters,

respectively. EJ and AA indicate the location "El Jimeneo" and "Aguas Amargas"

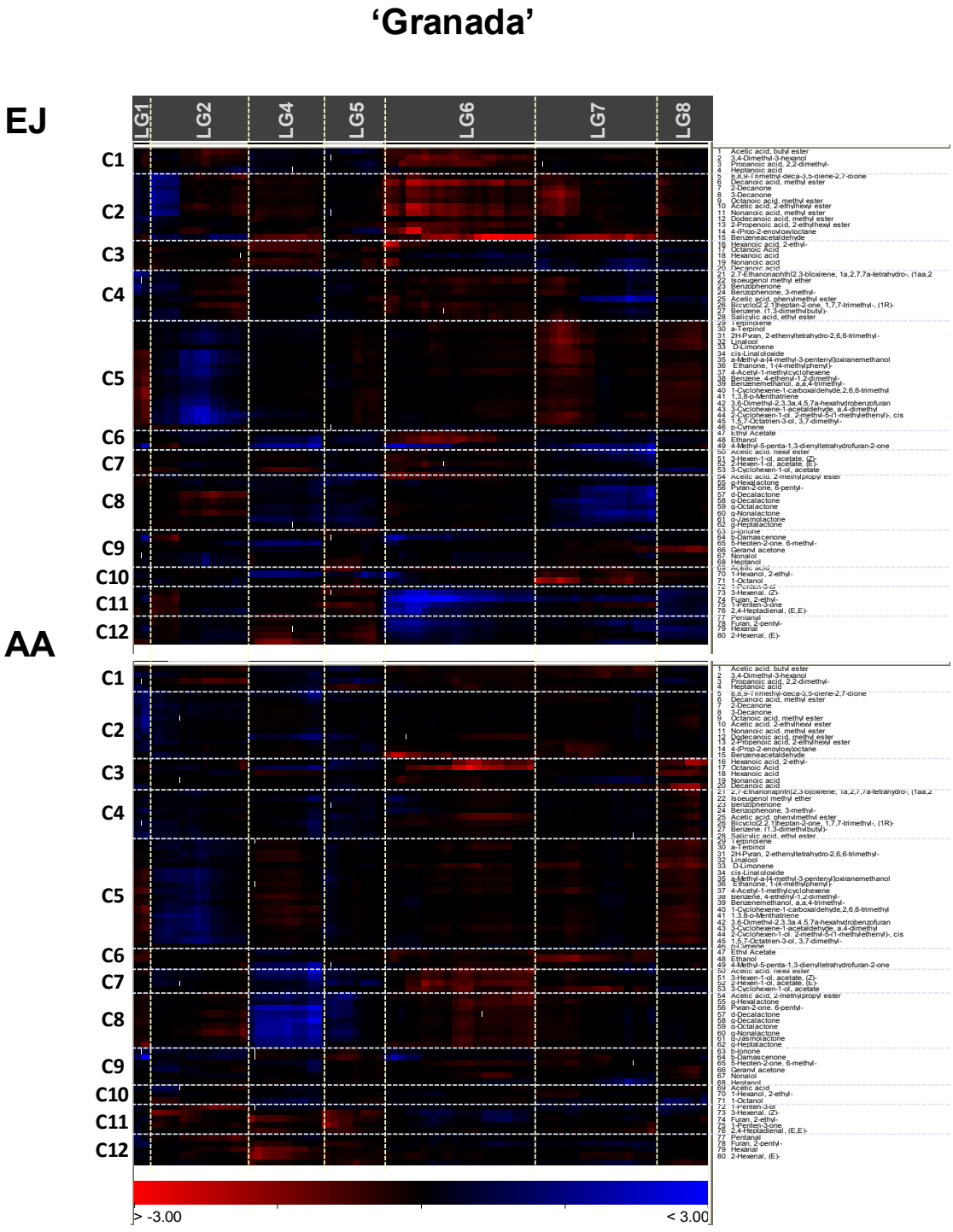



**Figure S4. Heatmap of likelihood for 'Granada' at EJ (top) and AA (bottom) locations.** The LOD score (computed by single correlation analysis) for each marker/volatile pair is presented in different color according to their additive effects (a), red for negatives a, and blue for positive a. The color intensity is according the LOD value, the higher the intensity the higher the LOD score. For

each linkage group (LG1-LG2, LG4-LG8) the markers are ordered from left to right according the position at the peach genome. The volatiles are ordered according the position on the HCA of Fig. 2. C1-C12 indicates the volatile clusters. Vertical and horizontal lines divide the linkage groups and the volatile clusters, respectively. EJ and AA indicate the location "El Jimeneo" and "Aguas Amargas"

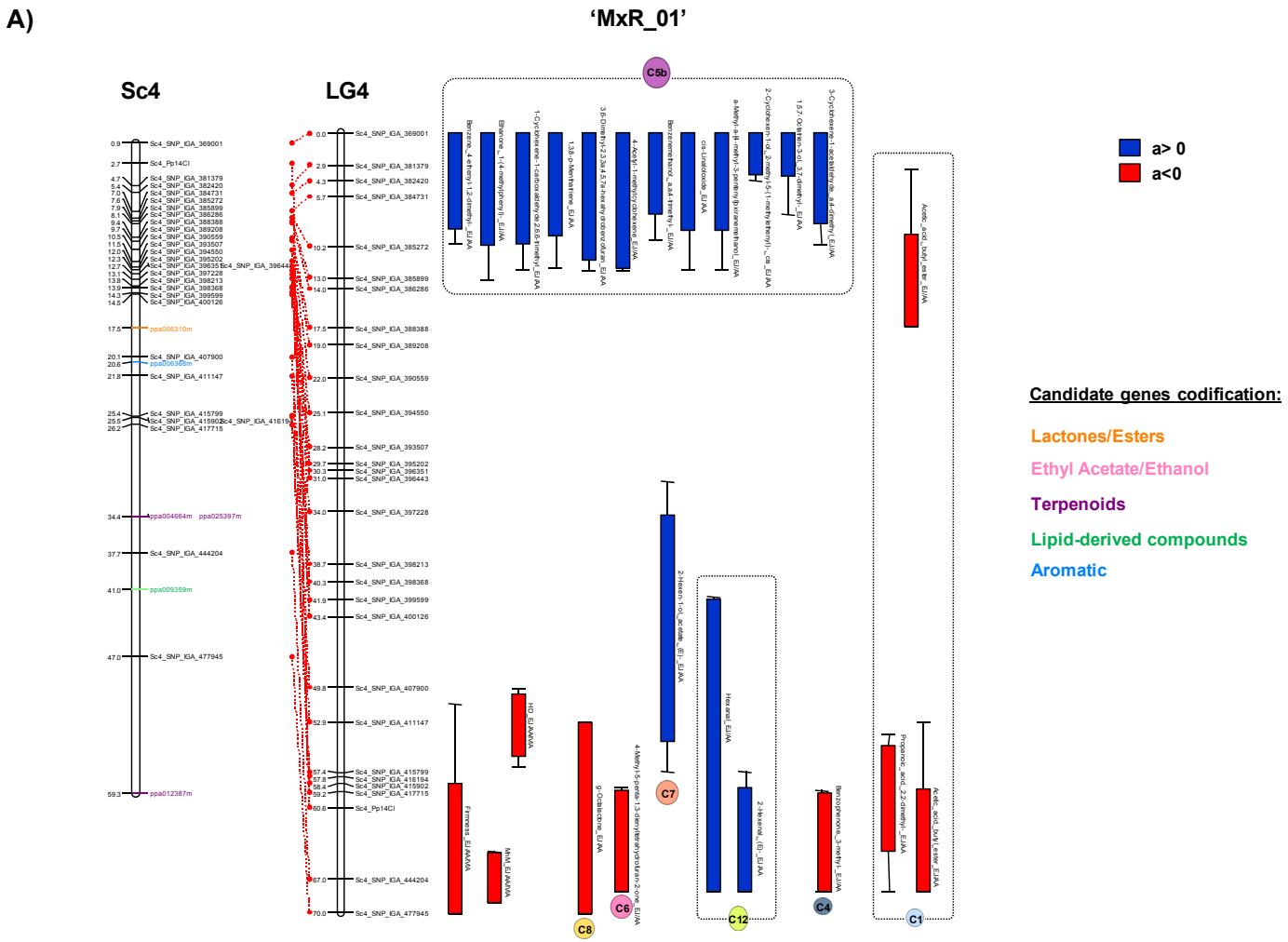



**Fig. S5 (see the legends below).**

B) ‘MxR_01’

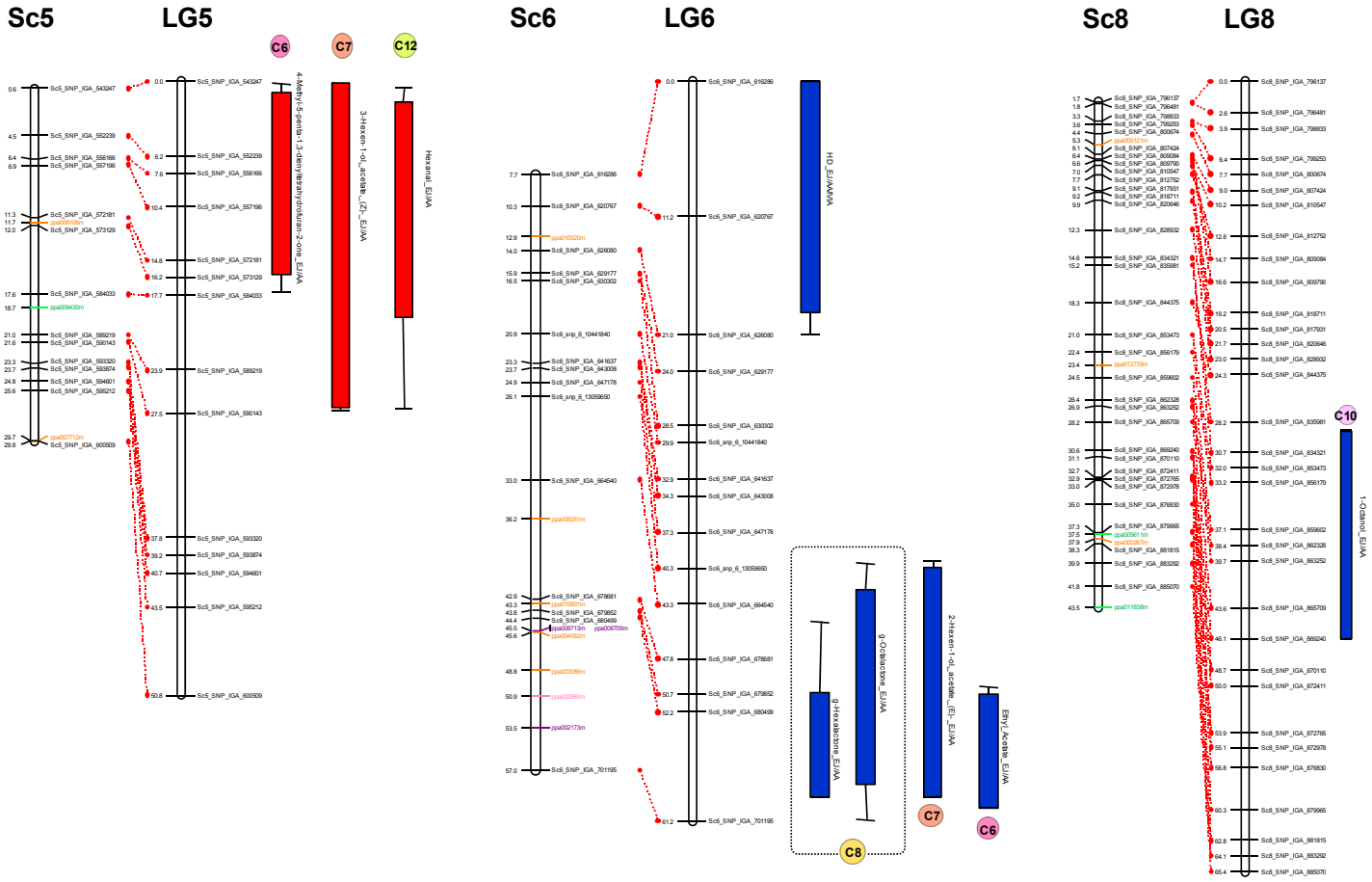



**Fig. S5-continued**

C) ‘Granada’

