



**CENTRO DE INVESTIGACIÓN Y ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO
NACIONAL**

**UNIDAD IRAPUATO
DEPARTAMENTO DE INGENIERÍA GENÉTICA**

**“Enfoques metabólicos para fenotipo y
selección de plantas de maíz”**

Tesis que presenta

M.C. Martín García Flores

Para obtener el grado de

DOCTOR EN CIENCIAS

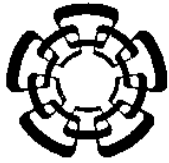
En la especialidad de Biotecnología de Plantas

Director de Tesis

Dr. Axel Tiessen Favier

Irapuato, México

Fecha de Examen:
24 de Julio 2017



**CENTRO DE INVESTIGACIÓN Y ESTUDIOS
AVANZADOS DEL INSTITUTO POLITÉCNICO
NACIONAL**

**UNIDAD IRAPUATO
DEPARTAMENTO DE INGENIERÍA GENÉTICA**

**“Enfoques metabolómicos para fenotipo y
selección de plantas de maíz”**

Tesis que presenta

M.C. Martín García Flores

Para obtener el grado de

DOCTOR EN CIENCIAS

En la especialidad de Biotecnología de plantas

COMITÉ DE TESIS

Director

Dr. Axel Tiessen Favier

Firma

Asesor

Dr. Robert Winkler

Firma

Asesor

Dr. John Paul Délano Frier

Firma

Asesor

Dr. Ruairidh J. H. Sawers

Firma

Asesora

Dra. Graciela Ma. De La Luz Ruiz Aguilar

Firma

El presente trabajo fue realizado en el **Laboratorio de Metabolómica y Fisiología Molecular** del Departamento de Ingeniería Genética de Plantas del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Unidad Irapuato, bajo la asesoría del Doctor Axel Tiessen.

Declaración:

Por este medio declaro que yo he realizado los experimentos y preparado este trabajo de tesis de forma independiente y sin ayuda externa. Especialmente declaro que he citado de forma correcta y explícita a los autores y trabajos en los que esta tesis se apoya, así como las contribuciones de las personas que coadyuvaron en su desarrollo.

MC. Martín García Flores

Irapuato, Guanajuato a 24 de Julio de 2017

Se agradece al Consejo Nacional de Ciencia y Tecnología (Conacyt) por el apoyo económico otorgado (Beca No. 208598) para estudios de Doctorado, durante el período comprendido entre septiembre de 2011 y agosto de 2015. Así como, al Centro de Investigación y de Estudios Avanzados del IPN, Unidad Irapuato, y en particular al Departamento de Ingeniería Genética por proporcionar los recursos humanos, materiales y económicos necesarios para el desarrollo y terminación de este trabajo. Se agradece a los investigadores y al personal de campo del Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT) que nos apoyaron durante las cosechas y los muestreos en las estaciones experimentales. Se agradece también al Laboratorio Nacional PlanTECC por la infraestructura brindada para la realización exitosa de este proyecto.

AGRADECIMIENTOS

Al Doctor Axel Tiessen por haberme brindado la oportunidad única de participar en sus proyectos de investigación científica en maíz, por todo el aprendizaje que me facilitó durante mi estancia en su laboratorio, por transmitir su entusiasmo en las actividades del día a día, por su compromiso con el rigor analítico de hipótesis y descubrimientos en investigación, por participarme de sus experiencias en las Ciencias Biológicas y por facilitarme el reencuentro con la Naturaleza.

A los integrantes de mi Comité de Tesis: Dr. Robert Winkler, Dr. John Délano Frier, Dr. Ruairidh J.H. Sawers y Dra. Graciela Ma. De La Luz Ruiz Aguilar por su acompañamiento, orientación y participación en las sesiones tutoriales que me permitieron alcanzar los frutos académicos de mi experiencia en el programa de Biotecnología de Plantas en el Cinvestav Unidad Irapuato.

Al comité de ingreso de la generación 2011-2015 del programa de posgrado, en especial a los Doctores Juan José Peña Cabriales y Alfredo Herrera Estrella.

A todo el personal académico, administrativo, auxiliares, técnicos, laboratoristas y de operación del Cinvestav Unidad Irapuato por su actitud y disposición a participar y colaborar.

A los integrantes del cuerpo académico del posgrado por su entusiasmo y dedicación con las causas que significan comprometerse con el reto de elevar la calidad educativa de nuestra Patria.

A los integrantes del Laboratorio de Metabolómica y Fisiología Molecular por la constancia, disciplina y desempeño que aplican en el logro de sus metas. Agradezco el apoyo administrativo de Lety Chong y de Alejandro Hernández y el apoyo técnico del Dr. Andrés Estrada y Norma Cecilia Morales para compras y asuntos del laboratorio.

CONTENIDO

RESUMEN	xi
SUMMARY	xii
1. INTRODUCCIÓN GENERAL	1
2. ANTECEDENTES	4
2.1 Metabolómica	4
2.1.1 Definición e importancia	4
2.1.2 Estrategias en metabolómica	4
2.1.3 Metabolómica y fisiología de plantas	6
2.2 Metabolismo	6
2.2.1 Características del metabolismo de plantas	6
2.2.2 Metabolismo de plantas	6
2.2.3 Vínculo del metaboloma con el fenotipo.....	7
2.3 Experimentos de estrés hídrico y bajo nitrógeno en maíz	9
2.3.1 Estrés hídrico	9
2.3.2 Bajo contenido de nitrógeno o “Bajo Nitrógeno”	10
2.4 Cromatografía de gases	13
2.5 Espectrometría de masas	13
2.6 Métodos estadísticos para análisis de la información.....	15
2.6.1 Regresión lineal “paso a paso” (o “step wise”).	16
3. HIPÓTESIS GENERAL.....	17
4. OBJETIVOS	18
4.1 Objetivo General.....	18
4.2 Objetivos específicos	18
5. MATERIALES Y MÉTODOS.....	19
5.1 Preguntas biológicas	19
5.2 Material biológico	20
5.2.1 Cultivo del Maíz y preparación de muestras de jugo del tallo	20
5.2.1.1 Experimento de estrés hídrico.....	20
5.2.1.2 Experimento de estrés por bajo nitrógeno	20
5.2.1.3 Experimento fuente-sumidero (o “sink-source”)	21
5.2.1.4 Experimento etapas de desarrollo 2010	22
5.2.1.5 Experimento etapas de desarrollo 2011	22
5.3 Experimentos preliminares	22
5.4 Recolección del jugo de tallo de maíz	24
5.5 Análisis de metabolitos.	25
5.5.1 Ensayos enzimáticos espectrofotométricos.	25
5.5.2 Espectrometría de masas con inyección directa ionización por electrospray (DIESI-MS)	25
5.5.4 Cromatografía de Gases (GC-FID). Condiciones de operación.	26
5.6 Análisis estadístico de datos de DIESI-MS con programa R	27
5.7 Diseño experimental en campo y en laboratorio	29
5.7 Definición de jugo	29
6. RESULTADOS	31
6.1 Experimentos control sobre DIESI-MS.....	31
6.1.1 Linealidad	33
6.1.2 Experimentos principales con DIESI-MS	36
6.1.2 Bajo contenido de nitrógeno en Cimmyt Batán.....	42
6.1.3 Experimento Fuente-Sumidero (Sink-Source) en Irapuato	48
6.1.4 Etapas de desarrollo 2010 en Irapuato.....	55

6.1.5 Etapas de desarrollo 2011 en Irapuato.....	60
6.2 Experimentos principales con GC	64
6.2.1 Estrés hídrico en Tlaltizapán y bajo contenido de nitrógeno en El Batán.....	64
6.3 Comparación de características agronómicas y mediciones bioquímicas.....	67
6.3.2 Comparación de niveles de azúcares no-estructurales, aminoácidos y características agronómicas.....	69
6.3.3 Producción de biomasa seca de los experimentos de estrés hídrico en Tlaltizapán y bajo nitrógeno en Batán.....	71
6.3.4 Producción de grano del experimento de bajo contenido de nitrógeno en El Batán.....	72
6.3.5 Ecuaciones de regresión lineal múltiple para predicción de rendimiento de biomasa fresca y grano.....	73
6.3.6 Validación de las ecuaciones de regresión múltiple para predicción de rendimiento de biomasa fresca y grano.....	74
7. DISCUSIÓN.....	76
7.1 Experimentos control.....	77
7.2 Fenotipo	77
7.3 Selección.....	78
7.4 Ecuaciones lineales de predicción	78
7.5 Experimento de estrés hídrico	79
7.6 Experimento de bajo contenido de nitrógeno	80
7.7 Experimento fuente-sumidero	80
7.8 Experimento etapas de desarrollo.....	80
7.10 Oportunidades en fisiología de plantas.....	81
8. CONCLUSIONES.....	83
9. BIBLIOGRAFÍA	87
10. ANEXOS	92
10.1 Publicación No.1	92
10.2 Publicación No. 2	95
10.3 Anexos.....	106

Índice de Cuadros

Cuadro 2.1 Estrategias analíticas en investigación metabolómica (adaptado de Koek et al., 2011) 5

Índice de Figuras

Figura 2.1 Esquema de las relaciones entre el fenotipo y el genotipo (Tiessen, 2009). Las funciones de la fisiología y los metabolitos están más próximos al fenotipo (por ejemplo, el rendimiento agronómico) que la secuencia del ADN. 5

Figura 2.2 Síntesis de alcoholes bajo condiciones limitantes de oxígeno. El etanol, ácido láctico y ácido acético son reciclados para suministrar al metabolismo central de el NADPH requerido en la etapa citosólica de la glucólisis (Tiessen, 2009). 12

Figura 2.3 En comparación con el aumento exponencial de los métodos de genotipo, las estrategias de fenotipo se están quedando atrás en las ciencias agrícolas (Tiessen, 2009). 8

Figura 2.4 Operación típica de un espectrómetro de masas (Dunn, 2008). 14

Figura 2.5 Esquema de un sistema de ionización ESI: del inglés *electrospray ionization* (Martín, 2012). 15

Figura 5.1 Estrategia experimental en investigaciones metabólicas (adaptado de Koek et al., 2011). 19

Figura 5.2 Diseño experimental con el que se analizó el efecto combinado de los absorbentes y nivel de dilución.. 24

Figura 5.3 Metodología para el tratamiento de las muestras de jugo de tallo de maíz en la prueba analítica de cromatografía de gases. 26

Figura 5.4 Diagrama para procesar los archivos de Toppas MS (Sturm et al., 2008; Winkler, 2010). 27

Figura 6.1 *Heat map* o diagrama de calor representativo de un experimento preliminar. 33

Figura 6.2 En las comparaciones con gráficas tipo caja (box plots) de cada tratamiento o condición, se obtuvo un patrón similar en la respuesta en intensidad para los iones 129.05 y 250.14 m/z (gráficas del lado izquierdo). Se observó un nivel más elevado en intensidad para la condición carbón activado (CAC) en comparación con resina R18 (C18), control (NOM) y polivinilpirrolidona (PVP)..

Figura 6.3 En la comparación de gráficas tipo caja (box plots) para cada repetición biológica, se tuvo un patrón parecido en la respuesta en intensidad de los iones 245.15 y 104.17 m/z (gráficas del lado izquierdo). Se observó un nivel más elevado en intensidad para la repetición (R2) en comparación con las repeticiones R1 y R3..

Figura 6.4 En las comparaciones con gráficas tipo caja (box plots) de cada tratamiento o condición para los iones más significativos, se tuvo un patrón similar en la respuesta en intensidad para los iones -781.34 y -173.08 m/z (gráficas del lado izquierdo). Se observó un nivel menor en intensidad para la condición carbón activado (CAC) en comparación con resina R18 (C18), control (NOM) y polivinilpolipirrolidona (PVP)..

Figura 6.5 Mapa de intensidades-m/z con agrupación de los 50 iones más intensos del experimento CT incluyendo iones positivos y negativos. Se observó una agrupación que no es uniforme en función del experimento, condición o tratamiento y nivel de dilución: función heatmap en R. (n = 456).. 36

Figura 6.6 Mapa de intensidades-m/z que agrupó los 80 iones más intensos del experimento LN incluyendo iones positivos y negativos. Se observó una representación uniforme en función del experimento, condición y nivel de dilución porque no se usó método alguno de clasificación, por nivel de intensidad en relación a la m/z ó tratamiento, en la elaboración de la gráfica; (n = 456)..

Figura 6.7 Gráfica comparativa de intensidades tipo caja (box plot) del ion 93.3 m/z entre los tratamientos WS y CT del experimento de estrés hídrico de Tlaltizapán. La intensidad del ión de cada genotipo fue mayor para el tratamiento de sequía WS (n = 456).. 38

Figura 6.8 Gráficas comparativas de intensidades tipo caja (box plot) de los iones 120.07, 314.22, 453.18 y 272.14 m/z, seleccionados de un grupo como los más significativos, entre los genotipos por cada tratamiento del experimento de estrés hídrico de Tlaltizapán (n = 456).. 39

Figura 6.9 Gráficas comparativas de los iones aproximados a la variable m/z de los aminoácidos prolina (116.13 m/z) y valina (118.15 m/z). En las gráficas de la izquierda se mostraron los valores de intensidad de los iones 116.3 y 118.3 m/z por tratamiento y genotipo.. 40

Figura 6.10 Mapa de intensidades-m/z agrupando los 50 iones seleccionados con mayor intensidad en el experimento de estrés hídrico en Tlaltizapán incluyendo los iones positivos y negativos.. 40

Figura 6.11 Mapa de intensidades-m/z agrupando los 80 iones más intensos del experimento de estrés hídrico en Tlaltizapán incluyendo iones positivos y negativos..

Figura 6.12 Clasificación de los genotipos y su tratamiento respectivo del experimento de estrés hídrico de Tlaltizapán mediante una representación tipo dendograma, agrupados por su proximidad en el nivel de intensidad.. 42

Figura 6.13 Gráfica comparativa de intensidades tipo caja (box plot) del ion 175.3 m/z entre los genotipos, maíces blancos y amarillos, y tratamientos LN y CT del experimento de bajo nitrógeno de Batán (n = 456).. 43

Figura 6.14 Gráficas comparativas de intensidades tipo caja (box plot) de los iones -160, -158.97, -157.99 y 157 m/z, seleccionados del grupo completo ordenado de iones

positivos y negativos, entre los genotipos por cada tratamiento del experimento de bajo nitrógeno de Batán (n= 456).. 44

Figura 6.15 Gráficas comparativas de intensidades tipo caja (box plot) de los iones -133.02 y -100.96 m/z, seleccionados del grupo de los 50 iones negativos más significativos, entre los genotipos por cada tratamiento del experimento de bajo nitrógeno de Batán (n = 456).. 45

Figura 6.16 Gráficas comparativas de los iones aproximados a la variable m/z de los aminoácidos valina (118.15 m/z) y treonina (120.12 m/z). En las gráficas de la izquierda se mostraron los valores de intensidad de los iones 118.3 y 120.3 m/z por tratamiento y genotipo.. 45

Figura 6.17 Mapa de intensidades-m/z que agrupó a los 50 iones seleccionados con mayor intensidad en el experimento de bajo nitrógeno en Batán incluyendo los iones positivos y negativos. 46

Figura 6.18 Clasificación de los genotipos y su tratamiento respectivo del experimento de bajo nitrógeno de Batán mediante un dendograma, agrupados por su proximidad en el nivel de intensidad.. 46

Figura 6.19 Clasificación de los genotipos (maíces amarillos) y su tratamiento respectivo del experimento de bajo nitrógeno de Batán mediante una representación tipo dendograma, agrupados por su proximidad en el nivel de intensidad.. 47

Figura 6.20 Clasificación de los genotipos (maíces blancos) y su tratamiento respectivo del experimento de bajo nitrógeno de Batán mediante una representación tipo dendograma, agrupados por su proximidad en el nivel de intensidad.. 47

Figura 6.21 Mapa de intensidades-m/z agrupando los 80 iones más intensos del experimento de bajo nitrógeno en Batán incluyendo iones positivos y negativos..

Figura 6.22 Gráfica comparativa de intensidades tipo caja (box plot) del ion 219.3 m/z entre los genotipos, Dow (D) y Puma (P), y tratamientos: (CT = control, WE = sin elotes, WL = sin hojas, WM = sin elotes y sin hojas) del experimento fuente-sumidero de Irapuato (n = 504). 49

Figura 6.23 Gráfica comparativa de intensidades tipo caja (box plot) del ion 219.3 m/z entre los genotipos, Dow (D) y Puma (P), y tratamientos: (CT = control, WE = sin elotes, WL = sin hojas, WM = sin elotes y sin hojas) del experimento fuente-sumidero de Irapuato, sin la inclusión del factor tiempo (n = 504). 50

Figura 6.24 Gráfica comparativa de intensidades tipo caja (box plot) de los iones 73.02 y 115.02 m/z entre los genotipos, Dow (D) y Puma (P), tratamientos (CT = control, WE = sin elotes, WL = sin hojas, WM = sin elotes y sin hojas) y tiempo (0, 1, 2, 3, 4 y 10 días) del experimento fuente-sumidero de Irapuato (n = 504). Gráfica complementaria a la figura 6.20. 51

Figura 6.25 Clasificación de los genotipos Dow (D), Puma (P) y su tratamiento respectivo del experimento de fuente-sumidero de Irapuato mediante una representación tipo dendograma, agrupados por su proximidad en el nivel de intensidad.. 52

Figura 6.26 Mapa de intensidades-m/z que agrupó a los 40 iones seleccionados con mayor intensidad en el experimento fuente-sumidero en Irapuato incluyendo los iones positivos y negativos.. 52

Figura 6.27 Mapa de intensidades-m/z que agrupó a los 20 iones seleccionados con mayor intensidad en el experimento fuente-sumidero en Irapuato incluyendo los iones positivos y negativos.. 52

Figura 6.28 Mapa de intensidades-m/z agrupando los 80 iones más intensos del experimento fuente-sumidero en Irapuato incluyó iones positivos y negativos.

Figura 6.29 Gráfica comparativa de intensidades tipo caja (box plot) del ión 99.02 m/z entre los meses de Mayo (M) y Junio (J) con los genotipos (H1, H2, H3, H4, y H5), del experimento etapas de desarrollo 2010 en Irapuato, sin la inclusión del factor tiempo (n = 420)..

Figura 6.30 Gráfica comparativa de intensidades tipo caja (box plot) del ion 99.02 m/z entre los meses de Mayo (M) y Junio (J) con los genotipos (H1, H2, H3, H4, y H5), del experimento etapas de desarrollo 2010 en Irapuato, con la inclusión del factor tiempo (46, 53, 60, 67, 74, 81 y 88 días en el mes de Junio; 60, 67, 74, 81, 89, 96, y 102 días en el mes de Mayo) (n = 420).

Figura 6.31 Gráfica comparativa de intensidades tipo caja (box plot) de los iones 56.06, 97.1 y 383.23 m/z entre los genotipos (H1, H2, H3, H4, y H5), tratamientos y tiempo (0, 1, 2, 3, 4 y 10 días) del experimento etapas de desarrollo 2010 de Irapuato (n = 420)..

Figura 6.32 Gráficas comparativas de los iones aproximados a la variable m/z de los aminoácidos ácido aspártico (134.1 m/z), glutamina (147.14 m/z) y lisina (147.19 m/z). En las gráficas de la izquierda se muestran los valores de intensidad de los iones 134.3 y 147.3 m/z por meses, mayo y junio, y genotipo (H1, H2, H3, H4, y H5).. 59

Figura 6.33 Clasificación de los genotipos (H1, H2, H3, H4, y H5), mes de siembra mayo y junio en función del tiempo de cosecha del experimento etapas de desarrollo de Irapuato mediante una representación tipo dendograma, agrupándose por su proximidad en el nivel de intensidad.. 60

Figura 6.34 Mapa de intensidades-m/z que agrupó a los 50 iones seleccionados con mayor intensidad en el experimento etapas de desarrollo 2010 en Irapuato incluyó los iones positivos y negativos. Se pudo observar una agrupación uniforme de los iones en función del mes de siembra (M) y (J) y genotipo (H1, H2, H3, H4, y H5), al usar el método scale como opción de la función heatmap en R, en la elaboración de la gráfica.. 60

Figura 6.35 Gráfica comparativa de intensidades tipo caja (box plot) de los iones 97.07 y 159.05 m/z entre plantas con vaina (CV) y sin vaina (SV), los sitios de siembra dentro y fuera del invernadero (DI, FI), con los genotipos (HP, HD), del experimento etapas de

desarrollo 2011 en Irapuato, con la inclusión del factor tiempo de cosecha (39, 46, 53, 60, 67, 74, 81, 95, 102, 109 y 116 días) (n = 416).. 61

Figura 6.36 Clasificación de los genotipos (HP y HD), sitio de siembra dentro del invernadero (DI) y fuera del invernadero (FI) en función del tiempo de cosecha del experimento, con plantas con vaina (CV) y sin vaina (SV) y etapas de desarrollo 2011 en Irapuato mediante una representación tipo dendograma, agrupándose por su proximidad en el nivel de intensidad. 62

Figura 6.37 Mapa de intensidades-m/z que agrupó a los 50 iones seleccionados con mayor intensidad en el experimento etapas de desarrollo 2011 en Irapuato que incluyó los iones positivos y negativos. Se observó una agrupación uniforme de los iones en función del sitio de siembra, genotipo (HP y HD) y plantas con vaina (CV) y sin vaina (SV), al usar el método scale como opción de la función heatmap en R, en la elaboración de la gráfica.. 63

Figura 6.38 Mapa de intensidades-m/z agrupando los 30 iones seleccionados con mayor intensidad en el experimento etapas de desarrollo 2011 en Irapuato que incluyó los iones positivos y negativos. Se pudo observar una agrupación uniforme de los iones en función del sitio de siembra, genotipo (HP y HD) y plantas con vaina (CV) y sin vaina (SV) en función del tiempo de cosecha (39, 46, 53, 60, 67, 74, 81, 95, 102, 109 y 116 días), al usar el método scale como opción de la función heatmap en R, en la elaboración de la gráfica.. 64

Figura 6.39 Gráfica de los niveles en el contenido de etanol por genotipo para los experimentos de estrés hídrico y bajo nitrógeno. Las barras de error fueron estimadas con el programa Excel: error típico, $p < 0.05$ (n = 456). 65

Figura 6.40 Gráfica de los niveles en el contenido de ácido acético por genotipo para los experimentos de estrés hídrico y bajo nitrógeno. Las barras de error fueron estimadas con el programa Excel: error típico, $p < 0.05$, (n = 456). 65

Figura 6.41 Gráfica de los niveles en el contenido de ácido láctico por genotipo para los experimentos de estrés hídrico y bajo nitrógeno. Las barras de error fueron estimadas con el programa Excel: error típico, $p < 0.05$ (n = 456).. 66

Figura 6.42 Clasificación de los genotipos (1-19 LN, 20-38 CT, bajo nitrógeno; 39-57 WS, 58-76 CT, estrés hídrico) de los experimentos estrés hídrico en Tlaltizapán y bajo nitrógeno en Batán mediante una representación tipo dendograma, agrupados por su proximidad en el nivel de intensidad. 67

Figura 6.43 Niveles comparativos de concentración de azúcares solubles y aminoácidos para cada experimento y tratamiento representados. En los cuatro tratamientos la glucosa presentó mayor concentración, siendo la fructosa y la sacarosa menores en ese orden, excepto en el tratamiento de estrés hídrico (WS) en el que la fructosa fue menor a la sacarosa, destacándose su papel como osmolito. 70

Figura 6.44 La producción de biomasa fresca fue mayor en los tratamientos control de los experimentos de estrés hídrico y bajo nitrógeno (n = 456). Las barras de error fueron estimadas con el programa Excel: error típico, $p < 0.05$ (n = 456). 70

Figura 6.45 Los niveles de las características agronómicas presentadas como, peso total de planta, altura de planta, peso total de jilotes y peso de tallo de planta, fueron mayores en los tratamientos control de los dos experimentos, estrés hídrico (WS) y bajo nitrógeno (LN). Las barras de error fueron estimadas con el programa Excel, $p < 0.05$ ($n = 45$).. 71

Figura 6.46 Los niveles de rendimiento en biomasa seca del tratamiento de bajo nitrógeno (LN) fueron superiores en el genotipo (CP1: 56.4%): G-3, G-7, G-9, G-6, G-10, G-4, G-8, G-2, G-1, de acuerdo al resultado del análisis de medias con el método estadístico de Tukey ($n = 19$). 72

Figura 6.47 Los niveles de rendimiento en biomasa seca del tratamiento de estrés hídrico (WS) fueron superiores variaron de de acuerdo al genotipo (CP1: 56.4%) siguiendo el siguiente orden descendente: E-37, E-114, E-104, E-99, E 69, E-34, de acuerdo al resultado del análisis de medias con el método estadístico de Tukey ($n = 19$).

Figura 6.48 Los niveles de rendimiento de grano del tratamiento de bajo nitrógeno (LN) fueron superiores en el variaron de acuerdo al genotipo (CP1 84.4%) siguiendo el siguiente orden descendente: G-7, G-5, G-4-, G-9, G-3, G-10, G-2, G-11 de acuerdo al resultado del análisis de medias con el método estadístico de Tukey ($n = 19$). 73

Índice de tablas

Tabla 5.1 Lista de genotipos usados en el experimento de estrés hídrico. 20

Tabla 5.2 Lista de genotipos usados en el experimento de bajo nitrógeno. 21

Tabla 6.1 Niveles de correlación entre factores del experimento de estrés hídrico en Tlaltizapán mediante el método de PEARSON en R ($n = 456$).

Tabla 6.2 Niveles de correlación entre factores del experimento de bajo nitrógeno en Batán mediante el método de PEARSON en R ($n = 456$).

Tabla 6.3 Ecuaciones de regresión múltiple usadas para predecir el rendimiento de grano y biomasa fresca del experimento de Batán en sus dos tratamientos..

Tabla 6.4 Ecuaciones de regresión múltiple validadas en la predicción del rendimiento de grano y biomasa fresca del experimento de Batán en sus dos tratamientos..

ABREVIACIONES

aa, aminoácidos; Ala, alanina; amu, unidad de masa atómica; ANOVA, analisis de varianza; APCI, ionización química a presión atmosférica; Arg, arginina; ASI, intervalo antesis-silking; Asn, asparagina; Asp, ácido aspartico; BLAST, basic local alignment search tool; C₄, fotosíntesis de baja fotorespiración; CE, electrofóresis capilar; CI, ionización química; CIMMYT, Centro Internacional para el mejoramiento del Maíz y Trigo; CT, ensayo control; Cys, cisteína; DIESI-MS, espectrometría de masas por inyección directa con ionización por electrospray; DIMS, espectrometría de masas por infusión directa; DS, etapa de desarrollo; DTMA, Iniciativa Africana pro Maíz tolerante a sequía; E, ambiente; EI, impacto electrónico; ESI, ionización por electrospray; F, valor de Fisher; FAB, bombardeo de átomos acelerados; FAO, Organización para la Alimentación y la Agricultura de la Naciones Unidas; FD, campo de desorción; FFWL, floración femenina; FI, campo de ionización; FISH, hibridación de fluorescencia en sitio; FTICRMS, espectrometría de masas por resonancia ión ciclotrón con transformada de Fourier; FTMS, espectrometría de masas con transformada de Fourier; G, genotipo; g, gravedad; GC, cromatografía de gases; GC-FID, cromatografía de gases con detector de ionización por flama; Gln, glutamina; Glu, ácido glutámico; Gly, glicina; GPS, Sistema de Posicionamiento Global; GWS, selección genómica completa; GY, rendimiento de grano; ha/año, hectáreas por año; HCA, Análisis por agrupamiento jerárquico; HEPES, ácido N-(2-Hidroxietil) piperazina-N'-(2-etanosulfónico); His, histidina; HM, masa alta; HP, alta presión; HPLC-RID, cromatografía líquida de alto rendimiento con detector de índice de refracción; HTTPs, plataformas de fenotipo a gran escala; Hz, Hertz; Ile, isoleucina; IMIC, Consorcio internacional para mejoramiento del Maíz; KEGG, Enciclopedia Kyoto de Genes y Genomas; kg/ha, kilogramos por hectárea; kV, kilovoltio; L, litro; Lab, laboratorio; LC, cromatografía líquida; Leu, leucina; LM, masa baja; LN, bajo nitrógeno; LSD, menor diferencia significativa de Tukey; Lys, lisina; m/z, razón masa/carga; MALDI-TOF-MS, espectrometría de masas con ionización/desorción láser con matriz asistida y tiempo de vuelo; MasAgro, Modernización sustentable de la Agricultura tradicional; MASL, metros sobre el nivel del mar; Mesh, filtro; Met, metionina; mQTLs, loci de características cuantitativas; MS, espectrometría de masas; NADPH, nicotinamida adenina dinucleótido fosfato reducido; NAE, eficiencia en la asimilación de nitrógeno; NALDI-TOF-MS, espectrometría de masas con ionización/desorción laser nano asistida con tiempo de vuelo; NIRS, espectroscopía por reflectancia en el infrarojo cercano; NMR, resonancia magnética nuclear; NRE, eficiencia en la remobilización del nitrógeno; NUE, eficiencia en el uso del nitrógeno; NUpE, eficiencia en la absorción del nitrógeno; NUtE, eficiencia en la utilización del nitrógeno; OECD, Organización para la Cooperación y el Desarrollo Económico; PANGAEA, editor de datos para las ciencias ambientales y de la tierra; PCA, análisis de componentes principales; Phe, fenilalanina; Pro, prolina; pvdf, polifluoruro de vinilideno; QPM, maíz de calidad proteica; QTL, loci con características cuantitativas; R, Proyecto para la computación estadística en lenguaje R; R², coeficiente de correlación; rf, radio frecuencia; RIL, líneas endogámicas recombinadas; UPLC, cromatografía líquida de ultra rendimiento; RT, tiempos de retención; S/N, razón señal/ruido; SAGARPA, Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación; SDS-PAGE, electrofóresis en gel de poliacrilamida con dodecilsulfato sódico; Ser, serina; SIMS, espectrometría de masas de iones secundarios; SNPs, polimorfismo de nucleótidos sencillos; TCA, ácido tricloroacético; TFA, ácido trifluoroacético; Thr, treonina; TMS, trimetilsilil; TPE, ambientes de la población objetivo; Trp, triptófano; Tyr, tirosina; UV, ultravioleta. V, voltio; v/v, volumen en volumen; Val, valina; VPD, caída de presión de vapor; ; 3d, tres dimensiones.

RESUMEN

El maíz (*Zea mays* L.) es un cultivo importante por su productividad y diversidad de usos, tanto para la alimentación humana como animal. El rendimiento, así como la calidad nutricional del maíz, dependen de la composición bioquímica de diversas células y tejidos. Lo que se planteó como proyecto medular fue el estudio bioquímico-analítico de diversos órganos de la planta de maíz que se utilizan como forraje. En la presente tesis doctoral analizamos el perfil y el contenido de metabolitos en el jugo (la savia extraída) de tallo de maíz. Lo anterior con el fin de esclarecer la función fisiológica del tallo en relación con los demás tejidos productores y consumidores de fotosintatos. Utilizamos diversas herramientas metabolómicas para cuantificar biomoléculas (metabolitos y proteínas) en las diferentes etapas de crecimiento, en diferentes variedades y bajo distintas condiciones de cultivo como son la sequía y el bajo contenido de nitrógeno. Utilizamos cromatografía tanto de gases (GC) como de líquidos (LC), así como espectrofotometría de luz ultravioleta (UV). También implementamos protocolos de fenotipo bioquímico usando espectrometría de masas (MS) con inyección directa de ionización con spray de electrones (DIESI-MS). Los resultados bioquímicos obtenidos se compararon con los datos agronómicos y parámetros fisiológicos secundarios (p. ej. floración femenina (FFWL), intervalo de floración (ASI), rendimiento de grano (GY), etc). La pregunta medular fue si el perfil bioquímico del jugo de tallo en la etapa de floración refleja el manejo agronómico de las plantas y si esto permite predecir el rendimiento final de grano (GY) en un ensayo de campo. Por la cantidad de datos tan grande, se usó el lenguaje de programación R para implementar y entrelazar diversas técnicas estadísticas como el análisis de varianza (ANOVA), las pruebas Tukey-LSD y la correlación Pearson y la regresión lineal en diversas modalidades. También optimizamos *scripts* para el procesamiento de los espectros DIESI-MS y la generación de diversas figuras como boxplots, dendogramas de similitud y heatmaps biclusters para facilitar la interpretación de los resultados. Se encontraron correlaciones significativas de los metabolitos (biomarcadores) con el estado fisiológico de la planta, que por una parte permitieron discriminar variedades y por otro predecir tempranamente el desempeño agronómico (rendimiento final de grano o de forraje). Los resultados de metabolitos y las correlaciones observadas en diversas condiciones de cultivo y con diferentes genotipos, servirán para posteriormente convencer a los mejoradores clásicos que incluyan las variables metabólicas en sus índices de selección. Finalmente podemos concluir que logramos desarrollar protocolos confiables y métodos costeables de fenotipo bioquímico basados en DIESI-MS que serán usados en los programas de mejoramiento genético de maíz, en el marco de la iniciativa MasAgro financiada por la SAGARPA.

Palabras clave: metabolómica, fisiología, maíz, aminoácidos, carbohidratos, espectrometría de masas, regresión lineal, programación R, bioinformática, nitrógeno, sequía.

SUMMARY

Maize (*Zea mays* L.) is an important crop because of its productivity and diversity of uses for both human and animal consumption. The nutritional quality and grain yield (GY) of maize depends on the biochemical composition of diverse cells and tissues. The project started as a bioanalytical study of various organs of corn plants that are used as fodder. In this thesis we analyzed the biochemical profile and content of metabolites in the soluble extract of the stalk (stem juice). This was done to clarify the physiological function of the stem relative to other sink and source organs. We used various metabolomic tools to quantify biomolecules (metabolites and proteins) at different growth stages, different varieties under various growth conditions such as drought (WS) and low nitrogen stress (LN). Different analytical techniques were used including gas (GC) and liquid (LC) chromatography and ultra violet (UV) light spectrophotometry. We also implemented a protocol of biochemical phenotyping using mass spectrometry (MS) via direct injection electrospray ionization (DIESI-MS). The biochemical results were compared with the agronomic data and secondary physiological parameters (e.g. female flowering (FFWL), anthesis-silking interval (ASI), and grain yield (GY)). The key question was whether the biochemical profile of the stem juice at the time of flowering reflects the agronomic management and if this allows to predict final grain yield (GY) in a field trial. Due to the large amount of data generated, R programming was used to implement and interlace various statistical techniques such as ANOVA, Tukey-LSD, Pearson correlation, and various linear regression methods. Also, we optimized scripts for processing DIESI-MS spectra and the generation of various figures such as boxplots, dendrograms and heat maps biclusters to facilitate biological interpretation of the results. We found significant correlation of metabolites (biomarkers) with the physiological state of the plant, which permitted the discrimination of genotypes on the one hand, and on the other, it allowed early prediction of the agronomic performance (final GY or fodder). The metabolite results and the observed correlations in various culture conditions and with different genotypes will help to convince classic breeders to include metabolic variables in their selection indices. Finally, we conclude that our group successfully developed reliable protocols and cost-efficient methods of biochemical phenotyping based on DIESI-MS that are more affordable for maize breeding programs in the framework of the MasAgro initiative funded by SAGARPA.

Keywords: metabolomics, physiology, maize, aminoacids, carbohydrates, mass spectrometry, linear regression, R software, bioinformatics, nitrogen, drought.

1. INTRODUCCIÓN GENERAL

Los programas de mejoramiento genético de plantas incluyen diferentes líneas de investigación científica. Existen las aproximaciones clásicas muy efectivas desde hace muchos siglos, pero también han surgido los enfoques modernos (Spiertz et al, 2014). El paradigma holístico de la biología de sistemas ha llevado a incluir las ciencias ómicas dentro del campo de la agronomía y la hibridación de especies vegetales como el maíz (Bänziger et al., 2002; Fernie y Schauer, 2009). El dilema del fenotipo-genotipo ha generado temas diversos en la biología molecular, entre ellas la genómica, transcriptómica, proteómica y, recientemente, la metabolómica vegetal. A manera de analogía, se puede decir que los genes representan el “potencial”, los transcritos la “estrategia”, las proteínas realizan el “trabajo” y la metabolómica refleja los “resultados” (Figura 1.1).

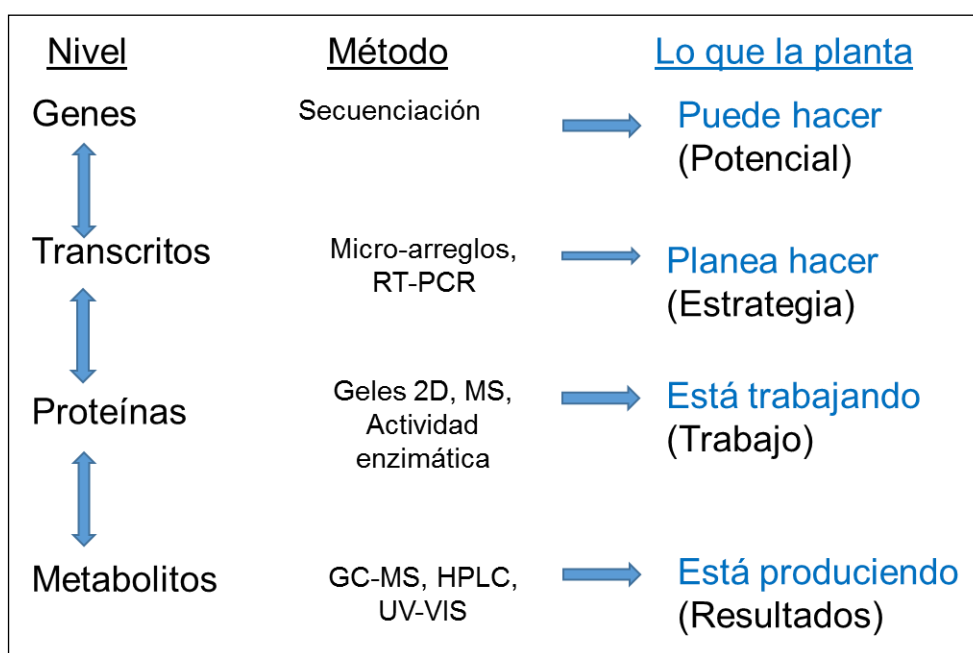


Figura 1.1 Flujo bidireccional de información celular entre los genes y los metabolitos. En cada uno de los niveles se utilizan diferentes métodos científicos para su estudio. La secuenciación de un genoma nos dice el potencial de lo que pudiera hacer, pero no necesariamente siempre lo logra. Los transcritos, nos indican el plan estratégico, mientras que las proteínas y las enzimas nos revelan en qué está trabajando la célula en ese momento específico. Finalmente, los resultados se reflejan de manera integrada y acumulada en los metabolitos. Adaptado de Tiessen (2009).

Las dos primeras ciencias ómicas han experimentado un auge importante por los avances metodológicos de la biotecnología y las nanociencias que las posicionan de manera preponderante en los trabajos actuales de ciencia básica. La genómica se ha vuelto tan popular y masiva por la facilidad técnica de secuenciar todos los ácidos nucleicos. La transcriptómica ha vivido un auge similar iniciando con los microarreglos, el PCR en tiempo real y ahora con la secuenciación directa de ARN (RNA Seq). La que se ha quedado más atrás por la falta de innovaciones técnicas y el alto costo de los equipos analíticos es la metabolómica y la fenómica agrícola. Los estudios de metabolómica amplían el nivel y grado de información acerca de los compuestos biológicos que son importantes en el funcionamiento, la regulación y homeóstasis de la célula (Hall, 2011). La investigación sistémica de las

ciencias ómicas genera una riqueza de datos que nos conducen a la necesidad de facilitar su análisis con programas bioinformáticos. El objetivo es encontrar relaciones causales entre los factores que controlan un proceso metabólico o una función fisiológica en los diversos experimentos diseñados en el campo de la agricultura (Bänziger et al., 2002; Fernie y Schauer, 2009). La robustez de los métodos estadísticos aplicados a las ciencias agrícolas, combinados con la practicidad de la informática computacional, nos permite la visualización de los datos numéricos y facilitan la interpretación de los resultados. Por medio de gráficas e imágenes (p. ej. *Heatmaps* o mapas de calor) se comparan los niveles de cientos de moléculas con fines de caracterizar y seleccionar a los mejores genotipos en determinados ambientes sembrados en múltiples campos experimentales.

La ingeniería genética y el mejoramiento vegetal es un proceso iterativo de entrecruzamiento, recombinación y selección para finalmente obtener un mejor genotipo (Bänziger et al., 2000). La hibridación sexual y la segregación durante la meiosis generan diversas variantes génicas, con características predeterminadas para la planta o el cultivo. En el caso del maíz, la hibridación lleva a un interesante efecto de heterosis que por medio del vigor híbrido produce mayores rendimientos de grano y biomasa. Se busca que los híbridos sean resistentes a plagas y enfermedades además de ser tolerantes a estreses abióticos como son la sequía, calor, salinidad, bajo contenido de nitrógeno en el suelo, inundación, y contaminación de metales pesados, entre otros (Bänziger et al., 2000). Debido a que el fenotipo de las plantas es determinado por el genotipo y sus interacciones con el ambiente, el estudio de estos fenómenos fisiológicos y moleculares es importante para los mejoradores. También es importante para los agrónomos que optimizan y establecen nuevas prácticas culturales para potenciar los rendimientos de grano y de biomasa fresca (Bänziger et al., 2000).