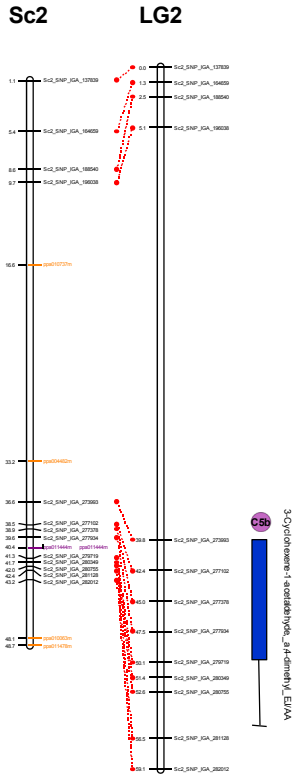



**Fig. S5 (see the legends below)**

**Figure S5. Co-localization of volatile QTL with candidate genes identified previously.** The physical (left) and linkage (right) maps were volatile QTL were indentified are shown. The QTL are colored according the additive effect (a) that exerts. Bars and lines represent 1-LOD and 2-LOD support intervals. The candidate genes previously associated to different volatile groups (Sánchez et al., 2013) are indicated with different color. The position of SNPs and candidate genes in the scaffolds of the peach genome v1 is indicated at the left of the map in arbitrary units (map position in base pare/ $5x10^5$). SNPs positions in linkage map are indicated at the right of the map in cM. A) QTL for LG4 of 'MxR' and the corresponding scaffold are shown. B) QTL for LG5 and LG6 of 'MxR' and the corresponding scaffolds are shown. C) QTL for LG2 of 'Granada' and the corresponding scaffold are shown.

| Nº | Name | Cluster | Family | Location | LG | Position | LOD | Additive | R2 | LOD2-left | LOD2-right |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 36 | Ethanone, 1-(4-methylphenyl)- | C5b | un. | EJ | LG1 | 53.3 | 3.83 | 0.52 | 0.12 | 34.3 | 55.9 |
| 41 | 1,3,8-p-Menthatriene | C5b | Terpenoid | AA | LG1 | 63.7 | 3.00 | 1.88 | 0.10 | 63.0 | 73.5 |
| 45 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5b | un. | EJ | LG1 | 49.3 | 3.86 | 1.85 | 0.13 | 32.6 | 55.7 |
| 45 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5b | un. | AA | LG1 | 63.7 | 3.45 | 2.09 | 0.10 | 57.5 | 73.5 |
| 70 | 1-Hexanol, 2-ethyl- | C10 | Alcohol | AA | LG1 | 56.0 | 11.47 | 0.24 | 0.07 | 43.2 | 71.0 |
| 75 | 1-Penten-3-one | C11 | Lipid-derived | EJ | LG1 | 73.5 | 3.70 | 0.91 | 0.17 | 59.9 | 74.5 |
| 44 | 2-Cyclohexen-1-ol, 2-methyl-5-(1-methylethenyl)-, cis | C5b | un. | EJ | LG1 | 45.6 | 2.70 | 1.33 | 0.08 | 36.6 | 55.3 |
| 52 | 2-Hexen-1-ol, acetate, (E)- | C7 | Ester | EJ | LG1 | 36.6 | 2.75 | -1.18 | 0.10 | 28.9 | 45.6 |
| 80 | 2-Hexenal, (E)- | C12 | Lipid-derived | EJ | LG1 | 63.1 | 3.60 | 0.29 | 0.14 | 59.4 | 65.8 |
| 80 | 2-Hexenal, (E)- | C12 | Lipid-derived | EJ | LG1 | 73.5 | 3.85 | 0.32 | 0.15 | 66.9 | 74.5 |
| 31 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | C5a | Terpenoid | EJ | LG1 | 36.6 | 2.96 | 1.52 | 0.08 | 25.7 | 55.8 |
| 31 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | C5a | Terpenoid | AA | LG1 | 67.8 | 5.44 | 1.94 | 0.16 | 59.0 | 73.5 |
| 42 | 3,6-Dimethyl-2,3,3a,4,5,7a-hexahydrobenzofuran | C5b | un. | EJ | LG1 | 52.3 | 3.64 | 1.97 | 0.14 | 31.7 | 55.8 |
| 42 | 3,6-Dimethyl-2,3,3a,4,5,7a-hexahydrobenzofuran | C5b | un. | AA | LG1 | 61.0 | 3.09 | 1.88 | 0.11 | 57.5 | 63.1 |
| 53 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | AA | LG1 | 63.7 | 3.06 | -1.21 | 0.14 | 63.1 | 65.8 |
| 73 | 3-Hexenal, (Z)- | C11 | Lipid-derived | EJ | LG1 | 73.5 | 3.78 | 0.56 | 0.14 | 64.7 | 74.5 |
| 37 | 4-Acetyl-1-methylcyclohexene | C5b | Terpenoid | EJ | LG1 | 52.3 | 3.30 | 1.18 | 0.11 | 43.8 | 64.5 |
| 50 | Acetic acid, hexyl ester | C7 | Ester | EJ | LG1 | 36.6 | 2.63 | -1.12 | 0.12 | 25.5 | 55.8 |
| 50 | Acetic acid, hexyl ester | C7 | Ester | AA | LG1 | 56.0 | 2.70 | 2.60 | 0.15 | 46.9 | 63.1 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 35 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | EJ | LG1 | 48.3 | 6.63 | 1.66 | 0.19 | 34.4 | 55.4 |
| 35 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | AA | LG1 | 61.0 | 3.67 | 1.20 | 0.11 | 57.5 | 63.1 |
| 30 | a-Terpinol | C5a | Terpenoid | AA | LG1 | 65.8 | 3.08 | 0.69 | 0.10 | 52.7 | 73.5 |
| 38 | Benzene, 4-ethenyl-1,2-dimethyl- | C5b | un. | EJ | LG1 | 53.3 | 3.46 | 1.18 | 0.13 | 31.7 | 55.8 |
| 38 | Benzene, 4-ethenyl-1,2-dimethyl- | C5b | un. | AA | LG1 | 63.7 | 4.31 | 1.38 | 0.16 | 63.1 | 73.5 |
| 63 | b-Ionone | C9 | Carotenoid-derived | EJ | LG1 | 61.0 | 4.32 | 1.32 | 0.24 | 45.9 | 71.0 |
| 34 | cis-Linaloloxide | C5b | Terpenoid | EJ | LG1 | 46.3 | 6.75 | 2.06 | 0.19 | 35.9 | 55.4 |
| 34 | cis-Linaloloxide | C5b | Terpenoid | AA | LG1 | 63.7 | 3.85 | 1.55 | 0.12 | 63.1 | 73.5 |
| 20 | Decanoic acid | C3 | Carboxilic acid | EJ | LG1 | 65.8 | 2.87 | 1.06 | 0.14 | 59.5 | 73.5 |
| 47 | Ethyl Acetate | C6 | Ester | EJ | LG1 | 72.5 | 3.69 | -2.28 | 0.15 | 71.8 | 73.5 |
| 74 | Furan, 2-ethyl- | C11 | Lipid-derived | EJ | LG1 | 72.5 | 6.61 | 0.83 | 0.26 | 63.7 | 73.5 |
| 79 | Hexanal | C12 | Lipid-derived | EJ | LG1 | 57.5 | 4.04 | 0.51 | 0.17 | 46.0 | 59.0 |
| 79 | Hexanal | C12 | Lipid-derived | EJ | LG1 | 63.1 | 4.43 | 0.53 | 0.18 | 59.0 | 65.8 |
| 32 | Linalool | C5a | Terpenoid | AA | LG1 | 67.8 | 4.13 | 2.13 | 0.13 | 59.1 | 72.5 |
| 46 | p-Cymene | C5c | Terpenoid | AA | LG1 | 57.5 | 2.92 | 0.77 | 0.16 | 46.0 | 63.0 |
| 77 | Pentanal | C12 | Lipid-derived | EJ | LG1 | 63.7 | 2.49 | 0.89 | 0.10 | 57.1 | 65.8 |
| 77 | Pentanal | C12 | Lipid-derived | EJ | LG1 | 73.5 | 4.04 | 1.10 | 0.16 | 65.8 | 74.5 |
| 56 | Pyran-2-one, 6-pentyl- | C8 | Lactone | AA | LG1 | 31.3 | 2.66 | 1.36 | 0.14 | 20.6 | 44.1 |
| 76 | 2,4-Heptadienal, (E,E)- | C11 | Lipid-derived | AA | LG3 | 8.9 | 2.94 | -0.81 | 0.18 | 0.0 | 16.4 |
| 7 | 2-Decanone | C2 | Long Ketone/Ester | AA | LG3 | 39.5 | 4.75 | 0.39 | 0.22 | 28.4 | 45.6 |
| 80 | 2-Hexenal, (E)- | C12 | Lipid-derived | AA | LG3 | 1.0 | 4.08 | -0.19 | 0.18 | 0.0 | 7.4 |
| 80 | 2-Hexenal, (E)- | C12 | Lipid-derived | AA | LG3 | 8.9 | 4.72 | -0.21 | 0.20 | 7.4 | 14.7 |
| 80 | 2-Hexenal, (E)- | C12 | Lipid-derived | AA | LG3 | 45.9 | 6.08 | 0.25 | 0.27 | 38.4 | 50.6 |
| 8 | 3-Decanone | C2 | Long Ketone/Ester | AA | LG3 | 39.5 | 3.39 | 0.29 | 0.16 | 28.2 | 50.8 |
| 49 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | C6 | Lactone | AA | LG3 | 72.7 | 3.32 | -0.82 | 0.14 | 64.3 | 83.3 |
| 54 | Acetic acid, 2-methylpropyl ester | C8 | Ester | AA | LG3 | 71.2 | 5.17 | -2.67 | 0.27 | 65.8 | 80.4 |
| 50 | Acetic acid, hexyl ester | C7 | Ester | AA | LG3 | 62.3 | 2.97 | -2.16 | 0.18 | 61.5 | 63.8 |
| 50 | Acetic acid, hexyl ester | C7 | Ester | AA | LG3 | 72.7 | 5.08 | -2.78 | 0.28 | 68.9 | 82.1 |
| 25 | Acetic acid, phenylmethyl ester | C4 | Aromatic Ester | EJ | LG3 | 86.0 | 2.60 | -0.11 | 0.14 | 73.8 | 87.0 |
| 30 | a-Terpinol | C5a | Terpenoid | AA | LG3 | 16.8 | 2.35 | 0.63 | 0.08 | 7.4 | 23.8 |
| 24 | Benzophenone, 3-methyl- | C4 | Aromatic Ketone | AA | LG3 | 39.5 | 3.13 | 0.53 | 0.14 | 31.5 | 49.8 |
| 26 | Bicyclo[2.2.1]heptan-2-one, 1,7,7-trimethyl-, (1R)- | C4 | Ketone | AA | LG3 | 28.5 | 2.69 | 0.19 | 0.15 | 19.0 | 39.5 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 57 | d-Decalactone | C8 | Lactone | AA | LG3 | 71.2 | 3.16 | -1.68 | 0.17 | 68.7 | 82.8 |
| 12 | Dodecanoic acid, methyl ester | C2 | Long Ketone/Ester | AA | LG3 | 37.2 | 2.93 | 0.43 | 0.18 | 26.6 | 39.5 |
| 48 | Ethanol | C6 | Alcohol | AA | LG3 | 59.5 | 3.14 | -1.66 | 0.17 | 54.5 | 62.1 |
| 47 | Ethyl Acetate | C6 | Ester | AA | LG3 | 71.2 | 4.09 | -2.14 | 0.20 | 63.9 | 81.3 |
| 81 | Eugenol | C12b | Aromatic alcohol | AA | LG3 | 7.4 | 2.43 | -0.66 | 0.13 | 0.0 | 14.0 |
| 58 | g-Decalactone | C8 | Lactone | AA | LG3 | 74.1 | 3.58 | -1.34 | 0.19 | 66.2 | 83.1 |
| 66 | Geranyl acetone | C9 | Carotenoid-derived | AA | LG3 | 6.0 | 4.35 | -1.70 | 0.23 | 0.0 | 9.5 |
| 66 | Geranyl acetone | C9 | Carotenoid-derived | AA | LG3 | 26.0 | 3.39 | 1.62 | 0.19 | 22.5 | 31.4 |
| 62 | g-Heptalactone | C8 | Lactone | AA | LG3 | 74.1 | 3.52 | -1.19 | 0.18 | 63.6 | 84.9 |
| 55 | g-Hexalactone | C8 | Lactone | AA | LG3 | 65.2 | 4.85 | -0.71 | 0.23 | 63.8 | 69.8 |
| 55 | g-Hexalactone | C8 | Lactone | AA | LG3 | 73.7 | 4.66 | -0.71 | 0.22 | 72.7 | 80.8 |
| 61 | g-Jasmolactone | C8 | Lactone | AA | LG3 | 71.2 | 3.31 | -1.54 | 0.17 | 63.8 | 83.1 |
| 60 | g-Nonalactone | C8 | Lactone | AA | LG3 | 71.2 | 3.80 | -1.06 | 0.18 | 56.6 | 80.8 |
| 59 | g-Octalactone | C8 | Lactone | AA | LG3 | 71.2 | 3.15 | -1.09 | 0.14 | 69.8 | 84.8 |
| 22 | Isoeugenol methyl ether | C4 | Aromatic Ether | AA | LG3 | 38.2 | 4.02 | 0.42 | 0.21 | 26.6 | 48.9 |
| 11 | Nonanoic acid, methyl ester | C2 | Long Ketone/Ester | AA | LG3 | 31.5 | 3.14 | 0.27 | 0.16 | 23.9 | 45.8 |
| 9 | Octanoic acid, methyl ester | C2 | Long Ketone/Ester | AA | LG3 | 39.5 | 4.08 | 0.38 | 0.21 | 28.4 | 43.9 |
| 46 | p-Cymene | C5c | Terpenoid | EJ | LG3 | 50.7 | 2.59 | 0.66 | 0.10 | 42.3 | 52.3 |
| 77 | Pentanal | C12 | Lipid-derived | EJ | LG3 | 54.5 | 4.30 | 1.53 | 0.17 | 49.8 | 60.0 |
| 56 | Pyran-2-one, 6-pentyl- | C8 | Lactone | AA | LG3 | 74.1 | 3.03 | -1.37 | 0.16 | 63.8 | 83.2 |
| 28 | Salicylic acid, ethyl ester | C4 | Aromatic Ester | AA | LG3 | 28.5 | 2.63 | 0.26 | 0.14 | 27.8 | 31.5 |
| 28 | Salicylic acid, ethyl ester | C4 | Aromatic Ester | AA | LG3 | 37.2 | 2.89 | 0.27 | 0.15 | 31.5 | 52.0 |
| 33 | D-Limonene | C5b | Terpenoid | EJ | LG4 | 0.0 | 7.57 | 1.65 | 0.29 | 0.0 | 4.3 |
| 33 | D-Limonene | C5b | Terpenoid | AA | LG4 | 30.9 | 3.79 | 0.58 | 0.18 | 29.6 | 34.0 |
| 36 | Ethanone, 1-(4-methylphenyl)- | C5b | un. | EJ | LG4 | 4.3 | 11.95 | 0.96 | 0.37 | 0.0 | 12.3 |
| 36 | Ethanone, 1-(4-methylphenyl)- | C5b | un. | AA | LG4 | 6.7 | 5.46 | 1.33 | 0.20 | 0.0 | 13.2 |
| 36 | Ethanone, 1-(4-methylphenyl)- | C5b | un. | AA | LG4 | 52.9 | 5.98 | 1.22 | 0.20 | 44.3 | 56.4 |
| 41 | 1,3,8-p-Menthatriene | C5b | Terpenoid | AA | LG4 | 2.8 | 5.80 | 3.28 | 0.19 | 0.0 | 10.2 |
| 41 | 1,3,8-p-Menthatriene | C5b | Terpenoid | EJ | LG4 | 2.8 | 8.05 | 3.19 | 0.34 | 0.0 | 12.2 |
| 41 | 1,3,8-p-Menthatriene | C5b | Terpenoid | AA | LG4 | 25.1 | 3.33 | -2.92 | 0.11 | 21.5 | 27.1 |
| 41 | 1,3,8-p-Menthatriene | C5b | Terpenoid | AA | LG4 | 52.9 | 9.01 | 4.14 | 0.34 | 49.5 | 56.3 |
| 45 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5b | un. | AA | LG4 | 0.0 | 4.57 | 2.99 | 0.13 | 0.0 | 4.3 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 45 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5b | un. | EJ | LG4 | 0.0 | 11.36 | 3.26 | 0.40 | 0.0 | 7.3 |
| 45 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5b | un. | AA | LG4 | 52.9 | 8.87 | 4.27 | 0.28 | 47.3 | 56.4 |
| 40 | 1-Cyclohexene-1-carboxaldehyde,2,6,6-trimethyl | C5b | Terpenoid | AA | LG4 | 2.8 | 4.20 | 3.29 | 0.18 | 0.0 | 4.3 |
| 40 | 1-Cyclohexene-1-carboxaldehyde,2,6,6-trimethyl | C5b | Terpenoid | EJ | LG4 | 4.3 | 6.89 | 3.47 | 0.31 | 0.0 | 12.3 |
| 40 | 1-Cyclohexene-1-carboxaldehyde,2,6,6-trimethyl | C5b | Terpenoid | AA | LG4 | 52.9 | 5.78 | 4.25 | 0.24 | 46.2 | 56.4 |
| 75 | 1-Penten-3-one | C11 | Lipid-derived | AA | LG4 | 49.8 | 3.54 | -0.94 | 0.20 | 38.5 | 58.1 |