Para identificar a los genotipos que mejor se desarrollan bajo diferentes condiciones ambientales y prácticas culturales proponemos enfocarnos al estudio metabólico del tallo vegetativo del maíz. Queremos entender mejor su papel dentro de la fisiología de la planta. La etapa de floración en el desarrollo de la planta es importante para tomar datos agronómicos y bioquímicos. Además, por ser una etapa previa al llenado y maduración de la mazorca, un fenotipo durante esta etapa temprana permitiría hacer una selección de variedades de forma más eficiente.

Una de las estrategias metabolómicas que escogimos es la conocida como *fingerprinting*, que se engloba en los estudios no dirigidos a una sola familia de metabolitos. Proponemos que por medio de la aplicación de la ionización por *electrospray* e inyección directa acoplada a espectrometría de masas con cuadrupolos (DIESI-MSQ, por sus siglas en inglés) es posible la obtención de perfiles y huellas metabólicas de extractos de jugo de tallo de maíz para conocer el estado fisiológico de la planta. Utilizando métodos bioestadísticos, con la combinación de factores y mediante relaciones lineales de dichos datos entre ellos, se obtienen ecuaciones de regresión para efectos de predecir los rendimientos agronómicos.

Los extractos del jugo de tallo de maíz se obtuvieron de los siguientes experimentos: “estrés hídrico”, “bajo nitrógeno”, “fuente-sumidero”, “diferentes etapas de desarrollo”, en 2010 (días de siembra) y en 2011 (plantas con y sin

vaina). Para hacer una adecuada interpretación de los resultados acudimos a diferentes niveles de análisis partiendo desde la morfometría de la planta hasta el contenido de metabolitos en el jugo del tallo usando diferentes técnicas analíticas. Para ello fue preciso ubicar el grado de amplitud y el nivel de detalle capaz de proporcionar la mejor respuesta a las preguntas biológicas que pretendíamos dilucidar.

El cambio climático, el deterioro de los recursos naturales y la reducción de las zonas de cultivo agrícola son las razones por las que en CIMMYT y en Cinvestav se ha intensificado la investigación para mejorar al maíz bajo condiciones de estrés abiótico en el marco de la iniciativa de MasAgro-SAGARPA (<http://masagro.mx>), que fue la que financió la mayor parte de los experimentos de esta tesis doctoral.

2. ANTECEDENTES

2.1 Metabolómica

2.1.1 Definición e importancia

Aunque hay antecedentes más antiguos, el metaboloma de plantas fue descrito por Fiehn (2001) como el conjunto completo de todos los compuestos moleculares sintetizados por un organismo autótrofo. El análisis del metaboloma fue considerado como “la medida del cambio en las concentraciones relativas de metabolitos como resultado de la eliminación o sobre expresión de un gen [...] que permite localizar el gen blanco de un producto nuevo en el mapa metabólico”. Fiehn (2001), hizo posteriores distinciones entre diferentes análisis metabólicos, y definió a la metabolómica, como el análisis exhaustivo y cuantitativo de todos los metabolitos de un organismo. En realidad lo que más interesa es medir las principales moléculas que son relevantes para el mantenimiento, crecimiento y el funcionamiento normal de la célula (Beecher, 2003).

En el campo de la metabolómica se usan diversos términos y conceptos. Sin embargo, las definiciones varían según los autores y algunas veces se combinan o se confunden. Aún cuando el estudio del metabolismo de plantas tuvo sus orígenes en la medicina naturista y abarca un amplio campo y trayectoria en la ciencia, para nuestro proyecto le dimos prioridad a los paradigmas de la fisiología molecular de plantas representado por los siguientes líderes de grupos científicos: O. Fiehn, U. Roessner, A. Fernie, L. Willmitzer y M. Stitt (Roessner et al., 2001).

2.1.2 Estrategias en metabolómica

Actualmente se utilizan diferentes estrategias para el análisis global de metabolitos (Cuadro 2.1 De acuerdo con las preguntas biológicas en estudio, se requiere utilizar diferentes enfoques porque existen diversos requisitos de desempeño analítico: límites de detección, precisión, exactitud, capacidad de procesamiento de muestras, etc. (Madalinski et al., 2008; Koek et al., 2011). La red metabólica en una célula viva es un complejo conjunto de reacciones que están fuertemente interconectadas, donde cualquier perturbación en alguna enzima puede causar cambios significativos en la concentración de diversos metabolitos (Nielsen, 2003). Los metabolitos tienen diversas estructuras químicas y por lo tanto, es necesario aprender a extraerlos con los solventes adecuados y a cuantificarlos por medio de plataformas robustas que puedan ser utilizadas con diferentes especies (Beecher, 2003).

En contraste con la estrategia de "huella metabólica" (o *fingerprinting*), la de "perfil metabólico" (o *metabolic profiling*) se orienta a identificar y cuantificar al menos algunos de los metabolitos presentes, enfocándose con frecuencia en aquellos que han sido identificados por medio de análisis multivariado, de forma discriminante entre varias muestras, tratamientos o genotipos (Cuadro 2.1 Dichos compuestos pueden entonces ser parte de las bases hipotéticas que los relacione con diferencias genéticas o fenotípicas (Hall, 2011). Para estudiar los efectos pleiotrópicos de las perturbaciones genéticas o ambientales, se necesita hacer un análisis completo del sistema biológico. La metabolómica proporciona un registro, podríamos decir “fotográfico”, de los eventos que sucedan en determinado punto del tiempo (Fiehn, 2001).

La metabolómica es “complementaria” a la genómica, transcriptómica y a la proteómica porque es una función que resulta de estos procesos moleculares y su relación proximal con el estado fisiológico de la planta es mayor (ver Figura 2.1).

Las interacciones entre el genotipo y el fenotipo pueden ser mejor entendidas al incluir en el análisis al metaboloma (Figura 2.1). Como consecuencia, el análisis metabolómico también puede ser utilizado en el fenotipo de organismos mutantes y transgénicos (Fiehn, 2001; Hall, 2011).

Cuadro 2.1 Estrategias analíticas de metabolómica (adaptado de Koek et al., 2011)

Análisis dirigido de Metabolitos	Análisis cuantitativo (absoluto o relativo) de uno o algunos pocos metabolitos. Estrategia típica: pretratamiento selectivo de la muestra seguido de un ensayo enzimático selectivo y detección por espectrofotometría UV-VIS. También, se usan ensayos de punto final o colorimétricos. En varios casos se hace una separación previa (GC, LC, CE) acoplada a detección selectiva y de alta sensibilidad (HPLC-DAD, HPLC-FLD, etc.).
Perfil metabólico	Análisis cuantitativo (absoluto o relativo) y cualitativo (identificación) seguido de análisis de multi-componentes que define o describe patrones metabólicos para un grupo de metabolitos relacionado de manera metabólica o analítica (Horning y Horning, 1971). Estrategia típica: pretratamiento selectivo de la muestra para cada metabolito o clases de metabolitos de alguna ruta metabólica seguido de la separación (GC, LC, CE) acoplada a detección por el espectrómetro de masas.
Huella metabólica	Manejo elevado en número de muestras para clasificarlas por similitud. Por lo general no hay cuantificación absoluta (ó sólo relativa) y no se hace identificación individual de metabolitos (Dunn y Ellis, 2005; Fiehn, 2002). Estrategia típica: pretratamiento simple de la muestra seguido de NMR, FTIR, o espectrometría de masas con infusión directa (DIMS).
Metabolómica GC-MS	Análisis cuantitativo (por lo general, relativo) y cualitativo del perfil completo de metabolitos presentes en un sistema biológico (células, fluidos corporales, tejidos). Estrategia típica: pretratamiento específico de la muestra (e.g., derivatización química) seguido de la separación por cromatografía de gases acoplada a espectrometría de masas con impacto electrónico (GC-EI-MS).

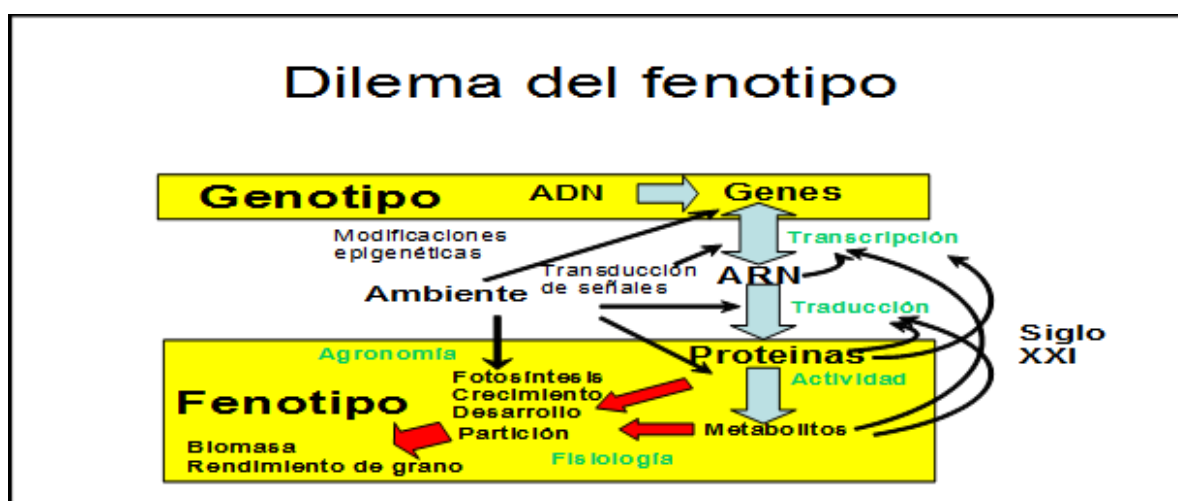


Figura 2.1 Esquema de las relaciones entre el fenotipo y el genotipo (adaptado de Tiessen, 2009). Los metabolitos se derivan de la actividad enzimática y las funciones de las proteínas, ya sea de la fotosíntesis, crecimiento o desarrollo. Los metabolitos están más próximos a la agronomía y al fenotipo (por ejemplo, el rendimiento). La secuencia del ADN está más cercana al genotipo, pero

mucho más lejana del fenotipo, ya que este depende del ambiente, las modificaciones epigenéticas y la transducción de señales.

2.1.3 Metabolómica y fisiología de plantas

Las principales áreas de aplicación de la metabolómica son para la fisiología e hibridación de plantas (Hall, 2011). Se dice que la metabolómica proporciona una inequívoca imagen de lo que está sucediendo a nivel celular (Hall, 2011). La naturaleza compleja de la metabolómica nos lleva a descubrir conexiones y relaciones entre metabolitos y otras variables que no son intuitivas. Una comprensión integral de la complejidad de esto requiere de procesamientos bioinformáticos que deben llevarse a cabo con un estricto rigor estadístico. Solo así mejoraremos nuestro entendimiento molecular de la fisiología de plantas que está influenciada por factores bióticos y abióticos. Al estudiar diversas variables (genética, cultivos, tejidos, aplicación de tratamientos, ambientes, etc.), aprendemos cómo la plasticidad metabólica permite la sobrevivencia de las plantas en un ambiente siempre cambiante y a veces hostil. Con dicha información, podríamos desarrollar novedosas estrategias para el mejoramiento de plantas en términos de su productividad, su adaptación a condiciones ecológicas específicas, calidad de productos, su tolerancia y resistencia a factores ambientales, etc. (Hall, 2011).

2.2 Metabolismo

2.2.1 Características del metabolismo de plantas

El metabolismo puede definirse como una actividad coordinada en la que participan muchos conjuntos de sistemas enzimáticos mutuamente relacionados intercambiando materia y energía entre los diversos compartimentos de la célula y su entorno.

Según Ap Rees y Stitt (1981), el metabolismo vegetal se puede comparar con la red de transporte de una gran ciudad (e. g., el “metro”). Es complejo por las diferentes rutas interrelacionadas implicadas en la síntesis de metabolitos, es robusto porque ante una mutación o un evento modificador, el sistema metabólico es capaz de seguir funcionando para que la planta sobreviva; es redundante por la disponibilidad de procesos paralelos y canales complementarios que evitan los cuellos de botella en el sistema; flexible por la variedad de flujos y estados metabólicos debido a las condiciones ambientales variables, y compartimentalizado porque las reacciones ocurren en diferentes organelos y células (Tiessen, 2009).

Las funciones específicas del metabolismo de plantas son cuatro: 1) la obtención de energía química de las moléculas combustibles o de la luz solar absorbida; 2) la conversión de dióxido de carbono en precursores de los componentes macromoleculares de la célula; 3) el ensamblaje de estos precursores para formar metabolitos primarios y secundarios, proteínas, ácidos nucleicos, lípidos, membranas, etc. y 4) la degradación de las biomoléculas necesarias para las funciones especializadas de las células (Tiessen, 2009).

2.2.2 Metabolismo de plantas

El metabolismo es un pre-requisito de la vida (Stryer 1975). Miles de reacciones químicas, en su mayoría catalizadas por enzimas, definen una red metabólica en la que está basada toda actividad biológica. En particular, se acoplan reacciones anabólicas que consumen energía con las que la generan, y dirigen la

biosíntesis de los polímeros y metabolitos que forman la fábrica celular. Las actividades enzimáticas son fuertemente controladas y ajustadas de acuerdo a las condiciones ambientales cambiantes y patrones de crecimiento (Sweetlove, 2011). Las redes metabólicas de la planta son considerablemente más complejas que las de cualquier otro organismo (desde bacterias a animales). El crecimiento y sobrevivencia de la planta está íntimamente ligado al metabolismo primario (Smith y Stitt, 2007; Stitt et al., 2010). Es por ello que es útil predecir el comportamiento metabólico (Fornie y Schauer, 2009). En particular, existe una necesidad para conectar el genotipo con salidas metabólicas específicas de tal forma que los mejoradores de plantas y los ingenieros metabólicos puedan generar nuevas variedades de cultivos con mayor rendimiento o composición química modificada. Aunque ha habido algunos resultados notables en la ingeniería metabólica (Butelli et al., 2008; Naqvi et al., 2009), éstos están principalmente relacionados con la producción de metabolitos secundarios. En contraste, existen pocos ejemplos donde la síntesis de los principales metabolitos primarios hayan sido manipulados en forma predictiva. La habilidad contrastante para modificar las redes metabólicas se debe a la diferencia en conectividad entre el metabolismo secundario con el primario. Muchos metabolitos secundarios son sintetizados por reacciones que ocurren en la periferia de la red metabólica, con relativamente pocas interconexiones con otras secciones de la red. Como resultado, hay muy pocas restricciones regulatorias de flujo a través de estas rutas lineales básicas y el control puede estar en gran medida sobre una sola enzima, proporcionando un solo blanco para la manipulación genética (Fraser et al., 2002; Enfissi et al., 2005). En lugares donde el control está distribuido más homogéneamente, es posible identificar factores de transcripción que afectan coordinadamente la expresión de las enzimas en una ruta particular (Schwinn et al., 2006; Memelink y Gantet, 2007). En contraste, los componentes principales del crecimiento celular son sintetizados a través de un conjunto de reacciones altamente conectadas al que normalmente se le refiere como “metabolismo central”. Generalmente, no sólo es el control de flujo compartido entre algunas, sino todas las enzimas del metabolismo central (Raines, 2003; Geigenberger et al., 2004; Araujo et al., 2012). Debido al grado de conectividad, una perturbación en una parte de la red tiene consecuencias en otras regiones de la misma (Sweetlove, 2011).

2.2.3 Vínculo del metaboloma con el fenotipo

El fenotipo es lo que puede observarse o medirse (Fiehn 2002). Por lo regular, los agrónomos y los mejoradores solo consideran ciertos aspectos del fenotipo, principalmente relacionados con el desempeño agronómico, como la altura de la planta, la tolerancia a patógenos y/ o insectos plaga y el rendimiento de grano. Lo demás como la composición bioquímica parece ser irrelevante y por consiguiente rara vez se cuantifica en los programas de mejoramiento genético vegetal. El fenotipo bioquímico es el resultado de la interacción entre el genotipo, el ambiente, el tejido y la etapa de desarrollo (Figura 2.1). Por lo tanto, se necesita identificar y cuantificar cientos de metabolitos para estudiar su dinámica y analizar los flujos metabólicos (Figura 2.2). El reto de la metabolómica reside en encontrar cambios en las cantidades de metabolitos que se puedan correlacionar con el estado fisiológico y de desarrollo de un tejido o de un organismo. Estos cambios pueden ser sutiles y las respuestas no siempre directas.

La enorme diversidad bioquímica de las plantas (se estima que pueden llegar a tener hasta > 200,000 compuestos) hace que una aproximación de la metabolómica en gran escala sea un tremendo reto. Se requieren diversas tecnologías para caracterizar todos los metabolitos presentes en plantas. La deducción del estado fisiológico de una planta a partir del perfil y las concentraciones de metabolitos requiere, por tanto, que la identificación y cuantificación de los mismos se realice con gran fiabilidad (Tiessen, 2009).

Debe de existir un puente entre la metabolómica y las otras ramas de la biología integrativa: genómica, transcriptómica y proteómica (Fernie y Schauer, 2009). Es por ello que se requieren bases de datos que almacenen, integren y permitan establecer relaciones causales entre fenotipos, genes, transcriptos, proteínas y metabolitos. Para obtener un conocimiento global de cómo funciona un sistema biológico es esencial saber cómo responden los diversos niveles de expresión. Sólo así será posible deducir asociaciones relevantes entre macromoléculas, identificar correlaciones estadísticamente significativas con fenotipos (Fernie y Schauer, 2009). Se ha postulado que los análisis fenotípicos y bioquímicos a gran escala permitirán en un futuro predecir cómo será el comportamiento de una planta dada en un determinado contexto ambiental (Tiessen, 2009). Derivado del incremental cambio climático, es preciso considerar en las relaciones fenotipo-genotipo, algunos de los factores que afectarán las condiciones del particular hábitat de las plantas (Bänziger et al., 1999). La temperatura, humedad, salinidad, erosión, composición del suelo e intensidad de explotación agrícola, son algunos de ellos (Bänziger et al., 2000).

Enfoque del fenotipo

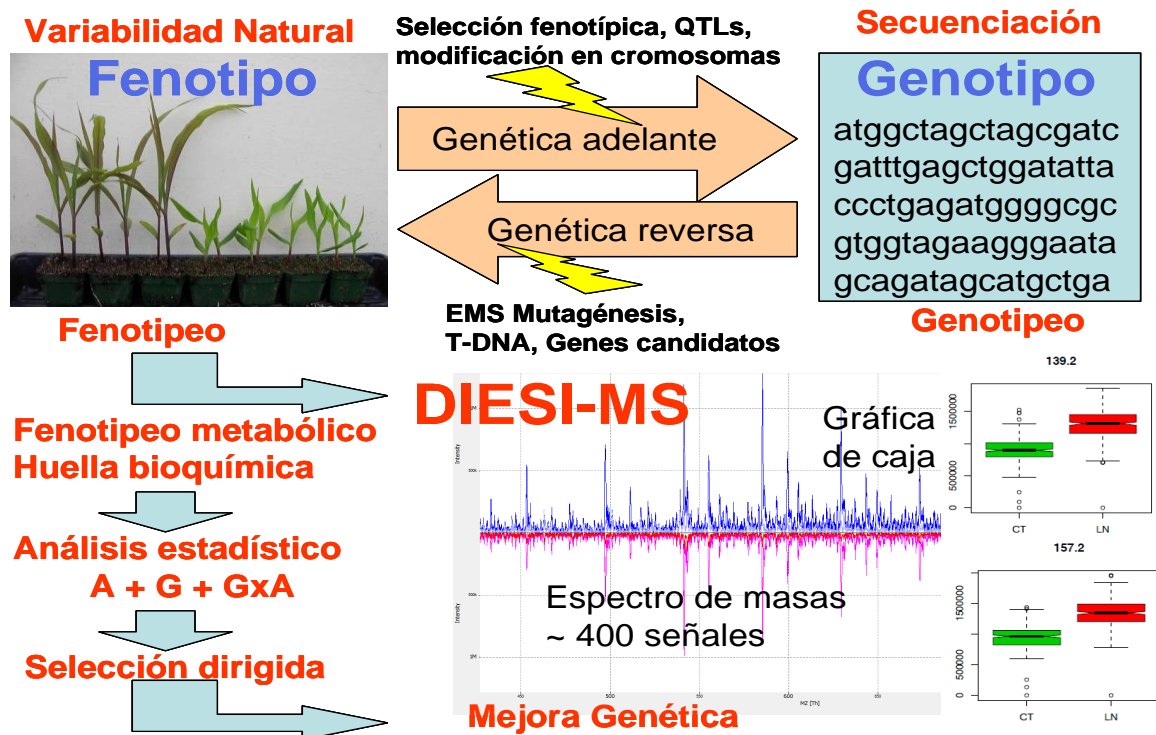


Figura 2.2 En comparación con el aumento exponencial de los métodos de genotipo, las estrategias de fenotipo se están quedando atrás en las ciencias agrícolas. El

mejoramiento genético depende de la abundancia de datos cuantitativos fenotípicos y la partición estadística de la varianza en los efectos ambientales, genéticos, y los debidos al azar. En este trabajo se adaptó una estrategia fenotipo-metabólica para aumentar el rendimiento de la muestra, ahorrando en reactivos, reduciendo costos y simplificando el análisis de datos (Tiessen, 2009).

2.3 Experimentos de estrés hídrico y bajo nitrógeno en maíz

En los últimos 20 años, los investigadores en Cimmyt han mejorado germoplasma de maíz para que sea tolerante a estrés por sequía y por bajo contenido de nitrógeno, utilizando un enfoque que es probablemente único (Bänziger et al., 2000). Poblaciones grandes se seleccionaron bajo condiciones controladas de estrés hídrico o bajo nitrógeno de tal forma que la tolerancia debida a la variación genética fuese revelada en la mayor medida posible (Worku et al., 2007). Los cultivos de maíz en los trópicos están continuamente expuestos a la sequía y al estrés por bajo nitrógeno (Bänziger et al., 1999). La incidencia de estrés puede aumentar debido, en parte, a los cambios climáticos globales (Adams et al., 2013), al desplazamiento de maíz a entornos de producción más difícil por cultivos de alto valor, y a la disminución de la materia orgánica del suelo, la reducción de la fertilidad del suelo y su capacidad de retención de agua (Bänziger et al., 2000). En el nivel micro (parcela), la fertilidad y la disponibilidad de agua varía mucho dentro de los campos agrícolas. Esto significa que una sola variedad debe ser capaz de soportar un amplio rango de estrés por sequía y disponibilidad de nitrógeno, y esta necesidad es más pronunciada en los trópicos que en campos con mayores aportes de las zonas templadas (Bänziger et al., 2000). Los experimentos antes del año 2000, se realizaron con siembra de plantas de maíz en campos experimentales a cielo abierto. Actualmente algunos experimentos bajo estrés por sequía también se están llevando a cabo bajo condiciones controladas de invernaderos (Witt et al., 2012).

2.3.1 Estrés hídrico

De acuerdo con Witt et al., (2012), la naturaleza multigénica de la tolerancia a la sequía (Lopes et al., 2011) representa un reto considerable para la aplicación de conocimientos a la mejora de los cultivos. En su trabajo de investigación ellos midieron una serie de parámetros morfo-fisiológicos como la altura de planta, tamaño del elote, la temperatura de la hoja, la conductancia estomática de la hoja, y el contenido de clorofila en la hoja, mientras que la composición metabólica se determinó por cromatografía de gases-espectrometría de masas (GC-MS) (Roessner et al., 2001) para la hoja, los elotes, los totomoxtles (hojas envolventes del elote), vaina, y pelos del elote de seis híbridos del programa de la fisiología de maíz del Cimmyt (Monneveux et al., 2008; Araus Ortega et al., 2012), en respuesta a la sequía impuesta bajo un ambiente experimental controlado de invernadero. Esto representó un análisis inicial muy completo proporcionando un análisis más detallado de una amplia gama de metabolitos entre tejidos, no llevado a cabo hasta la fecha, facilitando información importante sobre la utilidad de estos tejidos diversos para la identificación de biomarcadores (Bänziger et al., 2002; Fernie y Schauer, 2009). Adicionalmente, los datos obtenidos fueron evaluados estadísticamente de manera cuidadosa respecto a si las plantas podrían ser discriminadas por su genotipo o a las condiciones ambientales a que fueron

sometidas (Roessner et al., 2001; Worku et al., 2007). Al final se hicieron análisis de correlación entre metabolitos y características fisiológicas en un intento de identificar las relaciones que pudieran ser de interés desde una perspectiva de hibridación (Obata et al., 2015). La combinación de resultados se analizó en el contexto de la utilización de la metabolómica como una herramienta para apoyar a las estrategias en mejora de cultivos (Witt et al., 2012). De manera complementaria, Obata et al. (2015), centraron su trabajo de investigación en la relación que hay entre los metabolitos de la hoja y el rendimiento de grano bajo sequía y calor de manera simultánea en condiciones de campo. El efecto negativo de la sequía sobre el rendimiento de maíz es especialmente agudo durante la etapa reproductiva, entre la emergencia de la espiga y el llenado temprano del grano, en donde se considera que se induce la desecación prematura de semillas y se reduce el llenado del grano (Bänziger et al., 2000). El grano es más susceptible a estrés por sequía que los tejidos vegetativos y, por lo tanto, la predicción del rendimiento de grano a partir de parámetros fisiológicos de las hojas es un reto importante (Jones y Simmons, 1983). Sin embargo, el rendimiento del maíz depende tanto de suministro de asimilados al grano (desde un tejido “fuente”) y el potencial del grano para acomodar estos asimilados (como un tejido “sumidero”) (Bänziger et al., 2002). La cruce de híbridos modernos de clima templado se ha centrado más en el potencial sumidero particularmente bajo condiciones de estrés y debe haber, por tanto un considerable potencial restante para mejorar la capacidad de la fuente. Fue posible establecer de manera anticipada que la sequía y el estrés por calor afectan en gran medida el metabolismo de la hoja y, especialmente la fotosíntesis, comprometiendo la capacidad de las hojas de actuar como tejidos fuente. En consonancia con lo anterior, se encontró que la sequía tiene un efecto dramático en la composición metabólica de las hojas en comparación con otros órganos en los experimentos anteriores de invernadero (Witt et al., 2012). Puesto que la capacidad fuente está muy relacionada con el metabolismo de la hoja, el perfil metabólico de la hoja debe tener estrecha relación con el rendimiento de grano, particularmente bajo condiciones de estrés. Teniendo en cuenta que varios estudios recientes han mostrado la importancia de la pre-adaptación metabólica a diversos grados de tolerancia al estrés en plantas, también fue postulado que los niveles basales de metabolitos bajo condiciones de crecimiento óptimo podrían estar correlacionados con la tolerancia al estrés (Bänziger et al., 2000). Con el fin de probar lo anterior, se analizaron perfiles de metabolitos de hojas de diez híbridos en experimentos de campo realizados en la estación experimental subtropical del Cimmyt en 2010 y 2011 en los que las plantas fueron expuestas a estrés por sequía y al calor en forma individual o combinada. Los resultados se evaluaron tanto en el contexto de los modelos actuales de tolerancia al estrés, así como con respecto a sus implicaciones prácticas en las futuras estrategias de mejoramiento (Obata et al., 2015).

2.3.2 Bajo contenido de nitrógeno o “Bajo Nitrógeno”

El estrés por bajo nitrógeno (N) reduce la fotosíntesis de los cultivos, reduciendo el desarrollo del área foliar y la tasa de fotosíntesis de la hoja y acelerando la senescencia foliar. Alrededor del 50% del N de toda la hoja está directamente involucrado en la fotosíntesis, ya sea como parte de la enzima Rubisco o de la clorofila (Bänziger et al., 2000). Las tasas fotosintéticas bajo saturación luminosa muestran una fuerte dependencia del contenido de N en la

hoja ($r > 0,75$), resultando en una relación curvilínea entre la eficiencia en el uso de nitrógeno (NUE, por sus siglas en inglés) y el contenido de N de la hoja. Se muestra una saturación para el maíz en alrededor del 2% del contenido de N en la hoja. Cuando el N se vuelve escaso, las plantas remobilizan N a partir de tejido más viejo (ya sea hojas o tallo) a los tejidos más jóvenes (ya hojas en expansión, flores y/o granos), que conduce a la senescencia temprana del tejido más viejo y de las hojas con menor cantidad de tejido (Bänziger et al., 2000). Bänziger y Lafitte (1997), evaluaron el valor de las características secundarias para mejorar el maíz tropical de tierras bajas en ambientes con bajo contenido de N. Diecinueve experimentos sin aporte de N fueron cultivados en el Cimmyt entre los años de 1986 y 1995 y fueron analizados para el rendimiento de grano, intervalo de antesis-silking (ASI), número de mazorcas por planta, concentración de clorofila en las hojas y una estimación de la senescencia foliar. Se calculó la heredabilidad de las características, las correlaciones genéticas entre rasgos secundarios y el rendimiento final de grano y la respuesta de predicción del rendimiento de grano mediante la selección por rasgos únicos o múltiples. Múltiples rasgos se combinaron usando un índice de selección sin restricciones (Smith-Hazel) (Bänziger and Lafitte, 1997). La heredabilidad promedió fue de 0.46, para rendimiento de grano, de 0,52, para el intervalo “antesis-silking” (ASI), de 0,44, para mazorcas por planta, 0,35, para la concentración de clorofila de la hoja, y de 0,60, para la senescencia foliar. Las correlaciones genéticas de rasgos secundarios con rendimiento de grano promediaron: -0.47 para ASI, 0,78, para número de mazorcas por planta, 0,24, para la concentración de clorofila de las hojas y 0,42, para la senescencia foliar (Bänziger and Lafitte, 1997). La selección directa del rendimiento de grano fue superior a la selección de un rasgo secundario único en la mayoría de los experimentos. Cuando se combinaron todos los rasgos en un índice de Smith-Hazel, la eficiencia de selección mejoró en un 14%, en promedio, a la selección para rendimiento de grano por sí sola (Bänziger y Lafitte, 1997). Entre los rasgos secundarios que mejor discriminaron el rendimiento entre genotipos se encontraron el número de las mazorcas por planta y la senescencia foliar. A su vez, la concentración de clorofila de la hoja, y en algunos casos el ASI, proporcionaron información sobre la variación ambiental en los experimentos. Los autores llegaron a la conclusión de que los rasgos secundarios pueden aumentar la eficiencia de la selección para rendimiento de grano en los programas de mejoramiento de maíz dirigidos a ambientes de bajo nitrógeno (Bänziger y Lafitte, 1997).

En otro estudio, Bänziger et al. (1997), encontraron que las correlaciones genéticas entre los rendimientos de grano bajo condiciones de bajo y alto N fueron en general positivas, disminuyendo o aumentando el rendimiento relativo de acuerdo a los contenidos de N, bajos o altos, respectivamente, lo que indica que la adaptación específica ya sea a bajo o alto N resultó ser más importante cuando los experimentos de bajo N (LN) y de alto N (HN) tuvieron una mayor diferencia en rendimiento de grano. La predicción sobre la cual se basó la selección bajo HN para mejorar el desempeño bajo LN fue significativamente menos eficaz que la selección realizada a bajo LN cuando la reducción relativa en el rendimiento debido al estrés por N superó el 43%. Los programas de mejoramiento de maíz dirigidos a entornos de LN en los trópicos deben incluir ambientes de selección de bajo nitrógeno para maximizar las ganancias en la selección (Bänziger et al., 1997).

Cuando se cruzan variedades de maíz para un incremento estable en el rendimiento de grano en condiciones limitadas de N, una estrategia razonable sería desarrollar poblaciones de maduración temprana y tardía de variedades locales que exhiban gran captación de N, movilización de una gran proporción de materia seca y de N hacia el grano y el mantenimiento de una gran concentración de N en el grano aún bajo suministro reducido de N (Lafitte et al., 1997).

2.3.3 Estrés por limitación de oxígeno

El estrés altera la fotosíntesis de la planta, produce un daño a las células, limita el transporte de agua, inhibe el fotosistema II y altera el aporte de oxígeno a los tejidos vasculares (Tiessen, 2009). El oxígeno es abundante en las hojas que están expuestas al aire atmosférico con 21% de O₂, pero es mucho menos disponible en las raíces y en los tejidos vasculares y de reserva que respiran y que están envueltos en otras estructuras de la planta. El aporte de oxígeno es importante para el metabolismo energético de las células del tejido vascular (Figura 2.3), modificando la razón de metabolismo fermentativo al respiratorio, y por consiguiente, la razón de ATP y NADPH, de ácido málico y oxalacético (Saglio et al., 1983; Rivoal y Hanson, 1994; Geigenberger et al., 2004). Por lo tanto, se tomó la decisión de medir algunos alcoholes y ácidos orgánicos en el jugo de tallo de maíz obtenido en los diferentes experimentos. Nos basamos en la siguiente pregunta de trabajo: ¿el estrés por nitrógeno reduce el oxígeno y redirige el metabolismo primario, alterando los metabolitos de la fermentación? La información que se obtenga será parte complementaria de la respuesta metabólica de las plantas de maíz sujetas a estrés. Es por ello que se consideró la posibilidad de combinar los resultados de DIESI-MS con GC-FID (Roessner et al., 2001) para ampliar el estudio metabólico. La metabolómica vegetal no se basa en una sola técnica analítica de espectrometría de masas, sino en el uso combinado de varias plataformas, métodos y equipos, para el mejor entendimiento de todas las rutas metabólicas.

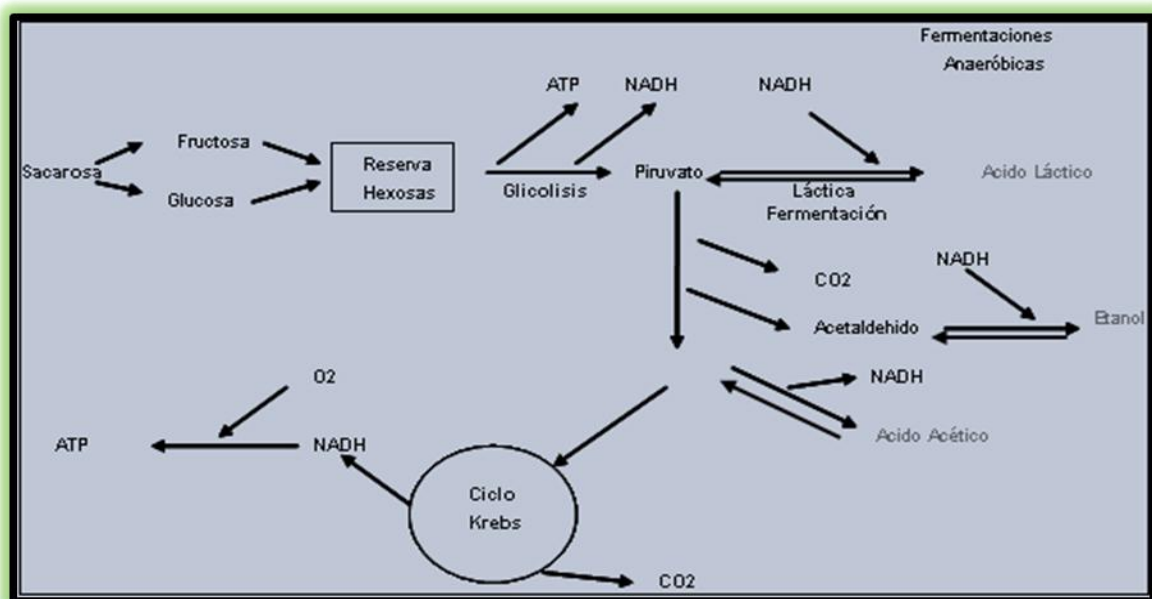


Figura 2.3 Síntesis de alcoholes y ácidos bajo condiciones limitantes de oxígeno. El etanol, y el ácido acético son reciclados para suministrar al metabolismo central el NADPH requerido en la fermentación a ácido láctico en el citoplasma (Tiessen, 2009).

2.4 Cromatografía de gases

La cromatografía de gases (GC, por sus siglas en inglés) permite altas resoluciones, sensibilidad y tiempos de corrida cortos, lo que ha contribuido a convertirla en la técnica de rutina en muchos laboratorios analíticos (Roessner et al., 2001; Liu y Rochfort, 2014). La GC tiene la ventaja de disponer de detectores universales (por ejemplo, el de ionización de flama o FID, por sus siglas en inglés). Además, para numerosas aplicaciones, los métodos GC son más simples, más rápidos y más sensibles que los de HPLC. La instrumentación requerida para GC también es más económica que la empleada en HPLC. Sin embargo, la GC presenta limitaciones en tres casos:

- a) Es ineficaz o inútil para el análisis de compuestos poco volátiles, generalmente los de peso molecular superior a 300 u.m.a.
- b) No se recomienda para el análisis de compuestos sensibles a elevadas temperaturas (muchos de ellos de alto interés biológico).
- c) No es funcional para compuestos que se encuentran en forma iónica en su forma nativa, ya que forman enlaces fuertes y son poco volátiles (moléculas de interés celular).

Por esta razón, la GC se emplea cuando los componentes de la muestra son volátiles y térmicamente estables a temperaturas de 150 a 400°C. Por ejemplo, es muy útil para el análisis de fragancias y olores. En cambio, cuando los compuestos a analizar son poco volátiles o termo-lábiles, la técnica preferida suele ser HPLC.

Tanto para GC como para HPLC se requiere la comparación de cromatogramas entre estándares (compuestos puros) y muestras complejas. Una de las dificultades es que no siempre se pueden adquirir estándares de forma comercial. Otra de las dificultades, es que puede haber diferentes compuestos con el mismo tiempo de retención, lo que conduce a errores cualitativos y cuantitativos. En consecuencia, las mejores técnicas analíticas son aquellas que combinan la capacidad de separación con la capacidad selectiva del detector, en el mejor de los casos usando equipos altamente sofisticados que permiten la identificación inequívoca del compuesto por medio de la fragmentación MS^n (conocidas como “técnicas acopladas”) (Agilent, 2008).

2.5 Espectrometría de masas

La cromatografía de líquidos (LC) está basada en las interacciones particulares de una muestra entre las fases móvil y la estacionaria. Tanto la electroforesis capilar (CE) como la espectrometría de masas (MS) no se clasifican dentro de los métodos cromatográficos (el contenido de las muestras no se separan por color), aunque realmente sí separan los compuestos (en muchos casos en forma ionizada) en base a sus propiedades químicas y físicas. En MS, en lugar de una fase móvil, se usa un campo eléctrico para acelerar los iones; En lugar de una fase estacionaria, la inercia física de las moléculas ayuda a separar los compuestos por su relación masa/ carga (m/z).

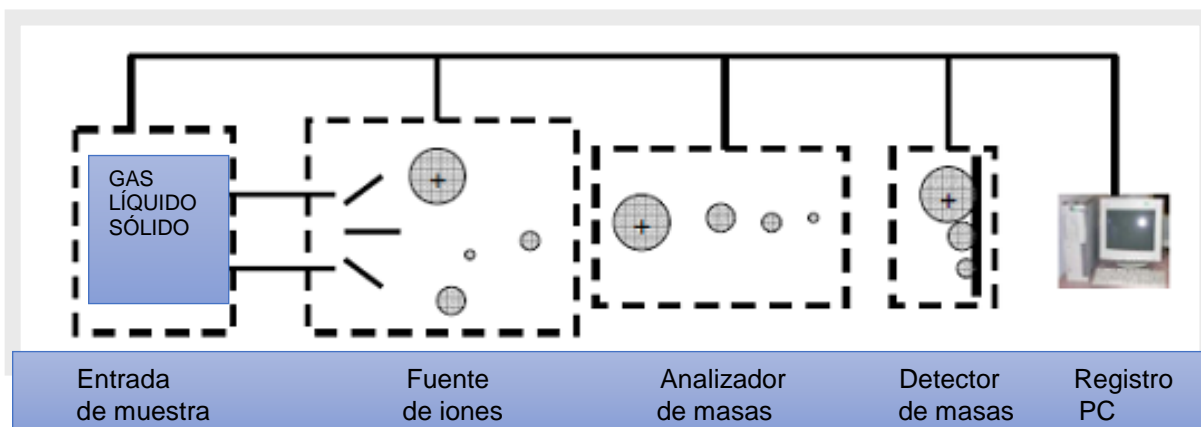


Figura 2.4 Operación típica de un espectrómetro de masas. La muestra puede ser ionizada en estado líquido sólido o gaseoso; a continuación se analizan las masas y después se detectan para finalmente ser registradas en la memoria de la PC (adaptado de Dunn, 2008).

La MS requiere una plataforma tecnológica considerable que permite la identificación de compuestos, la cuantificación de los mismos así como la determinación de la estructura de las moléculas. Se pueden analizar cantidades realmente pequeñas (en el rango fmolar) de muestra para obtener información característica como la masa molecular así como conformación estructural por medio de los patrones de fragmentación. Como requisito se requiere ionizar las moléculas, obteniéndose iones positivos o negativos en fase gaseosa que son acelerados hacia un analizador y separados en función de su relación m/z . La señal eléctrica generada por los iones que llegan al detector es procesada, ampliada y enviada a una computadora. El resultado que se obtiene es denominado “espectro de masas” y representa las abundancias iónicas según su m/z (Martín, 2012).