| 21 | 2,7-Ethanonaphth[2,3-b]oxirene, 1a,2,7,7a-tetrahydro-, (1aa,2a,7a,7aa)- | C4 | un. | AA | LG4 | 40.2 | 2.75 | 0.54 | 0.17 | 30.0 | 49.4 |
| 44 | 2-Cyclohexen-1-ol, 2-methyl-5-(1-methylethenyl)-, cis | C5b | un. | EJ | LG4 | 0.0 | 10.51 | 2.90 | 0.40 | 0.0 | 4.3 |
| 44 | 2-Cyclohexen-1-ol, 2-methyl-5-(1-methylethenyl)-, cis | C5b | un. | AA | LG4 | 2.8 | 4.34 | 2.66 | 0.12 | 0.0 | 4.3 |
| 44 | 2-Cyclohexen-1-ol, 2-methyl-5-(1-methylethenyl)-, cis | C5b | un. | AA | LG4 | 52.9 | 9.78 | 4.06 | 0.32 | 47.5 | 56.3 |
| 7 | 2-Decanone | C2 | Long Ketone/Ester | AA | LG4 | 52.8 | 5.39 | 0.62 | 0.32 | 50.0 | 63.9 |
| 7 | 2-Decanone | C2 | Long Ketone/Ester | AA | LG4 | 67.0 | 3.27 | -0.35 | 0.15 | 66.7 | 68.0 |
| 52 | 2-Hexen-1-ol, acetate, (E)- | C7 | Ester | AA | LG4 | 41.2 | 4.56 | 2.48 | 0.24 | 38.5 | 57.5 |
| 52 | 2-Hexen-1-ol, acetate, (E)- | C7 | Ester | EJ | LG4 | 41.8 | 6.72 | 1.92 | 0.25 | 31.3 | 49.8 |
| 80 | 2-Hexenal, (E)- | C12 | Lipid-derived | EJ | LG4 | 41.8 | 3.85 | -0.44 | 0.15 | 38.7 | 48.6 |
| 80 | 2-Hexenal, (E)- | C12 | Lipid-derived | AA | LG4 | 57.4 | 5.30 | 0.20 | 0.23 | 52.9 | 60.2 |
| 80 | 2-Hexenal, (E)- | C12 | Lipid-derived | AA | LG4 | 67.0 | 5.30 | 0.19 | 0.23 | 60.2 | 68.0 |
| 80 | 2-Hexenal, (E)- | C12 | Lipid-derived | EJ | LG4 | 67.0 | 6.12 | 0.51 | 0.25 | 57.3 | 68.0 |
| 31 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | C5a | Terpenoid | EJ | LG4 | 0.0 | 11.20 | 3.30 | 0.38 | 0.0 | 4.2 |
| 31 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | C5a | Terpenoid | AA | LG4 | 52.9 | 8.70 | 3.22 | 0.30 | 49.2 | 56.5 |
| 13 | 2-Propenoic acid, 2-ethylhexyl ester | C2 | Long Ketone/Ester | AA | LG4 | 52.9 | 3.99 | 0.47 | 0.24 | 49.5 | 58.4 |
| 42 | 3,6-Dimethyl-2,3,3a,4,5,7a-hexahydrobenzofuran | C5b | un. | AA | LG4 | 2.8 | 5.95 | 3.13 | 0.18 | 0.0 | 4.3 |
| 42 | 3,6-Dimethyl-2,3,3a,4,5,7a-hexahydrobenzofuran | C5b | un. | EJ | LG4 | 4.3 | 10.62 | 3.28 | 0.40 | 0.0 | 12.2 |
| 42 | 3,6-Dimethyl-2,3,3a,4,5,7a-hexahydrobenzofuran | C5b | un. | AA | LG4 | 25.1 | 3.55 | -2.91 | 0.11 | 19.8 | 28.1 |
| 42 | 3,6-Dimethyl-2,3,3a,4,5,7a-hexahydrobenzofuran | C5b | un. | AA | LG4 | 52.9 | 11.50 | 4.30 | 0.38 | 49.6 | 56.3 |
| 53 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | EJ | LG4 | 0.0 | 3.85 | -1.06 | 0.18 | 0.0 | 5.3 |
| 53 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | EJ | LG4 | 9.7 | 3.15 | -1.09 | 0.15 | 5.3 | 11.9 |
| 43 | 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl | C5b | Terpenoid | AA | LG4 | 2.8 | 5.64 | 3.19 | 0.22 | 0.0 | 5.3 |
| 43 | 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl | C5b | Terpenoid | EJ | LG4 | 2.8 | 7.24 | 3.19 | 0.29 | 0.0 | 10.1 |
| 43 | 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl | C5b | Terpenoid | AA | LG4 | 52.9 | 5.61 | 3.63 | 0.21 | 44.4 | 56.4 |
| 8 | 3-Decanone | C2 | Long Ketone/Ester | AA | LG4 | 52.8 | 4.49 | 0.56 | 0.26 | 49.8 | 61.6 |
| 73 | 3-Hexenal, (Z)- | C11 | Lipid-derived | AA | LG4 | 48.4 | 7.19 | -1.04 | 0.31 | 38.7 | 57.6 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 37 | 4-Acetyl-1-methylcyclohexene | C5b | Terpenoid | EJ | LG4 | 2.8 | 11.78 | 2.25 | 0.40 | 0.0 | 9.9 |
| 37 | 4-Acetyl-1-methylcyclohexene | C5b | Terpenoid | AA | LG4 | 7.7 | 5.45 | 2.58 | 0.21 | 6.1 | 12.4 |
| 37 | 4-Acetyl-1-methylcyclohexene | C5b | Terpenoid | AA | LG4 | 52.9 | 7.09 | 2.47 | 0.24 | 45.6 | 56.5 |
| 49 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | C6 | Lactone | AA | LG4 | 57.4 | 3.07 | -1.12 | 0.13 | 53.7 | 59.2 |
| 49 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | C6 | Lactone | AA | LG4 | 67.0 | 2.93 | -0.89 | 0.12 | 59.2 | 68.0 |
| 49 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | C6 | Lactone | EJ | LG4 | 67.0 | 5.89 | -0.69 | 0.21 | 59.0 | 68.0 |
| 5 | 8,8,9-Trimethyl-deca-3,5-diene-2,7-dione | C2 | Long Ketone/Ester | AA | LG4 | 52.8 | 2.12 | 0.57 | 0.14 | 48.3 | 64.7 |
| 5 | 8,8,9-Trimethyl-deca-3,5-diene-2,7-dione | C2 | Long Ketone/Ester | AA | LG4 | 67.0 | 2.49 | -0.40 | 0.15 | 66.7 | 68.0 |
| 10 | Acetic acid, 2-ethylhexyl ester | C2 | Long Ketone/Ester | AA | LG4 | 52.8 | 3.56 | 0.55 | 0.22 | 49.8 | 60.5 |
| 1 | Acetic acid, butyl ester | C1 | Ester | AA | LG4 | 16.0 | 4.21 | -0.27 | 0.25 | 6.3 | 20.3 |
| 1 | Acetic acid, butyl ester | C1 | Ester | AA | LG4 | 29.0 | 4.25 | 0.27 | 0.25 | 28.0 | 37.1 |
| 25 | Acetic acid, phenylmethyl ester | C4 | Aromatic Ester | AA | LG4 | 38.7 | 3.98 | 0.37 | 0.22 | 31.0 | 49.4 |
| 35 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | AA | LG4 | 2.8 | 3.83 | 1.54 | 0.11 | 0.0 | 12.2 |
| 35 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | EJ | LG4 | 2.8 | 12.14 | 1.83 | 0.37 | 0.0 | 12.3 |
| 35 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | AA | LG4 | 52.9 | 10.93 | 2.65 | 0.35 | 50.0 | 56.3 |
| 30 | a-Terpinol | C5a | Terpenoid | EJ | LG4 | 0.0 | 7.45 | 1.74 | 0.30 | 0.0 | 4.1 |
| 30 | a-Terpinol | C5a | Terpenoid | AA | LG4 | 52.9 | 6.71 | 1.17 | 0.29 | 49.6 | 57.0 |
| 64 | b-Damascenone | C9 | Carotenoid-derived | AA | LG4 | 18.0 | 3.95 | 0.75 | 0.25 | 17.0 | 28.0 |
| 64 | b-Damascenone | C9 | Carotenoid-derived | EJ | LG4 | 36.0 | 4.09 | -0.50 | 0.23 | 27.9 | 48.2 |
| 27 | Benzene, (1,3-dimethylbutyl)- | C4 | Aromatic | AA | LG4 | 38.7 | 2.64 | 0.37 | 0.16 | 38.0 | 49.7 |
| 38 | Benzene, 4-ethenyl-1,2-dimethyl- | C5b | un. | EJ | LG4 | 2.8 | 8.84 | 1.90 | 0.31 | 0.0 | 9.9 |
| 38 | Benzene, 4-ethenyl-1,2-dimethyl- | C5b | un. | AA | LG4 | 5.3 | 5.28 | 1.97 | 0.20 | 0.0 | 10.0 |
| 38 | Benzene, 4-ethenyl-1,2-dimethyl- | C5b | un. | AA | LG4 | 25.1 | 3.75 | -1.92 | 0.14 | 19.4 | 28.1 |
| 38 | Benzene, 4-ethenyl-1,2-dimethyl- | C5b | un. | AA | LG4 | 52.9 | 6.42 | 2.10 | 0.26 | 48.4 | 56.4 |
| 39 | Benzenemethanol, a,a,4-trimethyl- | C5b | un. | AA | LG4 | 2.8 | 5.63 | 3.08 | 0.26 | 0.0 | 9.6 |
| 39 | Benzenemethanol, a,a,4-trimethyl- | C5b | un. | EJ | LG4 | 2.8 | 8.89 | 2.14 | 0.31 | 0.0 | 9.3 |
| 39 | Benzenemethanol, a,a,4-trimethyl- | C5b | un. | AA | LG4 | 29.6 | 2.95 | -2.91 | 0.13 | 28.1 | 39.9 |
| 39 | Benzenemethanol, a,a,4-trimethyl- | C5b | un. | AA | LG4 | 52.9 | 4.46 | 3.28 | 0.22 | 48.8 | 56.5 |
| 23 | Benzophenone | C4 | Aromatic Ketone | AA | LG4 | 49.4 | 3.96 | 0.48 | 0.25 | 41.6 | 49.8 |
| 23 | Benzophenone | C4 | Aromatic Ketone | AA | LG4 | 67.0 | 2.77 | -0.34 | 0.15 | 66.7 | 68.0 |
| 24 | Benzophenone, 3-methyl- | C4 | Aromatic Ketone | AA | LG4 | 52.8 | 3.32 | 0.85 | 0.20 | 49.8 | 58.5 |
| 24 | Benzophenone, 3-methyl- | C4 | Aromatic Ketone | AA | LG4 | 67.0 | 5.20 | -0.62 | 0.24 | 66.9 | 68.0 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 24 | Benzophenone, 3-methyl- | C4 | Aromatic Ketone | EJ | LG4 | 67.0 | 3.33 | -0.26 | 0.17 | 58.7 | 68.0 |
| 63 | b-Ionone | C9 | Carotenoid-derived | AA | LG4 | 53.0 | 5.31 | 1.23 | 0.31 | 43.8 | 57.1 |
| 34 | cis-Linaloloxide | C5b | Terpenoid | AA | LG4 | 2.8 | 4.26 | 2.11 | 0.13 | 0.0 | 12.2 |
| 34 | cis-Linaloloxide | C5b | Terpenoid | EJ | LG4 | 2.8 | 11.87 | 2.19 | 0.38 | 0.0 | 12.3 |
| 34 | cis-Linaloloxide | C5b | Terpenoid | AA | LG4 | 25.1 | 2.68 | -1.96 | 0.09 | 17.4 | 28.1 |
| 34 | cis-Linaloloxide | C5b | Terpenoid | AA | LG4 | 52.9 | 11.32 | 3.34 | 0.37 | 49.9 | 56.3 |
| 57 | d-Decalactone | C8 | Lactone | EJ | LG4 | 57.4 | 3.16 | 1.50 | 0.14 | 49.8 | 58.6 |
| 20 | Decanoic acid | C3 | Carboxilic acid | AA | LG4 | 34.0 | 2.38 | 0.36 | 0.14 | 25.1 | 49.6 |
| 6 | Decanoic acid, methyl ester | C2 | Long Ketone/Ester | AA | LG4 | 49.4 | 2.59 | 0.55 | 0.19 | 41.4 | 49.8 |
| 6 | Decanoic acid, methyl ester | C2 | Long Ketone/Ester | AA | LG4 | 67.0 | 2.68 | -0.47 | 0.17 | 66.7 | 68.0 |
| 81 | Eugenol | C12b | Aromatic alcohol | EJ | LG4 | 57.4 | 5.91 | 0.66 | 0.26 | 47.1 | 68.0 |
| 74 | Furan, 2-ethyl- | C11 | Lipid-derived | AA | LG4 | 40.2 | 2.73 | -1.44 | 0.17 | 32.5 | 51.2 |
| 78 | Furan, 2-pentyl- | C12 | Lipid-derived | EJ | LG4 | 57.4 | 4.65 | 0.65 | 0.18 | 48.0 | 58.3 |
| 66 | Geranyl acetone | C9 | Carotenoid-derived | AA | LG4 | 38.0 | 3.03 | 1.09 | 0.15 | 27.8 | 48.0 |
| 60 | g-Nonalactone | C8 | Lactone | AA | LG4 | 67.0 | 3.73 | -1.03 | 0.17 | 55.3 | 68.0 |
| 59 | g-Octalactone | C8 | Lactone | AA | LG4 | 67.0 | 2.71 | -1.01 | 0.12 | 52.9 | 68.0 |
| 79 | Hexanal | C12 | Lipid-derived | AA | LG4 | 52.9 | 5.09 | 0.30 | 0.27 | 41.6 | 57.8 |
| 79 | Hexanal | C12 | Lipid-derived | EJ | LG4 | 52.9 | 4.22 | 0.80 | 0.17 | 52.8 | 68.0 |
| 18 | Hexanoic acid | C3 | Carboxilic acid | AA | LG4 | 38.7 | 3.20 | 7.64 | 0.20 | 29.9 | 41.6 |
| 22 | Isoeugenol methyl ether | C4 | Aromatic Ether | AA | LG4 | 52.8 | 3.01 | 0.56 | 0.18 | 44.8 | 63.9 |
| 22 | Isoeugenol methyl ether | C4 | Aromatic Ether | AA | LG4 | 67.0 | 2.87 | -0.38 | 0.15 | 66.7 | 68.0 |
| 32 | Linalool | C5a | Terpenoid | EJ | LG4 | 0.0 | 10.60 | 4.01 | 0.37 | 0.0 | 4.2 |
| 32 | Linalool | C5a | Terpenoid | AA | LG4 | 52.9 | 8.35 | 3.37 | 0.29 | 48.3 | 56.8 |
| 11 | Nonanoic acid, methyl ester | C2 | Long Ketone/Ester | AA | LG4 | 52.8 | 3.11 | 0.54 | 0.19 | 49.9 | 63.5 |
| 11 | Nonanoic acid, methyl ester | C2 | Long Ketone/Ester | AA | LG4 | 67.0 | 2.84 | -0.36 | 0.15 | 66.7 | 68.0 |
| 9 | Octanoic acid, methyl ester | C2 | Long Ketone/Ester | AA | LG4 | 52.8 | 5.23 | 0.59 | 0.31 | 50.0 | 61.8 |
| 46 | p-Cymene | C5c | Terpenoid | EJ | LG4 | 2.8 | 6.86 | 1.10 | 0.27 | 0.0 | 9.8 |
| 46 | p-Cymene | C5c | Terpenoid | EJ | LG4 | 42.8 | 3.87 | -0.83 | 0.15 | 34.0 | 56.9 |
| 77 | Pentanal | C12 | Lipid-derived | EJ | LG4 | 57.4 | 3.47 | 1.09 | 0.14 | 52.9 | 58.3 |
| 77 | Pentanal | C12 | Lipid-derived | EJ | LG4 | 65.6 | 3.96 | 1.13 | 0.16 | 58.3 | 68.0 |