Los componentes principales de un espectrómetro de masas son los siguientes: a) sistema de bombeo; b) sistema de entrada; c) fuente de iones (cámara de ionización); d) óptica iónica; e) analizador de masas, y f) detector (Figura 2.4).

Se han desarrollado diversos métodos de ionización: a) ionización en fase gaseosa, ya sea electro-ionización (EI, por sus siglas en inglés), ionización química (CI, por sus siglas en inglés); b) ionización y desorción: ionización de campo (FI, por sus siglas en inglés), o desorción de campo (FD, por sus siglas en inglés); c) bombardeo de partículas: bombardeo de átomos rápidos rápido (FAB, por sus siglas en inglés); espectrometría de ión másico secundario (SIMS, por sus siglas en inglés); d) ionización a presión atmosférica: electrospray (ESI, por sus siglas en inglés), “termospray”, ionización química a presión atmosférica (APCI, por sus siglas en inglés), y e) desorción láser: ionización por desorción láser en matriz asistida (MALDI, por sus siglas en inglés) (Wait, 1993). El analizador de masas resuelve los iones después que fueron formados en la fuente de ionización; los de mayor utilización son los siguientes: a) sector magnético; b) cuadrupolo; c) trampa de iones, y d) tiempo de vuelo.

Con el continuo desarrollo de las diversas técnicas de ionización (e. g., EI, CI, ESI, MALDI, etc.), la MS se ha convertido en un método esencial en la caracterización tanto de metabolitos como de proteínas. Las técnicas de MS se han popularizado tanto, que se tiende a olvidar que la metabolómica también requiere de métodos no basados en m/z , como ensayos enzimáticos, espectroscopia UV-VIS, NIRS y diversos métodos como NMR, GC-FID, HPLC-RID, HPLC-DAD, etc.

Como se mencionó anteriormente, la combinación de GC con MS tiene la severa limitante de que no todas las moléculas de interés son volátiles, y las que sí lo son, generalmente no ionizan fácilmente. Es por ello que la GC-MS requiere de modificaciones previas (i.e., derivatización química) y métodos de ionización fuertes. Además, la cantidad de muestras que se pueden procesar por día es muy limitada, y esto restringe su uso en programas de mejoramiento genético. Por lo anterior, se buscó una técnica basada en MS, pero que no tuviera el requerimiento de derivatización, ni tampoco requiriera de una separación cromatográfica previa. Además, se buscó una plataforma que fuera económica, robusta, confiable y sencilla, y que permitiera el procesamiento de un gran número de muestras por día. Si bien todos los métodos metabolómicos tienen ventajas y desventajas, la técnica de ionización por electrospray con inyección directa (DIESI-MS) (Figura 2.5) se presentó como la mejor alternativa dentro del Cinvestav, Irapuato, gracias a las facilidades brindadas por el Laboratorio del Dr. Robert Winkler.

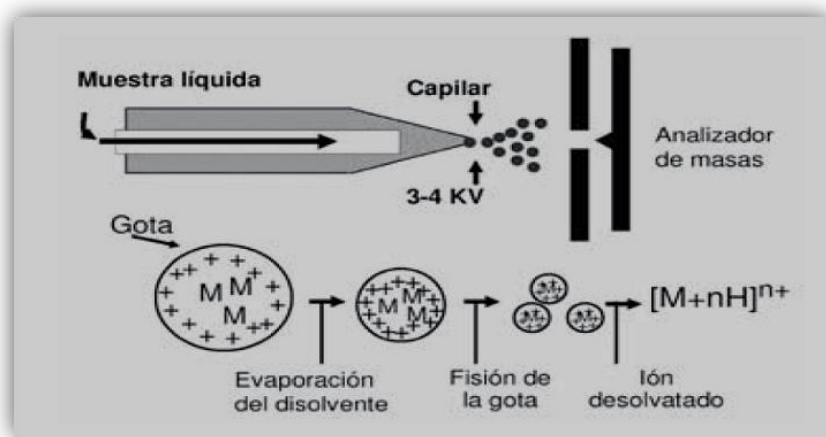


Figura 2.5 Esquema de un sistema de ionización por inyección directa y pulverización eléctrica a voltaje elevado; por efecto de la presión negativa y el contraflujo de gas nitrógeno se logra la desolvatación hasta obtener el ion positivo o negativo (ESI) (Martín, 2012).

2.6 Métodos estadísticos para análisis de la información

Los datos obtenidos tienen que ser interpretados en un contexto biológico. A la fecha, a fin de reconocer patrones en los datos obtenidos, hay herramientas estadísticas tales como son el conglomerado jerárquico (HCA) o el análisis de un componente principal (PCA) que son utilizadas rutinariamente por facilidad de comparación y visualización de similitudes y diferencias entre conjuntos de datos multivariados. Otro enfoque consiste en identificar dependencias y conexiones entre metabolitos y, más recientemente, entre genes, proteínas y metabolitos utilizando análisis de correlación lineal paso a paso, o “step wise”. Es interesante conocer que es posible la construcción de redes regulatorias cuando se detectan correlaciones significativas. La comparación de la conectividad de redes entre diferentes genotipos permite no sólo la identificación de nuevas rutas biosintéticas sino que también representa una opción para descubrir mutaciones silenciosas, las cuales no muestran un fenotipo obvio en ninguno de los parámetros usuales de análisis (Go, 2010).

El explosivo desarrollo de las computadoras en la última década ha facilitado el uso de regresión en el análisis estadístico. Regresión también es conocido como “Ajuste por cuadrados mínimos”, debido al método que se usa para estimar los

parámetros de regresión. Los siguientes son los principales usos de un modelo de regresión, aunque frecuentemente estos se dan al mismo tiempo en el análisis de un conjunto de datos: a) predicción, cuyo objetivo es pronosticar valores de la variable de respuesta para valores futuros de las variables predictoras, es decir para valores más allá del rango de valores de las variables de predicción en la muestra de referencia; b) descripción, cuyo objetivo es establecer una ecuación lineal o linealizable que describa la relación entre la variable dependiente y las variables predictoras; c) control, cuya función es controlar el comportamiento o variación de la variable de respuesta de acuerdo a los valores que asumen las variables predictoras y d) selección de variables, ya que inicialmente se pueden haber considerado muchas variables para explicar el comportamiento de la variable de respuesta a través de un modelo lineal. Sin embargo, la presencia de muchas variables de predicción puede afectar el rendimiento del modelo además de que la computación del mismo se puede volver lenta. Por lo tanto, es necesario usar técnicas para escoger solo las variables predictoras que sean más relevantes y aquellas que no sean redundantes para explicar la variación de la variable de respuesta (Castaño, 2003; Montgomery, 2009).

2.6.1 Regresión lineal “paso a paso” (o “step wise”).

La idea de estos métodos es elegir el mejor modelo en forma secuencial pero incluyendo (o excluyendo) una sola variable predictora en cada paso de acuerdo a ciertos criterios. El proceso secuencial termina cuando una regla, llamada “de parada”, se satisface. Hay tres algoritmos más comúnmente usados: a) eliminación hacia atrás; b) selección hacia adelante, y c) selección paso a paso. En la eliminación hacia atrás, se comienza con el modelo completo y en cada paso se va eliminando una variable. Si resultara que todas las variables predictoras son importantes, es decir tienen “*p-values*” pequeños para la prueba “*t*”, entonces no se modifica nada y el mejor modelo es el que tiene todas las variables predictoras disponibles. En caso contrario, en cada paso la variable que se elimina del modelo es aquella que satisface cualquiera de estos requisitos equivalentes entre sí: a) aquella variable que tiene el estadístico “*F*” parcial más pequeño; b) aquella variable que produce la menor disminución en el R^2 al ser eliminada del modelo, es decir, aquella variable que influye menos en la suma de cuadrados y c) aquella variable que tiene la correlación parcial (en valor absoluto) más pequeña con la variable de respuesta, tomando en cuenta las variables que quedarían en el modelo. La correlación parcial de Y con la variable X_i se define como la correlación entre los residuales de la regresión de Y con todas las variables predictoras, excepto X_i y los residuales de la regresión de X_i con todas las otras restantes variables predictoras. El método “*backward*” padece del efecto de anidamiento ya que toda variable que es eliminada del modelo ya no vuelve a ingresar al mismo (Acuña, 2011).

3. HIPÓTESIS GENERAL

El perfil metabólico de las muestras de jugo del tallo de maíz en floración está correlacionado con el estado fisiológico y el desempeño agronómico de la planta, y por lo tanto se puede usar para predecir el rendimiento final de forraje o de grano.

De ser cierta esta hipótesis, entonces el perfil metabólico pudiera ser utilizado con fines de mejoramiento genético, como un parámetro secundario de selección capaz de ser determinado en una etapa temprana del crecimiento de la planta. También puede servir para optimizar el manejo agronómico, o para determinar si un tratamiento incrementará o no el rendimiento final de grano.

4. OBJETIVOS

4.1 Objetivo General

Determinar un amplio perfil de metabolitos en el jugo del tallo de maíz y relacionar estos datos con los estados fisiológicos de la planta y con el desempeño agronómico.

4.2 Objetivos específicos

- a) Optimizar la técnica DIESI-MS y los algoritmos para realizar un fenotipo bioquímico de forma confiable y eficiente.
- b) Documentar y estar consciente de las ventajas y desventajas de DIESI-MS
- c) Cuantificar el perfil metabólico bajo las siguientes condiciones:
 - 1.- Diferentes etapas de desarrollo y fechas de siembra
 - 2.- Remoción de hojas (tejido “fuente”) y mazorca (tejido “sumidero”)
 - 3.- Condiciones de sequía y riego normal
 - 4.- Condiciones de bajo contenido de nitrógeno y fertilización normal
 - 5.- Utilización de diferentes genotipos
- d) Análisis por métodos estadísticos de regresión lineal, seleccionando factores con la metodología “paso a paso” (*step wise*).

5. MATERIALES Y MÉTODOS

5.1 Preguntas biológicas

Los estudios científicos pueden iniciar a partir de hipótesis, preguntas o problemas. De ahí deriva que el enfoque pueda ser de ciencia básica o de ciencia aplicada. Nuestras preguntas biológicas iniciales fueron las siguientes:

1. ¿Existen correlaciones significativas entre el desempeño agronómico de las plantas de maíz durante la etapa de cosecha, y los datos bioquímicos del jugo de tallo de las plantas cosechadas en etapa de floración?
2. ¿De qué manera influyen factores experimentales como son la fecha de siembra, la remoción de las hojas (tejido fuente) o de la mazorca (tejido sumidero) y las etapas de desarrollo de la planta en su perfil metabólico?
3. ¿Cuál es el comportamiento fisiológico y metabólico del tallo en condiciones de estrés hídrico o de bajo aporte de nitrógeno?
4. ¿Cuáles son las diferencias metabólicas en el jugo del tallo que presentan los diferentes genotipos de maíz?
5. ¿Cuáles procedimientos estadísticos y graficas multivariadas facilitan el análisis y la interpretación de los resultados que se obtengan de los diferentes experimentos?
6. ¿Cuál es el costo-beneficio de las diferentes técnicas metabolómicas?
7. ¿Qué es más importante para determinar el rendimiento de maíz, el ambiente o el genotipo?

El trabajo experimental se diseñó de acuerdo a los objetivos anteriormente planteados siguiendo pasos secuenciales estructurados (ver Figura 5.1). Nos enfocamos a estudiar los metabolitos del jugo de tallo de maíz con la finalidad de encontrar correlaciones con otros parámetros fisiológicos y agronómicos importantes para los mejoradores como la producción de biomasa verde, la calidad forrajera y el rendimiento de grano.

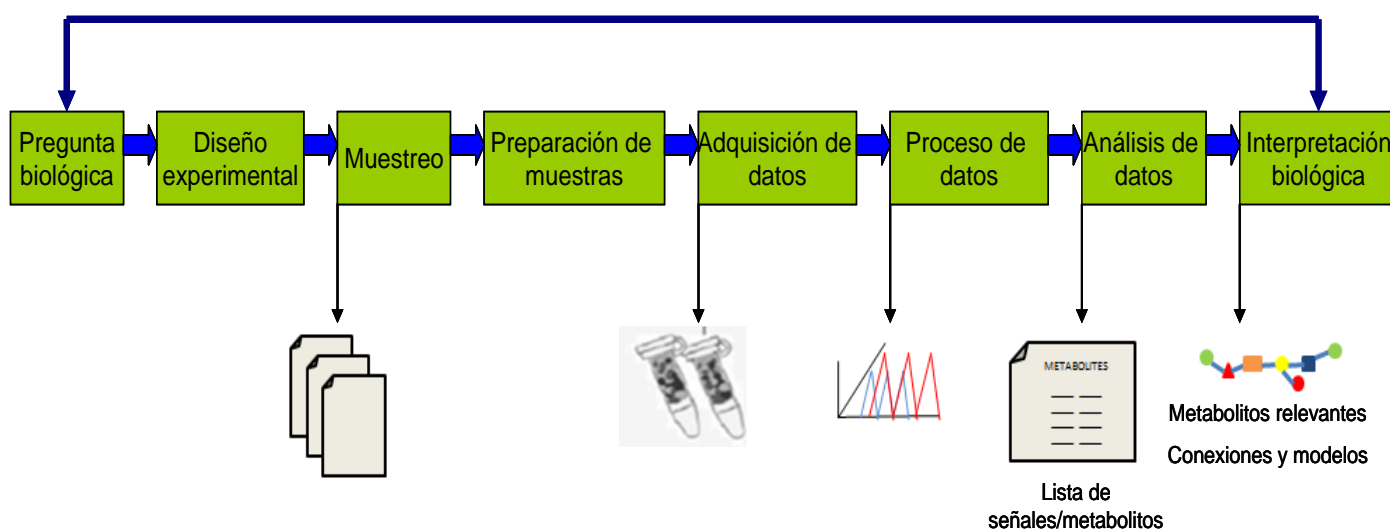


Figura 5.1 Estrategia experimental en investigaciones metabolómicas (Koek et al., 2011), similar a la empleada en este estudio.

5.2 Material biológico

5.2.1 Cultivo del Maíz y preparación de muestras de jugo del tallo

5.2.1.1 Experimento de estrés hídrico

El ensayo de sequía se llevó a cabo durante la temporada de invierno en la estación experimental del Cimmyt ubicada en Tlaltizapán, Morelos, México (18° 41' de latitud norte, 99° 07' de longitud O, 945 metros sobre el nivel del mar (msnm)). El régimen de riego se realizó de acuerdo a las normas de gestión de campo establecidas por el programa de fisiología del Cimmyt (Bänziger et al., 2000). El tratamiento de estrés (sequía) consistió en la supresión de riego durante tres semanas antes de la floración mientras que en el tratamiento control se regó con la frecuencia habitual (i. e., riego semanal) (Monneveux et al., 2008). Seis plantas por genotipo fueron muestreadas para cada tratamiento (sequía y control). Tres plantas por parcela se recogieron de cada experimento de campo. Los genotipos se cosecharon ~ 15 días después de la emisión de estigmas, a ~ 90 días después de la siembra para los controles y tratamientos de sequía (Tabla 5.1).

Tabla 5.1 Lista de genotipos usados en el experimento de estrés hídrico.

Genealogía	Entrada	Genotipo
DTPYC9-F143-5-4-1-2-B*5/CML312SR	8	G-008
DTPWC9-F24-2-3-1-1-B*5/CML312SR	20	G-020
CML323-BB/CML312SR	26	G-026
CML311/MBRC2BcF41-2-B*9/CML312SR	34	G-034
LaPostaSeqC7-F18-3-2-1-1-B*5/CML312SR	37	G-037
LaPostaSeqC7-F78-2-1-1-1-B*4/CML312SR	69	G-069
CML486-BB/CML312SR	72	G-072
CML311/MBRC3BcF95-2-2-1-B*7/CML312SR	87	G-087
LaPostaSeqC7-F96-1-2-1-3-B*4/CML312SR	89	G-089
DTPYC9-F69-3-5-1-1-B*4/CML312SR	91	G-091
LaPostaSeqC7-F64-2-6-2-2-BB/CML312SR	99	G-099
DTPWC9-F31-1-3-1-1-B*5/CML312SR	102	G-102
[KILIMAST94A]-30/MSV-03-1-10-B-1-BB	104	G-104
LaPostaSeqC7-F97-3-1-1-2-B*5/CML312SR	114	G-114
[GQL5]/[GQL5]/[MSRXPOOL9]	117	G-117
CML311xCML311/CML311	125	G-125
CLQ-RCWQ83=(CML146xCML150)	132	G-132
CLO2450Q/CML451QLocalCheckQPM2	143	G-143
LaPostaSeqC7-F64-2-6-2-2-BB	147	G-147

5.2.1.2 Experimento de estrés por bajo nitrógeno

El experimento de bajo contenido de nitrógeno (LN), junto con su respectivo experimento control (CT) se llevaron a cabo durante la temporada de verano de 2011 en El Batán, Texcoco, México [19° 30' de latitud norte, 98° 53' de longitud oeste, 2247 msnm. Se llevó a cabo de acuerdo con las normas de gestión de campo establecidos por el Centro Internacional de Mejoramiento de Maíz y Trigo (Cimmyt) (Monneveux et al., 2008). El tratamiento LN consistió en la supresión de la fertilización de nitrógeno durante los últimos 4 años (~ 8 ciclos), mientras que el

campo de control se fertilizó químicamente con urea con la frecuencia (3 veces por ciclo de crecimiento) y la dosis (~ 400 kg / ha) habitual. La densidad de siembra fue de 67 000 plantas / ha en dos hileras de 5 m distribuidos en bloques al azar (“ α lattice”) (Monneveux et al., 2008). Seis plantas se muestrearon por genotipo para cada tratamiento (LN y CT). Se cosecharon los genotipos del ensayo de campo CIMMYT HL2011-11-01-PVB-2 ~ 10 días después de la emisión de estigmas (~ 85 días después de la siembra), mientras que los genotipos de la prueba de campo CIMMYT HL2011-11- 01-PVB-1 fue cosechado ~ 12 días después de la emisión de estigmas (~ 85 días después de la siembra) (

Tabla 5.2

Tabla 5.2 Lista de genotipos usados en el experimento de bajo nitrógeno.

Genealogía	Entrada	Genotipo
Testigo Local	1	G-1
CV-702	2	G-2
HID-15	3	G-3
CMT099027	4	G-4
CMT099003	5	G-5
A-823	6	G-6
H57	7	G-7
(CML457/CML459)//IML6)	8	G-8
H55	9	G-9
P1684	10	G-10
CMT099001	11	G-11
CMT 099004	1	G-1
INSURGENTE	2	G-2
CMD 080001	3	G-3
P1832	4	G-4
CMT 099002	5	G-5
Testigo Local	6	G-6
TORITO AMARILLO 1	7	G-7
CMT 029502	8	G-8

5.2.1.3 Experimento fuente-sumidero (o “sink-source”)

Se realizaron dos experimentos fuente-sumidero (Knoblauch y Peters, 2013) con el fin de comprobar si la eliminación de órganos fuente y órganos sumidero, como son las hojas y los elotes, respectivamente (Paul and Foyer, 2001; Rolland et al., 2002), afectan los niveles de metabolitos en el tallo del maíz. Se definieron dos grupos de control en el tiempo, uno antes de la remoción de los órganos (CB, por sus siglas en inglés) y otro después de la remoción (CA, por sus siglas en inglés). El experimento se llevó a cabo con un número de plantas sembradas en el mes de Junio de 2010 y también utilizamos los mismos 5 genotipos como está descrito en el experimento de la etapa del desarrollo 2010. La primera parte del experimento se llevó a cabo el 20 de agosto de 2010, 75 días después de la siembra (DAS, por sus siglas en inglés), y unos diez días después de la emergencia de estigmas. Previamente, etiquetamos grupos de cinco plantas al azar de cada grupo para cada tratamiento e híbrido (Monneveux et al., 2008; Montgomery, 2008). En total se utilizaron 20 plantas sanas y vigorosas de cada genotipo. Los tratamientos se aplicaron en la mañana, entre las 9 y las 10 a. m. El

tratamiento WL se realizó mediante la reducción de todas las láminas de las hojas con unas tijeras; dejamos la vaina y el tallo intactos para garantizar la estabilidad mecánica del tallo. La justificación de este tratamiento fue para bloquear drásticamente el flujo de sacarosa originada por la fotosíntesis en las hojas. Por lo tanto, el tratamiento WL es un experimento de manipulación de la fuente (Rolland et al., 2002; Slewinski et al., 2009). En el tratamiento de eliminación de sumidero (WE) (Paponov et al., 2005), todos los elotes en desarrollo (cada planta podría contener 1-2 elotes) fueron retirados. Los grupos de control se mantuvieron intactos y se cosecharon en los dos puntos de tiempo relevantes. En la fecha de tratamiento, se colectaron únicamente las muestras de CB. Después de cuatro días, se procedió a probar el resto de los tratamientos, WE, WL, y CA (Paul y Driscoll, 1997; Juárez, 2012).

5.2.1.4 Experimento etapas de desarrollo 2010

El experimento se llevó a cabo en el campo experimental de la Universidad de Guanajuato, en Irapuato. Trabajamos con cinco híbridos de maíz comercial: 3 maíces blancos (H1 Puma, H2 Oso, H3 Leopardo: semillas donadas por INIFAP Celaya), y 2 maíces amarillos (H4 Dow2A, H5 Dow2B: semillas donadas por INIFAP Celaya). Con el fin de comparar las fechas de siembra, un primer grupo de plantas se sembró el 23 de mayo, 2010. El segundo se sembró aproximadamente dos semanas después (el 8 de junio). Ambos se sembraron en filas de 5m, tratando de tener 25 plantas en cada fila. Manejamos cinco bloques repetidos para cada genotipo, dando un total de al menos 125 plantas de cada genotipo para el muestreo. Colectamos muestras cada semana antes, durante y después del tiempo de la floración (Bänziger et al., 2000; Monneveux et al., 2008). Tratamos de elegir plantas saludables y representativas de diferentes bloques para hacer grupos aleatorios de tres plantas por híbridos y por fecha de siembra (Montgomery, 2008). El primer muestreo se realizó a 60 DAS, mientras que los otros grupos se tomaron muestras por primera vez 46 DAS. Los tiempos de colecta para el mes de Mayo fueron: 60, 67, 74, 81, 88, 95, y 102 DAS, mientras que para los sembrados en junio se colectaron a los 46, 53, 60, 67, 74, 81 y 88 DAS (Bänziger et al., 2000; Juárez, 2012).

5.2.1.5 Experimento etapas de desarrollo 2011

El experimento se llevó a cabo en un invernadero del Cinvestav, en Irapuato. Trabajamos con los cinco híbridos de maíz comercial que se usaron en el experimento de etapas de desarrollo 2010: 3 maíces blancos (H1 Puma, H2 Oso, H3 Leopardo), y 2 maíces amarillos (H4 Dow2A, H5 Dow2B). Se manejaron dos sitios de siembra, el primero se estableció dentro del invernadero y el segundo fuera del invernadero y a ambos se les aplicó los mismos tratamientos, con vaina y sin vaina (Rolland et al., 2002; Slewinski et al., 2009). Colectamos muestras cada semana antes, durante y después del tiempo de la floración (Monneveux et al., 2008). Tratamos de elegir plantas saludables y representativas de diferentes bloques para hacer grupos aleatorios de tres plantas por híbridos y por fecha de siembra. Los tiempos de cosecha fueron los siguientes: 39, 46, 53, 60, 67, 74, 81, 95, 102, 109 y 116 DAS (Bänziger et al., 2000; Juárez, 2012).

5.3 Experimentos preliminares

Los índices de los parámetros analíticos en espectrometría de masas son importantes para mantener bajo control el proceso experimental, entre algunos de

ellos tenemos al rango de sensibilidad en m/z , nivel de resolución, coeficiente de variación, exactitud, reproducibilidad técnica y la linealidad que tenga un nivel elevado de correlación. El nivel de dilución de muestra con el que se obtuvo una alta resolución de señal se seleccionó a partir de los niveles 1:100, 1:1000 y 1:10000. Con el programa *TOPPAS view* (Junker et al., 2012) se revisaron visualmente los espectros metabólicos de diferentes muestras de jugo de tallo, relacionando los valores puntuales de intensidad y en rangos de iones en función de su m/z . También se registraron lecturas con los espectros de las huellas metabólicas de carbohidratos, (sacarosa, glucosa y fructosa, etc.) y aminoácidos, (alanina, prolina, histidina, etc.) a diferentes niveles de dilución y con diferentes ionizantes (cloruro de litio, LiCl; fluoruro de sodio, NaF; fluoruro de amonio, NH_4F ; hidróxido de amonio, NH_4OH , y ácido fórmico, CH_2O_2) y absorbentes (carbón activado, resina C-18 y polivilpirrolidona) para probar su abundancia de iones y los niveles de intensidad de los mismos en los cuadrupolos Z/Q y SQD2. Posteriormente se obtuvieron espectros de las huellas metabólicas de los mismos carbohidratos antes mencionados en el triple cuadrupolo (TQD) (Chai et al., 2002; Banerjee et al., 2011); mediante fragmentación controlada se procesaron las huellas metabólicas con el fragmento particular del carbohidrato asociado a los iones derivados de las reacciones de deshidratación y descarboxilación del carbohidrato; dichos fragmentos se analizaron permitiendo la identificación del analito en muestras del jugo de tallo de maíz ionizadas con ácido fórmico al 5% (v/v).

En una primera aplicación de los *scripts* en el lenguaje R (R-Development-Core-Team, 2008), se establecieron comparaciones mediante *heatmaps*, *boxplots* y dendogramas entre los tratamientos de experimentos de estrés hídrico y con bajo contenido de N, y el nivel de dilución con absorbentes. En la Figura 5.2 se muestra el diseño experimental combinado de sustancias ionizantes y diferentes niveles de dilución para probar la linealidad de un cuadrupolo sencillo, en especial el SQD2. Mediante estos experimentos preliminares se pretendió solucionar las limitaciones de DIESI por linealidad, supresión de iones, efecto matriz y la asignación de señales a un metabolito específico (Tiessen, 2009). Los archivos y análisis de los experimentos preliminares se encuentran en la base de datos compartidos del Laboratorio de Metabolómica y Fisiología Molecular del Cinvestav-Irapuato.

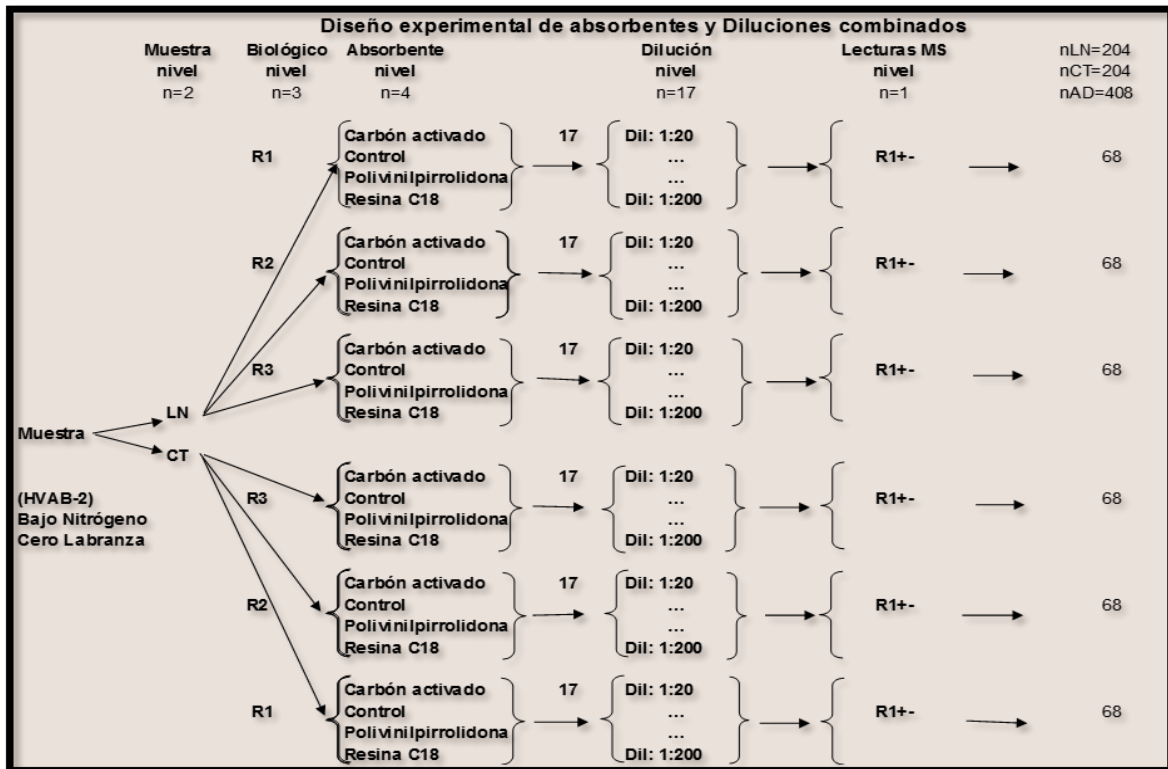


Figura 5.2 Diseño experimental para analizar el efecto combinado de los absorbentes: carbón activado, polivinilpirrolidona, resina C18 y una muestra control bajo diferentes niveles de dilución: 1:20, 1:30, 1:40, 1:60, 1:80, 1:90, 1:100, 1:110, 1:120, 1:130, 1:140, 1:150, 1:160, 1:170, 1:180, 1:190 y 1:200. Las lecturas se hicieron en el cuadrupolo sencillo (SQD) de Waters en modo positivo y negativo (R1+,-: una repetición en modo positivo y modo negativo). Las características de los absorbentes fueron las siguientes: el carbón activado retiene compuestos aromáticos, fenoles e hidrofóbicos; la resina C18 retiene a los compuestos hidrofóbicos y la polivinilpirrolidona retiene compuestos polifenólicos.

5.4 Recolección del jugo de tallo de maíz

En todos los experimentos, se seleccionaron plantas sanas representativas de cada genotipo y del tratamiento en el campo y procesadas dentro de ~ 15 minutos después de la cosecha. Los datos morfológicos y fisiológicos que se registraron fueron: peso total de la planta, altura de la planta (hasta la hoja bandera), número de mazorcas, el peso del elote y el peso de tallo. La espiga y las hojas laminares se retiraron del tallo de soporte (incluyendo la hoja de la vaina) se cortó en segmentos (~ 30 cm) para la homogeneización del tejido. El jugo del tallo (floema savia + xilema savia + extractos celulares) se obtuvo usando un extractor mecánico (modelo de extractor de jugo escala industrial EXS, Cooperación Internacional, México). El jugo de tallo (~ 80 ml) se recogió en un vaso de precipitados de plástico, ponderado con precisión ($\pm 0,5$ g), se dividió en alícuotas y se congeló inmediatamente en hielo seco (< 30 s). Alícuotas del jugo de tallo (1 ml) se almacenaron a -80 °C durante ~ 4 meses en placas selladas de pozos profundos (formato de 96 pozos). Para el análisis, el jugo de tallo se descongeló y se mantuvo en hielo durante cortos períodos de tiempo (~ 1 h). Las muestras se centrifugaron en primer lugar (10 min a 4000 rpm [$\sim 5000 \times g$] a 4 °C) y el sobrenadante se usó directamente para el análisis de proteínas. Para cuantificación de metabolitos por

DIESI-MS, el sobrenadante se filtró (0,4 μm) y después se diluyó apropiadamente (1: 100) y se acidificó con ácido fórmico (ver más abajo).

5.5 Análisis de metabolitos.

Los métodos de análisis se evaluaron teniendo en cuenta los costos, rendimiento de la muestra, y los requisitos de separación cromatográfica y la derivación química. También se evaluó el flujo de trabajo requerido para los análisis computacionales y estadísticos de los datos generados. Se excluyeron NIRS porque es bien sabido que se puede detectar simultáneamente / discriminar sólo 1-3 compuestos diferentes (Spielbauer et al., 2009), mientras que los métodos basados en MS producen información sobre más de 300 metabolitos diferentes. También se excluyeron la UPLC-MS de alto rendimiento y la resonancia magnética nuclear (NMR), ya que los costos del equipo son prohibitivos para el fitomejoramiento de rutina.

5.5.1 Ensayos enzimáticos espectrofotométricos.

Los azúcares solubles se midieron en formato de microplacas (Tiessen et al., 2002; Angeles-Nunez y Tiessen, 2010). El sedimento de almidón insoluble se disolvió en 0,5 ml de KOH 10 mM a 99 °C y en autoclave durante 30 min. El almidón se hidrolizó en 50 mM de N-2-hydroxyethylpiperazine-ácido-N'-2-etanosulfónico (HEPES) (pH 7,5) a 37 °C durante la noche por la adición de 10 unidades de α -amilasa (A3176, Sigma; San Luis, MO. Estados Unidos de Norteamérica) y 10 unidades de amiloglucosidasa (A9913, Sigma; San. Luis, MO. Estados Unidos de Norteamérica). Las muestras se centrifugaron (13000 \times g durante 5 min); el sobrenadante resultante se almacenó a 4 °C; y el sedimento se hidrolizó de nuevo durante 30 min a 37 °C. Ambos sobrenadantes se combinaron, y una alícuota se ensayó enzimáticamente para la glucosa, como se describe anteriormente.

5.5.2 Espectrometría de masas con inyección directa ionización por electrospray (DIESI-MS)

El ensayo DIESI-MS se realizó como se ha descrito previamente (García-Flores et al., 2012; Montero-Vargas et al., 2013) con algunos ajustes (ver abajo). El sobrenadante del jugo de tallo se filtró a través de un filtro de 0,45 μm de polivinilideno difluoruro y se diluyó 1:100 con agua desionizada para evitar la saturación del capilar y suciedad del cono. Para mejorar la ionización, a 475 μl de muestra diluida y 25 μl de ácido fórmico puro se añadieron [concentración final de 5% (v/v) de ácido fórmico]. Los experimentos de control se llevaron a cabo para determinar los parámetros óptimos de pretratamiento de la muestra y de dilución (datos no mostrados). Las muestras diluidas y acidificadas fueron analizadas por DIESI-MS en modo positivo y negativo, el empleo de un Micromass ZQ 2000 y un SQD2 con un analizador de cuadrupolo (Waters) y MassLynx 4.0 como software de control (García-Flores et al., 2012). Una tasa de flujo volumétrico constante (10 μl / min) se consiguió utilizando una bomba de jeringa cargada con 100 μl de muestra. La tensión capilar se fijó a 3 kV; el voltaje de cono se fijó a 60 V; la tensión de extractor se fijó a 3 V; y la frecuencia de radio (RF) de la lente se dejó a 0.5 V. La temperatura de la fuente (80 °C) y la temperatura de desolvatación (150 °C) se usa con un flujo de gas de nitrógeno de 250 L/h y un flujo de gas de cono de 50 L/h. En la sección de análisis, de masa baja (LM, por sus siglas en inglés) y de masa alta

(HM, por sus siglas en inglés), se estableció una resolución de 15.0 y una energía de iones de 0.5. El multiplicador se ajustó a un valor de 650. Se recogieron espectros continuos en una gama de m/z 15-2000, con una duración de ejecución de 1 min, un tiempo de exploración de 10 s, y un tiempo de *InterScan* de 0.1 s, produciendo seis espectros por muestra.

5.5.4 Cromatografía de Gases (GC-FID). Condiciones de operación.

Para cuantificación de compuestos volátiles, se realizó la inyección de la muestra en el cromatógrafo marca Agilent 7890a, columna HP Inowax. Las condiciones de operación del cromatógrafo así como otros datos del proceso de inyección se muestran a continuación:

Componentes del cromatógrafo de gases (GC)	Datos
Detector	FID (flama)
Temperatura del detector	250 °C
Temperatura del inyector	150 °C
Columna	0.25 mm x 30 m
Temperatura columna	Inicial 60 °C
	Final 180 °C
Rango de temperatura	12 °C / min
Tiempo	3 min
Rampa, Gradiente de temperatura	27 °C / min.
Flujo de Gas (H2), Aire, N2.	40, 440, 5 ml / min

Figura 5.3 Metodología para el tratamiento de las muestras de jugo de tallo de maíz en la prueba analítica de cromatografía de gases.



5.6 Análisis estadístico de datos de DIESI-MS con programa R

Análisis de datos y optimización de script. Los archivos primarios de los espectros de masas se convirtieron a *.mzXML usando masswolf, versión 1.4. El análisis de los espectros de masas se realizó con el software OpenMS / TOPP, versión 1.8.0 (Junker et al., 2012), y posteriormente procesados con el software estadístico R libre (<http://www.r-project.org>) (R-Development-Core-Team, 2008). Un diagrama de flujo en Toppas fue escrito para la ejecución de las siguientes tareas secuenciales: conversión de archivos a mzML, fusión espectros con un método de bloque (seis espectros de réplica), sgolay NoiseFilter con una longitud de trama de 21 y el orden polinomio de 4, seguido de un PeakPicker con una señalización relación a ruido (S/N) de 1 y una anchura de pico de 0.15 (

Figura 5.4 .

Por último, las tablas de datos se convirtieron en archivos Toppas (Junker et al., 2012) utilizando formato de texto, que contiene columnas separadas por tabuladores para los nombres de genotipo, factores de tratamiento, los valores m/z , y la intensidad de señal correspondientes de los iones individuales. Archivos de datos individuales fueron importados de forma automática con un guión adaptado en R (R-Development-Core-Team, 2008). Además de la recolección de la muestra y la optimización del flujo de trabajo en DIESI-MS, otras ventajas/novedades presentadas en este documento son los algoritmos mejorados para el procesamiento de datos. El nuevo guión R importa todos los archivos de texto desde un directorio dado y permitió la definición de análisis de variación (ANOVA) de factores experimentales basados en el nombre del archivo original. La asignación de variables se realizó dentro del entorno R y no requiere software adicional para la fusión de archivos o el cambio de nombre. El guión R también se ha optimizado para procesar los datos de cientos de muestras en 1/10 del tiempo de los guiones anteriores (Garcia-Flores et al., 2012; Montero-Vargas et al., 2013).

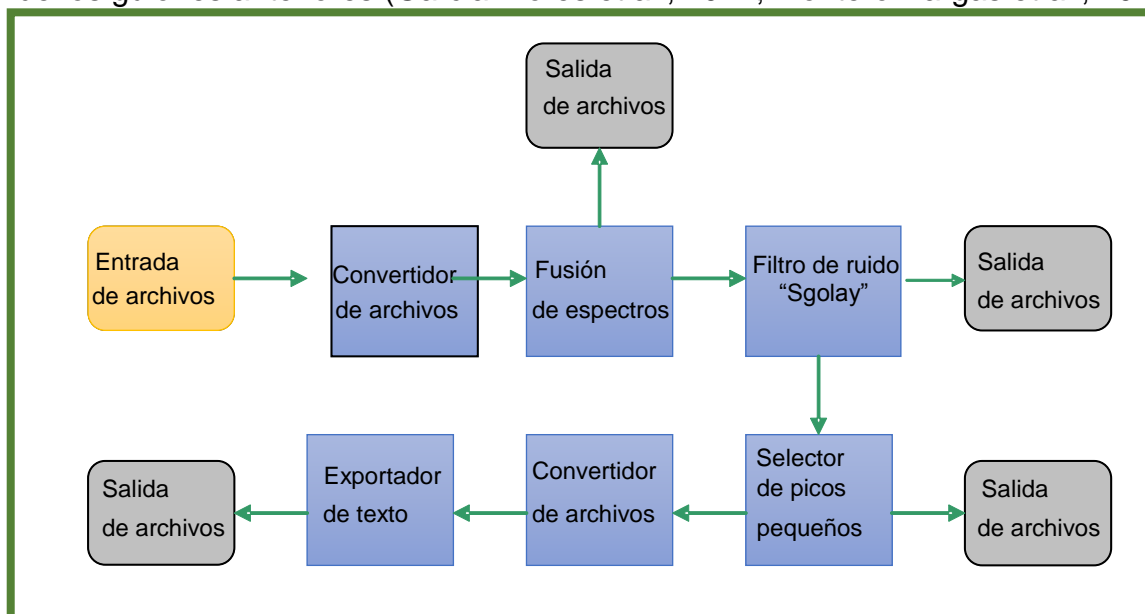


Figura 5.4 Diagrama para procesar los archivos de Toppas MS. El proceso inicia con la entrada de los archivos de lectura del espectro de masas y termina con la salida de archivos procesados en formato de Excel. Las flechas de color verde indican la dirección del flujo del procesamiento de los archivos. Los parámetros de proceso para cada etapa

son definidos en el método diseñado con el programa Topas MS (Adaptado de Sturm et al., 2008; Winkler, 2010).