| 3 | Propanoic acid, 2,2-dimethyl- | C1 | C1 | AA | LG4 | 58.3 | 2.87 | 0.82 | 0.17 | 54.9 | 68.0 |
| 3 | Propanoic acid, 2,2-dimethyl- | C1 | C1 | EJ | LG4 | 58.3 | 5.21 | -0.66 | 0.26 | 56.9 | 66.2 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 28 | Salicylic acid, ethyl ester | C4 | Aromatic Ester | AA | LG4 | 49.4 | 3.72 | 0.44 | 0.23 | 41.5 | 49.7 |
| 29 | Terpinolene | C5a | Terpenoid | EJ | LG4 | 0.0 | 6.91 | 1.62 | 0.27 | 0.0 | 4.2 |
| 29 | Terpinolene | C5a | Terpenoid | AA | LG4 | 52.9 | 6.83 | 1.34 | 0.27 | 44.6 | 57.1 |
| 53 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | AA | LG5 | 16.2 | 3.42 | -1.02 | 0.15 | 6.1 | 23.8 |
| 53 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | AA | LG5 | 26.8 | 3.03 | -0.98 | 0.14 | 23.8 | 36.4 |
| 51 | 3-Hexen-1-ol, acetate, (Z)- | C7 | Ester | EJ | LG5 | 0.0 | 2.56 | -0.64 | 0.14 | 0.0 | 7.5 |
| 51 | 3-Hexen-1-ol, acetate, (Z)- | C7 | Ester | AA | LG5 | 16.2 | 4.77 | -1.26 | 0.26 | 6.1 | 27.0 |
| 49 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | C6 | Lactone | EJ | LG5 | 7.5 | 3.24 | -0.54 | 0.11 | 0.0 | 17.2 |
| 49 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | C6 | Lactone | AA | LG5 | 8.5 | 5.14 | -1.10 | 0.23 | 0.0 | 10.4 |
| 50 | Acetic acid, hexyl ester | C7 | Ester | EJ | LG5 | 0.0 | 4.72 | -2.09 | 0.22 | 0.0 | 5.3 |
| 23 | Benzophenone | C4 | Aromatic Ketone | EJ | LG5 | 40.2 | 3.27 | 0.32 | 0.18 | 27.4 | 46.6 |
| 79 | Hexanal | C12 | Lipid-derived | AA | LG5 | 6.1 | 5.53 | -0.40 | 0.29 | 0.4 | 11.4 |
| 79 | Hexanal | C12 | Lipid-derived | EJ | LG5 | 16.2 | 2.94 | -0.42 | 0.12 | 7.5 | 26.9 |
| 79 | Hexanal | C12 | Lipid-derived | AA | LG5 | 17.7 | 3.21 | 0.39 | 0.18 | 15.8 | 23.8 |
| 79 | Hexanal | C12 | Lipid-derived | AA | LG5 | 27.4 | 3.32 | 0.32 | 0.18 | 23.8 | 37.4 |
| 56 | Pyran-2-one, 6-pentyl- | C8 | Lactone | EJ | LG5 | 0.0 | 2.84 | -1.08 | 0.15 | 0.0 | 6.1 |
| 75 | 1-Penten-3-one | C11 | Lipid-derived | EJ | LG6 | 47.7 | 3.46 | -0.88 | 0.16 | 40.0 | 50.6 |
| 75 | 1-Penten-3-one | C11 | Lipid-derived | EJ | LG6 | 56.1 | 3.83 | -0.95 | 0.19 | 50.6 | 59.1 |
| 76 | 2,4-Heptadienal, (E,E)- | C11 | Lipid-derived | EJ | LG6 | 11.1 | 2.73 | -0.53 | 0.10 | 0.0 | 20.9 |
| 7 | 2-Decanone | C2 | Long Ketone/Ester | EJ | LG6 | 0.0 | 5.76 | 0.22 | 0.30 | 0.0 | 11.0 |
| 52 | 2-Hexen-1-ol, acetate, (E)- | C7 | Ester | AA | LG6 | 50.6 | 2.71 | 2.08 | 0.15 | 39.6 | 59.1 |
| 52 | 2-Hexen-1-ol, acetate, (E)- | C7 | Ester | EJ | LG6 | 50.6 | 3.66 | 1.37 | 0.13 | 39.9 | 59.1 |
| 80 | 2-Hexenal, (E)- | C12 | Lipid-derived | EJ | LG6 | 52.1 | 2.59 | -0.24 | 0.10 | 47.7 | 59.1 |
| 8 | 3-Decanone | C2 | Long Ketone/Ester | EJ | LG6 | 0.0 | 2.54 | 0.16 | 0.13 | 0.0 | 11.1 |
| 51 | 3-Hexen-1-ol, acetate, (Z)- | C7 | Ester | EJ | LG6 | 29.9 | 2.64 | 0.62 | 0.14 | 24.0 | 32.5 |
| 73 | 3-Hexenal, (Z)- | C11 | Lipid-derived | EJ | LG6 | 32.8 | 3.09 | -0.44 | 0.12 | 31.9 | 34.3 |
| 73 | 3-Hexenal, (Z)- | C11 | Lipid-derived | EJ | LG6 | 43.2 | 3.82 | -0.48 | 0.14 | 40.2 | 50.6 |
| 5 | 8,8,9-Trimethyl-deca-3,5-diene-2,7-dione | C2 | Long Ketone/Ester | EJ | LG6 | 0.0 | 2.66 | 0.16 | 0.13 | 0.0 | 11.1 |
| 10 | Acetic acid, 2-ethylhexyl ester | C2 | Long Ketone/Ester | EJ | LG6 | 0.0 | 5.90 | 0.23 | 0.29 | 0.0 | 11.0 |
| 54 | Acetic acid, 2-methylpropyl ester | C8 | Ester | EJ | LG6 | 50.6 | 8.02 | 2.38 | 0.34 | 48.0 | 59.1 |
| 57 | d-Decalactone | C8 | Lactone | EJ | LG6 | 52.1 | 2.94 | 0.88 | 0.14 | 39.8 | 59.1 |
| 47 | Ethyl Acetate | C6 | Ester | EJ | LG6 | 58.1 | 5.41 | 1.94 | 0.23 | 50.4 | 59.1 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 47 | Ethyl Acetate | C6 | Ester | AA | LG6 | 61.0 | 2.98 | 1.68 | 0.15 | 50.0 | 60.0 |
| 74 | Furan, 2-ethyl- | C11 | Lipid-derived | EJ | LG6 | 43.2 | 3.88 | -0.59 | 0.14 | 34.3 | 47.7 |
| 55 | g-Hexalactone | C8 | Lactone | EJ | LG6 | 50.6 | 8.28 | 0.79 | 0.36 | 44.8 | 59.1 |
| 55 | g-Hexalactone | C8 | Lactone | AA | LG6 | 52.1 | 6.06 | 0.78 | 0.27 | 44.9 | 59.1 |
| 61 | g-Jasmolactone | C8 | Lactone | EJ | LG6 | 50.6 | 5.69 | 1.91 | 0.22 | 47.8 | 59.1 |
| 59 | g-Octalactone | C8 | Lactone | EJ | LG6 | 49.5 | 3.40 | 0.69 | 0.14 | 44.6 | 58.1 |
| 59 | g-Octalactone | C8 | Lactone | AA | LG6 | 52.1 | 2.90 | 1.02 | 0.13 | 39.8 | 59.1 |
| 4 | Heptanoic acid | C1 | Carboxilic acid | EJ | LG6 | 11.0 | 4.50 | -2.28 | 0.23 | 0.0 | 23.4 |
| 4 | Heptanoic acid | C1 | Carboxilic acid | EJ | LG6 | 47.0 | 3.28 | 1.98 | 0.17 | 33.7 | 52.0 |
| 4 | Heptanoic acid | C1 | Carboxilic acid | AA | LG6 | 52.0 | 10.53 | -0.15 | 0.07 | 28.0 | 59.0 |
| 11 | Nonanoic acid, methyl ester | C2 | Long Ketone/Ester | EJ | LG6 | 0.0 | 3.64 | 0.20 | 0.18 | 0.0 | 11.0 |
| 9 | Octanoic acid, methyl ester | C2 | Long Ketone/Ester | EJ | LG6 | 0.0 | 4.17 | 0.17 | 0.22 | 0.0 | 11.0 |
| 73 | 3-Hexenal, (Z)- | C11 | Lipid-derived | AA | LG7 | 1.3 | 2.70 | 0.61 | 0.10 | 0.0 | 26.4 |
| 49 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | C6 | Lactone | EJ | LG7 | 50.9 | 3.67 | -0.54 | 0.12 | 44.1 | 66.2 |
| 47 | Ethyl Acetate | C6 | Ester | EJ | LG7 | 50.9 | 3.38 | -1.51 | 0.13 | 43.4 | 60.4 |
| 58 | g-Decalactone | C8 | Lactone | EJ | LG7 | 67.6 | 2.85 | -0.61 | 0.15 | 60.4 | 68.6 |
| 62 | g-Heptalactone | C8 | Lactone | EJ | LG7 | 54.5 | 3.94 | -0.70 | 0.19 | 54.5 | 57.4 |
| 55 | g-Hexalactone | C8 | Lactone | EJ | LG7 | 50.9 | 2.52 | -0.32 | 0.11 | 41.9 | 63.1 |
| 61 | g-Jasmolactone | C8 | Lactone | EJ | LG7 | 50.9 | 4.58 | -0.98 | 0.18 | 44.6 | 58.1 |
| 60 | g-Nonalactone | C8 | Lactone | EJ | LG7 | 59.2 | 3.32 | -0.54 | 0.13 | 54.5 | 63.4 |
| 60 | g-Nonalactone | C8 | Lactone | EJ | LG7 | 66.2 | 3.85 | -0.57 | 0.15 | 63.4 | 68.6 |
| 59 | g-Octalactone | C8 | Lactone | EJ | LG7 | 59.2 | 4.23 | -0.75 | 0.17 | 54.5 | 63.4 |
| 59 | g-Octalactone | C8 | Lactone | EJ | LG7 | 67.6 | 4.68 | -0.77 | 0.18 | 63.4 | 68.6 |
| 68 | Heptanol | C9 | Alcohol | AA | LG7 | 64.0 | 1.83 | 0.17 | 0.09 | 60.0 | 66.0 |
| 56 | Pyran-2-one, 6-pentyl- | C8 | Lactone | EJ | LG7 | 59.2 | 4.74 | -1.00 | 0.24 | 54.5 | 62.4 |
| 56 | Pyran-2-one, 6-pentyl- | C8 | Lactone | EJ | LG7 | 66.2 | 5.26 | -1.01 | 0.26 | 62.4 | 68.6 |
| 36 | Ethanone, 1-(4-methylphenyl)- | C5b | un. | EJ | LG8 | 7.7 | 5.97 | -0.91 | 0.17 | 6.3 | 10.0 |
| 36 | Ethanone, 1-(4-methylphenyl)- | C5b | un. | EJ | LG8 | 20.4 | 4.49 | 1.13 | 0.13 | 19.8 | 31.6 |
| 45 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5b | un. | EJ | LG8 | 8.9 | 3.71 | -2.19 | 0.12 | 0.0 | 18.7 |
| 71 | 1-Octanol | C10 | Alcohol | AA | LG8 | 30.0 | 2.31 | 0.32 | 0.11 | 28.8 | 45.0 |
| 71 | 1-Octanol | C10 | Alcohol | EJ | LG8 | 43.0 | 2.63 | 0.39 | 0.13 | 32.6 | 46.0 |
| 31 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | C5a | Terpenoid | EJ | LG8 | 7.7 | 4.49 | -2.41 | 0.13 | 0.0 | 18.6 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 31 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | C5a | Terpenoid | EJ | LG8 | 38.4 | 3.48 | 2.13 | 0.10 | 35.6 | 47.1 |
| 53 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | AA | LG8 | 9.9 | 6.41 | -3.61 | 0.25 | 8.9 | 12.4 |
| 53 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | AA | LG8 | 20.4 | 7.37 | 3.89 | 0.29 | 20.1 | 21.6 |
| 37 | 4-Acetyl-1-methylcyclohexene | C5b | Terpenoid | EJ | LG8 | 21.7 | 3.20 | 2.17 | 0.11 | 18.9 | 28.1 |
| 1 | Acetic acid, butyl ester | C1 | Ester | AA | LG8 | 43.0 | 15.94 | -0.09 | 0.06 | 0.0 | 55.9 |
| 35 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | EJ | LG8 | 7.7 | 5.34 | -1.38 | 0.14 | 0.0 | 11.6 |
| 35 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | EJ | LG8 | 28.1 | 2.67 | 1.24 | 0.08 | 27.2 | 30.7 |
| 35 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | EJ | LG8 | 44.5 | 2.64 | 0.99 | 0.08 | 43.5 | 48.1 |
| 38 | Benzene, 4-ethenyl-1,2-dimethyl- | C5b | un. | EJ | LG8 | 7.7 | 4.17 | -1.29 | 0.14 | 6.4 | 12.8 |
| 39 | Benzenemethanol, a,a,4-trimethyl- | C5b | un. | EJ | LG8 | 8.9 | 5.56 | -1.71 | 0.19 | 1.4 | 12.8 |
| 39 | Benzenemethanol, a,a,4-trimethyl- | C5b | un. | EJ | LG8 | 14.6 | 4.52 | -1.66 | 0.17 | 12.8 | 18.6 |
| 39 | Benzenemethanol, a,a,4-trimethyl- | C5b | un. | EJ | LG8 | 53.8 | 2.78 | 1.17 | 0.09 | 48.7 | 60.2 |
| 63 | b-Ionone | C9 | Carotenoid-derived | EJ | LG8 | 33.0 | 4.61 | 1.92 | 0.26 | 29.4 | 37.0 |
| 34 | cis-Linaloloxide | C5b | Terpenoid | EJ | LG8 | 7.7 | 4.83 | -1.62 | 0.14 | 0.2 | 12.2 |
| 34 | cis-Linaloloxide | C5b | Terpenoid | EJ | LG8 | 28.1 | 2.91 | 1.59 | 0.09 | 25.8 | 30.7 |
| 34 | cis-Linaloloxide | C5b | Terpenoid | EJ | LG8 | 38.4 | 3.21 | 1.35 | 0.10 | 30.7 | 48.1 |
| 81 | Eugenol | C12b | Aromatic alcohol | EJ | LG8 | 3.8 | 3.56 | -0.53 | 0.15 | 0.0 | 12.4 |
| 74 | Furan, 2-ethyl- | C11 | Lipid-derived | EJ | LG8 | 8.9 | 2.74 | -0.64 | 0.10 | 0.0 | 16.6 |
| 78 | Furan, 2-pentyl- | C12 | Lipid-derived | EJ | LG8 | 6.4 | 3.45 | -0.80 | 0.14 | 0.0 | 10.2 |
| 68 | Heptanol | C9 | Alcohol | AA | LG8 | 3.0 | 1.75 | -0.16 | 0.09 | 0.0 | 7.3 |
| 68 | Heptanol | C9 | Alcohol | AA | LG8 | 20.0 | 2.03 | 0.28 | 0.09 | 16.0 | 22.0 |
| 32 | Linalool | C5a | Terpenoid | EJ | LG8 | 8.9 | 3.07 | -2.54 | 0.10 | 0.0 | 12.8 |
| 32 | Linalool | C5a | Terpenoid | EJ | LG8 | 16.6 | 2.75 | -2.42 | 0.09 | 12.8 | 18.6 |
| 32 | Linalool | C5a | Terpenoid | EJ | LG8 | 38.4 | 2.76 | 2.32 | 0.09 | 33.2 | 48.3 |
| 67 | Nonalol | C9 | Alcohol | EJ | LG8 | 37.0 | 3.41 | 0.38 | 0.18 | 27.6 | 45.0 |
| 67 | Nonalol | C9 | Alcohol | EJ | LG8 | 49.0 | 2.49 | -0.32 | 0.13 | 47.7 | 59.0 |
| 19 | Nonanoic acid | C3 | Carboxilic acid | AA | LG8 | 58.8 | 2.98 | 1.42 | 0.19 | 53.4 | 62.8 |
| 46 | p-Cymene | C5c | Terpenoid | EJ | LG8 | 3.6 | 3.04 | -0.69 | 0.12 | 0.0 | 9.9 |