En lugar de nombrar picos por límites de intervalo, los picos m/z fueron nombrados por su centro de intensidades máximas. En lugar de definir rangos de masas en toda la gama m/z (García-Flores et al., 2012; Montero-Vargas et al., 2013), sólo m/z picos reales fueron transferidos a la matriz rectangular resultante utilizada para los gráficos de mapas de calor (*Heatmaps*). En lugar de definir límites "bin" centrados a 0, se introdujo un cambio de masa para reducir los saltos de rango. Los límites de rango m/z se ajustaron de acuerdo con el patrón de m/z observado (espacio metabólico) en diversos extractos: en el rango de masa de 50-400 Da, los picos de m/z se centraron en 0.2 Da de la masa unitaria. En el rango de masa de 400 a 700 Da, un cambio de masa de 0.3 Da fue más preciso. Los iones fueron menos frecuentes en el rango de masa de 800-2000 y tenían más variabilidad en desplazamiento de masa, por lo que era más conveniente definir rangos centrados basados en la aparición de picos de iones específicos (espacio metabólico).

El script R mejorado también nos permitió definir de forma variable la tolerancia de asignación de picos m/z (tamaño *bin*). Sin embargo, para nuestras muestras de plantas, había poca ganancia de información mediante la definición de "bins" de masa menor que ± 0.25 Da, y por lo tanto, el tamaño de rango se definió convenientemente en intervalos de 1 Da (± 0.5 Da desde el centro) en toda la gama de m/z (30-2000 Da).

Cuando aparecieron dobles picos en el mismo rango (por ejemplo, picos isotópicos), las intensidades fueron resumidas y no promediadas ni eliminadas. La secuencia de comandos R optimizada produjo tablas separadas que documentan el número de picos por rango, la intensidad de los picos individuales, y se suma la intensidad de picos múltiples. El control automático de la calidad se llevó a cabo para detectar dobles picos de incidencia en los rangos de masas individuales.

Además de presentar los datos metabólicos como mapas de calor de colores, el script R optimizado produjo "*heatmaps bicluster*" grises y varias figuras con diagramas de caja que muestran las estadísticas y controles de calidad. Por ejemplo, además de informar los valores de intensidad para m/z dados en rango de iones, la secuencia de comandos también computan el valor exacto m/z para cada grupo de muestra. El m/z residual (diferencia entre la masa real observado y la masa unitaria) se calculó para cada pico de la muestra y se representó en paralelo a los diagramas de caja de intensidad (por ejemplo, paneles D-F de la figura 8, Publicación No.1). Desde este enfoque, los usuarios pueden decidir si un valor m/z dado en una muestra de la planta corresponde al mismo metabolito en otra muestra de la planta. Lo ideal es que los valores de m/z del mismo compuesto a través de muestras deben ser similares. Pero en la práctica, lo hicieron variar ± 0.2 Da con nuestro equipo de cuadrupolo sencillo MS.

En general, el algoritmo optimizado permite determinar la importación automática de miles de archivos, los cuales se admiten más rápido para el procesamiento de datos, y también aumentó significativamente la fiabilidad al disminuir "saltos pico" y la eliminación de la exigencia de post-procesamiento manual de los datos.

Procedimientos estadísticos. Todos los experimentos se realizaron utilizando diseños de bloques completos al azar con un número adecuado de controles y

réplicas (ver las leyendas). Se aplicaron pruebas ANOVA y Tukey con la opción de mínima diferencia significativa (LSD) con $p \leq 0,05$ ($\alpha = 0,05$). Se utilizaron modelos lineales de efectos mixtos para el análisis de datos, con factores independientes para el efecto de la G, E, etapa de desarrollo (DS), y sus respectivas interacciones. La corrección de Bonferroni se utilizó para hacer frente a la tasa de falsos positivos. Para el análisis estadístico, elaboración de los dendogramas y mapas de calor, se empleó R versión 2.14.1 con las bibliotecas: *vegetariana* y *pheatmap* (R-Development-Core-Team, 2008).

La secuencia de comandos R se proporciona como datos suplementarios. Los datos de espectrometría de masas se importan, ordenados en “bins” con un intervalo de 1 m/z y organizados en una matriz. Después, los datos se normalizaron con respecto a la intensidad de la señal de los espectros. La agrupación jerárquica se limita a los 100 rangos de masas que muestran las intensidades de señal más altas en cualquiera de las filas. Finalmente, la agrupación jerárquica se realiza con filas escala, usando el algoritmo de Euclides para ambos, filas y columnas.

5.7 Diseño experimental en campo y en laboratorio

El experimento de campo fue diseñado y supervisado por los mejoradores del Cimmyt. Por lo general, los ensayos corresponden a diseños *alfa lattice* de bloques incompletos. Se cosechan al menos 6 plantas diferentes de cada genotipo o tratamiento (6 réplicas biológicas que se cosecharon independientemente en dos grupos de 3). La obtención del jugo del tallo de cada planta de maíz se hizo por duplicado. Las lecturas en cromatografía de gases y líquida de alta presión se realizaron en una sola lectura, siguiendo el orden de ubicación de las muestras en placa *deep well*, en el cual se consideran repeticiones biológicas. Las repeticiones técnicas se preparan en el momento de la dilución de la muestra. En DIESI-MS, las inyecciones se hicieron al menos por triplicado. De cada inyección que dura >1 min se realizaron 6 lecturas de espectros MS, que después se promediaron con el software TOPPAS. Para las mediciones en el espectrofotómetro UV-VIS, se siguió el orden de la muestra en placas y la repetición técnica se preparó en el momento de la medición. Todas y cada una de las técnicas analíticas se desarrollaron de acuerdo a los protocolos establecidos en el Laboratorio de Metabolómica y Fisiología Molecular por el Dr. Axel Tiessen.

5.7 Definición de jugo

Según las normas mexicanas, un jugo es un extracto líquido puro de algún tejido vegetal sin contener ningún aditivo extra. El néctar que se vende en los supermercados por ejemplo contiene azúcares, espesantes y conservadores. Un jugo debe prepararse libre de gomas, sólidos y de bagazo. Por lo regular el jugo se prepara de alguna fruta (jugo de naranja, manzana), pero también hay jugos de raíz (jugo de zanahoria), o de hojas y tallos (jugo de apio). El tallo del maíz y los tejidos vasculares contienen diferentes tipos de células, cada una con funciones fisiológicas diferentes. Por ejemplo, la células del floema contienen los foto asimilados (sacarosa y amino ácidos) que son transportados de un tejido fuente a un tejido demanda. Algunos insectos (mosquitos y áfidos) son capaces de extraer la savia del floema por medio de su fino estilete. En el xilema y el apoplasto se transporta agua y minerales, que constituyen un segundo tipo de savia (savia bruta). En este trabajo doctoral se llama jugo de tallo de maíz a esa mezcla de la

savia del floema y la savia del xilema, pero que incluye también el contenido de las células del mesófilo, del apoplasto y de la epidermis. El extracto que preparamos del tallo de maíz no es savia en el sentido botánico, sino jugo del tallo de maíz en su sentido más puro según las leyes internacionales.

6. RESULTADOS

La meta principal de este trabajo fue el establecimiento de una técnica confiable y eficiente para el fenotipo metabólico de plantas de maíz. Usando extractos de los tejidos vasculares del maíz (jugo de tallo) se trató de clasificar el estado fisiológico de las plantas usando espectrometría de masas y otras técnicas analíticas (GC-FID y UV-VIS). Para lograr implementar la metodología DIESI-MS, primero fue necesario llevar a cabo experimentos exploratorios y ajustar las condiciones generales y específicas. También era necesario optimizar los algoritmos estadísticos y los scripts de R. En la presente tesis de doctorado se dió especial lugar a incluir todos aquellos resultados, figuras y gráficas que por limitación de espacio no fueron incorporados a las publicaciones resultantes de este trabajo doctoral (García-Flores et al 2012; García-Flores et al 2014). Los resultados control son importantes para comprender la técnica de *fingerprinting*, describir sus ventajas y limitantes. Algunos colegas científicos que prefieren la cromatografía de gases (GC-MS) o la cromatografía líquida (LC-MS) ven con escepticismo la técnica DIESI-MS debido a diversos problemas como la complejidad extrema de la matriz (muestra biológica). Los datos DIESI-MS se consideran poco confiables y son menos fáciles de publicar comparado con otras técnicas metabolómicas como NMR (*nuclear magnetic resonance*). En ese contexto resulta entonces relevante presentar algunos resultados DIESI-MS relacionados con la linealidad, la supresión de iones y la interacción con la matrix. Primero se mostrarán algunos resultados control y posteriormente se presentarán los resultados de los experimentos de campo con maíz, los cuales fueron estrés de sequía y bajo contenido de N. Finalmente, en el anexo se incluyen las figuras y las publicaciones en las que participé como primer autor (García-Flores et al 2012; García-Flores et al 2014).

6.1 Experimentos control sobre DIESI-MS

Se obtuvieron espectros de masas (perfiles ionómicos o huellas metabólicas) de diferentes muestras de jugo de tallo encontrándose diferencias globales y puntuales de intensidades de iones en función de su m/z . La forma preferencial para mostrar los datos DIESI-MS es a través de *boxplots* (cajas de datos) y *heatmaps* (mapas de calor) (Figura 6.1). En el mapa de calor se agrupan las muestras por columna mientras que los iones están en cada fila (Figura 6.1). El dendograma (cluster herárquico) agrupa las muestras o los iones según alguna función de similitud, que por *default* en R es por distancia euclídeana (R-Development-Core-Team, 2008). También se pueden usar otras funciones para el agrupamiento, por ejemplo el coeficiente R de correlación (Tiessen et al 2016). Sin embargo, esto lo descubrimos posteriormente por lo que todas las figuras de la tesis se realizaron con las funciones *default* de R que nos permitió publicar 2 artículos internacionales (García-Flores et al 2012; García-Flores et al 2014).

Con la primera implementación de los *scripts* en TOPPAS y en el lenguaje R, se establecieron comparaciones mediante mapas de calor, gráficas tipo caja y dendogramas entre los tratamientos de experimentos de estrés hídrico, con bajo contenido de N, y nivel de dilución con absorbentes (Figuras 6.1 y 6.2; ver también García-Flores et al 2012; García-Flores et al 2014).

Se realizó un primer experimento de estrés de bajo nitrógeno con plantas de maíz, y el jugo del tallo se inyectó en el DIESI-MS (Figura 6.1). Las 9 replicas biológicas

de bajo nitrógeno se agruparon muy separadas de las demás muestras. Esto fue muy prometedor para un fenotipo metabólico. La Figura 6.1 comprueba que el estado fisiológico de la planta cambia el perfil ionómico del tallo de forma considerable (ver también García-Flores et al 2012; García-Flores et al 2014). El jugo de tallo de maíz contiene niveles altos de sacarosa y de glucosa, pero desafortunadamente estas moléculas no tienen carga ni ionizan de forma natural, por lo que son invisibles para el DIESI-MS. Sin embargo, se puede agregar aditivos para visualizar los azúcares. Se registraron los espectros de las huellas metabólicas de carbohidratos (sacarosa, glucosa, fructosa, etc.) y aminoácidos (alanina, prolina, histidina, etc.) a diferentes niveles de dilución y con diferentes ionizantes y absorbentes. El cloruro de litio (LiCl) resultó ser el aditivo más adecuado para la ionización de carbohidratos con lecturas en modo negativo en el cuadrupolo sencillo (SQD). Posteriormente, se obtuvieron espectros de los mismos carbohidratos en el triple cuadrupolo (TQD). Mediante fragmentación controlada se obtuvieron huellas metabólicas de un fragmento particular de cada carbohidrato asociado a los iones derivados de las reacciones de deshidratación. Dichos fragmentos permitieron la identificación del analito en muestras del jugo de tallo de maíz ionizadas con ácido fórmico al 5% (v/v). Los archivos se encuentran en la base de datos compartidos del Laboratorio de Metabolómica y Fisiología Molecular del Cinvestav-Irapuato.

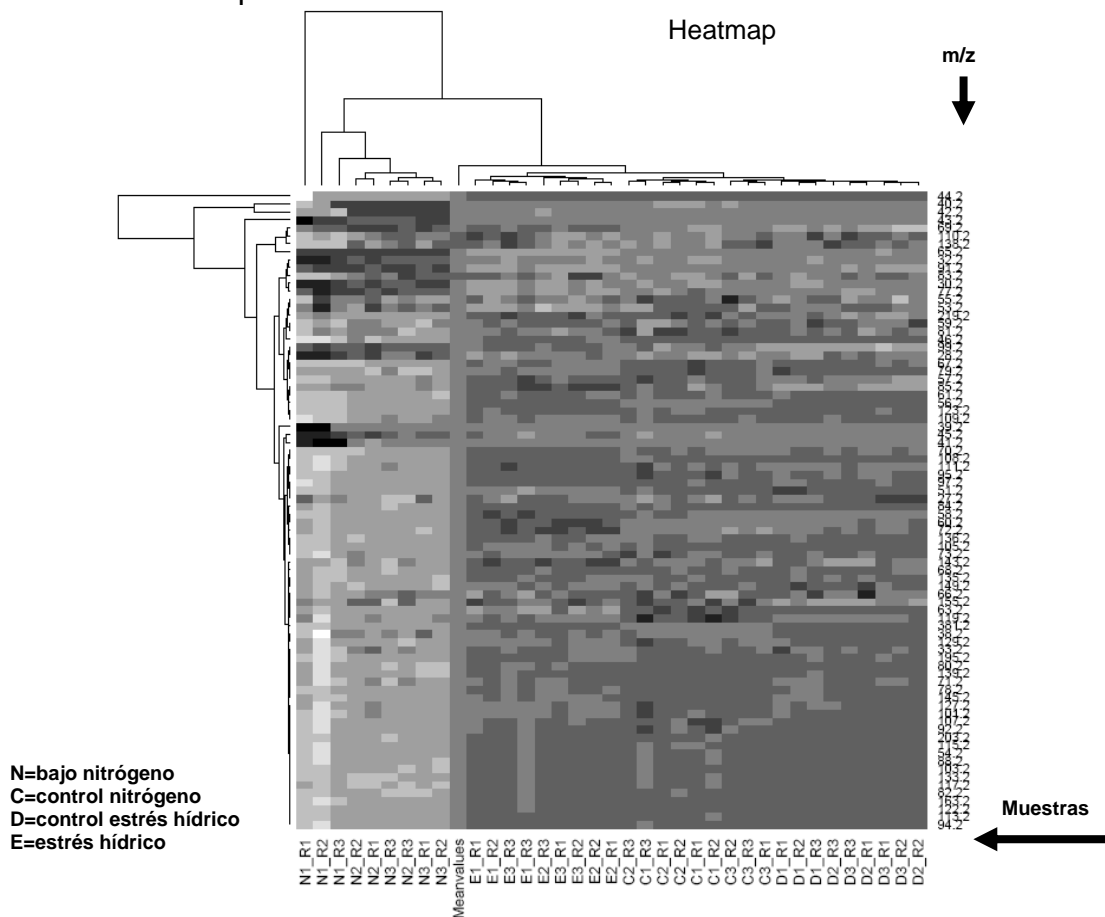


Figura 6.1 Mapa de calor DIESI-MS de un experimento con jugos de tallo de plantas expuestas a diferentes tratamientos de estrés fisiológico. Se muestran solo los 60 iones más intensos del experimento. El dendograma de arriba clasifica el tratamiento de estrés bajo nitrógeno (N) por separado de los tratamientos control (C), observándose muchos iones con intensidad diferencial ($n = 144$).

6.1.1 Linealidad

Con el jugo de tallo se realizó un experimento de linealidad combinando los factores de dilución y de absorbentes (Figura 6.2). Mediante este experimento se comprobó la capacidad analítica del cuadrupolo sencillo (SQD) para establecer ecuaciones de regresión lineal que permitieran cuantificar la concentración de compuestos en función de su intensidad y su masa / carga (m/z). Considerando que los experimentos *fingerprinting* no son dirigidos a un solo metabolito determinado, la linealidad es aplicable solamente a ciertos iones específicos como se muestra en las Figuras 6.3 y 6.4. Muchas veces la linealidad solo se presenta en ciertos rangos de diluciones de la muestra. Afortunadamente en nuestro proyecto usamos solo detectores tipo cuadrupolo que a diferencia de las trampas iónicas tienen un rango mucho más amplio de linealidad cuantitativa. En estas figuras se incluyen iones representativos que tienen mayor intensidad a mayor concentración de la muestra (Figuras 6.2-6.4).

Cabe mencionar que además de encontrar iones que disminuyen su intensidad a mayor dilución como es esperado, también encontramos iones que se comportan de manera opuesta y no lineal. Esto se debe a la supresión de iones que es más fuerte a mayor concentración de la muestra. Es por ello que para las aplicaciones DIESI-MS se usan diluciones muy altas para disminuir la interacción entre moléculas de la matriz. A diferencia del estado gaseoso en donde hay nula interacción entre moléculas, el estado líquido se caracteriza por interacciones fuertes y estables entre moléculas (puentes de hidrógeno y fuerzas de Van der Wals). Parece ser que las moléculas no viajan dentro del espectrómetro de masas (alto vacío) como iones desnudos sino que se asocian preferencialmente unas con otras dando como resultado aglomerados de iones y esferas solvatadas que tienen diferentes masa / carga (m/z). De ahí que el espectro de masas es diferente si se usa agua, etanol, metanol o acetonitrilo como solvente para diluir la muestra. Los agregados moleculares y las esferas solvatadas dependen mucho del solvente que se utiliza. Es por ello que es imposible con los datos DIESI-MS de un cuadrupolo simple hacer una inferencia de la identidad del metabolito en cuestión, y más que bautizar a los iones con nombres químicos, solo nos podemos limitar a dar el valor numérico del centroide de su m/z .

En las Figuras 6.5 y 6.6 se presentan heatmaps, clasificando por grado de similitud, y comparando los iones y su intensidad de acuerdo a los niveles de dilución y absorbentes ($n = 456$).

Para el jugo de tallo de maíz, la dilución óptima de muestra con la que se obtuvo poca supresión cruzada de iones, pero a la vez una intensidad aceptable, fue de 1:100 (Figura 6.2). Los índices de los parámetros analíticos que se obtuvieron en estas pruebas iniciales fueron los siguientes: rango de sensibilidad 30-2000 m/z , resolución alta, coeficiente de variación 95%, exactitud 99%, reproducibilidad 95%, ($n = 72$).

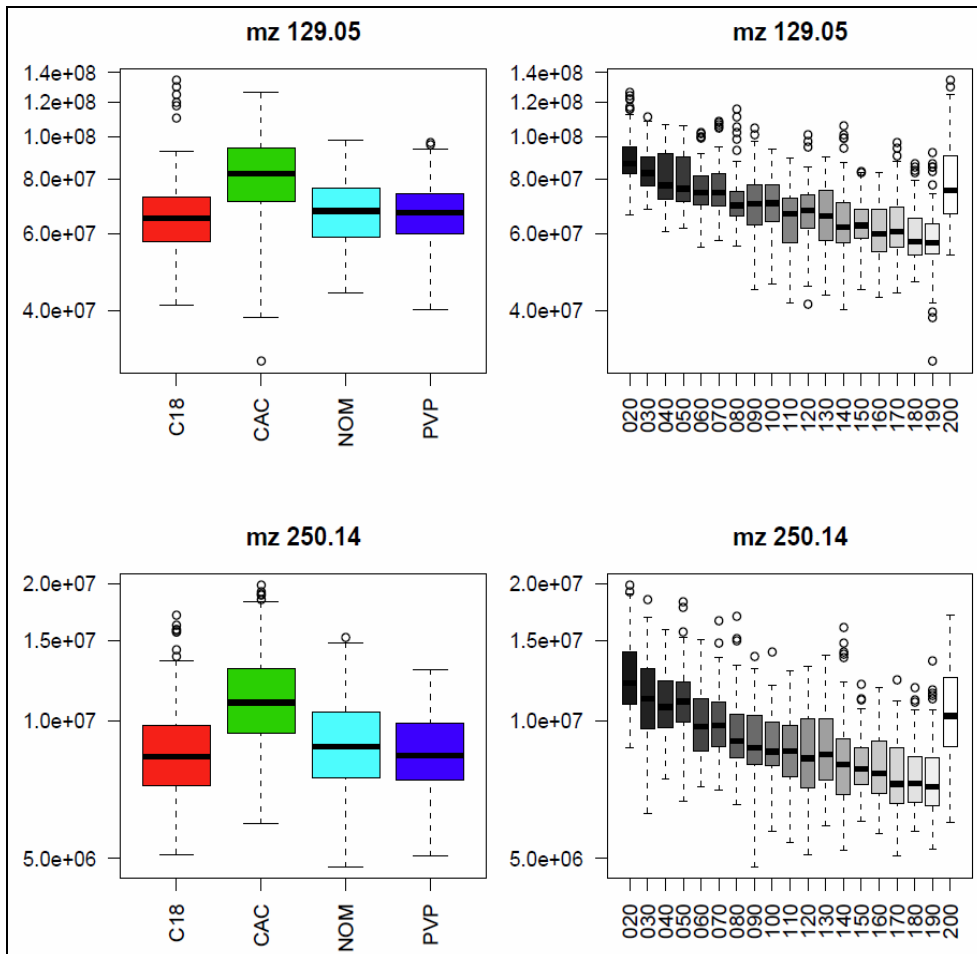


Figura 6.2 En las comparaciones con gráficas tipo caja de cada tratamiento o condición, se obtuvo un patrón similar en la respuesta en intensidad para los iones 129.05 y 250.14 m/z . En las gráficas del lado izquierdo, se observa un nivel más elevado en intensidad para la condición carbón activado (CAC) en comparación con resina R18 (C18), control (NOM) y polivinilpolipirrolidona (PVP). En los gráficos del lado derecho, se observó una linealidad en función de la dilución de las muestras (1:20 a 1:200). Se muestran los resultados de un experimento CT ($n = 456$). Se presentan estos iones porque son afectados en menor medida por el fenómeno de supresión de iones.

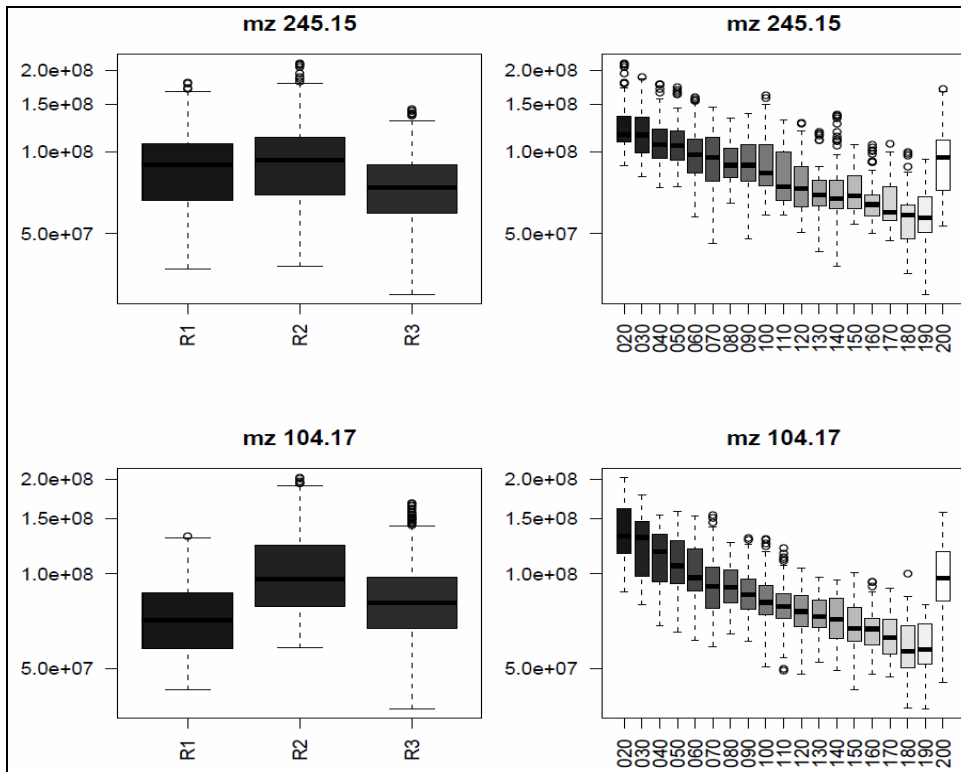


Figura 6.3 En la comparación de gráficas tipo caja para cada repetición biológica, se tuvo un patrón parecido en la respuesta en intensidad de los iones 245.15 y 104.17 m/z . En las gráficas del lado izquierdo, se observan intensidades similares para la repeticiones R1, R2 y R3. En los gráficos del lado derecho, se observó una linealidad en función de la dilución de las muestras (1:20 a 1:200). Experimento CT ($n = 456$). Se presentan estos iones porque son afectados en menor medida por el fenómeno de supresión de iones.

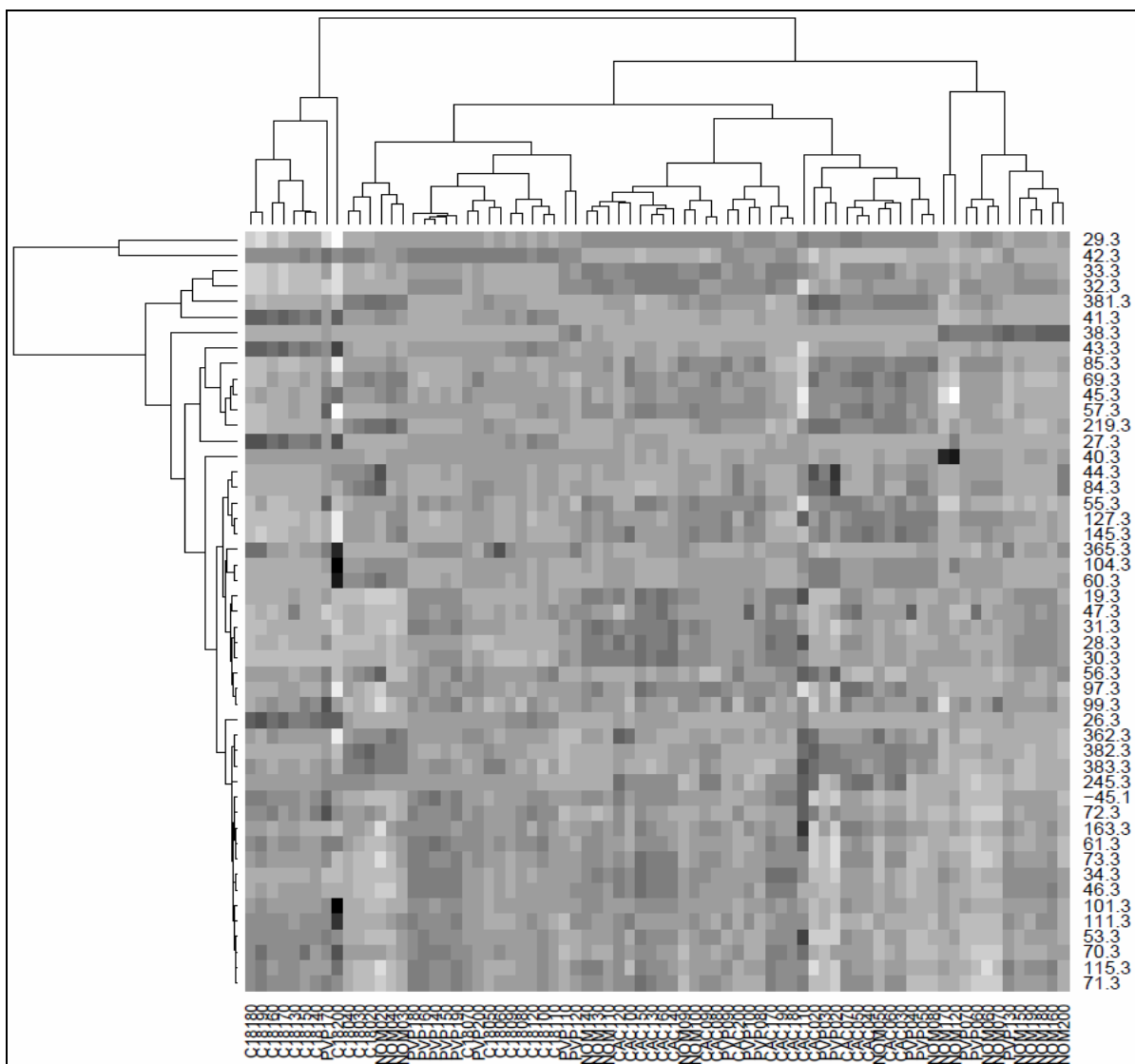


Figura 6.4 Mapa de intensidades m/z con agrupación de muestras y de los 50 iones más intensos del experimento de linealidad (CT) incluyendo iones positivos y negativos. La agrupación (dendogramas) se hicieron con la función por default de heatmap en R que se basa en distancias euclidianas. Los datos no fueron normalizados sino presentados crudos con sus valores numéricos. El dendograma a la izquierda agrupa los iones mas intensos (29.3 y 42.3) hasta arriba separados de los siguientes iones intensos (32.3 y 33.3). El único ion negativo que tuvo una intensidad similar a la de los iones positivos fue -45.1. Esta figura incluye más iones positivos que negativos, y esto demuestra que la intensidad de los iones positivos es más grande que la de los iones negativos. En el dendograma de arriba se observó que hay una agrupación parcial de algunas muestras del experimento, condición o tratamiento (absorbente) y nivel de dilución. Sin embargo, el perfil ionomico cambia en base a la dilucion como en base al adsorbente.

6.12 Experimentos principales con DIESI-MS

Con la aplicación de los scripts en R proporcionados inicialmente por el Dr. Winkler, y después optimizados por mi asesor, el Dr. Tiessen, logamos publicar 2 artículos (García-Flores et al 2012; García-Flores et al 2014). Considerando que los experimentos son no dirigidos a metabolitos determinados a un grado de identificación, de los archivos de las lecturas se generaron gráficos de algunos iones ejemplo que pudieran servir de biomarcadores. Las gráficas comparativas

(heatmaps, boxplots) que se obtuvieron variaron de acuerdo a la definición de factores en el script: genotipo (entrada), tratamiento (condición), repetición biológica, el factor tiempo y su combinación. Para las figuras se formaron grupos de iones con las siguientes etiquetas: todos los iones y los 100 más significativos según factores, los 400 iones más intensos aplicando la comparación de medias de Tukey, los 50 iones negativos más intensos. Se obtuvieron las gráficas comparativas de todos iones ordenados por la m/z .

6.1.1 Estrés hídrico en CIMMYT Tlaltizapan

En la Figura 6.5

Figura 6.5 se graficó, de forma ejemplar y porque presentó lecturas completas para cada tratamiento, el ión 93.3 m/z para cada genotipo y tratamiento siendo el tratamiento bajo estrés hídrico el que presentó una mayor intensidad en comparación con el control indicando un efecto del ambiente. En las siguientes Figuras se grafica uno de los iones más significativos; comparado con otros tres iones que no se grafican 314.22, 453.18 y 272.14 m/z , el 120.07 m/z presenta un patrón diferente ya que la intensidad en el tratamiento control es mayor, indicando un mayor efecto del genotipo. En la Figura 6.7 se presenta una aproximación de los iones a su posible identificación en el grupo de los aminoácidos. Los heatmaps de las Figura 6.8 (grupo de 50 iones) y Figura 6.5 (grupo de 80 iones) se clasificaron a los iones de acuerdo a su intensidad, destacando la agrupación por tratamiento en la primera figura y por genotipo en la segunda. Ambos heatmaps nos facilitaron el estudio de marcadores diferenciales entre genotipos y tratamientos. En el dendograma de la Figura 6. se aprecia la agrupación de los genotipos según su tratamiento y el nivel de intensidad grupal con una excepción: el genotipo T231WS, que se asocia con el genotipo T282CT.

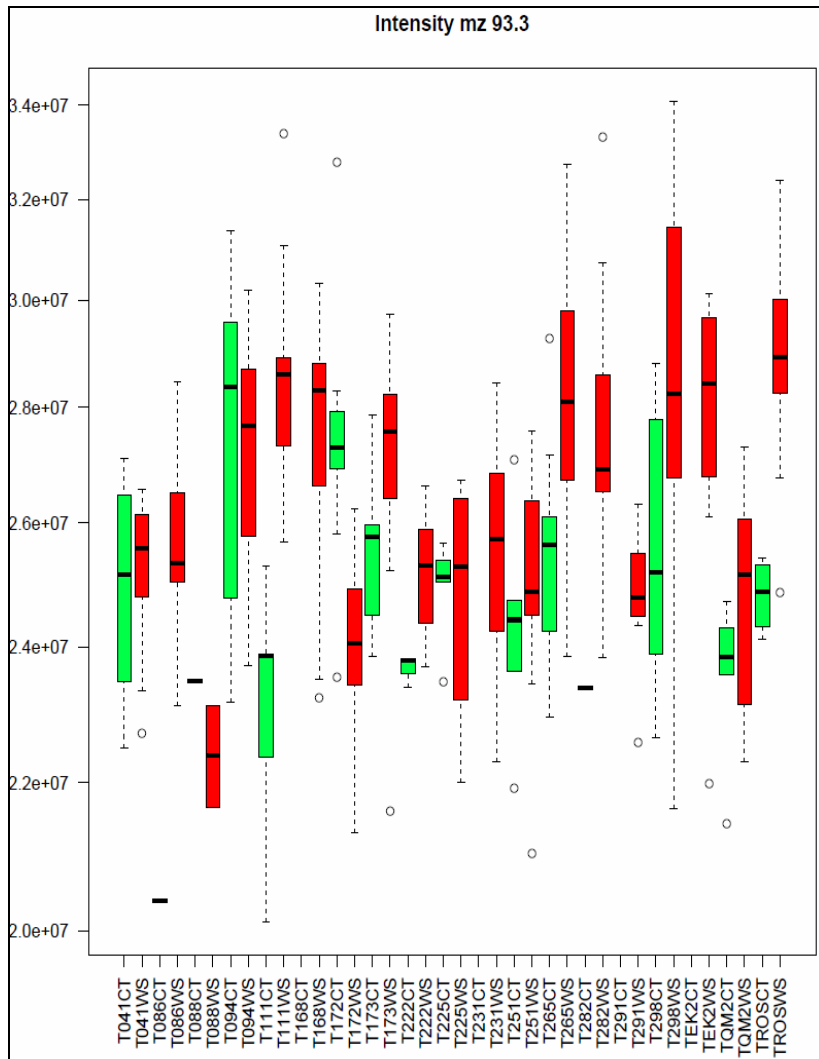


Figura 6.5 Gráfica comparativa de intensidades tipo caja del ion 93.3 m/z entre los tratamientos WS y CT del experimento de estrés hídrico de Tlaltizapán. La intensidad del ión de cada genotipo fue mayor para el tratamiento de sequía WS ($n = 456$). Color verde: control; color rojo: estrés hídrico.

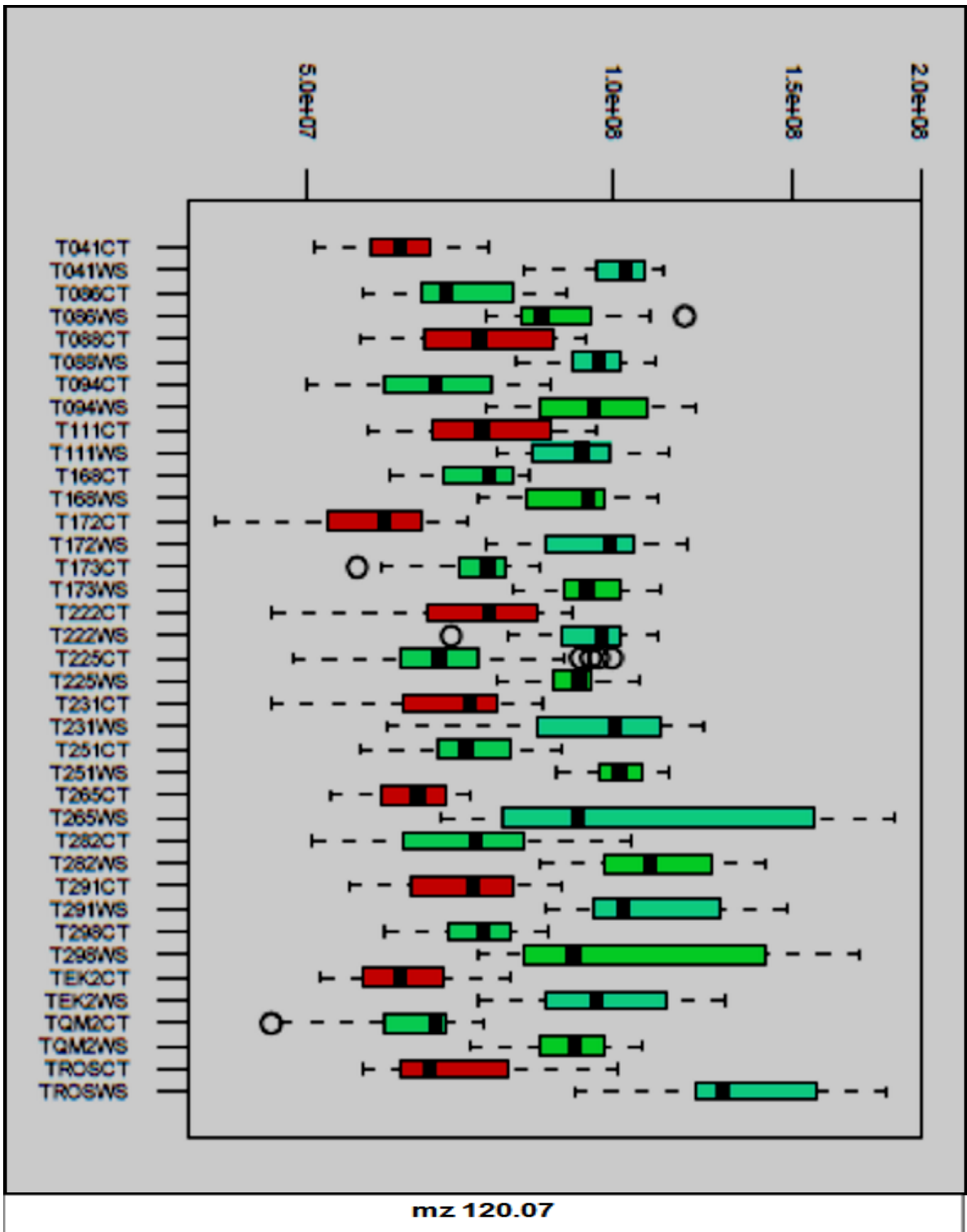


Figura 6.6 Gráfica comparativa de intensidades tipo caja del ion 120.07 *m/z*, seleccionado de un grupo como los más significativos, entre los genotipos por cada tratamiento del experimento de estrés hídrico de Tlaltizapán, CT vs WS ($n = 456$). “*n*” significa: número de muestras.