**Table S1. Volatile QTL detected for 'MxR_01' map.** For each QTL, the annotation of the volatile that controls (Nº, Name, Cluster and Family) and the location (EJ or AA) where was found, the linkage group (LG), the position in cM (position), the likelihood

(LOD), the additive effect (additive), the proportion of the phenotipic variance explained (R2) and the 2-LOD confidence interval are shown. All the QTL showed are significant assessed by 1000-permutation test at $\alpha$=0.05.

| Nº | Volatile | Cluster | Family | Location | LG | Position | LOD | Additive | R2 | LOD2-left | LOD2-right |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 3,4-Dimethyl-3-hexanol | C1 | Alcohol | EJ | LG2 | 47.5 | 2.65 | -0.14 | 0.13 | 42.30 | 57.50 |
| 2 | 3,4-Dimethyl-3-hexanol | C1 | Alcohol | EJ | LG6 | 34.2 | 2.72 | -0.13 | 0.13 | 23.30 | 49.90 |
| 70 | 1-Hexanol, 2-ethyl- | C10 | Alcohol | AA | LG8 | 5.6 | 3.10 | -0.92 | 0.17 | 0.00 | 8.30 |
| 71 | 1-Octanol | C10 | Alcohol | EJ | LG7 | 7.1 | 2.37 | -0.51 | 0.12 | 5.70 | 20.40 |
| 75 | 1-Penten-3-one | C11 | Lipid-derived | EJ | LG6 | 20.7 | 3.05 | 0.87 | 0.16 | 10.50 | 34.00 |
| 76 | 2,4-Heptadienal, (E,E)- | C11 | Lipid-derived | EJ | LG6 | 18.7 | 3.15 | 0.74 | 0.18 | 10.50 | 33.70 |
| 73 | 3-Hexenal, (Z)- | C11 | Lipid-derived | EJ | LG6 | 34.2 | 3.13 | 0.52 | 0.17 | 23.30 | 37.10 |
| 78 | Furan, 2-pentyl- | C12 | Lipid-derived | EJ | LG6 | 21.7 | 2.84 | 0.58 | 0.14 | 10.50 | 34.00 |
| 79 | Hexanal | C12 | Lipid-derived | AA | LG4 | 10.3 | 2.71 | -0.24 | 0.18 | 0.00 | 18.80 |
| 77 | Pentanal | C12 | Lipid-derived | EJ | LG6 | 26.7 | 2.49 | 1.20 | 0.14 | 10.40 | 34.10 |
| 13 | 2-Propenoic acid, 2-ethylhexyl ester | C2 | Long Ketone/Ester | EJ | LG6 | 23.0 | 2.27 | -0.45 | 0.12 | 10.70 | 34.10 |
| 5 | 8,8,9-Trimethyl-deca-3,5-diene-2,7-dione | C2 | Long Ketone/Ester | EJ | LG2 | 0.0 | 2.79 | 0.17 | 0.15 | 0.00 | 29.10 |
| 5 | 8,8,9-Trimethyl-deca-3,5-diene-2,7-dione | C2 | Long Ketone/Ester | AA | LG4 | 19.6 | 3.53 | 0.60 | 0.22 | 16.10 | 20.60 |
| 10 | Acetic acid, 2-ethylhexyl ester | C2 | Long Ketone/Ester | EJ | LG6 | 23.0 | 2.97 | -0.28 | 0.14 | 13.50 | 33.80 |
| 6 | Decanoic acid, methyl ester | C2 | Long Ketone/Ester | AA | LG1 | 10.0 | 2.46 | 0.32 | 0.17 | 0.00 | 11.00 |
| 6 | Decanoic acid, methyl ester | C2 | Long Ketone/Ester | EJ | LG6 | 23.0 | 3.04 | -0.37 | 0.15 | 16.10 | 33.70 |
| 20 | Decanoic acid | C3 | Carboxilic acid | EJ | LG7 | 23.6 | 4.72 | -2.01 | 0.23 | 21.60 | 27.70 |
| 20 | Decanoic acid | C3 | Carboxilic acid | EJ | LG7 | 41.3 | 5.02 | 1.68 | 0.24 | 37.50 | 46.20 |
| 18 | Hexanoic acid | C3 | Carboxilic acid | AA | LG8 | 4.6 | 2.50 | -3.75 | 0.15 | 0.00 | 5.90 |
| 16 | Hexanoic acid, 2-ethyl- | C3 | Carboxilic acid | AA | LG6 | 54.7 | 3.32 | -0.85 | 0.12 | 50.50 | 61.70 |
| 16 | Hexanoic acid, 2-ethyl- | C3 | Carboxilic acid | AA | LG8 | 7.2 | 2.75 | -0.59 | 0.16 | 2.60 | 14.70 |
| 19 | Nonanoic acid | C3 | Carboxilic acid | EJ | LG7 | 23.6 | 2.47 | -1.08 | 0.11 | 13.00 | 29.30 |
| 21 | 2,7-Ethanonaphth[2,3-b]oxirene, 1a,2,7,7a-tetrahydro-, (1aa,2a,7a,7aa)- | C4 | un. | EJ | LG1 | 1.0 | 2.93 | 0.17 | 0.14 | 0.00 | 10.00 |
| 21 | 2,7-Ethanonaphth[2,3-b]oxirene, 1a,2,7,7a-tetrahydro-, (1aa,2a,7a,7aa)- | C4 | un. | EJ | LG6 | 10.7 | 2.67 | 0.21 | 0.12 | 0.90 | 21.70 |
| 25 | Acetic acid, phenylmethyl ester | C4 | Aromatic Ester | EJ | LG7 | 41.3 | 4.05 | 0.15 | 0.21 | 29.80 | 48.90 |
| 22 | Isoeugenol methyl ether | C4 | Aromatic Ether | EJ | LG7 | 41.3 | 2.17 | 0.14 | 0.11 | 29.40 | 48.90 |
| 28 | Salicylic acid, ethyl ester | C4 | Aromatic Ester | EJ | LG7 | 45.3 | 2.34 | 0.18 | 0.10 | 29.50 | 48.90 |
| 31 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | C5a | Terpenoid | AA | LG5 | 0.0 | 2.66 | 3.00 | 0.16 | 0.00 | 9.30 |
| 36 | Ethanone, 1-(4-methylphenyl)- | C5b | un. | AA | LG5 | 0.0 | 3.86 | 1.56 | 0.22 | 0.00 | 9.20 |
| 36 | Ethanone, 1-(4-methylphenyl)- | C5b | un. | AA | LG5 | 20.8 | 2.83 | -1.36 | 0.16 | 20.40 | 32.50 |
| 45 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5b | un. | AA | LG5 | 0.0 | 3.06 | 4.09 | 0.17 | 0.00 | 9.80 |
| 45 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5b | un. | EJ | LG7 | 9.1 | 3.30 | -2.53 | 0.15 | 0.00 | 10.40 |
| 44 | 2-Cyclohexen-1-ol, 2-methyl-5-(1-methylethenyl)-, cis | C5b | un. | EJ | LG7 | 29.3 | 3.97 | 4.68 | 0.19 | 26.60 | 37.20 |
| 44 | 2-Cyclohexen-1-ol, 2-methyl-5-(1-methylethenyl)-, cis | C5b | un. | EJ | LG7 | 42.3 | 3.45 | -3.52 | 0.14 | 41.20 | 48.90 |
| 43 | 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl | C5b | Terpenoid | EJ | LG2 | 46.9 | 4.87 | 3.75 | 0.24 | 42.40 | 51.60 |
| 43 | 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl | C5b | Terpenoid | AA | LG2 | 47.5 | 2.59 | 1.98 | 0.13 | 39.80 | 55.30 |
| 43 | 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl | C5b | Terpenoid | AA | LG5 | 0.0 | 4.09 | 3.85 | 0.22 | 0.00 | 9.20 |
| 43 | 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl | C5b | Terpenoid | AA | LG5 | 21.8 | 2.63 | -3.09 | 0.16 | 13.30 | 33.80 |
| 37 | 4-Acetyl-1-methylcyclohexene | C5b | Terpenoid | EJ | LG2 | 47.5 | 2.75 | 1.30 | 0.15 | 41.80 | 52.50 |
| 35 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | AA | LG5 | 0.0 | 3.80 | 2.57 | 0.22 | 0.00 | 8.50 |
| 34 | cis-Linaloloxide | C5b | Terpenoid | AA | LG5 | 0.0 | 3.43 | 2.92 | 0.20 | 0.00 | 9.20 |
| 53 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | AA | LG7 | 29.3 | 5.21 | -2.69 | 0.29 | 28.20 | 34.90 |
| 53 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | AA | LG7 | 41.3 | 4.56 | 2.39 | 0.23 | 40.90 | 44.10 |
| 50 | Acetic acid, hexyl ester | C7 | Ester | AA | LG4 | 19.6 | 4.31 | 2.20 | 0.22 | 15.20 | 20.60 |
| 57 | d-Decalactone | C8 | Lactone | EJ | LG2 | 49.0 | 2.98 | -0.62 | 0.14 | 44.00 | 57.00 |
| 57 | d-Decalactone | C8 | Lactone | AA | LG4 | 19.0 | 2.36 | 1.23 | 0.13 | 14.00 | 20.00 |
| 58 | g-Decalactone | C8 | Lactone | AA | LG4 | 19.0 | 3.63 | 1.06 | 0.17 | 10.20 | 20.00 |
| 62 | g-Heptalactone | C8 | Lactone | EJ | LG1 | 10.0 | 2.29 | -0.32 | 0.12 | 0.00 | 11.00 |
| 55 | g-Hexalactone | C8 | Lactone | EJ | LG6 | 3.0 | 1.22 | 0.09 | 0.07 | 0.00 | 21.90 |
| 56 | Pyran-2-one, 6-pentyl- | C8 | Lactone | EJ | LG7 | 16.0 | 2.27 | 0.77 | 0.17 | 9.10 | 22.00 |
| 65 | 5-Hepten-2-one, 6-methyl- | C9 | Carotenoid-derived | EJ | LG4 | 15.5 | 2.65 | 0.25 | 0.11 | 0.00 | 19.60 |
| 65 | 5-Hepten-2-one, 6-methyl- | C9 | Carotenoid-derived | EJ | LG7 | 37.9 | 2.76 | -0.41 | 0.13 | 32.60 | 38.90 |
| 64 | b-Damascenone | C9 | Carotenoid-derived | AA | LG1 | 10.0 | 3.71 | 0.59 | 0.18 | 0.00 | 11.00 |
| 64 | b-Damascenone | C9 | Carotenoid-derived | AA | LG6 | 3.3 | 2.53 | 0.46 | 0.11 | 0.00 | 21.70 |
| 68 | Heptanol | C9 | Alcohol | AA | LG7 | 40.4 | 2.58 | 0.49 | 0.17 | 29.60 | 48.90 |

**Table S2. Volatile QTL detected for 'Granada' map.** For each QTL, the annotation of the volatile that controls (Nº, Name, Cluster and Family) and the location (EJ or AA) where was found, the position in cM (position), the likelihood (LOD), the additive effect (additive), the proportion of the phenotipic variance explained (R2) and the 2-LOD confidence interval are shown. All the QTL showed are significant assessed by 1000-permutation test at $\alpha$=0.05.

| MxR | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Trait | location | LG | Position | LOD | Additive | R2 | LOD2-left | LOD2-right |
| C | EJ | LG6 | 11.0 | 1.26 | -4.07 | 0.04 | 0.0 | 18.7 |
| C | EJ | LG7 | 59.2 | 3.01 | -3.30 | 0.11 | 54.5 | 64.8 |
| C | EJ | LG7 | 67.6 | 3.70 | -3.54 | 0.13 | 64.8 | 68.6 |
| Firmness | AA | LG3 | 72.7 | 2.80 | 11.73 | 0.14 | 68.2 | 83.4 |
| Firmness | AA | LG4 | 60.2 | 6.17 | -18.07 | 0.32 | 51.2 | 66.6 |
| Firmness | EJ | LG4 | 60.2 | 12.54 | -18.77 | 0.43 | 58.3 | 67.0 |
| Firmness | IVIA | LG4 | 68.0 | 4.36 | -16.20 | 0.27 | 65.0 | 69.0 |
| Firmness | EJ | LG6 | 52.1 | 3.44 | -9.09 | 0.10 | 41.7 | 59.1 |
| Firmness | EJ | LG7 | 50.9 | 2.83 | 8.74 | 0.09 | 46.4 | 56.4 |
| Firmness | IVIA | LG8 | 8.9 | 2.85 | -11.80 | 0.13 | 7.7 | 12.8 |
| H | AA | LG3 | 74.1 | 3.81 | 7.07 | 0.20 | 65.1 | 85.9 |
| H | EJ | LG7 | 66.2 | 5.04 | 6.25 | 0.20 | 60.4 | 68.6 |
| L | EJ | LG3 | 65.2 | 2.82 | -1.90 | 0.10 | 55.1 | 67.2 |
| L | EJ | LG3 | 74.1 | 2.66 | -1.88 | 0.09 | 72.7 | 83.1 |
| L | AA | LG4 | 57.4 | 3.81 | -2.44 | 0.18 | 52.9 | 59.2 |
| L | EJ | LG6 | 3.0 | 5.88 | -2.95 | 0.24 | 0.0 | 11.0 |
| L | EJ | LG8 | 22.9 | 3.86 | -2.89 | 0.15 | 19.4 | 28.1 |
| L | EJ | LG8 | 30.7 | 3.68 | -3.23 | 0.15 | 28.1 | 31.8 |
| HD | AA | LG1 | 73.1 | 10.58 | -11.30 | 0.16 | 60.0 | 75.0 |
| HD | EJ | LG1 | 73.2 | 6.73 | -15.80 | 0.18 | 64.0 | 75.0 |
| HD | IVIA | LG1 | 73.5 | 8.55 | -14.87 | 0.19 | 65.2 | 75.0 |
| HD | EJ | LG4 | 51.8 | 23.38 | -20.27 | 0.52 | 49.8 | 56.9 |
| HD | AA | LG4 | 52.9 | 18.05 | -27.33 | 0.51 | 50.0 | 56.9 |
| HD | IVIA | LG4 | 52.9 | 16.40 | -22.47 | 0.45 | 50.4 | 56.9 |
| HD | IVIA | LG6 | 6.0 | 4.65 | 11.01 | 0.11 | 0.0 | 19.1 |
| HD | AA | LG6 | 11.1 | 3.29 | 11.28 | 0.08 | 0.0 | 20.9 |
| HD | EJ | LG6 | 11.1 | 5.50 | 8.40 | 0.09 | 1.2 | 19.9 |
| HD | AA | LG7 | 4.0 | 3.55 | 11.56 | 0.09 | 0.0 | 5.3 |
| HD | EJ | LG8 | 12.8 | 3.48 | 6.84 | 0.06 | 12.2 | 31.9 |
| MnM | AA | LG4 | 64.6 | 18.46 | -0.84 | 0.66 | 64.5 | 68.0 |
| MnM | EJ | LG4 | 64.6 | 18.46 | -0.84 | 0.66 | 64.5 | 68.0 |
| MnM | IVIA | LG4 | 68.0 | 30.37 | -1.86 | 0.89 | 67.5 | 69.0 |
| SSC | IVIA | LG3 | 54.3 | 2.75 | -2.62 | 0.15 | 49.2 | 58.0 |
| SSC | AA | LG4 | 54.9 | 12.36 | -3.28 | 0.46 | 50.0 | 57.3 |
| SSC | AA | LG8 | 46.1 | 3.34 | 1.47 | 0.10 | 45.5 | 53.8 |
| Weight | EJ | LG4 | 67.0 | 5.34 | -36.25 | 0.25 | 58.3 | 68.0 |

| Granada | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Trait | location | LG | Position | LOD | Additive | R2 | LOD2-left | LOD2-right |
| Weight | AA | LG4 | 14.1 | 2.75 | 68.34 | 0.14 | 10.3 | 18.9 |
| Weight | EJ | LG6 | 12.7 | 2.72 | -24.75 | 0.14 | 0.0 | 21.9 |
| Weight | AA | LG6 | 6.3 | 3.08 | -21.66 | 0.16 | 0.0 | 21.8 |
| SSC | IVIA | LG6 | 22.0 | 2.45 | -1.99 | 0.13 | 10.4 | 33.7 |
| SSC | IVIA | LG6 | 53.2 | 2.64 | 2.05 | 0.14 | 41.3 | 65.1 |
| L | EJ | LG6 | 0.0 | 3.07 | 2.25 | 0.13 | 0.0 | 10.4 |
| L | AA | LG6 | 12.7 | 2.80 | 3.64 | 0.16 | 10.4 | 33.7 |
| C | AA | LG6 | 10.7 | 4.08 | 4.80 | 0.20 | 10.3 | 30.5 |