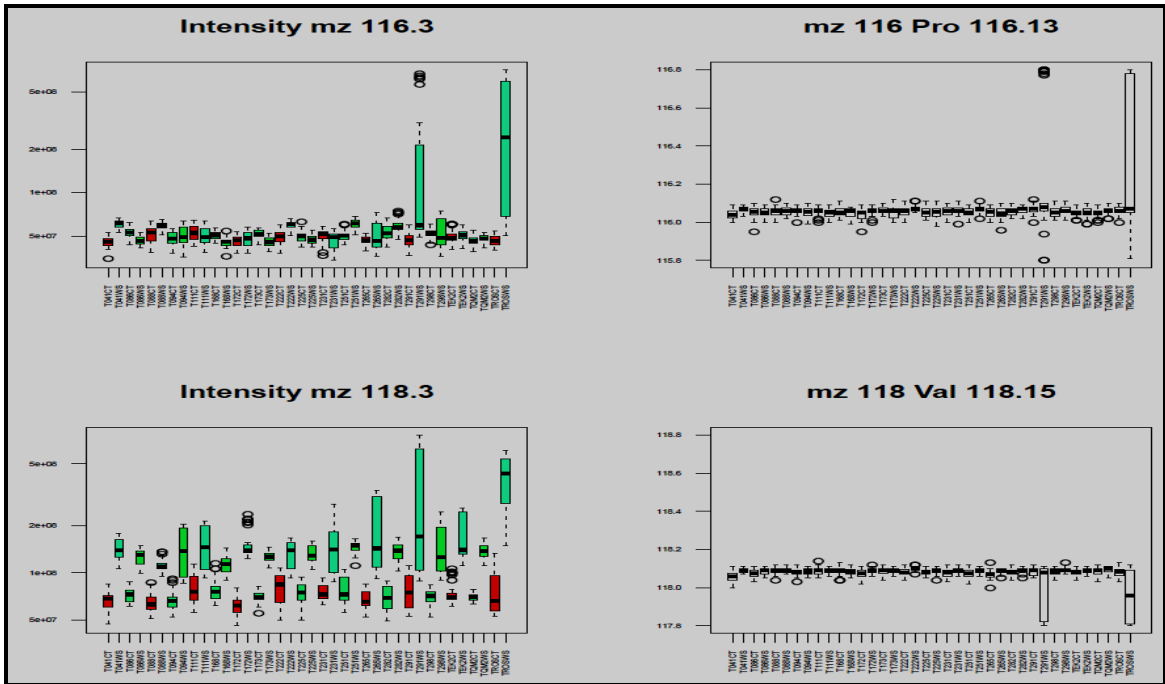
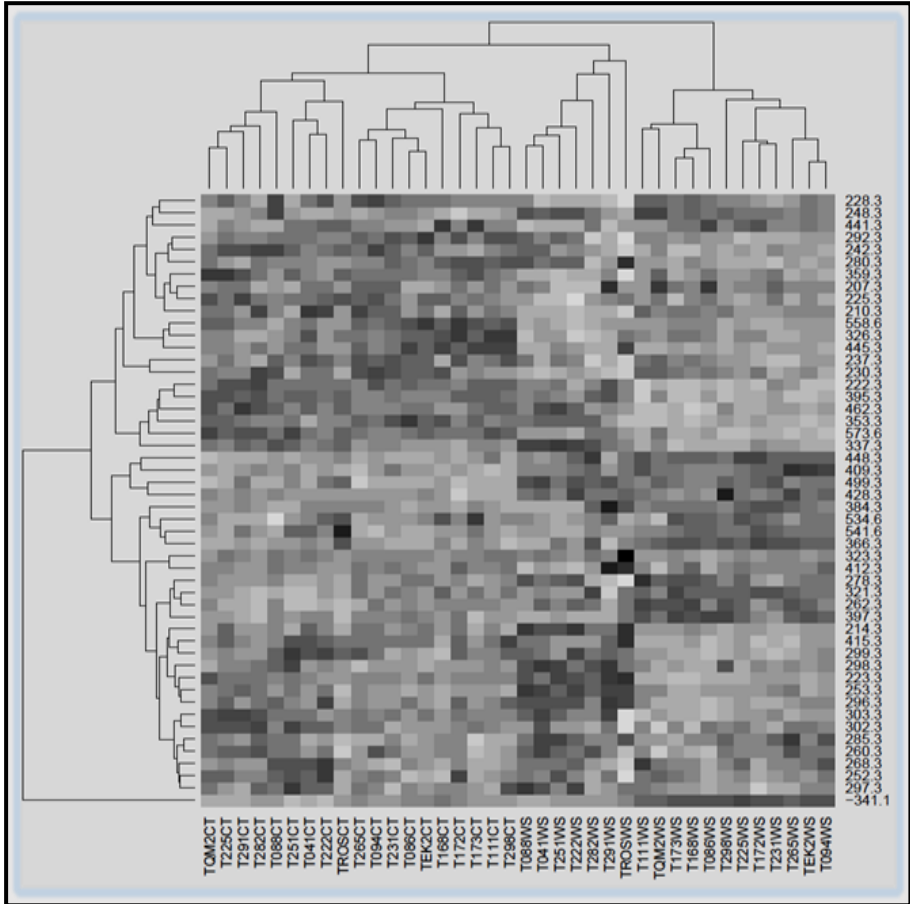


Figura 6.7 Gráficas comparativas de los iones aproximados a la variable m/z de los aminoácidos prolina (116.13 m/z) y valina (118.15 m/z). En las gráficas de la izquierda se muestran los valores de intensidad de los iones 116.3 y 118.3 m/z por tratamiento, CT vs WS, y genotipo. En la derecha se representaron los gráficos de los valores m/z estimados del aminoácido comparado con el valor de la señal m/z detectado en cada tratamiento CT vs WS, y genotipo. El orden de los genotipos en el eje horizontal corresponde al de la figura 6.2, del T041 CT y WS hasta el TROSS CT y WS respectivamente ($n = 456$).

Figura 6.8 Mapa de intensidades- m/z agrupando los 50 iones seleccionados con mayor intensidad en el experimento de estrés hídrico en Tlaltizapán incluyendo los iones positivos y negativos. Se puede observar una agrupación uniforme de los iones en función del tratamiento y genotipo porque se usó el método *scale* como opción de la función heatmap en R, en la elaboración de la gráfica ($n = 456$).



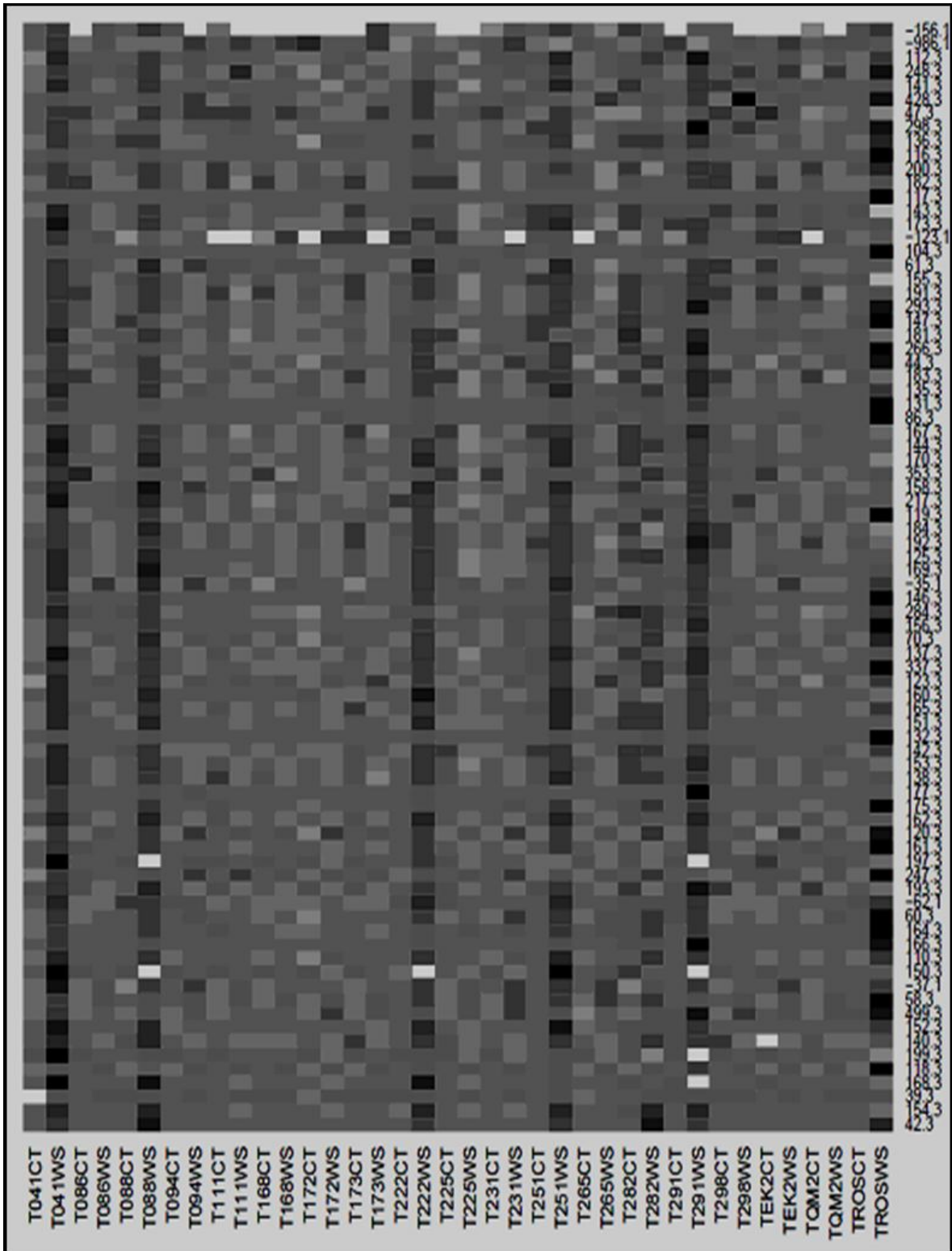
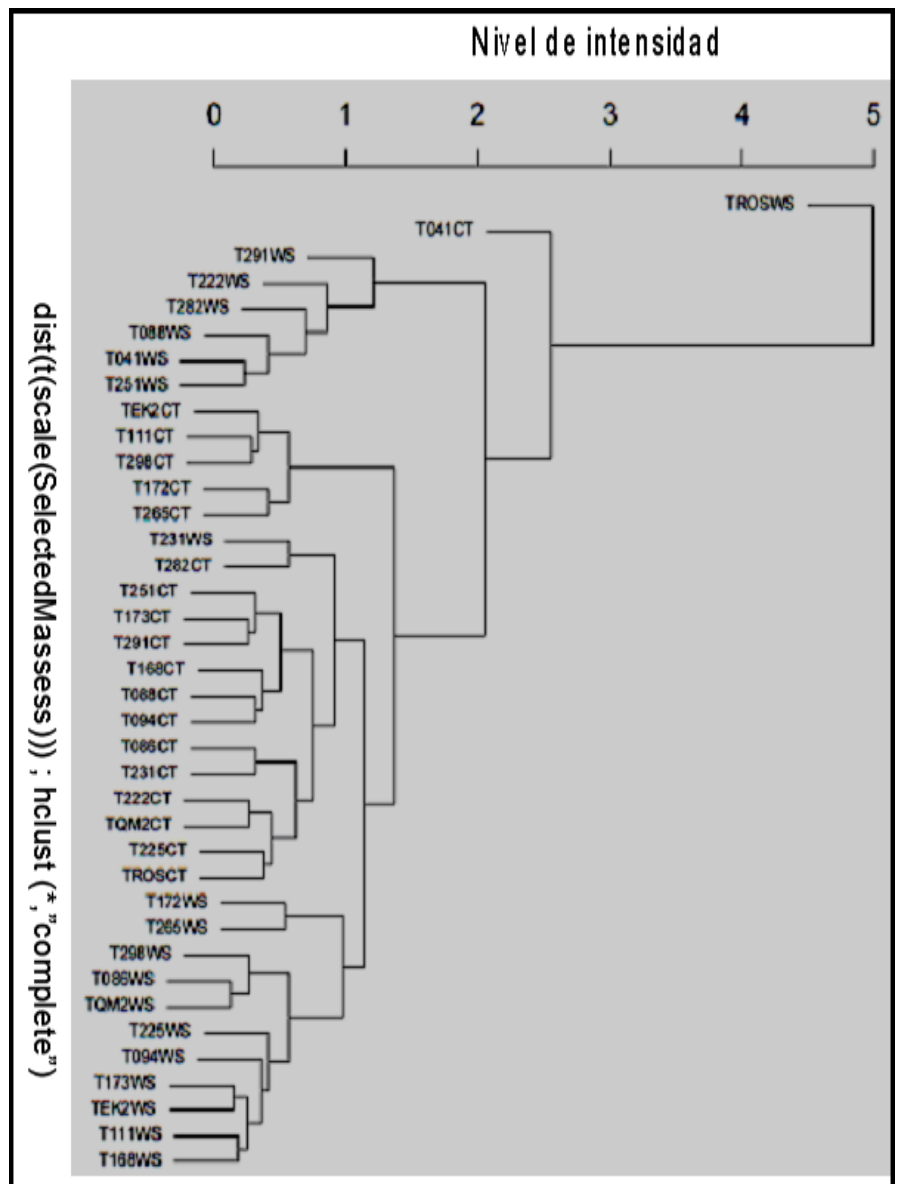


Figura 6.9 Mapa de intensidades- m/z con los 80 iones más intensos del experimento de estrés hídrico en Tlaltizapán incluyendo iones positivos y negativos. Se observa una representación uniforme y poco informativa del mapa debido a que no se usó método alguno de clasificación. Esta figura demuestra que el agrupamiento automático generado por los dendogramas de R si permite ver con más claridad los resultados DIESI-MS. ($n = 456$).

Figura 6.10 Clasificación de los genotipos y su tratamiento respectivo del experimento de estrés hídrico de Tlaltizapán mediante una representación tipo dendograma, agrupados por su proximidad en el nivel de intensidad. Se puede observar una agrupación uniforme de los genotipos en función del tratamiento porque se usó el método $\text{dist}(t(\text{scale}))$ que es una opción para calcular la proximidad en valor de las lecturas con la función $\text{hclust}(\text{dendograma con agrupamiento})$ en R, en la elaboración de la gráfica. Los tratamientos T041CT, muestra control y genotipo 041 y TROSWS, muestra de estrés hídrico y genotipo ROS son los que mayor nivel de intensidad presentaron ($n = 456$).



Los resultados de varianza y de correlación entre algunas variables agronómicas y metabolitos específicos, se analizaron a detalle en Juárez, (2012). Los metabolitos analizados en dicho estudio metabólico dirigido fueron: glucosa, fructosa, sacarosa, almidón y aminoácidos totales.

6.1.2 Bajo contenido de nitrógeno en Cimmyt Batán

Los resultados siguientes se conceptualizan de manera similar como se hizo en el experimento de estrés hídrico. En la Figura 6.11 se compararon las intensidades del ion 175.3 m/z de acuerdo a la variedad de maíces blancos y amarillos, resultando ser variable la intensidad de acuerdo al genotipo y al tratamiento. Resultados similares se muestran en las Figura 6.12 y 6.9 para grupos de iones negativos. En la Figura 6.10 se presenta iones que pudieran pertenecer al

grupo de los aminoácidos. También es posible determinar la exactitud de la lectura respecto a la fórmula molecular teórica del aminoácido ionizado. Los heatmaps de las Figura 6.8 (grupo de 50 iones) y 6.9 (grupo de 80 iones) clasifican a los mismos de acuerdo a su intensidad, destacando la agrupación por variedad de maíz blanco o amarillo a excepción del genotipo AE14CT que se agrupó con los maíces blancos en la primera Figura y por genotipo en la segunda; ambos heatmaps nos facilitan el estudio de marcadores diferenciales entre genotipos y tratamientos. En el dendograma de la Figura 6.12 se aprecia la agrupación de los genotipos según su variedad y tratamiento con su nivel de intensidad grupal con una excepción: la variedad AE12CT se asoció con la variedad BE10CT indicando similitud entre ellas. Las

Figura 6.133 y 6.14 agrupan las variedades maíces blancos y amarillos por tratamiento, exceptuando a la variedad AE05CT que se asoció al tratamiento LN y la variedad BE10CT que por su nivel de intensidad se asoció a ambos tratamientos.

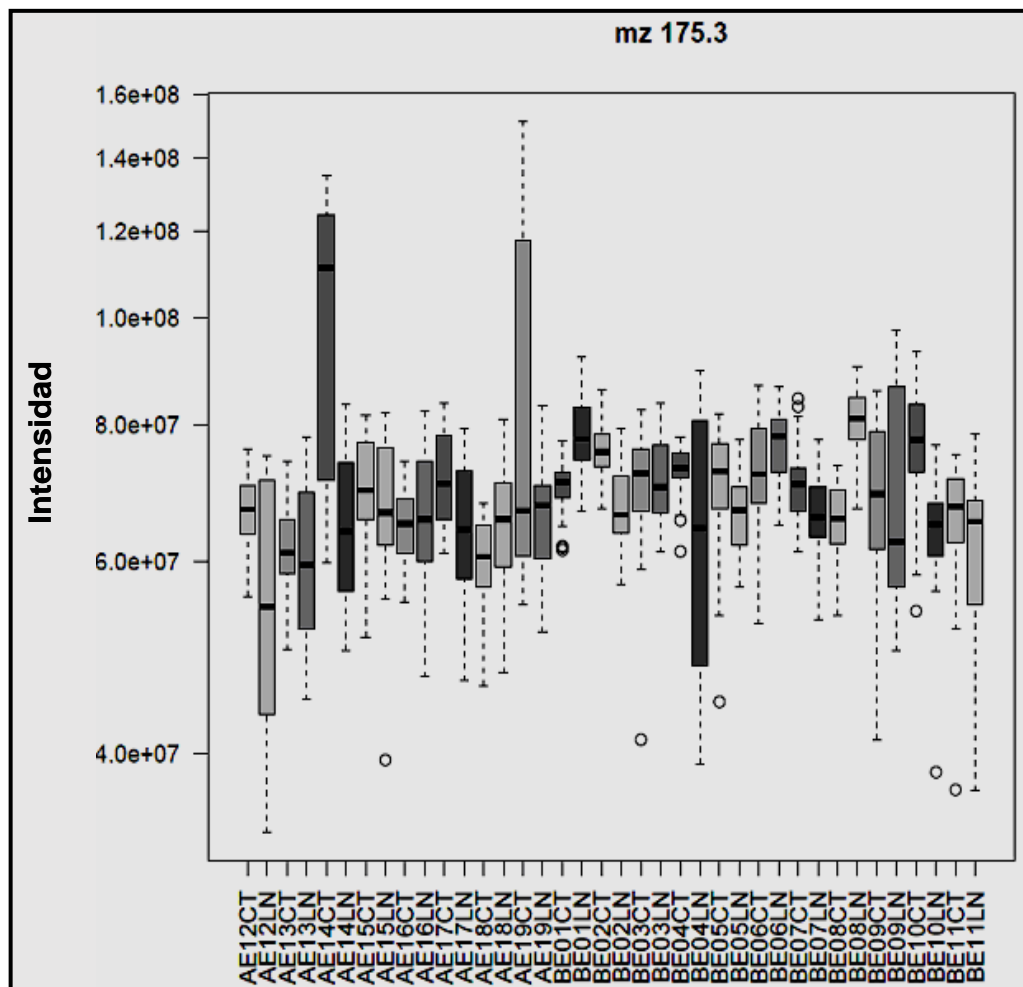


Figura 6.11 Gráfica comparativa de intensidades tipo caja del ion 175.3 m/z entre los genotipos, maíces blancos y amarillos, y tratamientos LN y CT del experimento de bacontenido de nitrógeno realizado en el Batán ($n = 456$). A: amarillo; B: blanco. Se constata que no existen diferencias significativas para los tratamientos ni para las variedades de maíz. Además es posible observar la presencia de valores que se encuentran fuera de los límites de control.

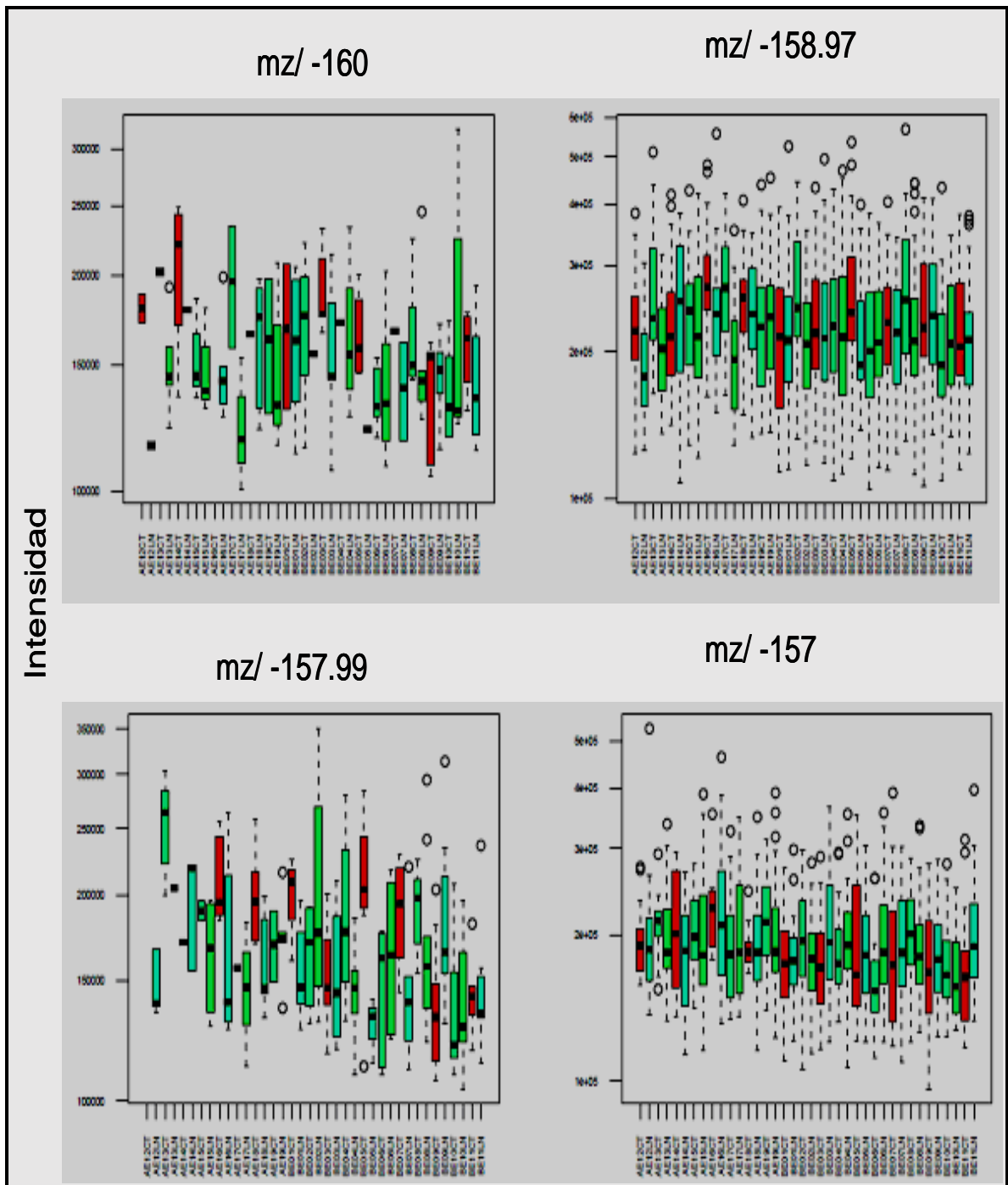


Figura 6.12 Gráficas comparativas de intensidades tipo caja de los iones -160, -158.97, -157.99 y -157 m/z , seleccionados del grupo completo ordenado de iones positivos y negativos, entre los genotipos por cada tratamiento del experimento de bajo contenido de nitrógeno, realizado en el Batán. El orden de los genotipos en el eje horizontal corresponde al de la figura 6.7, del AE12 CT y LN hasta el BE11 CT y LN respectivamente (n = 456).

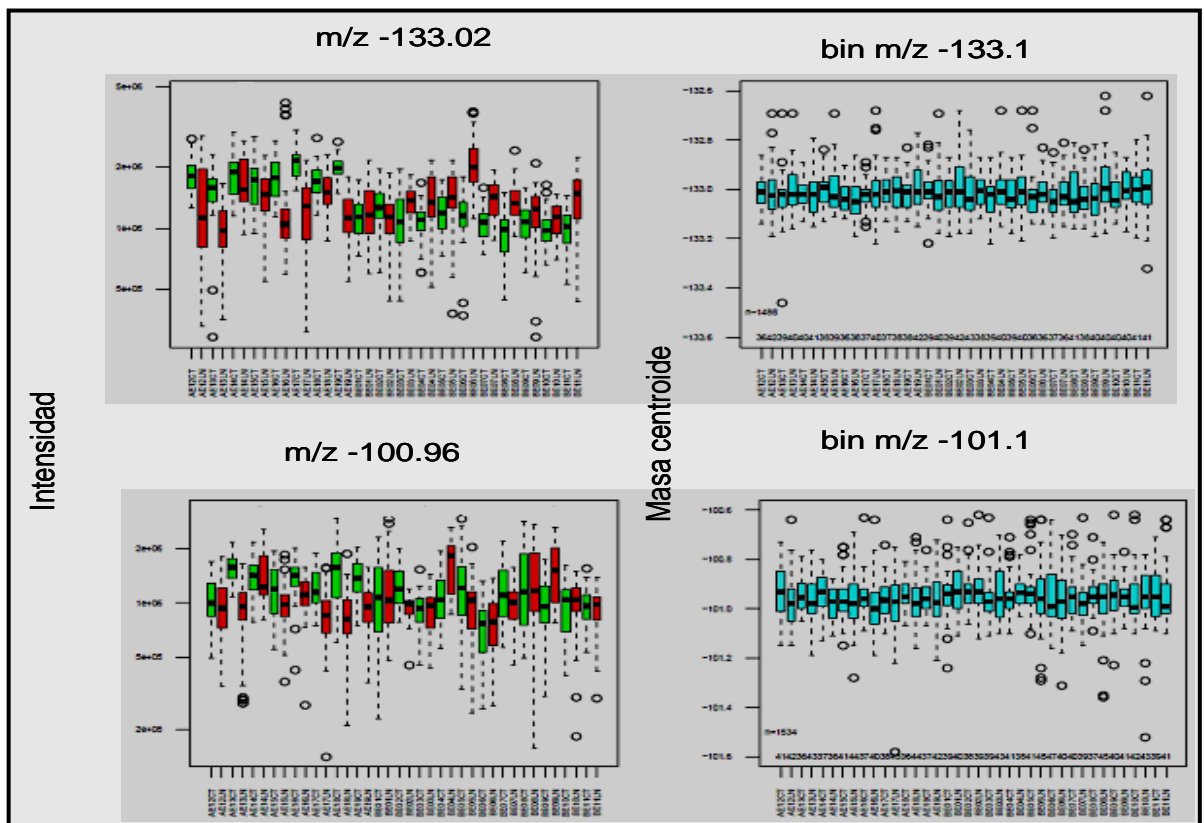


Figura 6.13 Gráficas comparativas de intensidades tipo caja de los iones -133.02 y -100.96 m/z , seleccionados del grupo de los 50 iones negativos más significativos, entre los genotipos por cada tratamiento del experimento de bajo contenido de nitrógeno, realizado en el Batán. El orden de los genotipos en el eje horizontal corresponde al de la figura 6.7, del AE12 CT y LN hasta el BE11 CT y LN respectivamente ($n = 456$).

Figura 6.14 Gráficas comparativas de los iones aproximados a la variable m/z correspondiente a los aminoácidos valina (118.15 m/z) y treonina (120.12 m/z). En las gráficas de la izquierda se muestran los valores de intensidad de los iones 118.3 y 120.3 m/z por tratamiento y genotipo. En la derecha se representaron los gráficos de los valores m/z estimados del aminoácido comparado con el valor de la señal m/z detectado en cada tratamiento y genotipo. El orden de los genotipos en el eje horizontal corresponde al de la figura 6.7, del AE12 CT y LN hasta el BE11 CT y LN respectivamente ($n = 456$).

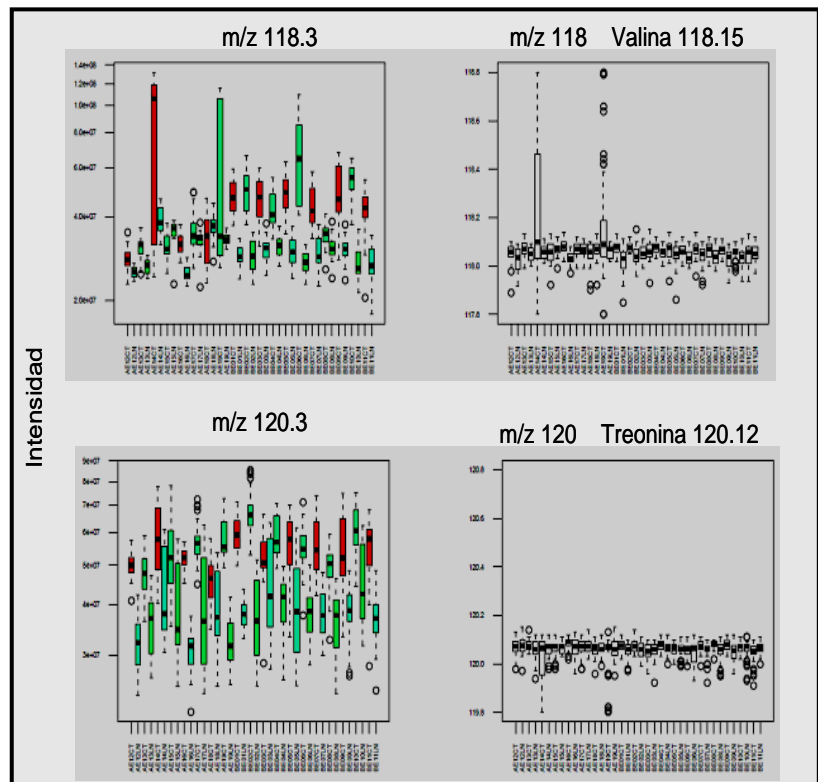


Figura 6.15 Mapa de intensidades- m/z que agrupó a los 50 iones seleccionados con mayor intensidad en el experimento de bajo contenido de nitrógeno realizado en Batán incluyendo los iones positivos y negativos. Se puede observar una agrupación uniforme de los iones en función del tratamiento y genotipo, maíces blancos y amarillos, porque se usó el método *scale* como opción de la función heatmap en R, en la elaboración de la gráfica (n = 456).

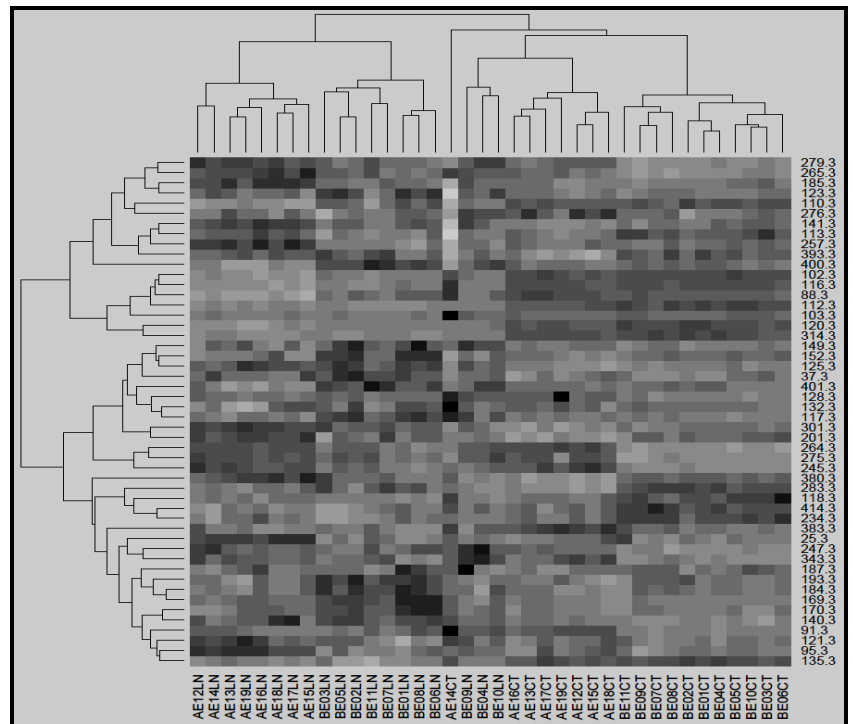
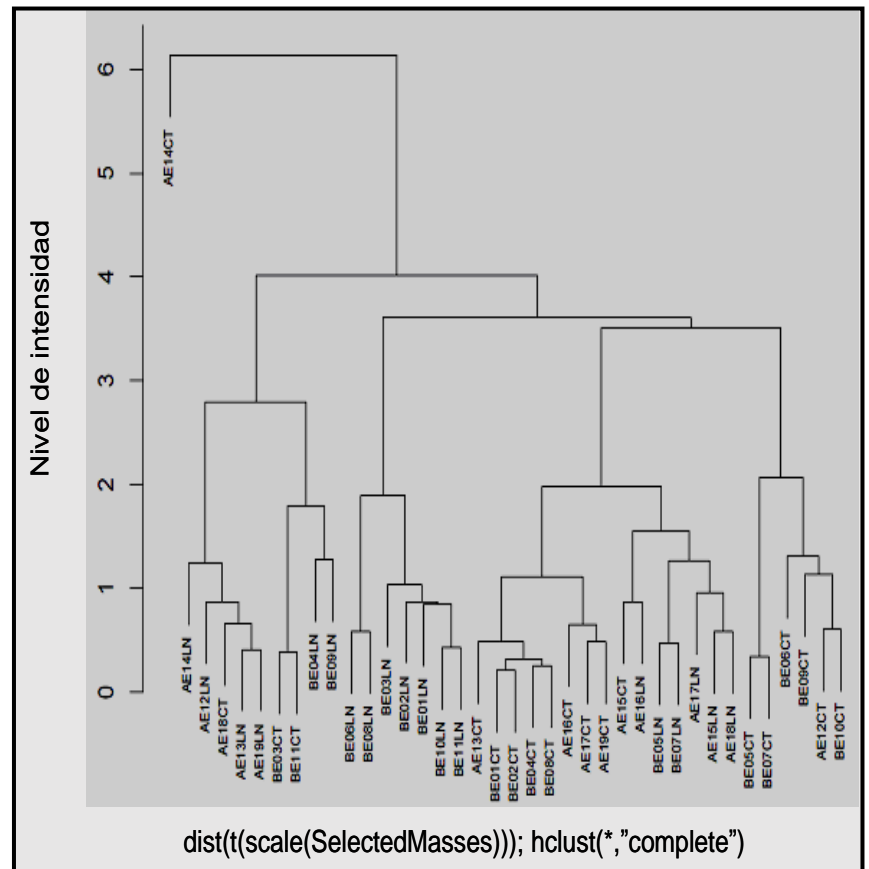


Figura 6.12 Clasificación de los genotipos y su tratamiento respectivo del experimento de bajo contenido de nitrógeno realizado en El Batán mediante un dendrograma, agrupados por su proximidad en el nivel de intensidad. Se observó una agrupación uniforme de los genotipos en función del tratamiento porque se usó el método $\text{dist}(t(\text{scale}))$ que es una opción de proximidad en valor de la función hclust (dendrograma por agrupamiento) en R. Se observa (en el eje vertical del dendrograma) que el tratamiento AE14CT es el que mayor nivel de intensidad presentó (n = 456).



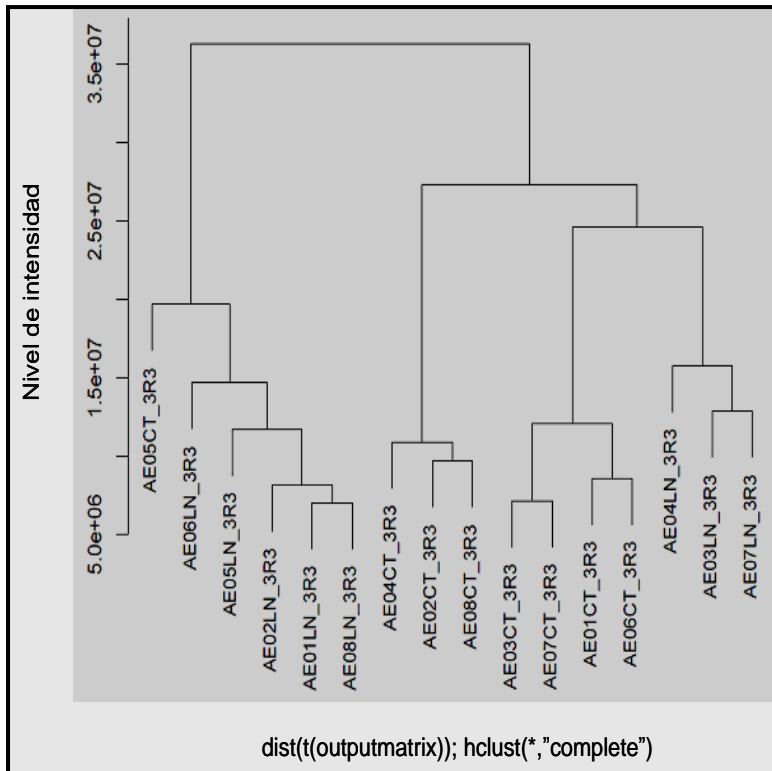
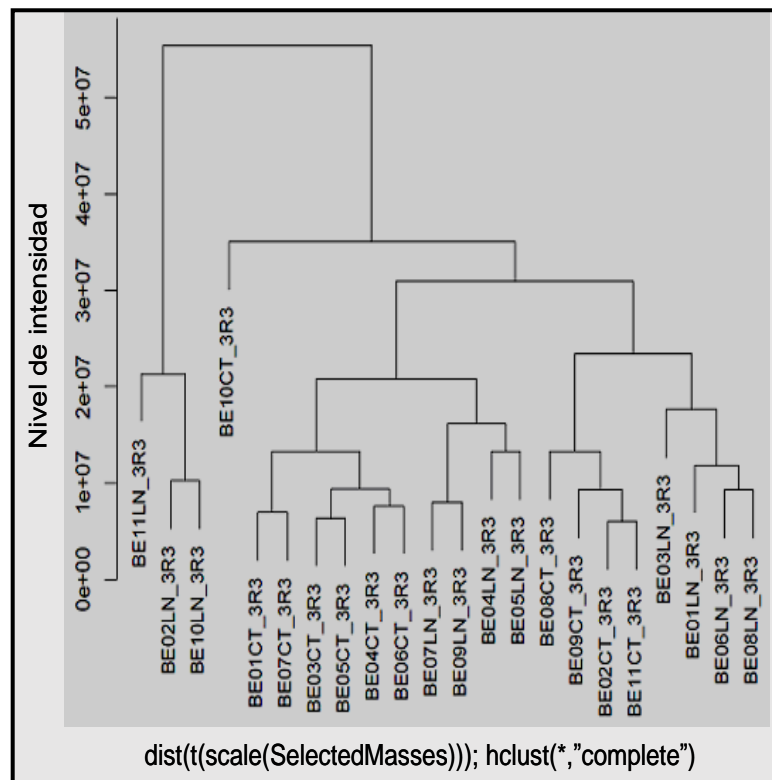


Figura 6.13 Clasificación de los genotipos (maíces amarillos) y su tratamiento respectivo del experimento de bajo contenido de nitrógeno realizado en El Batán mediante una representación tipo dendograma, donde éstos son agrupados por su proximidad en el nivel de intensidad. Se puede observar una agrupación uniforme de los genotipos en función del tratamiento porque se usó el método $\text{dist}(t(\text{scale}))$ que es una opción de proximidad en valor de la función hclust (dendograma por agrupamiento) en R, en la elaboración de la gráfica. El tratamiento AE05CT fue el que mayor nivel de intensidad presentó ($n = 16$).

Figura 6.14 Clasificación de los genotipos (maíces blancos) y su tratamiento respectivo del experimento de bajo contenido de nitrógeno realizado en El Batán mediante una representación tipo dendograma, donde los genotipos se agruparon por su proximidad en el nivel de intensidad. Se puede observar una agrupación uniforme de los genotipos en función del tratamiento porque se usó el método $\text{dist}(t(\text{scale}))$ que es una opción de proximidad hclust (dendograma por agrupamiento) en R, en la elaboración de la gráfica. El tratamiento AE05CT fue el que mayor nivel de intensidad presentó ($n = 22$).



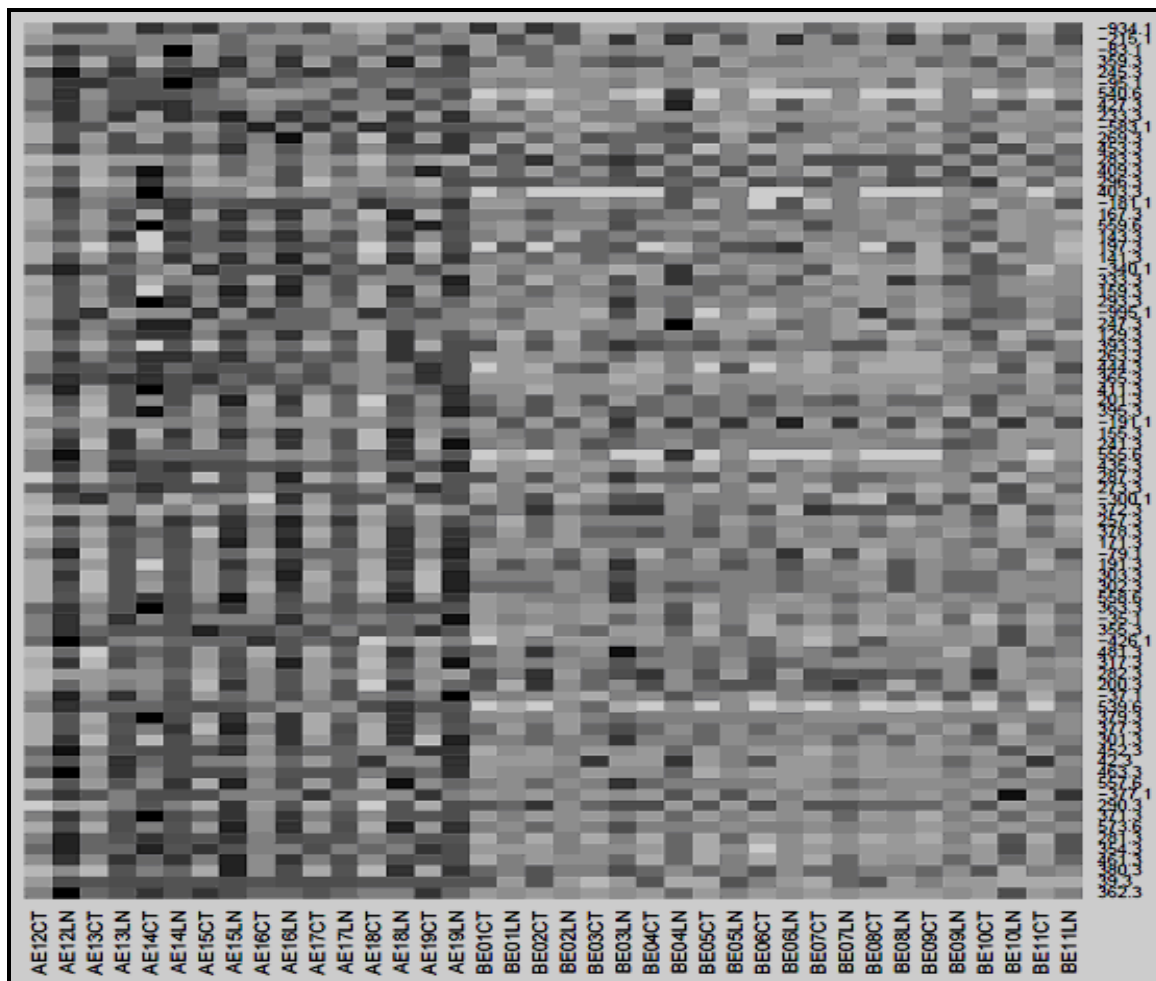


Figura 6.15 Mapa de intensidades- m/z agrupando los 80 iones más intensos del experimento de bajo contenido de nitrógeno realizado en El Batán, incluyendo iones positivos y negativos. Se observa una representación uniforme en función del genotipo (maíces blancos y amarillos) y tratamiento porque no se usó método alguno de clasificación ($n = 456$).