**Table S3. QTL for fruit type and maturity related traits.** For each QTL, the location (EJ, AA or IVIA) where was found (indicated in Trait), the linkage group (LG), the likelihood (LOD), the additive effect (additive), the proportion of the phenotipic variance explained (R2) and 2-LOD confidence interval are shown. All the QTL showed are significant assessed by 1000-permutation test at

α=0.05 are shown. The traits analyzed are: Melting/noMelting fruit type (MnM), flesh firmness (Firmness), fruit weight (weight), solid soluble content (SSC), peel ground color parameters (L, lightness, C, chroma and H, color measured in Hue degree), and maturity date (MD). The QTL detected in 'MxR_01' and 'Granada' maps are listed at the top and the bottom of the table, respectively.

| N | id | Name | Cluster | Family | Melting | No-Melting | p |
|---|---|---|---|---|---|---|---|
| 1 | 43_7136 | Acetic acid, butyl ester | C1 | Ester | -0.124 | 0.051 | 0.006 |
| 3 | 41_6638 | Propanoic acid, 2,2-dimethyl- | C1 | Carboxilic acid | -0.424 | -0.009 | 0.024 |
| 15 | 91_11117 | Benzeneacetaldehyde | C2b | Aromatic | -0.603 | -1.727 | 0.000 |
| 29 | 121_11693 | Terpinolene | C5 | Terpenoid | -1.296 | -2.064 | 0.004 |
| 30 | 121_13280 | a-Terpinol | C5 | Terpenoid | -1.461 | -2.126 | 0.009 |
| 31 | 139_9925 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | C5 | Terpenoid | -2.031 | -3.541 | 0.002 |
| 32 | 93_11733 | Linalool | C5 | Terpenoid | -3.787 | -5.796 | 0.001 |
| 34 | 59_11450 | cis-Linaloloxide | C5 | Terpenoid | -0.641 | -2.180 | 0.000 |
| 35 | 59_11676 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5 | un. | -0.518 | -1.817 | 0.000 |
| 36 | 119_13188 | Ethanone, 1-(4-methylphenyl)- | C5 | un. | 0.293 | -0.255 | 0.002 |
| 37 | 138_12371 | 4-Acetyl-1-methylcyclohexene | C5 | Terpenoid | -0.448 | -1.868 | 0.000 |
| 38 | 117_11758 | Benzene, 4-ethenyl-1,2-dimethyl- | C5 | un | 0.051 | -0.821 | 0.007 |
| 39 | 43_13124 | Benzenemethanol, a,a,4-trimethyl- | C5 | un. | -0.739 | -1.721 | 0.028 |
| 40 | 123_10728 | 1-Cyclohexene-1-carboxaldehyde,2,6,6-trimethyl | C5 | Terpenoid | 0.086 | -1.570 | 0.006 |
| 41 | 134_12449 | 1,3,8-p-Menthatriene | C5 | Terpenoid | -0.764 | -2.455 | 0.002 |
| 42 | 137_13215 | 3,6-Dimethyl-2,3,3a,4,5,7a-hexahydrobenzofuran | C5 | un. | -0.945 | -2.909 | 0.000 |
| 43 | 94_13611 | 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl | C5 | Terpenoid | -2.470 | -3.832 | 0.011 |
| 44 | 69_13073 | 2-Cyclohexen-1-ol, 2-methyl-5-(1-methylethenyl)-, cis | C5 | un. | -1.818 | -3.570 | 0.001 |
| 45 | 71_11820 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5 | un. | -3.053 | -4.958 | 0.001 |
| 49 | 68_15955 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | C6b | Lactone | 0.489 | 1.307 | 0.000 |
| 52 | 43_10351 | 2-Hexen-1-ol, acetate, (E)- | C7 | Ester | -0.433 | -1.490 | 0.012 |
| 53 | 80_10583 | 3-Cyclohexen-1-ol, acetate | C7 | Ester | -0.253 | 0.195 | 0.041 |
| 54 | 43_6475 | Acetic acid, 2-methylpropyl ester | C8 | Ester | -1.210 | -0.286 | 0.022 |
| 56 | 95_16515 | Pyran-2-one, 6-pentyl- | C8 | Lactone | 0.094 | 0.853 | 0.002 |
| 57 | 99_16910 | d-Decalactone | C8 | Lactone | -0.126 | 0.834 | 0.001 |
| 58 | 85_16556 | g-Decalactone | C8 | Lactone | -0.145 | 0.427 | 0.010 |
| 59 | 85_14019 | g-Octalactone | C8 | Lactone | -0.104 | 0.845 | 0.000 |
| 60 | 85_15326 | g-Nonalactone | C8 | Lactone | 0.175 | 1.087 | 0.000 |
| 61 | 68_16405 | g-Jasmolactone | C8 | Lactone | -0.185 | 0.916 | 0.000 |
| 62 | 85_12613 | g-Heptalactone | C8 | Lactone | -0.144 | 0.459 | 0.001 |
| 63 | 177_16736 | b-Ionone | C9 | Carotenoid-derived | 1.801 | 1.183 | 0.012 |
| 77 | 44_5291 | Pentanal | C12 | Lipid-derived | 1.914 | 1.430 | 0.019 |
| 79 | 56_6998 | Hexanal | C12 | Lipid-derived | -0.077 | -0.297 | 0.034 |
| 80 | 83_7941 | 2-Hexenal, (E)- | C12 | Lipid-derived | 0.039 | -0.172 | 0.000 |
| 81 | 164_15243 | Eugenol | C12b | Aromatic alcohol | 0.809 | 0.491 | 0.011 |

**Table S4. Difference in volatile levels between melting vs no-melting peaches.** The mean of levels (expressed as a Log2 sample/reference) of melting and no-melting genotypes are shown. The difference were stated by ANOVA analysis, the p- value (p) obtained for each volatile is showed.

| | Harvest | | | |
|---|---|---|---|---|
| **Fruit type** | **Early** | **Medium** | **Late** | **Total** |
| **Melting** | 83% | 67% | 21% | **54%** |
| **no-Melting** | 17% | 33% | 79% | **46%** |

**Table S5. Percentage of melting/no-melting peaches in early, medium and late genotypes.**

| Nº | id | Name | Cluster | Family | monoterpene-rich | Rest of genotypes | p |
|---|---|---|---|---|---|---|---|
| 1 | 43_7136 | Acetic acid, butyl ester | C1 | Ester | -0.134 | 0.018 | 0.006 |
| 29 | 121_1169 | Terpinolene | C5a | Terpenoid | -0.831 | -2.205 | 0.000 |
| 30 | 121_1328 | a-Terpinol | C5a | Terpenoid | -1.011 | -2.301 | 0.000 |
| 31 | 139_9925 | 2H-Pyran, 2-ethenyltetrahydro-2,6,6-trimethyl- | C5a | Terpenoid | -1.092 | -3.723 | 0.000 |
| 32 | 93_11733 | Linalool | C5a | Terpenoid | -2.672 | -5.987 | 0.000 |
| 33 | 93_10867 | D-Limonene | C5b | Terpenoid | -0.199 | -1.141 | 0.000 |
| 34 | 59_11450 | cis-Linaloloxide | C5b | Terpenoid | -0.113 | -2.202 | 0.000 |
| 35 | 59_11676 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5b | un. | -0.075 | -1.811 | 0.000 |
| 36 | 119_1318 | Ethanone, 1-(4-methylphenyl)- | C5b | un. | 0.698 | -0.382 | 0.000 |
| 37 | 138_1237 | 4-Acetyl-1-methylcyclohexene | C5b | Terpenoid | 0.265 | -2.028 | 0.000 |
| 38 | 117_1175 | Benzene, 4-ethenyl-1,2-dimethyl- | C5b | un | 0.743 | -0.935 | 0.000 |
| 39 | 43_13124 | Benzenemethanol, a,a,4-trimethyl- | C5b | un. | 0.259 | -1.961 | 0.000 |
| 40 | 123_1072 | 1-Cyclohexene-1-carboxaldehyde,2,6,6-trimethyl | C5b | Terpenoid | 1.310 | -1.880 | 0.000 |
| 41 | 134_1244 | 1,3,8-p-Menthatriene | C5b | Terpenoid | 0.282 | -2.662 | 0.000 |
| 42 | 137_1321 | 3,6-Dimethyl-2,3,3a,4,5,7a-hexahydrobenzofuran | C5b | un. | -0.064 | -2.996 | 0.000 |
| 43 | 94_13611 | 3-Cyclohexene-1-acetaldehyde, a,4-dimethyl | C5b | Terpenoid | -1.298 | -4.380 | 0.000 |
| 44 | 69_13073 | 2-Cyclohexen-1-ol, 2-methyl-5-(1-methylethenyl)-, cis | C5b | un. | -0.918 | -3.803 | 0.000 |
| 45 | 71_11820 | 1,5,7-Octatrien-3-ol, 3,7-dimethyl- | C5b | un. | -1.991 | -5.212 | 0.000 |
| 46 | 119_1078 | p-Cymene | C5b | Terpenoid | 0.433 | -0.100 | 0.002 |
| 67 | 56_12707 | Nonalol | C9 | Alcohol | 0.130 | -0.141 | 0.000 |
| 71 | 56_11249 | 1-Octanol | C10 | Alcohol | -0.154 | -0.395 | 0.022 |

**Table S6. Difference in volatile levels between monoterpene-rich ideotype vs the rest of the genotype.** The mean of levels (expressed as a Log2 sample/reference) of the ideotype and the rest of genotypes are shown. The difference were stated by ANOVA analysis, the p- value (p) obtained for each volatile is showed.

| N | id | Name | Cluster | Family | Lactone-rich | Rest of genotype | p |
|---|---|---|---|---|---|---|---|
| 18 | 60_9627 | Hexanoic acid | C3 | Carboxilic acid | -3.354 | 0.599 | 0.029 |
| 35 | 59_11676 | a-Methyl-a-[4-methyl-3-pentenyl]oxiranemethanol | C5 | un. | -2.165 | -0.973 | 0.047 |
| 44 | 69_13073 | 2-Cyclohexen-1-ol, 2-methyl-5-(1-methylethenyl)-, cis | C5 | un. | -4.456 | -2.417 | 0.037 |
| 49 | 68_15955 | 4-Methyl-5-penta-1,3-dienyltetrahydrofuran-2-one | C6b | Lactone | 1.606 | 0.798 | 0.024 |
| 59 | 85_14019 | g-Octalactone | C8 | Lactone | 1.169 | 0.218 | 0.020 |
| 60 | 85_15326 | g-Nonalactone | C8 | Lactone | 1.328 | 0.483 | 0.016 |
| 61 | 68_16405 | g-Jasmolactone | C8 | Lactone | 1.280 | 0.183 | 0.035 |
| 62 | 85_12613 | g-Heptalactone | C8 | Lactone | 0.867 | 0.020 | 0.032 |
| 68 | 56_9747 | Heptanol | C9b | Alcohol | -0.404 | -0.039 | 0.013 |
| 77 | 44_5291 | Pentanal | C12 | Lipid-derived | 0.891 | 1.746 | 0.034 |

**Table S7. Difference in volatile levels between lactone-rich ideotype vs the rest of the genotype.** The mean of levels (expressed as a Log2 sample/reference) of the ideotype and the rest of genotypes are shown. The difference were stated by ANOVA analysis, the p- value (p) obtained for each volatile is showed