6.1.3 Experimento Fuente-Sumidero (Sink-Source) en Irapuato

Como resultados de éste experimento, las plantas control tienden a incrementar su intensidad con el tiempo, de uno a diez días. Las plantas (D) sin elote (WE), aumentan de intensidad al día dos mientras que las plantas (P) disminuyeron su intensidad en el mismo día. El efecto inverso sucede al tercer día. En el cuarto día, los dos genotipos disminuyen y aumentan nuevamente en el día diez. Las plantas (D) y (P) sin hoja aumentan en el día dos y disminuyen en los días sucesivos. Las plantas sometidas al tratamiento sin elote y sin hoja (WM) presentan una tendencia a aumentar su intensidad (de los iones) con el tiempo, en los dos genotipos. Ya que no se tiene demanda del metabolito en las plantas sin elote, el efecto fuente-sumidero se hace presente incrementando la intensidad del ion. En las plantas sin hojas, el efecto se manifiesta en la disminución de la intensidad del ion debido a la falta de síntesis del metabolito. El incremento de la intensidad del ion en las plantas sin elote y sin hojas, establece la presencia de la fotosíntesis en el tallo y en las hojas del elote; se concluye lo anterior por la falta de demanda de tejidos sumidero y la ausencia de fotosíntesis en las hojas laminares; por lo tanto, el origen del incremento de la intensidad del ion se encuentra en la fotosíntesis de los tejidos con clorofila que, en el caso del maíz, son el tallo y las

hojas del elote (totomoxtle). Un análisis individual de cada metabolito permitió conocer el patrón de variación de un genotipo y tratamiento con respecto al tiempo (Figura 6.16).

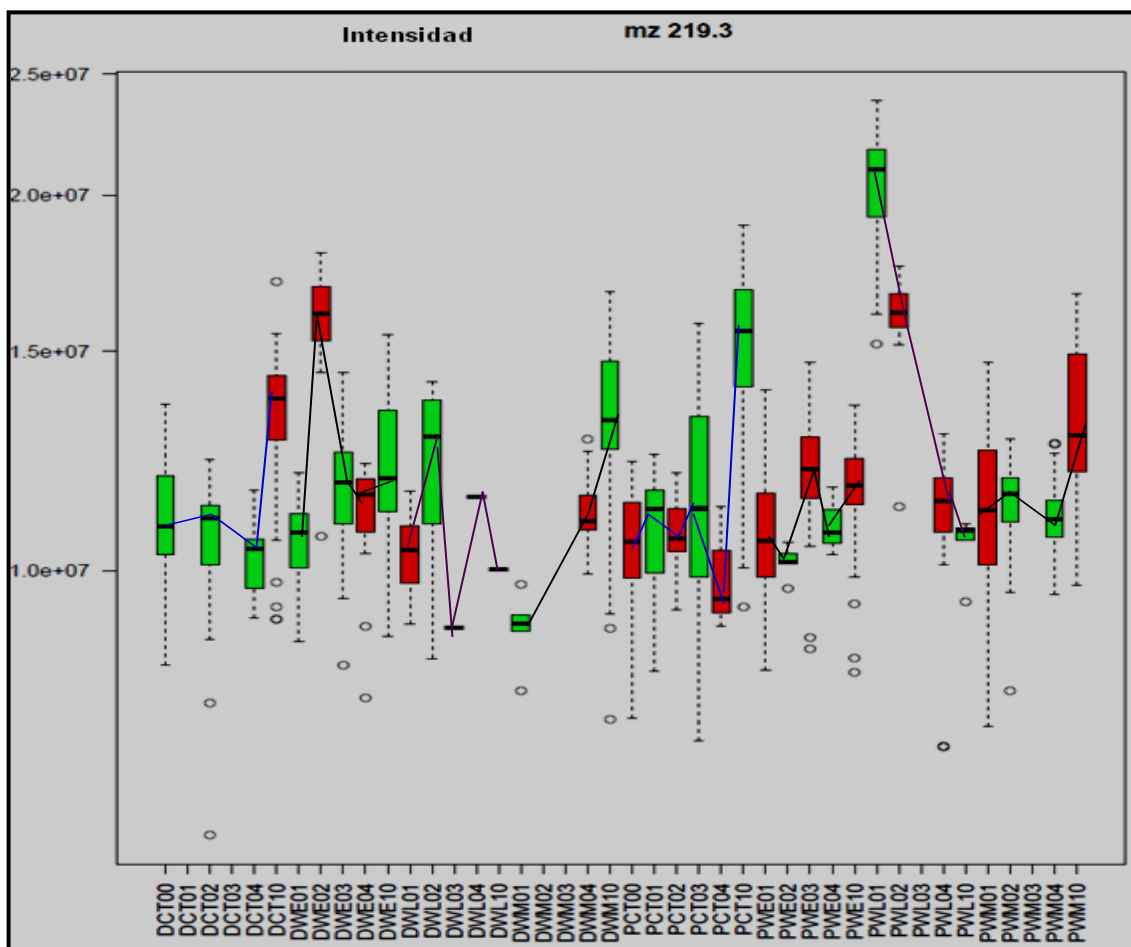


Figura 6.16 Gráfica comparativa de intensidades tipo caja del ion 219.3 m/z entre los genotipos, Dow (D) y Puma (P), y tratamientos: (CT = control, WE = sin elotes, WL = sin hojas, WM = sin elotes y sin hojas) del experimento fuente-sumidero de Irapuato ($n = 504$). Los colores rojo y verde indican cambios entre una muestra con respecto a otra. “n”: representa el número de muestras. Los círculos exteriores son valores extremos que se encuentran fuera de los límites de control.

En la comparación del tratamiento por genotipo se encontró que las plantas control tienden a mantener constante su nivel de intensidad durante el experimento. Las plantas (D) y (P) sin elote (WE), disminuyen su intensidad en relación a las plantas control. Las plantas (D) sin hoja disminuyen su intensidad y las plantas (P) la aumentan, presentándose un efecto contrario en los genotipos. El efecto fuente-sumidero se manifiesta en las variaciones de intensidad del ion 219.3 m/z ya que no se tiene demanda del metabolito en las plantas sin elote. En las plantas sin hojas del genotipo Dow (D), el efecto se manifiesta en la disminución de la intensidad del ion 219.3 m/z debido a la reducción de síntesis del metabolito. Un efecto inverso se presenta en las plantas (P) sin hojas en las que la intensidad se incrementa posiblemente por la mayor tasa de síntesis del metabolito en el genotipo Puma (P) o al número de hojas de las plantas (P). El incremento de la intensidad del ion en las plantas sin elote y sin hojas, establecieron la presencia de

la fotosíntesis en el tallo y en las hojas del elote; concluimos lo anterior por no haber demanda de los tejidos sumidero (elotes) y la faltante fotosíntesis en las hojas laminares; por lo tanto, el origen del incremento de la intensidad del ion 219.3 m/z se encuentra en la fotosíntesis de los tejidos con clorofila que, en el caso del maíz, son el tallo y las hojas del elote (totomoxtle). Un análisis individual de cada metabolito es requerido para conocer el patrón de variación de un genotipo con respecto al tratamiento (Figura 6.177).

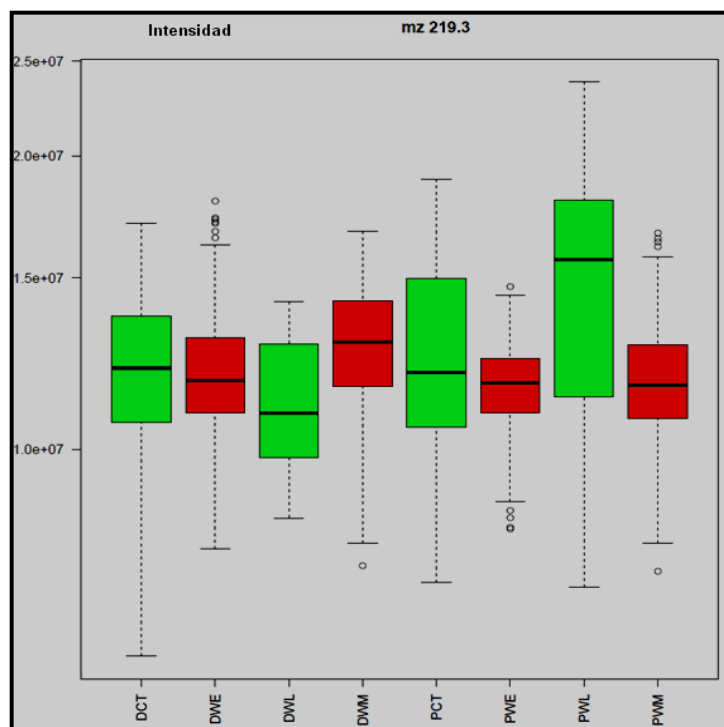


Figura 6.17 Gráfica comparativa de intensidades tipo caja del ion 219.3 m/z entre los genotipos, Dow (D) y Puma (P), y tratamientos: (CT = control, WE = sin elotes, WL= sin hojas, WM = sin elotes y sin hojas) del experimento fuente-sumidero de Irapuato, sin la inclusión del factor tiempo (n = 504).

De la siguiente gráfica individual del ion 73.02 (Figura 6.18 8) se determinó el patrón de variación por genotipo y tratamiento con respecto al tiempo, y se observó variabilidad biológica en los controles del genotipo Dow (D), aumento de intensidad en el tratamiento WE con un declinación al final del tiempo, disminución gradual en el tratamiento WL y aumento al final del tiempo, aumento de intensidad en el tratamiento WM con declinación al final del tiempo. Respecto al genotipo Puma (P) los controles presentaron menor variabilidad biológica y el comportamiento fue similar al genotipo Dow (D) excepto para el tratamiento WL que tuvo un patrón contrario. Para el ion 115.02 m/z también presenta un patrón de variabilidad por genotipo y tratamiento con respecto al tiempo, observándose variabilidad biológica en los controles de ambos genotipos Dow (D) y Puma (P). Aumento de intensidad del ion con respecto al tiempo para los tres tratamientos: WE, WL y WM de los dos genotipos, presentando un incremento gradual de intensidad a un máximo a los 10 días.

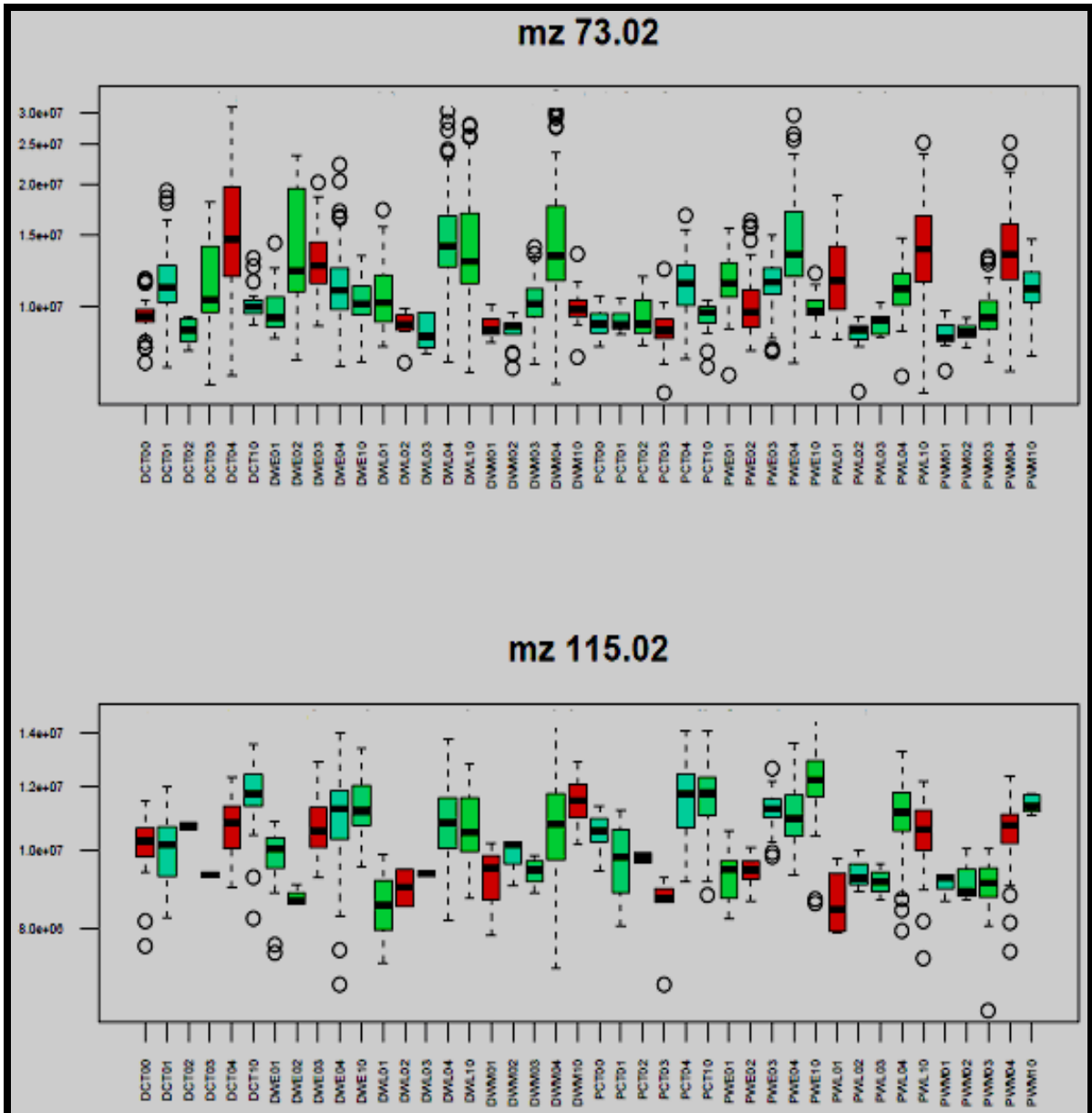


Figura 6.18 Gráfica comparativa de intensidades tipo caja (box plot) de los iones 73.02 y 115.02 m/z entre los genotipos, Dow (D) y Puma (P), tratamientos (CT=control, WE= sin elotes, WL=sin hojas, WM=sin elotes y sin hojas) y tiempo (0, 1, 2, 3, 4 y 10 días) del experimento fuente-sumidero de Irapuato ($n = 504$). "n" representa: el número de muestras. Gráfica complementaria a la figura 6.20.

Figura 6.19 Clasificación de los genotipos Dow (D), Puma (P) y su tratamiento respectivo del experimento de fuente-sumidero de Irapuato mediante una representación tipo dendograma, agrupados por su proximidad en el nivel de intensidad. Se puede observar una agrupación uniforme de los genotipos en función del tratamiento porque se usó el método $\text{dist}(t(\text{scale}))$ (cálculo de valor proximal) como opción de la función `hclust` (dendograma con agrupamiento) en R, en la elaboración de la gráfica. El tratamiento DWM (Dow sin elote y sin hojas) es el que mayor nivel de intensidad presentó ($n = 504$).

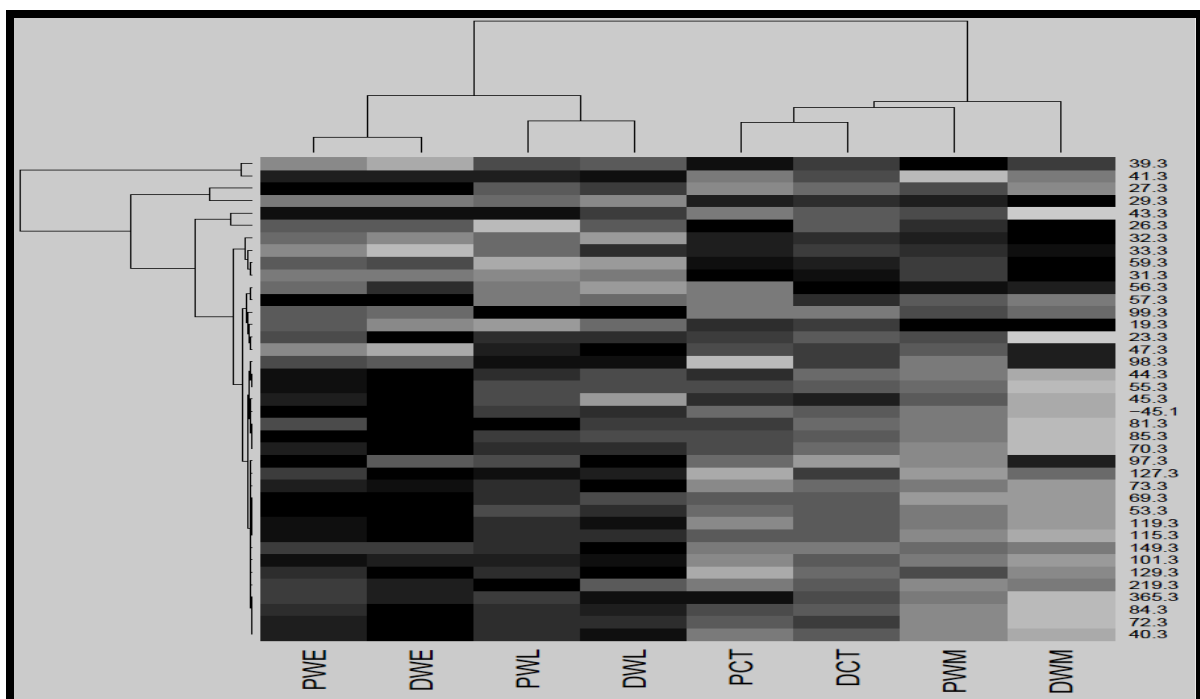
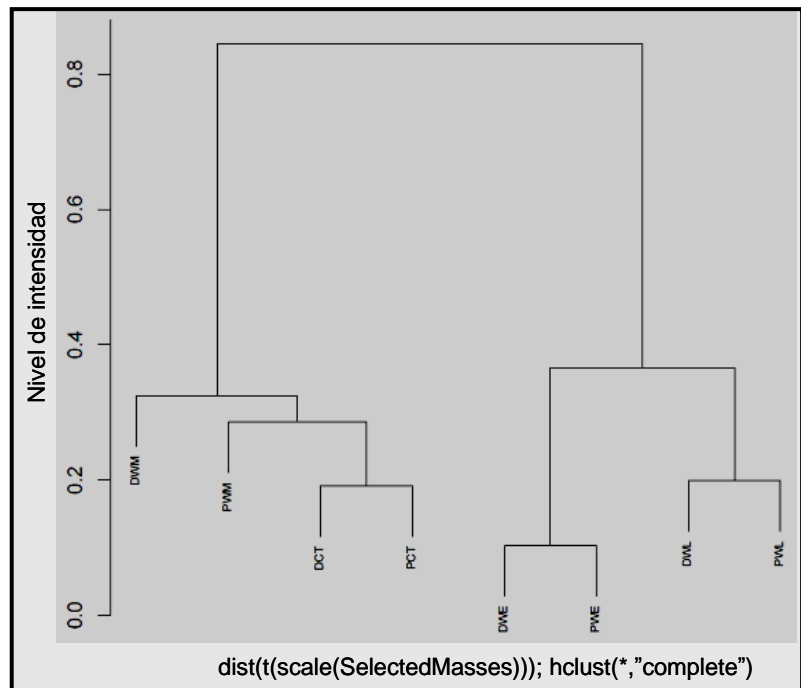


Figura 6.20 Mapa de intensidades- m/z que agrupó a los 40 iones seleccionados con mayor intensidad en el experimento fuente-sumidero en Irapuato incluyendo los iones positivos y negativos. Se puede observar una agrupación uniforme de los iones en función del tratamiento y genotipo, porque se usó el método `scale` como opción de la función `heatmap` en R, en la elaboración de la gráfica. Lo relevante en el heatmap es la agrupación de los tratamientos (PCT y DCT, PWE y DWE, PWL y DWL, PWM y DWM) independientemente del genotipo. La observación detallada de la gráfica permite identificar marcadores diferenciales en los niveles de intensidad en función de la m/z ($n = 504$).

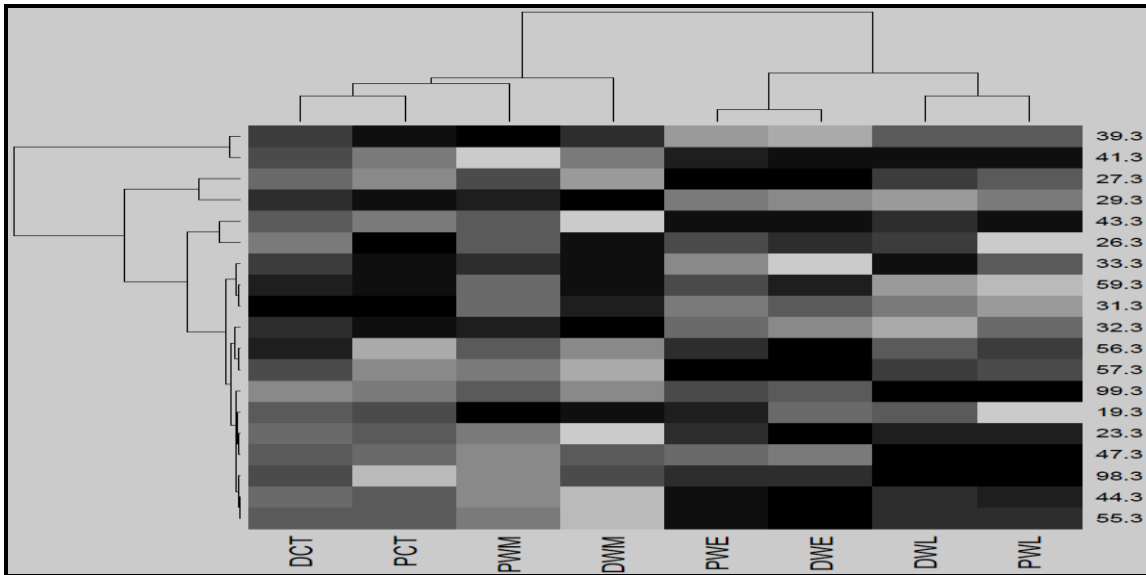


Figura 6.21 Mapa de intensidades- m/z agrupando los 20 iones seleccionados con mayor intensidad en el experimento fuente-sumidero en Irapuato incluyendo los iones positivos y negativos. Se puede observar una agrupación uniforme de los iones en función del tratamiento y genotipo, porque se usó el método *scale* como opción de la función *heatmap* en R, en la elaboración de la gráfica. Lo relevante en el heatmap es la agrupación de los tratamientos (PCT y DCT, PWE y DWE, PWL y DWL, PWM y DWM) independientemente con los tratamientos y del genotipo. Al graficar diferentes iones (como en la gráfica 6.22), la agrupación es consistente con el tratamiento y difiere en el orden debido a los niveles de intensidad de cada ion. La observación detallada de la gráfica permite identificar marcadores diferenciales en los niveles de intensidad en función de la m/z ($n = 504$).

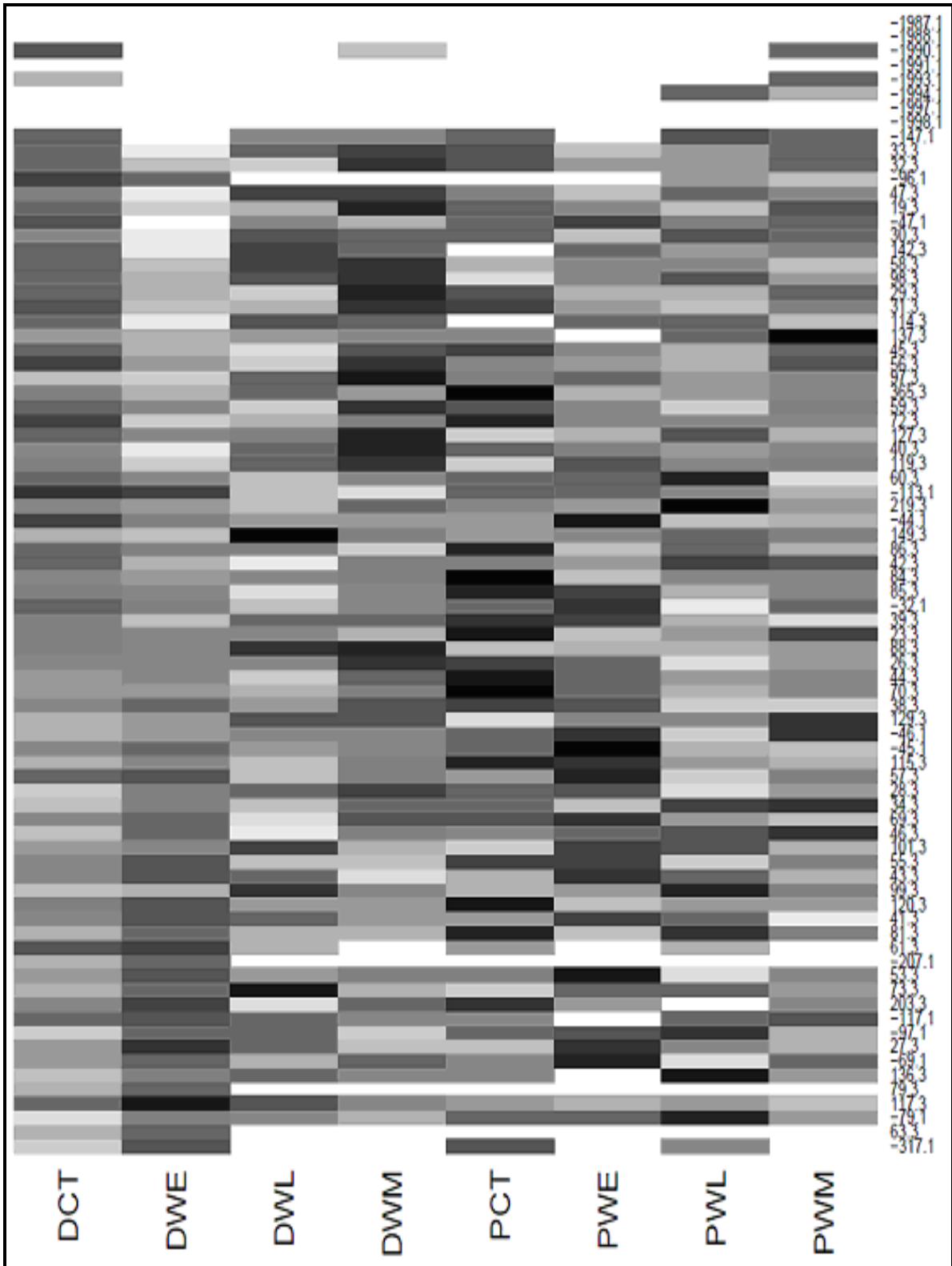


Figura 6.22 Mapa de intensidades- m/z agrupando los 80 iones más intensos del experimento fuente-sumidero en Irapuato incluyó iones positivos y negativos. Se observó una representación uniforme en función del genotipo (Dow y Puma) y tratamiento, (CT = control, WE = sin elotes, WL = sin hojas, WM = sin elotes y sin hojas) porque no se usó método alguno de clasificación ($n = 504$), función heatmap en R.

En la Figura 6.20 se puede observar una agrupación uniforme de los genotipos en función del tratamiento. Lo relevante en el heatmap de la Figura 6.21 es la agrupación de los tratamientos (PCT y DCT, PWE y DWE, PWL y DWL, PWM y DWM) independientemente del genotipo. La observación detallada de la gráfica permite identificar marcadores diferenciales en los niveles de intensidad en función de la relación masa/carga (m/z). De los heatmaps de las Figura 6.21 y 6.22 se desprende que la clasificación por similitud es importante para agrupar los genotipos por tratamiento en la primera, mientras que en la segunda solamente se agruparon los iones de acuerdo a su genotipo.

Los resultados de varianza y de correlación entre algunas variables agronómicas y metabolitos específicos, se analizaron a detalle en Juárez (2012). Los metabolitos analizados en dicho estudio metabólico dirigido fueron: glucosa, fructosa, sacarosa, almidón y aminoácidos totales.

6.1.4 Etapas de desarrollo 2010 en Irapuato

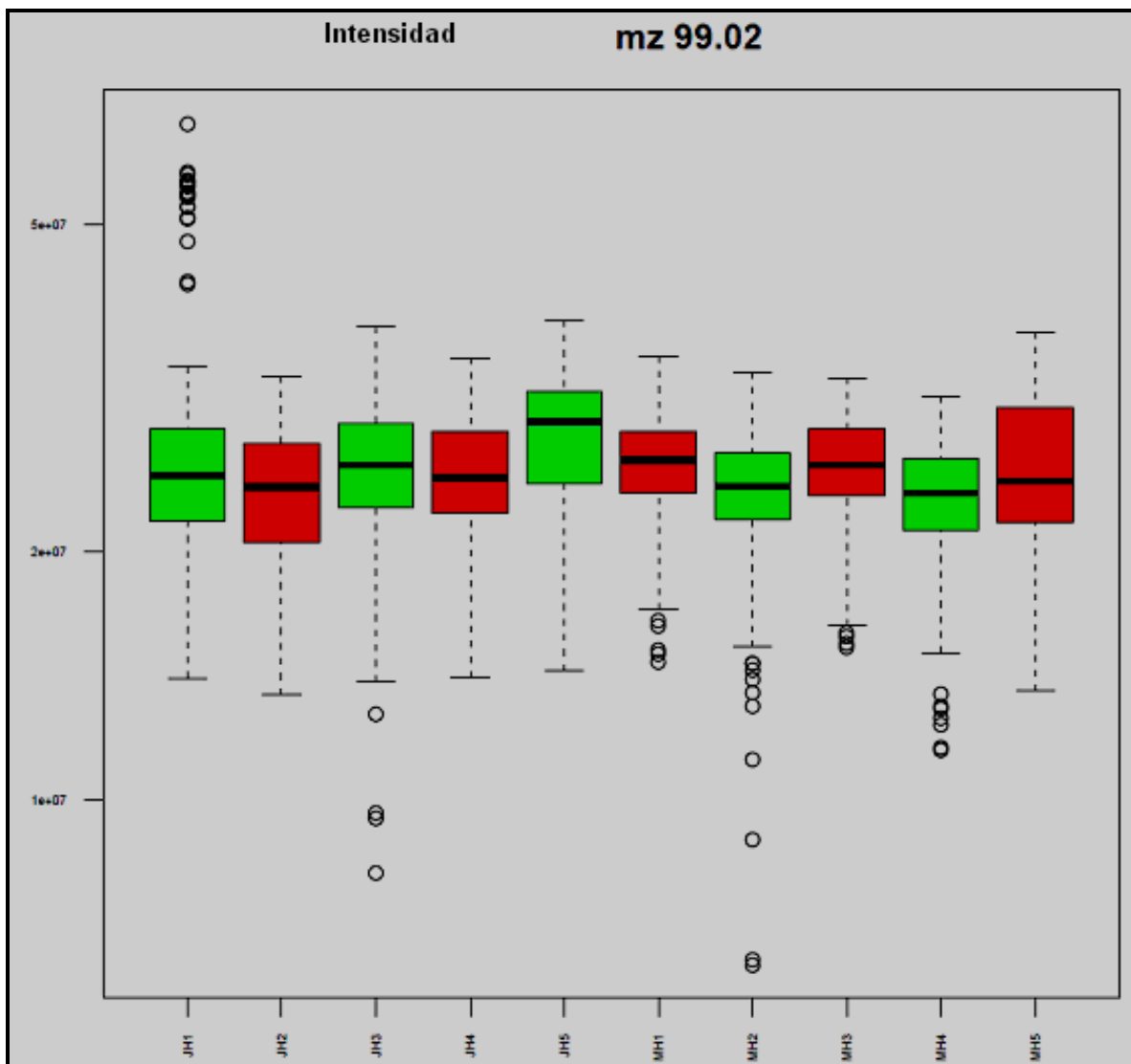


Figura 6.23 Gráfica comparativa de intensidades tipo caja del ion 99.02 m/z entre los meses de Mayo (M) y Junio (J) con los genotipos (H1, H2, H3, H4, y H5), del experimento etapas de desarrollo 2010 en Irapuato, sin la inclusión del factor tiempo ($n = 420$). Las plantas de cada híbrido mantuvieron su nivel de intensidad durante el experimento, presentando una tendencia similar en variación de los niveles de intensidad en cada uno de los híbridos comparado con su correspondiente mes (M y J). Las cinco primeras cajas representan a los genotipos en el mes de Junio y las siguientes cinco a los genotipos durante el mes de Mayo. Los círculos externos a la caja, representan valores extremos.

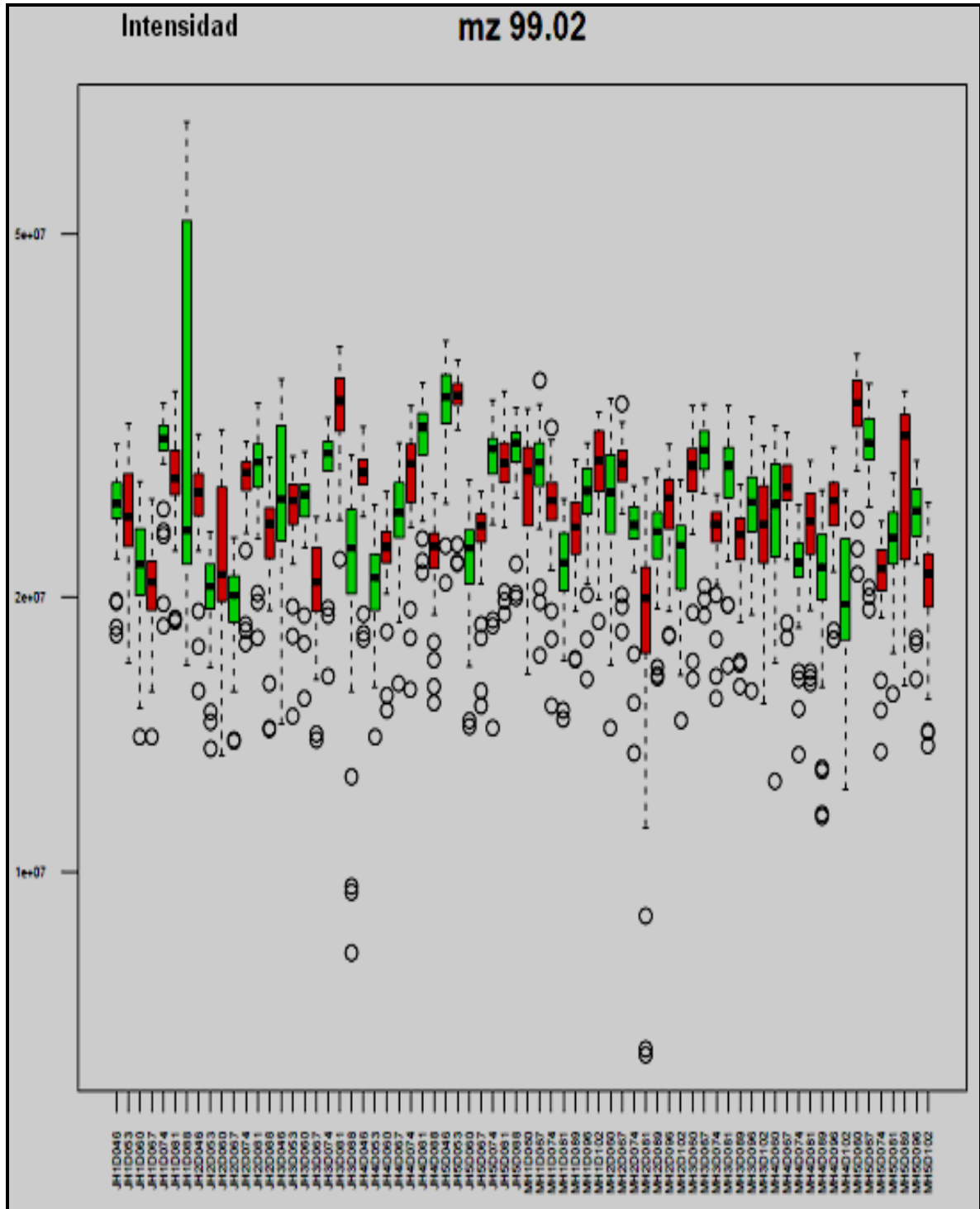


Figura 6.24 Gráfica comparativa de intensidades tipo caja del ion 99.02 m/z entre los meses de Mayo (M) y Junio (J) con los genotipos (H1, H2, H3, H4, y H5), del experimento etapas de desarrollo 2010 en Irapuato, con la inclusión del factor tiempo (46, 53, 60, 67, 74, 81 y 88 días en el mes de junio; 60, 67, 74, 81, 89, 96, y 102 días en el mes de mayo) (n = 420). Aquí se presenta un análisis individual de cada ión que permitió conocer el patrón de variación de un genotipo y tratamiento con respecto al tiempo. Gráfica complementaria a la figura 6.29. Los colores rojo y verde diferencian el nivel de intensidad de los genotipos con respecto a los días de desarrollo. Los círculos representan valores extremos con respecto a la mediana.

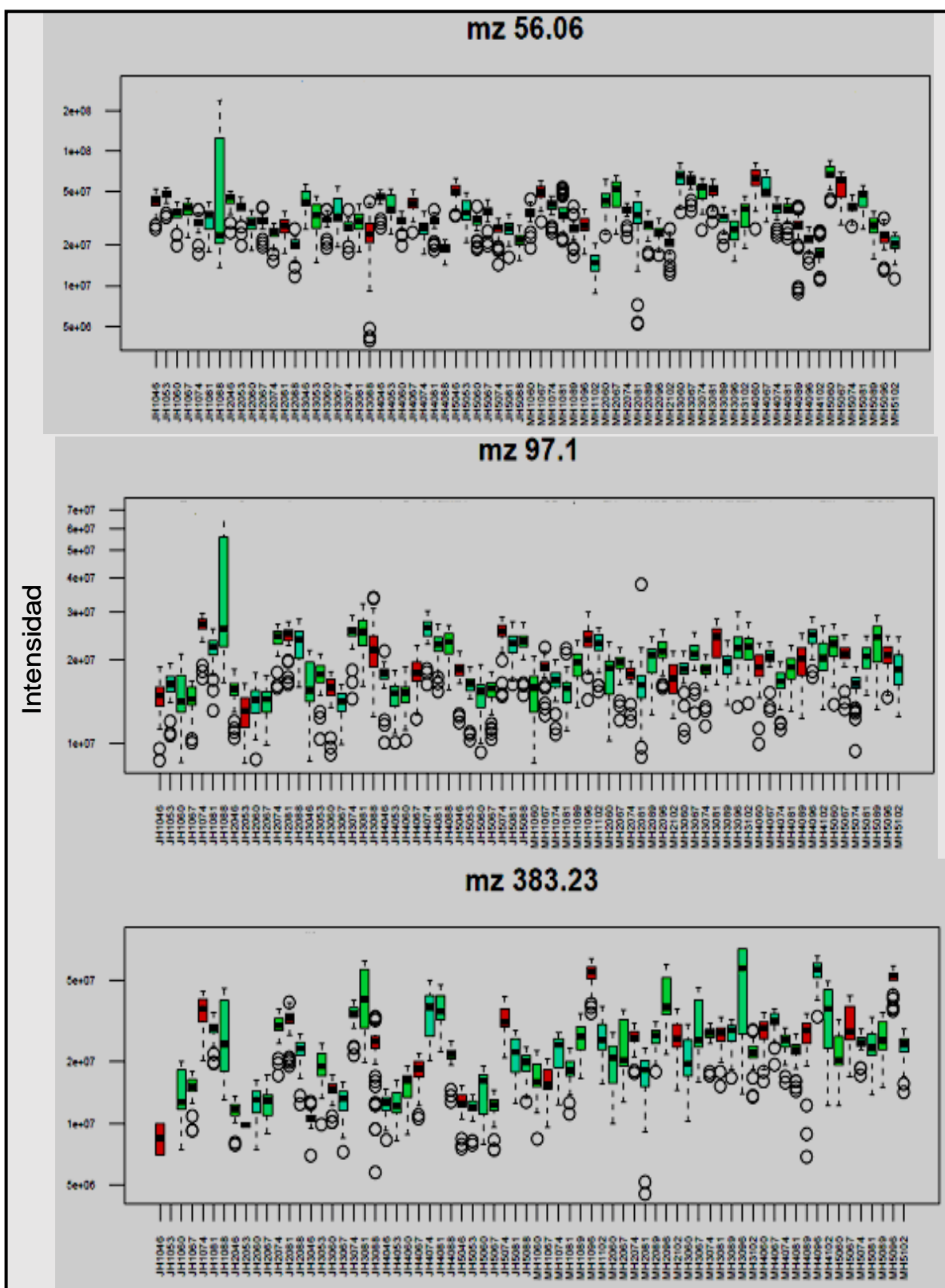


Figura 6.25 Gráfica comparativa de intensidades tipo caja de los iones 56.06, 97.1 y 383.23 m/z entre los genotipos (H1, H2, H3, H4, y H5), tratamientos y tiempo (0, 1, 2, 3, 4 y 10 días) del experimento etapas de desarrollo 2010 de Irapuato ($n = 420$). Aquí se presenta un análisis individual de cada ión para conocer el patrón de variación de un genotipo con respecto al tiempo de cosecha de muestras y meses de siembra: mayo (M) y junio (J). Gráfica complementaria a la figura 6.28.

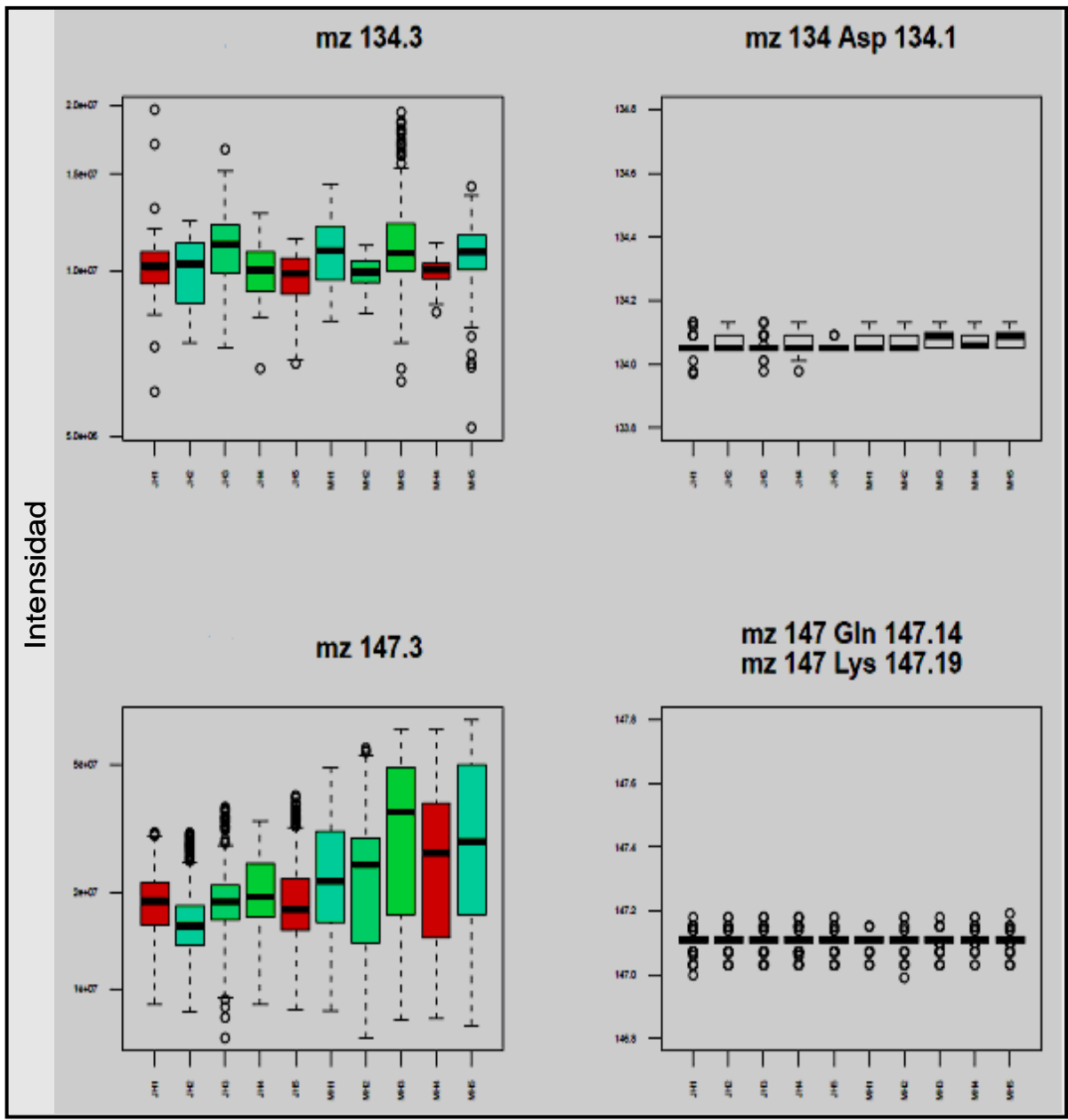


Figura 6.26 Gráficas comparativas de los iones aproximados a la variable m/z de los aminoácidos ácido aspártico (134.1 m/z), glutamina (147.14 m/z) y lisina (147.19 m/z). En las gráficas de la izquierda se muestran los valores de intensidad de los iones 134.3 y 147.3 m/z que corresponden a los aa: glutamina y/o lisina por meses mayo y junio, y genotipo (H1, H2, H3, H4, y H5). En la derecha se representaron los gráficos de los valores m/z del aminoácido comparado con el valor de la señal m/z detectado en cada tratamiento y genotipo. Se observó un incremento en la intensidad del ion en función del tiempo, siendo más elevada en el mes de Mayo (n = 420).

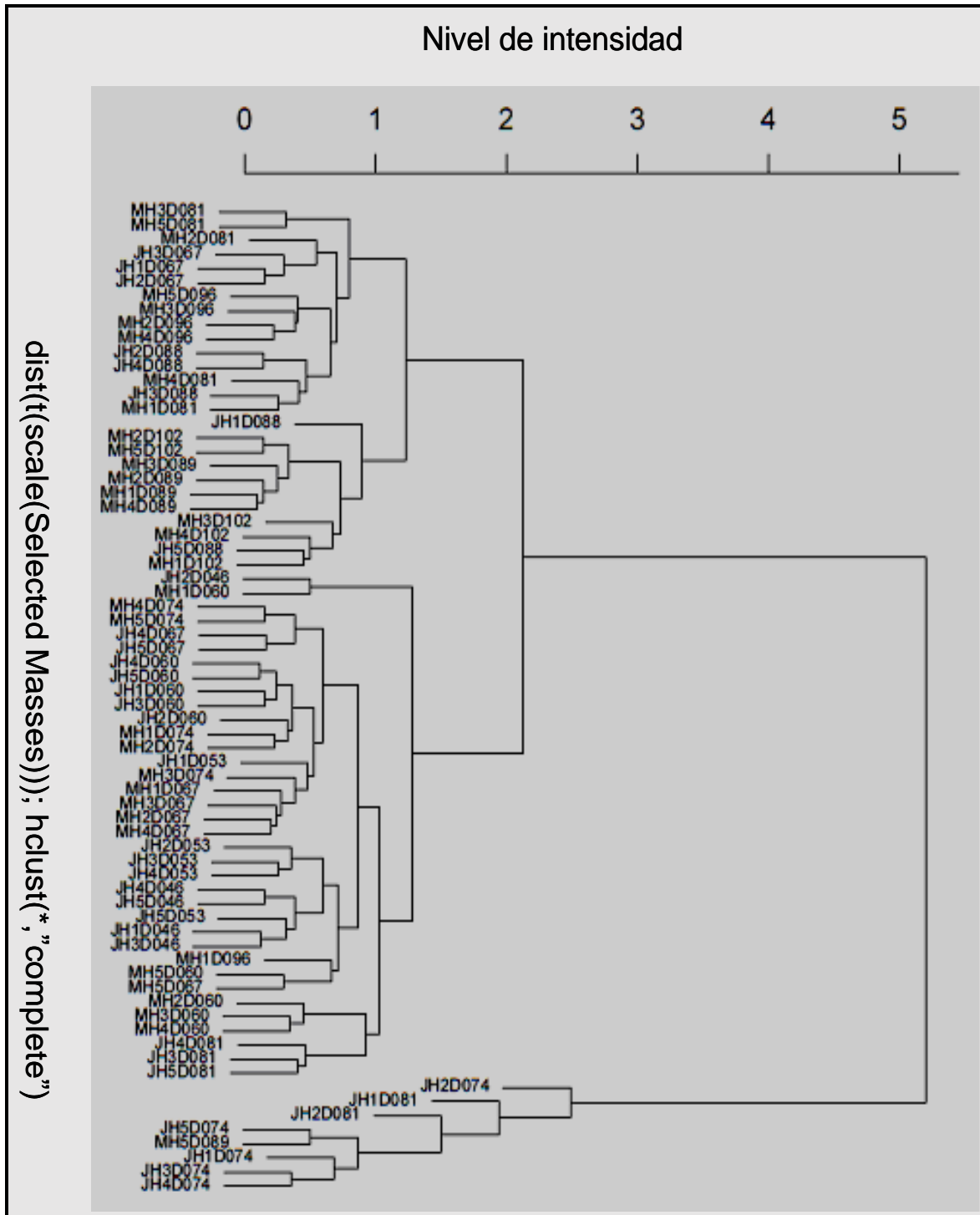
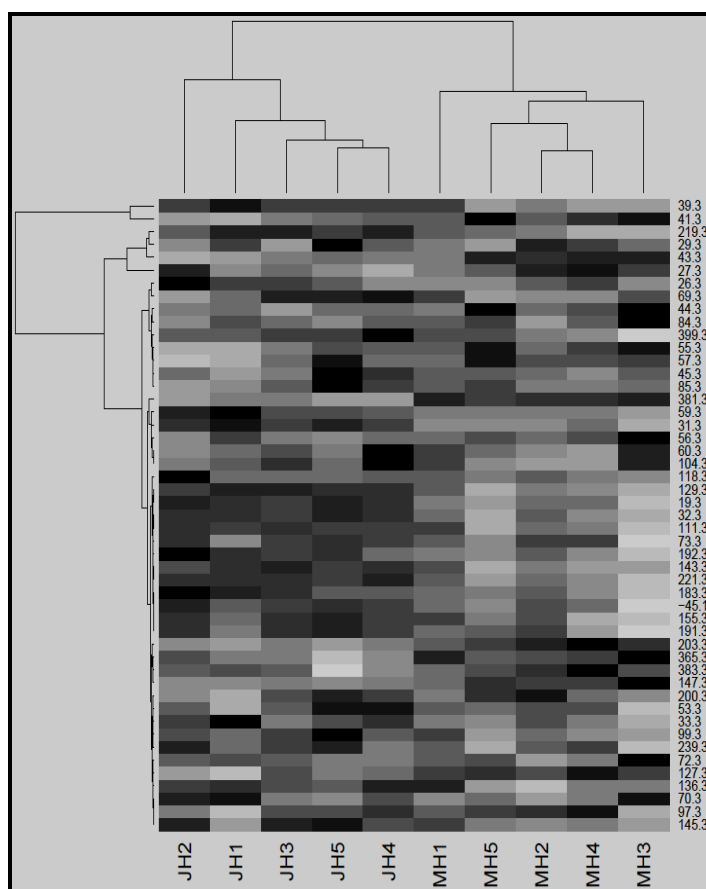


Figura 6.27 Clasificación de los genotipos (H1, H2, H3, H4, y H5), mes de siembra mayo y junio en función del tiempo de cosecha del experimento etapas de desarrollo de Irapuato mediante una representación tipo dendograma, agrupados por su proximidad en el nivel de intensidad. Se puede observar una agrupación uniforme de los genotipos en función del mes de siembra y tiempo de cosecha porque se usó el método $\text{dist}(t(\text{scale}))$ (calcula la proximidad en valor de las lecturas) como opción de la función `hclust` (dendograma con agrupamiento) en R, en la elaboración de la gráfica. El tratamiento JH2D074 (Mes de siembra, junio, híbrido 2, tiempo de cosecha, 74 días) es el que mayor nivel de intensidad presenta ($n = 420$).

Figura 6.28 Mapa de intensidades- m/z que agrupó a los 50 iones seleccionados con mayor intensidad en el experimento etapas de desarrollo 2010 en Irapuato incluyó los iones positivos y negativos. Se puede observar una agrupación uniforme de los iones en función del mes de siembra (M) y (J) y genotipo (H1, H2, H3, H4, y H5), porque se usó el método `scale` como opción de la función `heatmap` en R, en la elaboración de la gráfica. Lo relevante en el heatmap fue la agrupación de los meses de siembra (mayo, M y junio, J) independientemente de los genotipos (H1, H2, H3, H4, y H5). La observación detallada de la gráfica permitió identificar marcadores diferenciales en los niveles de intensidad en función de la m/z ($n = 420$).



6.1.5 Etapas de desarrollo 2011 en Irapuato

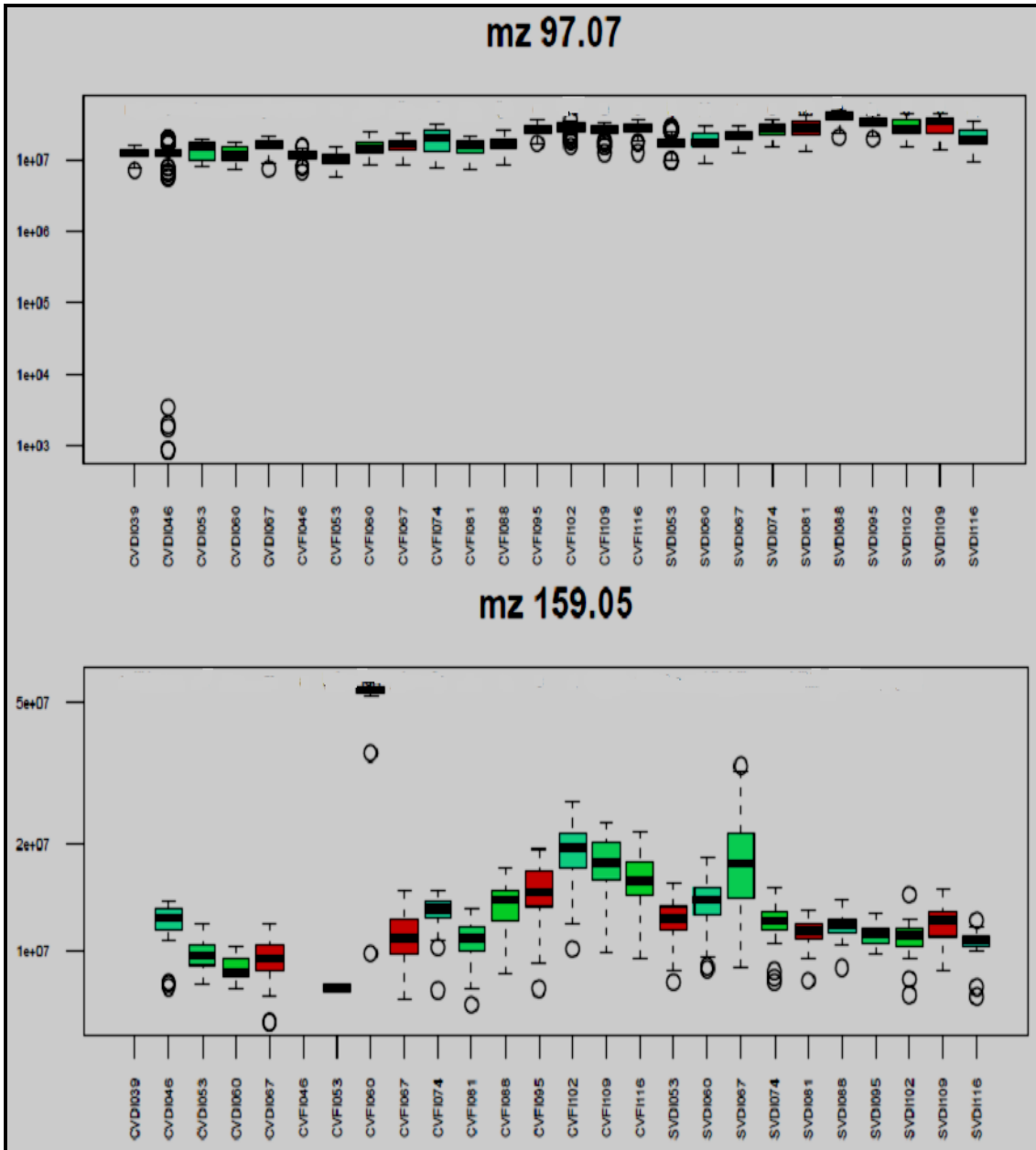


Figura 6.29 Gráfica comparativa de intensidades tipo caja de los iones 97.07 y 159.05 m/z entre plantas con vaina (CV) y sin vaina (SV), los sitios de siembra dentro y fuera del invernadero (DI, FI), con los genotipos (HP, HD), del experimento etapas de desarrollo 2011 en Irapuato, con la inclusión del factor tiempo de cosecha (39, 46, 53, 60, 67, 74, 81, 95, 102, 109 y 116 días) ($n = 416$). Del análisis individual de cada ion se puede observar que una planta con vaina dentro del invernadero para el ion 97.07 permanece estable y el ion 159.05 disminuye. Las plantas con vaina y fuera del invernadero tienen una tendencia a aumentar su intensidad hasta un máximo a los 102 días. Las plantas sin vaina y dentro del invernadero tuvieron un máximo a los 67 días para el ion 159.05 y a los 88 días para el ion 97.07. El análisis de cada uno de los iones permitió conocer el patrón de variación de un genotipo y tratamiento con respecto al tiempo.

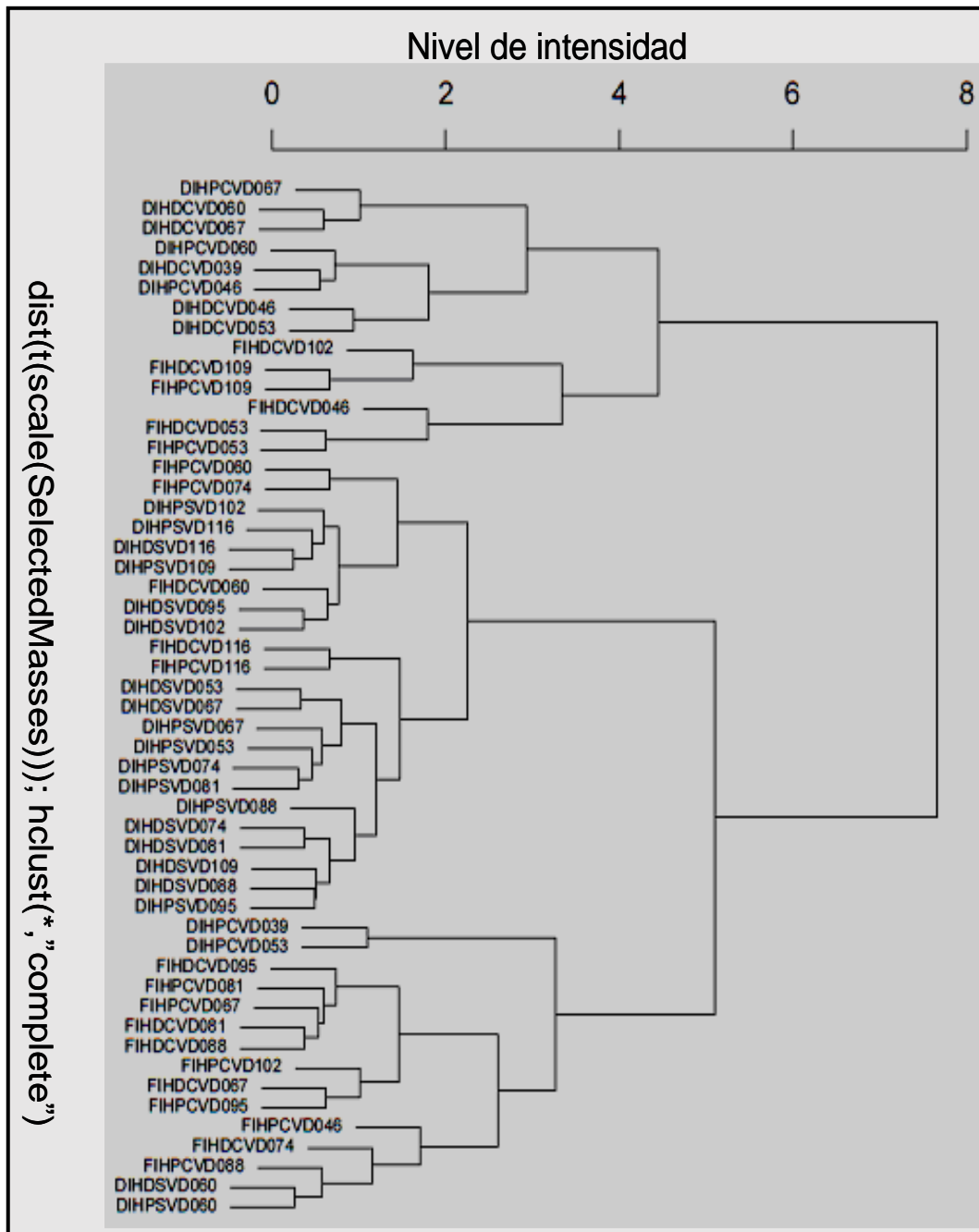


Figura 6.30 Clasificación de los genotipos (HP y HD), sitio de siembra dentro del invernadero (DI) y fuera del invernadero (FI) en función del tiempo de cosecha del experimento, con plantas con vaina (CV) y sin vaina (SV), etapas de desarrollo 2011 en Irapuato mediante una representación tipo dendrograma, agrupados por su proximidad en el nivel de intensidad. Se puede observar una agrupación uniforme de los genotipos en función del sitio de siembra, genotipo, plantas con y sin vaina y tiempo de cosecha porque se usó el método $\text{dist}(t(\text{scale}))$ (evaluación por proximidad) como opción de la función hclust (dendrograma con agrupamiento) en R, en la elaboración de la gráfica. El tratamiento FIHPCVD074 (Sitio de siembra, fuera del invernadero, genotipo Puma, plantas con vaina y tiempo de cosecha, 74 días) fue el que mayor nivel de intensidad presentó ($n = 416$).

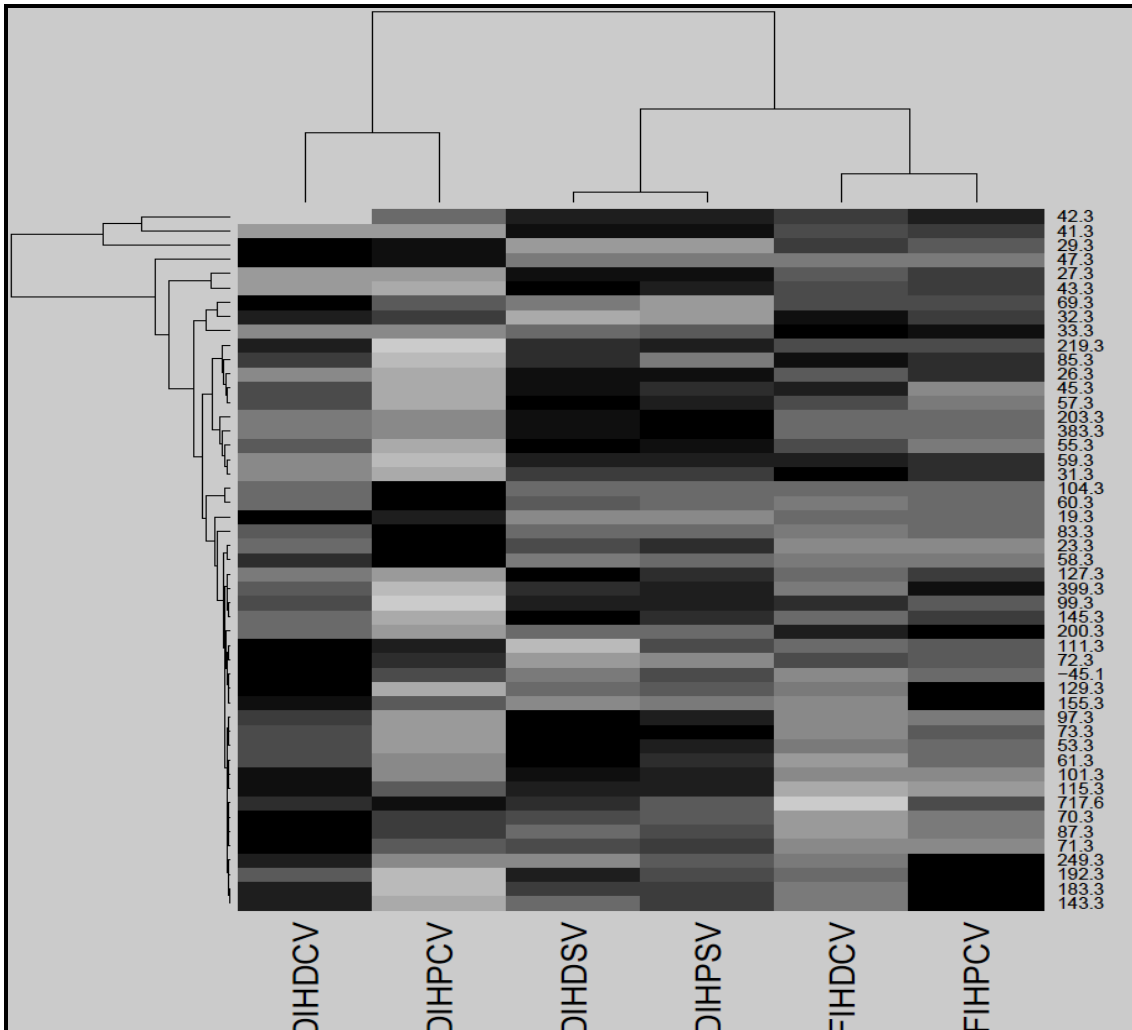


Figura 6.31 Mapa de intensidades- m/z que agrupa los 50 iones seleccionados con mayor intensidad en el experimento etapas de desarrollo 2011 en Irapuato incluyendo los iones positivos y negativos. Se puede observar una agrupación uniforme de los iones en función del sitio de siembra, genotipo (HP y HD) y plantas con vaina (CV) y sin vaina (SV), porque se usó el método scale (valores a escala) como opción de la función heatmap en R, en la elaboración de la gráfica. Lo relevante en el mapa de calor fue la agrupación de los sitios de siembra, dentro del invernadero (DI) y fuera del invernadero (FI) independientemente de los genotipos (HP y HD) así como con y sin vaina. La observación detallada de la gráfica permitió identificar marcadores diferenciales en los niveles de intensidad en función de la m/z ($n = 416$).

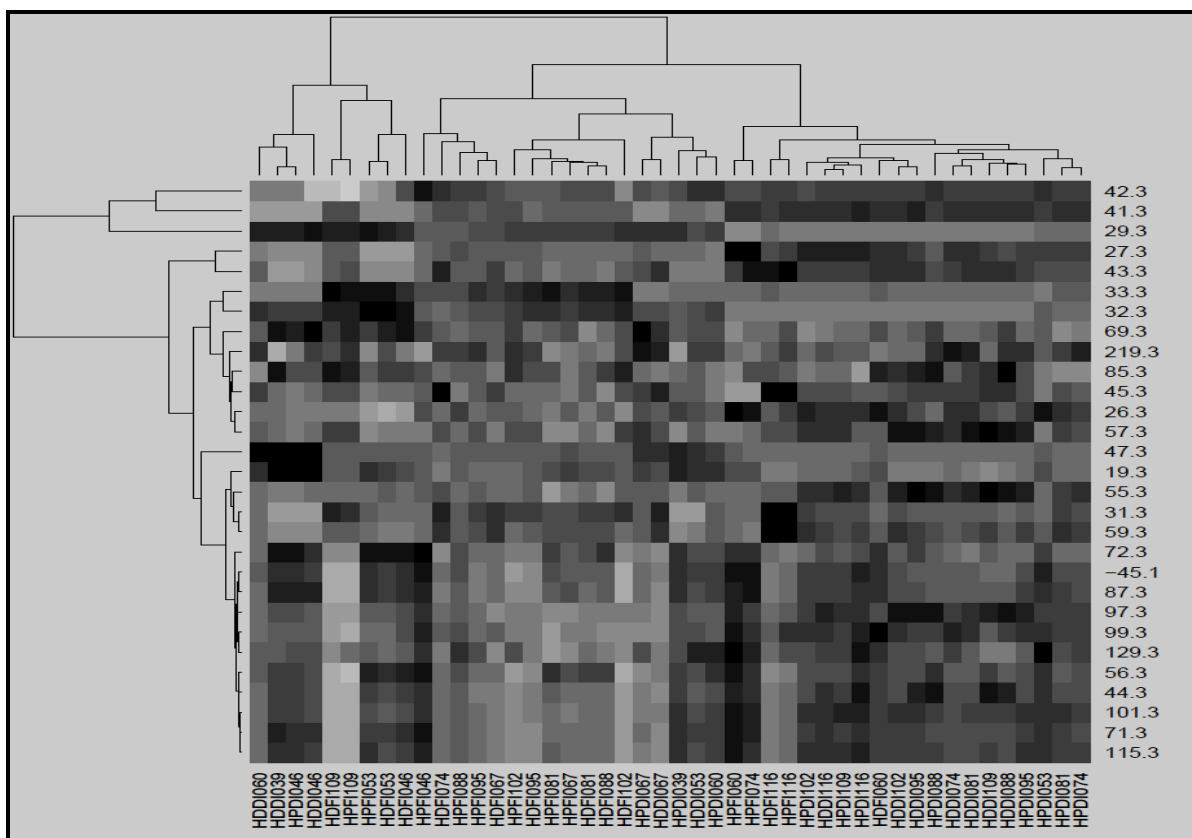


Figura 6.32 Mapa de intensidades- m/z que agrupa los 30 iones seleccionados con mayor intensidad en el experimento etapas de desarrollo 2011 en Irapuato, en el cual se incluyeron los iones positivos y negativos. Se puede observar una agrupación uniforme de los iones en función del sitio de siembra, genotipo (HP y HD) y plantas con vaina (CV) y sin vaina (SV) en función del tiempo de cosecha (39, 46, 53, 60, 67, 74, 81, 95, 102, 109 y 116 días), porque se usó el método scale (valores escalados) como opción de la función heatmap en R, en la elaboración de la gráfica. Lo relevante en el mapa de calor fue la agrupación de los sitios de siembra, dentro del invernadero (DI) y fuera del invernadero (FI) independientemente de los genotipos (HP y HD) así como plantas con y sin vaina y los tiempos de cosecha. La observación detallada de la gráfica permite identificar marcadores diferenciales en los niveles de intensidad en función de la m/z ($n = 416$).

6.2 Experimentos principales con GC

6.2.1 Estrés hídrico en Tlaltizapán y bajo contenido de nitrógeno en El Batán.

Como resultado del proceso de estrés al que estuvieron sujetas las plantas, derivado de la limitación de oxígeno causada por sequía y calor, las plantas que se trataron con estrés hídrico y bajo contenido nitrógeno tuvieron los niveles de síntesis de metabolitos del grupo de alcoholes que se muestran en las Figura 6.33, Figura 6.34 y Figura 6.35. Dichos metabolitos son el etanol, el ácido láctico y el ácido acético. Los datos generados en la cromatografía de gases (GC) resultaron un complemento en el estudio metabolómico de los estreses a que fueron sometidas las plantas; de manera particular, las concentraciones por metabolito y genotipo nos permitieron clasificar por el grado de similitud los experimentos y tratamientos mediante un dendrograma (Figura 6.36).

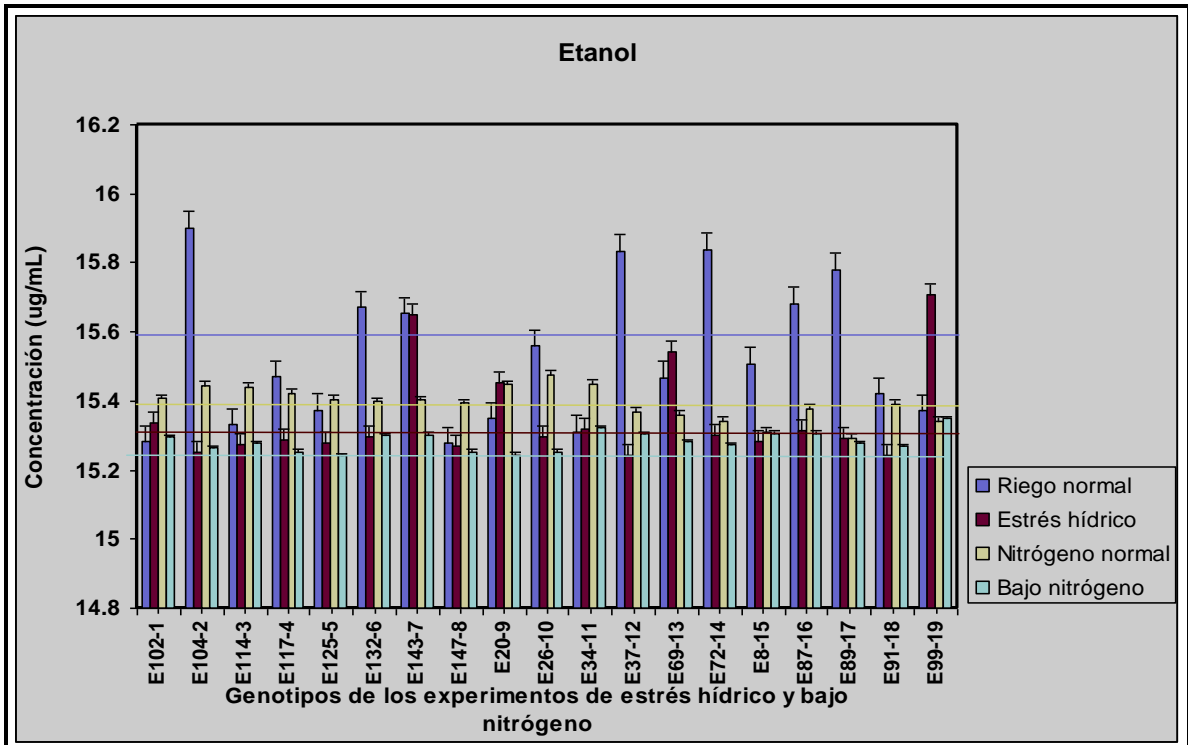


Figura 6.33 Gráfica de los niveles en el contenido de etanol por genotipo para los experimentos de estrés hídrico y bajo contenido de nitrógeno. Las barras de error fueron estimadas con el programa Excel: error típico, $p < 0.05$ ($n = 456$).

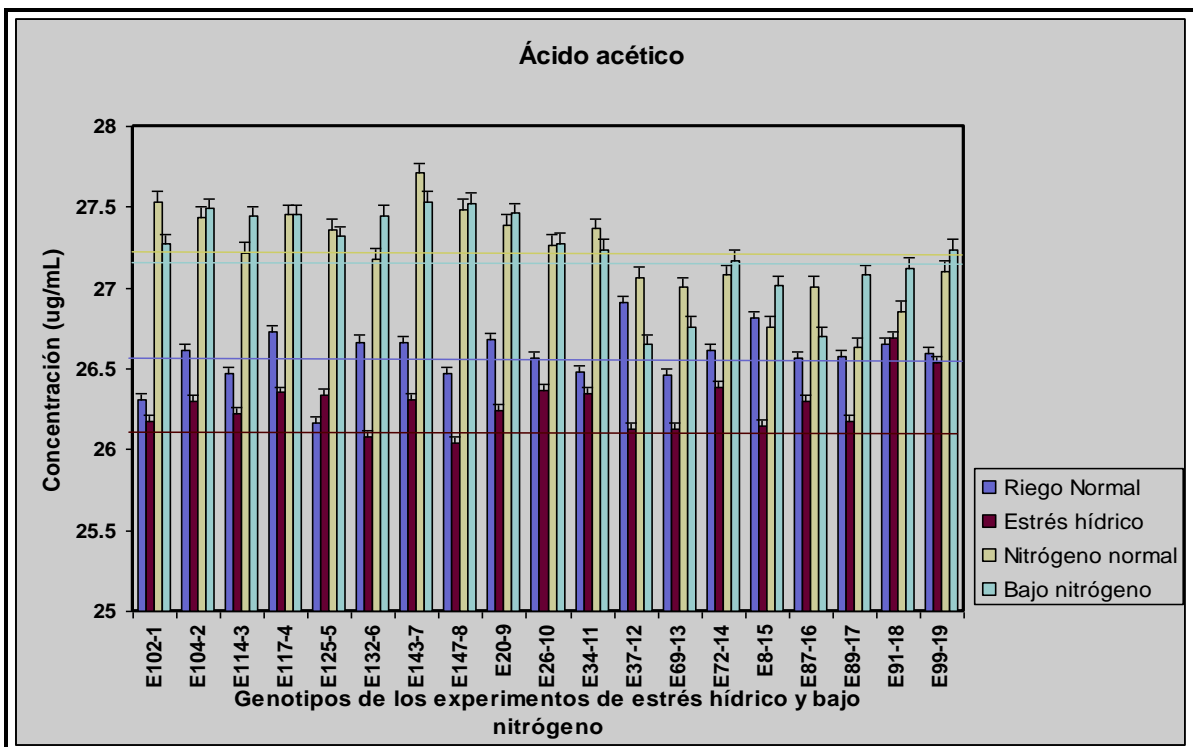


Figura 6.34 Gráfica de los niveles en el contenido de ácido acético por genotipo para los experimentos de estrés hídrico y bajo contenido de nitrógeno. Las barras de error fueron estimadas con el programa Excel: error típico, $p < 0.05$ ($n = 456$).

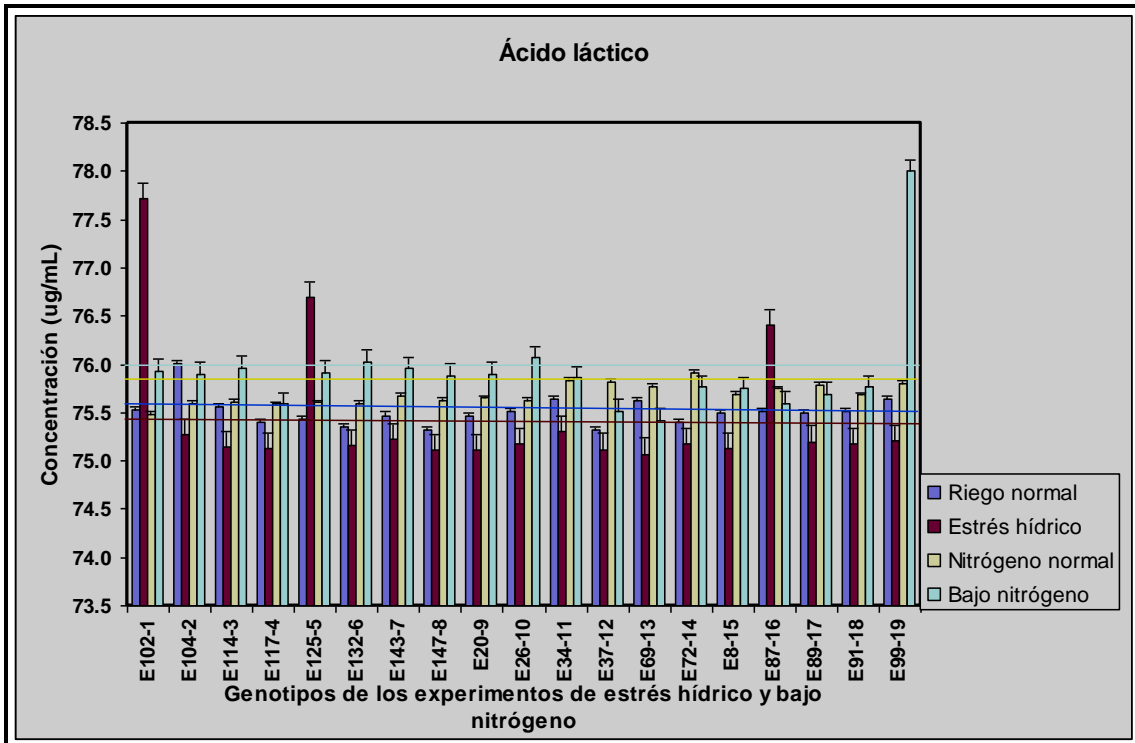


Figura 6.35 Gráfica de los niveles en el contenido de ácido láctico por genotipo para los experimentos de estrés hídrico y bajo contenido de nitrógeno. Las barras de error fueron estimadas con el programa Excel: error típico, $p < 0.05$ ($n = 456$).

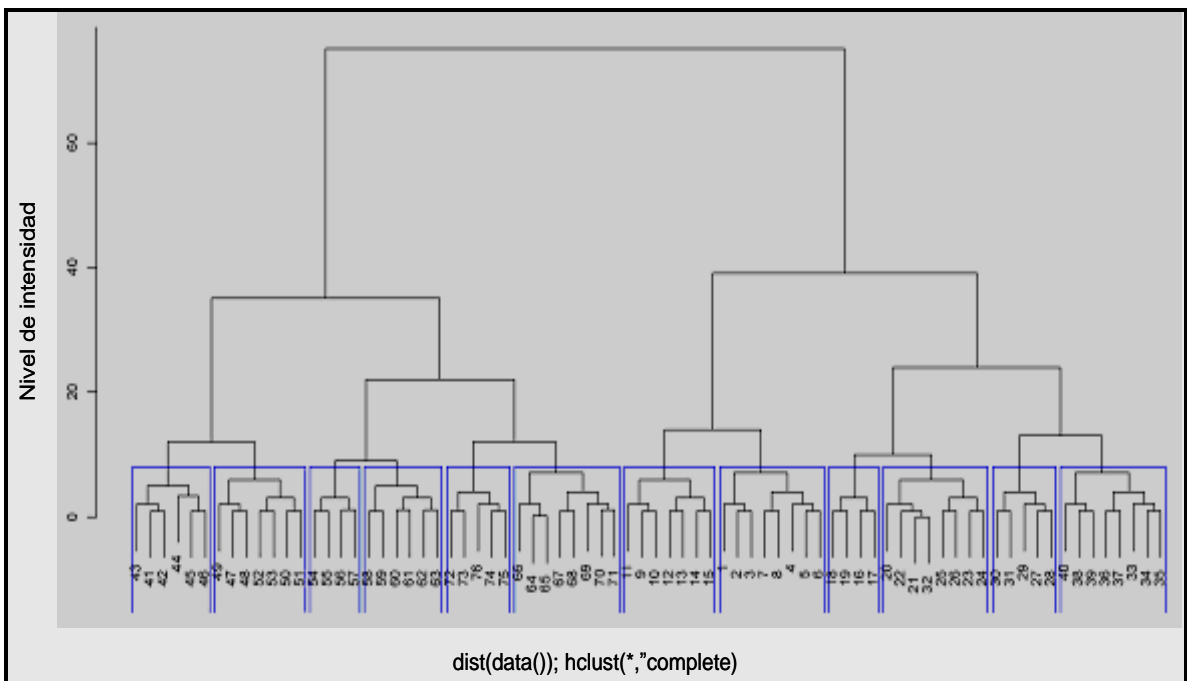


Figura 6.36 Clasificación de los genotipos (1-19 LN, 20-38 CT, bajo contenido de nitrógeno; 39-57 WS, 58-76 CT, estrés hídrico) de los experimentos estrés hídrico en Tlaltizapán y bajo contenido de nitrógeno en El Batán mediante una representación tipo dendograma, agrupados por su proximidad en el nivel de intensidad. Se puede observar una agrupación homogénea de los genotipos en función del experimento y condición o tratamiento (n = 76). La gráfica nos permite conocer la proximidad de valores en concentración (ug/mL) de etanol de los genotipos lo que nos permite identificar y asociar los genotipos con su tratamiento o condición.

6.3 Comparación de características agronómicas y mediciones bioquímicas.

6.3.1 Correlaciones en experimentos de estrés hídrico y bajo nitrógeno con el método estadístico de Pearson. Tabla 6.1 Niveles de correlación entre factores del experimento de estrés hídrico en Tlaltizapán. Método de Pearson en R. (n= 456). Podemos observar variables con valores de correlación positivos y negativos. Estableciendo un 30% como un nivel de correlación a considerar aceptable, se presentan de color verde los valores mayores y en azul los valores menores a dicho valor en consideración. La presente tabla permitió considerar a los factores iniciales en la construcción del modelo lineal de regresión.

\$Correlacion	Tlaltizapan Estrés Hídrico			WS					
	W_total	Largo	jilotesPPlanta	W_total_jilotes	PesoPjilote	W_tallo	W_tallo_sec	W_jugo_sec	W_jugo
W_total	1								
Largo	0.45	1							
jilotesPPlanta	0.29	0.18	1						
W_total_jilotes	0.65	0.35	0.41	1					
PesoPjilote	0.56	0.33	0.03	0.91	1				
W_tallo	0.85	0.42	0.13	0.24	0.20	1			
W_tallo_sec	0.65	0.27	0.00	0.25	0.26	0.76	1		
W_jugo_sec	0.47	0.21	0.00	0.20	0.23	0.53	0.52	1	
W_jugo	0.49	0.24	0.04	0.16	0.17	0.56	0.32	0.93	1
Glc	-0.04	-0.13	0.13	-0.07	-0.12	-0.04	-0.08	0.13	0.15
Frc	-0.05	-0.13	0.04	0.00	-0.01	-0.11	-0.15	0.12	0.16
Suc	-0.03	0.08	0.25	0.21	0.12	-0.19	-0.05	-0.02	-0.07
AA	-0.14	-0.18	0.02	-0.08	-0.11	-0.15	-0.01	-0.04	-0.09
Etanol	0.20	0.10	0.05	0.16	0.15	0.17	0.29	0.15	0.07
Ac.Acetico	0.16	0.23	0.18	0.15	0.08	0.16	0.05	0.08	0.11
Ac.Lactico	-0.21	-0.30	-0.03	-0.17	-0.16	-0.14	-0.14	-0.13	-0.11
Biomasa_Fresca	0.59	0.37	0.13	0.39	0.36	0.56	0.42	0.26	0.25
	Glc	Frc	Suc	AA	Etanol	Ac.Acetico	Ac.Lactico	Biomasa_Fresca	
W_total									
Largo									
jilotesPPlanta									
W_total_jilotes									
PesoPjilote									
W_tallo									
W_tallo_sec									
W_jugo_sec									
W_jugo									
Glc	1								
Frc	0.86	1							
Suc	0.18	0.05	1						
AA	0.61	0.54	0.17	1					
Etanol	0.03	0.03	0.01	0.04	1				
Ac.Acetico	0.05	0.03	-0.11	0.01	0.08	1			
Ac.Lactico	-0.06	-0.05	0.02	-0.11	-0.02	0.12	1		
Biomasa_Fresca	-0.31	-0.37	-0.10	-0.29	0.01	0.10	-0.17	1	

Tabla 6.2 Niveles de correlación entre factores del experimento de bajo contenido de nitrógeno en El Batán (n = 456). Es posible observar variables con valores de correlación positivos y negativos. Estableciendo un 30% como un nivel de correlación a considerar aceptable, se presentan de color verde los valores mayores y en azul los valores menores a dicho valor en consideración. La presente tabla permitió incluir a los factores iniciales en la construcción del modelo lineal de regresión.

	Batan Bajo Nitrogeno					LN					
	W_total	Wsin_jil	Largo	W_espiga	jilotesPPlanta	W_total_jilotes	W_toto	W_elotes	Nudos	W_tallo	Tallo_moler
W_total	1										
Wsin_jil	0.95	1									
Largo	0.70	0.73	1								
W_espiga	0.43	0.55	0.22	1							
jilotesPPlanta	0.03	-0.02	-0.02	-0.08	1						
W_total_jilotes	0.36	0.04	0.05	-0.24	0.17	1					
W_toto	0.41	0.13	0.12	-0.21	0.18	0.93	1				
W_elotes	0.23	-0.06	-0.04	-0.26	0.16	0.91	0.74	1			
Nudos	-0.05	-0.20	-0.09	-0.42	0.10	0.42	0.44	0.30	1		
W_tallo	0.92	0.98	0.76	0.54	-0.02	0.00	0.07	-0.09	-0.23	1	
Tallo_moler	0.83	0.88	0.64	0.51	-0.07	0.02	0.12	-0.07	-0.15	0.88	1
W_jugo_sec	0.72	0.83	0.56	0.58	-0.07	-0.17	-0.08	-0.23	-0.32	0.83	0.90
W_jugo	0.78	0.90	0.63	0.59	-0.05	-0.18	-0.10	-0.24	-0.34	0.91	0.84
Glc	-0.01	0.07	0.16	0.19	-0.11	-0.25	-0.28	-0.19	-0.22	0.08	0.05
Frc	0.14	0.26	0.33	0.31	-0.08	-0.32	-0.30	-0.30	-0.29	0.27	0.24
Suc	-0.29	-0.43	-0.27	-0.57	0.21	0.37	0.30	0.39	0.37	-0.43	-0.33
AA	0.02	-0.05	0.04	-0.20	-0.05	0.22	0.26	0.13	0.38	-0.09	-0.09
Etanol	-0.03	-0.01	-0.18	0.26	-0.12	-0.06	-0.06	-0.10	-0.08	-0.04	-0.04
Ac.Acetico	0.18	0.24	0.26	0.31	0.04	-0.14	-0.09	-0.18	-0.09	0.26	0.22
Ac.Lactico	0.09	0.10	0.05	0.03	-0.02	0.01	0.07	-0.06	0.06	0.06	0.07
Yield	0.07	0.02	0.29	-0.38	0.02	0.16	0.20	0.08	0.09	0.02	0.05
Biom_fresca	0.75	0.83	0.76	0.52	0.00	-0.08	-0.03	-0.17	-0.18	0.84	0.69
	W_jugo_sec	W_jugo	Glc	Frc	Suc	AA	Etanol	Ac.Acetico	Ac.Lactico	Yield	Biom_fresca
W_total											
Wsin_jil											
Largo											
W_espiga											
jilotesPPlanta											
W_total_jilotes											
W_toto											
W_elotes											
Nudos											
W_tallo											
Tallo_moler											
W_jugo_sec	1										
W_jugo	0.96	1									
Glc	0.19	0.19	1								
Frc	0.35	0.36	0.77	1							
Suc	-0.47	-0.53	-0.20	-0.35	1						
AA	-0.22	-0.22	-0.06	0.02	0.32	1					
Etanol	0.03	0.02	-0.02	-0.07	0.01	0.14	1				
Ac.Acetico	0.18	0.19	0.02	0.28	-0.17	0.23	-0.02	1			
Ac.Lactico	0.10	0.08	-0.08	-0.03	-0.08	0.21	0.18	0.28	1		
Yield	-0.01	-0.02	0.17	0.28	0.23	0.22	-0.17	0.07	0.01	1	
Biom_fresca	0.67	0.76	0.15	0.39	-0.50	0.00	0.00	0.28	0.03	0.09	1

El nivel estadístico de correlación de Pearson entre las características agronómicas y los metabolitos de los dos experimentos de estrés, hídrico y bajo contenido de nitrógeno, permitió evaluar la importancia de cada variable en relación a las demás, para correlacionar efectos entre los valores de cada una de ellas.

6.3.2 Comparación de niveles de azúcares no-estructurales, aminoácidos y características agronómicas.

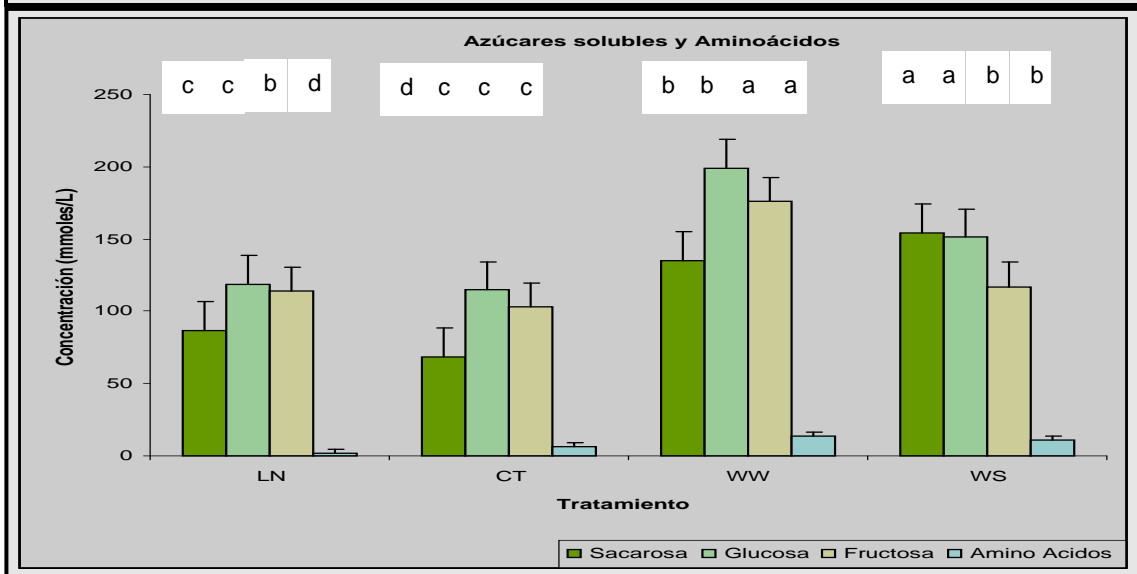
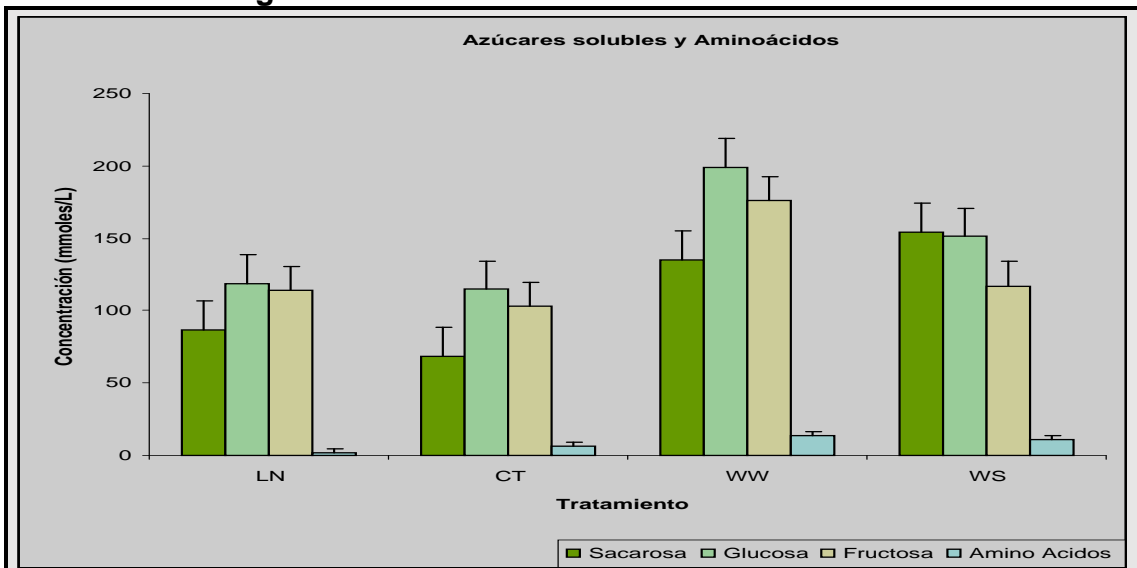


Figura 6.37 Niveles comparativos de concentración de azúcares solubles y aminoácidos totales para cada experimento y tratamiento representados. En los cuatro tratamientos la glucosa presenta mayor concentración, siendo la fructosa y la sacarosa menores en ese orden, excepto en el tratamiento de estrés hídrico (WS) en el que la fructosa fue menor a la sacarosa, destacándose su papel como osmolito. Los niveles de aminoácidos fueron mayores en los tratamientos de estrés hídrico y su control y menores en el experimento de bajo contenido de nitrógeno. Las barras de error fueron estimadas con el programa Excel: error típico, $p < 0.05$ ($n = 456$).

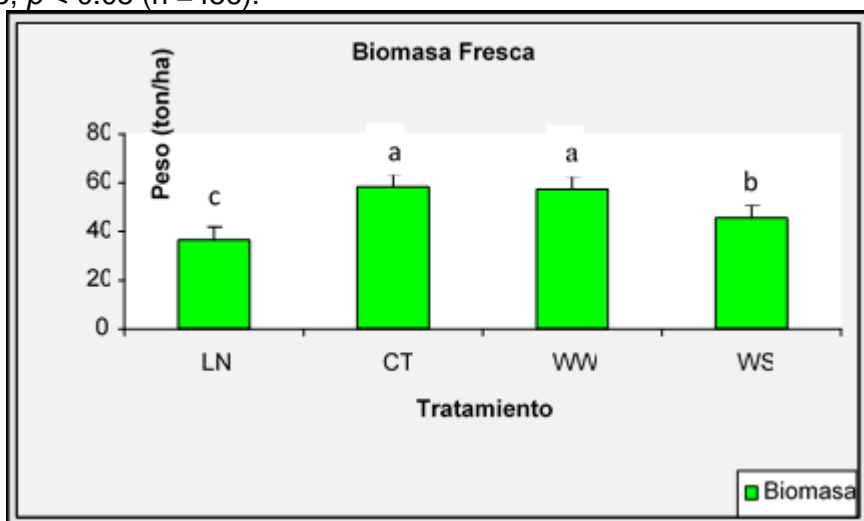


Figura 6.38 La producción de biomasa fresca es mayor en los tratamientos control de los experimentos de estrés hídrico y bajo contenido de nitrógeno ($n = 456$); como era de esperarse, se confirma la importancia del suministro de agua y nitrógeno en el desarrollo de las plantas. Las barras de error fueron estimadas con el programa Excel: error típico, $p < 0.05$, ($n = 456$).

En las Figura 6.377, Figura 6.38 y Figura 6.39 se establecieron los niveles comparativos entre algunas características agronómicas y metabólicas. Los dos experimentos con sus dos tratamientos se compararon para conocer su desempeño global elucidando el significado de las diferencias entre los experimentos con respecto a sus controles.

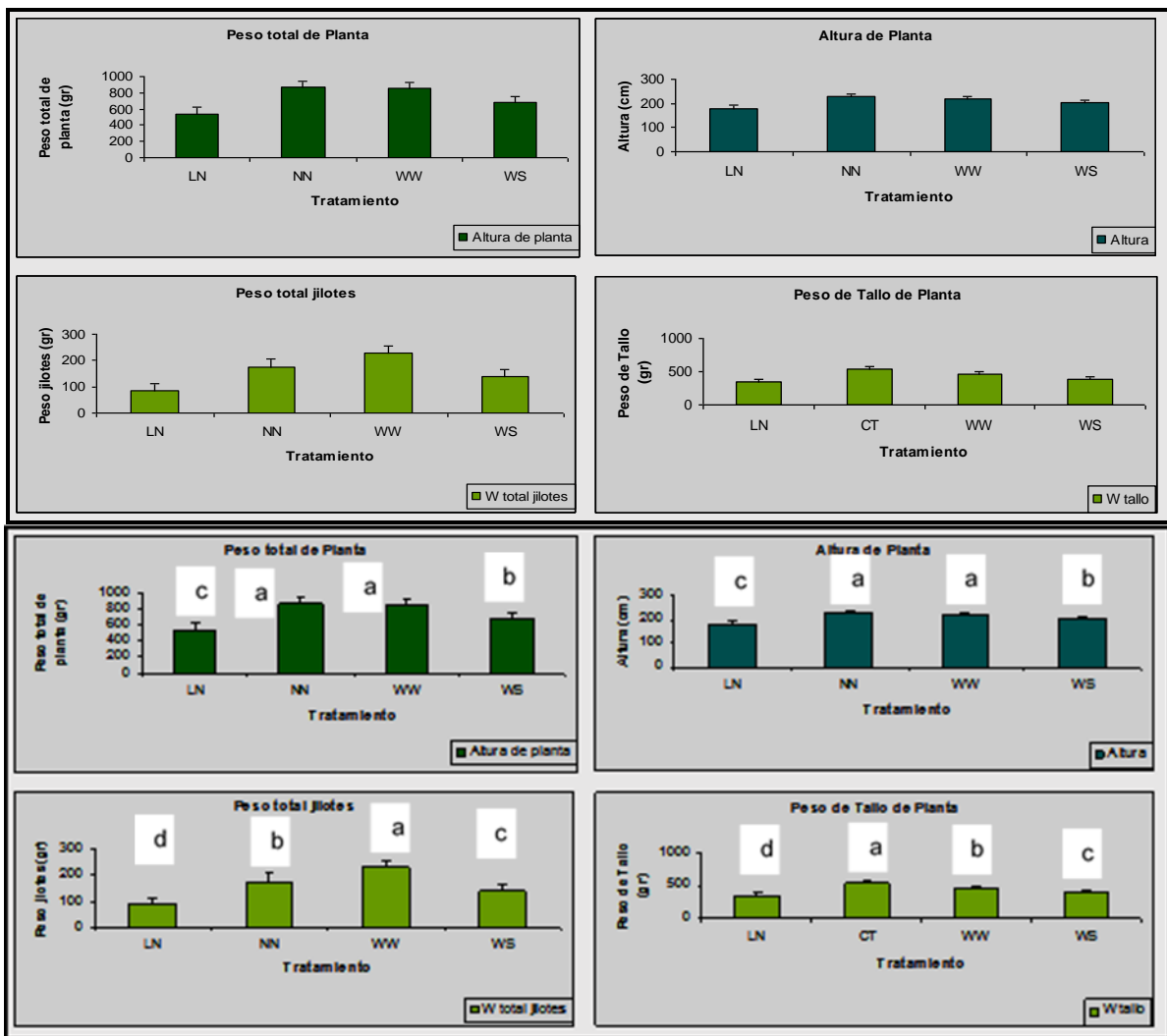


Figura 6.39 Los niveles de las características agronómicas presentadas, (A) peso total de planta, (B) altura de planta, (C) peso total de jilotes y (D) peso de tallo de planta, son mayores en los tratamientos control de los dos experimentos, estrés hídrico y bajo con nitrógeno, disminuyendo en los tratamientos (WS) y (LN). Las barras de error fueron estimadas con el programa Excel: error típico, $p < 0.05$, ($n = 45$).

6.3.3 Producción de biomasa seca de los experimentos de estrés hídrico en Tlaltizapán y bajo nitrógeno en Batán.

En la Figura 6.40 se presentaron los niveles de rendimiento en biomasa seca para los tratamientos bajo estrés hídrico y bajo contenido de nitrógeno agrupados mediante el análisis de medias de Tukey.

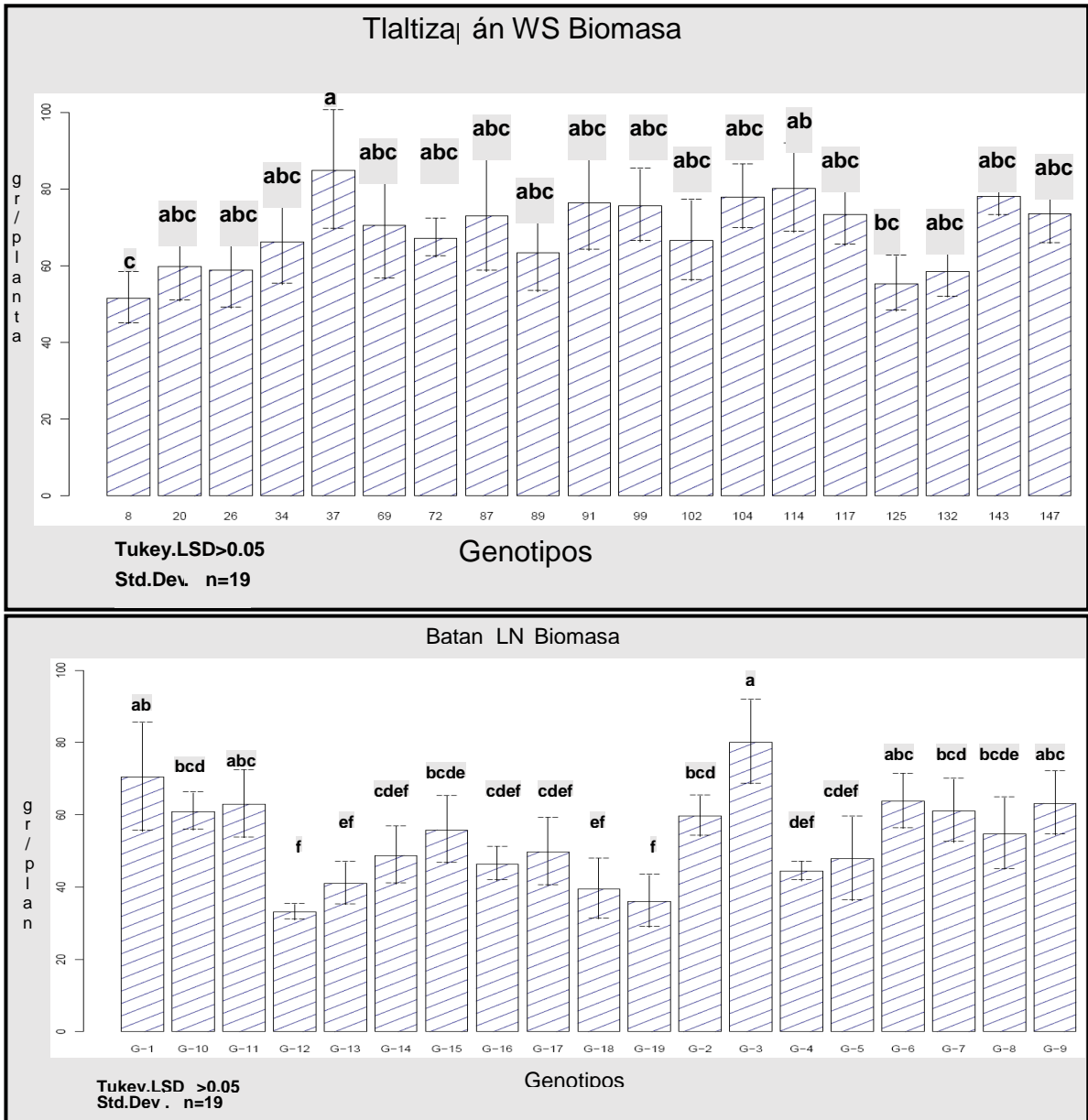


Figura 6.40 Los niveles de rendimiento en biomasa seca del tratamiento de bajo contenido de nitrógeno (LN) son superiores en el genotipo (CP1: 56.4%): G-3, G-7, G-9, G-6, G-10, G-4, G-8, G-2, G-1, de acuerdo al resultado del análisis de medias con el método estadístico de Tukey (n = 19).

6.3.4 Producción de grano del experimento de bajo contenido de nitrógeno en El Batán.

En la Figura 6.42 se registró el nivel de rendimiento de grano para el experimento de bajo contenido de nitrógeno, agrupando cada genotipo mediante el análisis de medias de Tukey

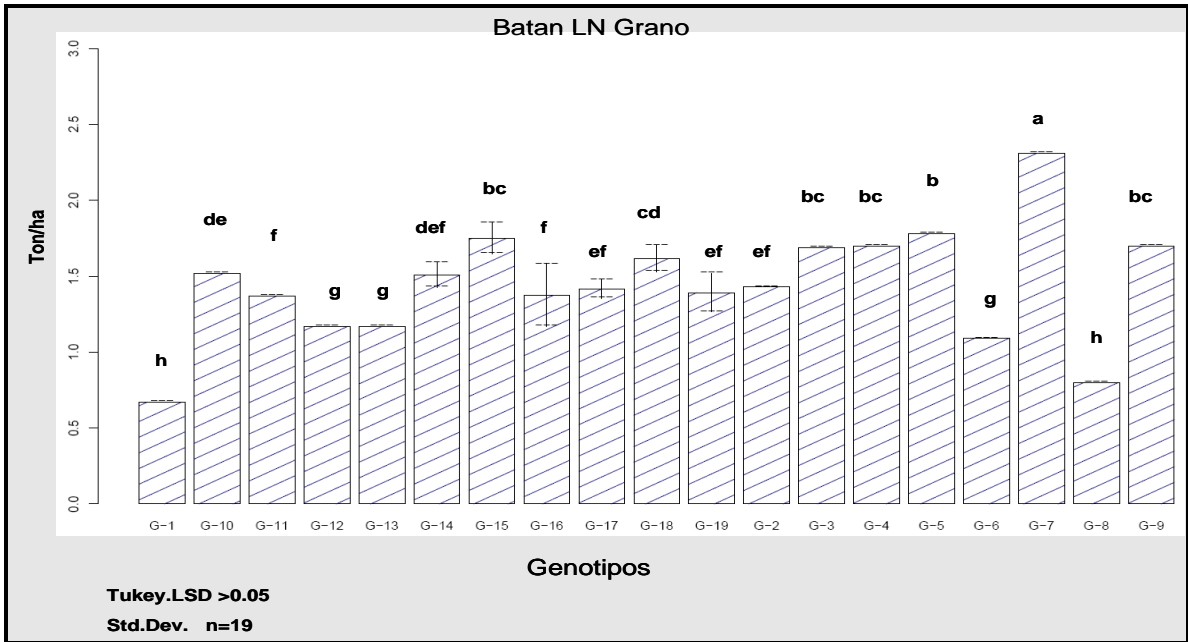


Figura 6.42 Los niveles de rendimiento de grano del tratamiento de bajo contenido de nitrógeno (LN) fueron superiores en los genotipos (CP1 84.4%): G-7, G-5, G-4-, G-9, G-3, G-10, G-2, G-11 de acuerdo al resultado del análisis de medias con el método estadístico de Tukey (n = 19).

6.3.5 Ecuaciones de regresión lineal múltiple para predicción de rendimiento de biomasa fresca y grano.

Tabla 6.3 Ecuaciones de regresión múltiple para predecir el rendimiento de grano y biomasa fresca del experimento de El Batán en sus dos tratamientos. También se incluye la ecuación de regresión para predecir el rendimiento de biomasa fresca del experimento de Tlaltizapán. Se observan las ecuaciones, sus R cuadradas y métodos de selección de variables con el método *step wise* y de regresión múltiple con valores α (n = 456).

Experimento	Batan CT-LN Grano	
Ecuaciones de regresión	Grano ^{0.286216} = 0.00328724 Altura + 0.182176 Etanol - 0.030292 Acido Lactico - 0.00207409 Peso elotes + 0.00192277 Peso jilotes	
R-cuadrada	R-ajustada	R-predicción
98.99%	98.96%	98.26%
Experimento	Batan CT-LN Biomasa	
Ecuaciones de regresión	Biomasa Fresca ^{0.5} = 0.00775786 Altura + 0.125606 Jilotes/planta + 0.00155075 Peso/tallo - 0.00334229 Sacarosa + 0.512114 Etanol + 0.040408 AAs + 0.00104665 Peso planta + 0.0292363 Nudos - 0.0585434 Acido Lactico	
R-cuadrada	R-ajustada	R-predicción
99.70%	99.69%	99.64%
Experimento	Tlalti CT-LN Biomasa	
Ecuaciones de regresión	Biomasa Fresca = 0.051337 Pesojilotes + 0.0738892 Peso tallo/seccion - 0.0352369 Sacarosa - 0.5837 Etanol + 1.70654 Acido Lactico + 0.0156569 Fructosa	
R-cuadrada	R-ajustada	R-predicción
99.07%	99.05%	98.89%
Regresion Paso a paso para selección de variables del modelo		
	α ingreso = 0.10 α salida = 0.15	
Regresion Multiple	α = 0.05	

6.3.6 Validación de las ecuaciones de regresión múltiple para predicción de rendimiento de biomasa fresca y grano.

Tabla 6.4 Ecuaciones de regresión múltiple validadas en la predicción del rendimiento de grano y biomasa fresca del experimento de El Batán en sus dos tratamientos. También se validó la ecuación de regresión para predecir el rendimiento de biomasa fresca del experimento de Tlaltizapán. Es interesante observar que las ecuaciones predicen individualmente en cada tratamiento y en la combinación de los dos.

Batan CT-LN Grano	Grano ^{0.286216} = 0.00328724 Altura + 0.182176 Etanol - 0.030292 Acido Láctico - 0.00207409 Peso elotes + 0.00192277 Peso jilotes													
Genotipo	Altura	Etanol	Acido Láctico	Peso elote	Peso jilote	Real	Predicción	Grano	%					0.9826
	0.0033	0.1822	0.0303	0.0021	0.0019									
G-Total	203.94	15.34	75.81	45.72	128.43	2.99	1.32	2.64	0.88					0.90
G-NN	227.22	15.39	75.69	57.18	173.85	4.53	1.47	3.88	0.86					0.87
G-LN	180.46	15.34	75.81	34.25	83.00	1.44	1.18	1.78	1.23					1.25
													1.0081	
Batan CT-LN Biomasa	Biomasa fresca ^{0.5} = 0.00775786 Altura + 0.125606 Peso jilotes + 0.00155075 Peso tallo - 0.00334229 Sacarosa + 0.512114 Etanol + 0.040408 AA + 0.00104665 Peso total + 0.0292363 Nudos - 0.0585434 Acido Láctico													
Genotipo	Altura	Peso elotes	Peso tallo	Sacarosa	Etanol	Amino Acidos	Peso total	Nudos	Acido Láctico	Real	Predicción	Biomasa	%	0.9964
	0.0078	0.1256	0.0016	0.0033	0.5121	0.0404	0.0010	0.0292	0.0585					
G-Total	203.94	1.48	439.93	77.47	15.34	4.30	706.13	9.75	75.81	47.31	6.81	46.31	0.98	0.98
G-NN	227.22	1.82	533.60	68.44	15.39	6.44	869.83	10.61	75.69	58.28	7.52	56.60	0.97	0.97
G-LN	180.46	1.12	345.45	86.58	15.28	2.14	541.00	8.88	75.93	36.25	6.08	36.98	1.02	1.02
													0.9937	
Tlalti CT-LN Biomasa	Biomasa fresca = 0.051337 Peso total tallo + 0.0738892 Peso tallo sección - 0.0352369 Sacarosa - 0.5837 Etanol + 1.70654 Acido Acético + 0.0156569 Fructosa													
Genotipo	Peso jilotes	Peso tallo/sec	Sacarosa	Etanol	Lactico	Fructosa	Real	Predicción	Biomasa	%				0.9964
	0.0513	0.0739	0.0352	0.5837	1.7065	0.0157								
G-Total	182.24	118.00	144.99	15.51	75.48	146.80	51.32	135.02	50.62	0.99				0.99
G-WW	227.03	131.43	134.78	15.53	75.51	176.13	57.32	139.17	51.86	0.90				0.91
G-WS	137.06	104.46	155.29	15.48	75.45	117.22	45.27	130.84	49.36	1.09				1.09
													0.9974	

7. DISCUSIÓN

La medición de metabolitos en estudios agronómicos de maíz apenas está siendo utilizada en trabajos de mejoramiento de plantas (Witt et al., 2012; Obata et al., 2015). La información que proporcionó el análisis fenotípico en nuestro proyecto va desde el nivel morfológico y fisiológico al estadístico, siendo la partición de la varianza en sus componentes ambientales y genotípicos (Leflon et al., 2005) una de las ventajas que pudimos comprobar mediante las comparaciones gráficas obtenidas con los scripts en R. Los programas que se escribieron para el procesamiento de la información probaron ser versátiles y confiables, así como su adaptabilidad a los requerimientos de análisis en el manejo de los datos de cada uno de los experimentos que se efectuaron. Sobre todo resaltamos el uso de software gratuito y libre (codigo abierto) y la posibilidad de modificar el script por parte del usuario (Garcia-Flores et al., 2012; Garcia-Flores et al., 2015). La correlación y la regresión lineal múltiple, permitieron identificar las relaciones entre los rasgos fenotípicos y metabólicos, su interacción e influencia entre ellos para establecer ecuaciones lineales de predicción del rendimiento de grano y biomasa en la cosecha a partir de los datos adquiridos en la etapa de floración de la planta (Obata et al., 2015).

En lo que se refiere a las características agronómicas, las mediciones se estandarizaron en unidades de cantidad, volumen, peso y longitud permitiendo establecer diferencias significativas. Todo esto condujo a determinar qué tratamiento experimental presentó un mejor desempeño en biomasa y grano bajo las condiciones de estrés impuestas (Bänziger et al., 2000; Monneveux et al., 2008; Araus Ortega et al., 2012). El fenotipo metabólico resultó ser un complemento a los datos agronómicos de los diferentes órganos y tejidos que se midieron en campo, enriqueciendo la información para dilucidar algunas funciones del tallo de la planta. Entre tales funciones fisiológicas está la fotosíntesis y la restricción de oxígeno en el tejido vascular, la acumulación transitoria de hexosas durante la floración, el almacenamiento intermediario de almidón del tallo, la afectación de las enzimas de invertasas vacuolares por el estrés, la regulación y partición de fotosintatos en las diferentes etapas de crecimiento de la planta incluyendo el llenado del grano (Bänziger y Lafitte, 1997; Paul y Driscoll, 1997; Sherson et al., 2003). Las modificaciones en las respuestas de las plantas, debido a su genotipo y tratamiento, son consecuencia de la plasticidad de las plantas de maíz. Cuando son sometidas a estrés adaptan su metabolismo y fisiología a las condiciones ambientales. De ahí que el *fingerprint* DIESI-MS del jugo de tallo no sea el mismo en todas las condiciones y etapas, sino que se modifica debido a cientos de señales e iones ms/z como se comprobó con múltiples figuras *boxplot* y *heatmap*.

El uso de espectrometría de masas con la técnica DIESI-MS es de bajo costo, tiene capacidad de manejar un elevado número de muestras y replicas por día. La inyección al equipo es directa, se utiliza poca cantidad de muestra y la muestra no requiere ser previamente fraccionada. La ionización se realiza a presión atmosférica y genera espectros del total de iones, tanto de iones desnudos como de aductos, conglomerados de moléculas y fragmentos de metabolitos. En modo positivo detecta un alto número de señales iónicas que permiten la clasificación del estado fisiológico de las plantas (Garcia-Flores et al., 2012; Garcia-Flores et al., 2015; Montero-Vargas et al., 2013). En modo negativo se detecta un número menor de señales, pero estas amplian la evaluación del estado fisiológico del tallo de maíz. El analisis de datos se puede automatizar una vez que se tiene montado el

script optimizado de R, facilitando el análisis de cientos de muestras por día. De forma automática se construyen miles de graficas boxplot y decenas de figuras heatmap-biclusters

Debido a que las ventajas de nuestra metodología DIESI-MS sobrepasan a las desventajas, es válido sugerir al fenotipo metabólico (*fingerprinting*) como una estrategia complementaria a los estudios de genómica, transcriptómica y proteómica con miras a la selección genética de germoplasma híbrido que permita potenciar el rendimiento agronómico y la calidad nutricional del Vitamaíz (García-Flores et al., 2012, 2015; Montero-Vargas et al., 2013).

7.1 Experimentos control

Los experimentos iniciales sirvieron para establecer los parámetros de calidad para el análisis de metabolitos y facilitar la comparación con los datos de otros grupos de investigación (Tiessen, 2009). El rango de detección, nivel de resolución, coeficiente de variación, exactitud, reproducibilidad y linealidad de los cuadrupolos sencillos (Z/Q y SQD) son índices primordiales que deberían ser siempre reportados (Douglas, 2009). Comparados con otros métodos y equipos de espectrometría de masas (p ej. cuadrupolos triples, la trampa lineal e inclusive el Orbitrap) los índices del DIESI-MS se consideran aceptables (Madalinski et al., 2008). Los experimentos de dilución, de absorbentes y las huellas metabólicas de aminoácidos y carbohidratos permitieron la optimización de la técnica DIESI-MS, del equipo y de los pipelines de ToppasMaize y ToppasView (Junker et al., 2012), y de los scripts en R para manejo y análisis de los archivos en formato de Excel (Tiessen, 2009). Si bien al inicio de la tesis teníamos que implementar varios programas de forma secuencial, al final del proyecto logramos simplificar el análisis informático. Por ejemplo, al inicio usábamos un programa especial para renombrar archivos (AntRenamer) y para apilar archivos de datos. Más tarde eliminamos la necesidad de usar esos programas incorporando algunas mejoras en el script de R para la importación automática de los datos a partir de diferentes archivos de texto sin tener que cambiarles el nombre.

El entrenamiento en los programas de análisis bioestadístico resultó un reto de aprendizaje por la complejidad de la información y la dificultad de entender las funciones y optimizar todos los parámetros. Sin embargo, con la guía de mis tutores logramos tener éxito en el establecimiento de la metodología de fenotipo bioquímico basada en DIESI-MS. Tan es así que no solo se generaron publicaciones internacionales y presentaciones en congresos, sino que también surgió una empresa *spin off* derivado del concurso ID4BIZ del Cinvestav.

7.2 Fenotipo

Los procedimientos para coleccionar plantas de maíz en el campo se llevaron a cabo de acuerdo a los estándares de trabajo de los campos experimentales del Cimmyt (Bänziger et al., 2000). Prestamos mucha atención y cuidado con la temperatura, la congelación y la conservación del jugo del tallo de maíz. También cuidamos los registros de los datos fenotípicos de la planta, y el traslado de las muestras al laboratorio en Irapuato. A partir de las alicuotas del jugo, y posterior a la medición de señales iónicas con DIESI-MS, mediante los scripts de Toppas (Junker et al., 2012; Montero-Vargas et al., 2013) se inspeccionaron visualmente las lecturas para identificar su huella metabólica. A partir de los archivos en formato

de texto realizamos análisis de varianza, heatmaps, dendogramas, gráficas box plots individuales y grupales de iones de interés combinando los factores experimentales (Tiessen, 2009). En el fenotipo gráfico observamos el cambio en los niveles de intensidad de las señales iónicas. A estas figuras heatmap y a los estos valores numéricos le llamamos ahora fenotipo, debido a que puede ser determinado de forma cualitativa y cuantitativa. Ya no nos limitamos solo al fenotipo macroscópico o morfológico, de altura de planta, color de la espiga o de rendimiento de grano, sino que estamos explorando el fenotipo bioquímico de las plantas de maíz bajo las condiciones del experimento en cuestión, favoreciendo una evaluación simultánea fisiológica y metabólica (García-Flores et al., 2012, 2015).

7.3 Selección

El proceso de selección de los mejores genotipos de maíz en los experimentos de estrés hídrico y bajo nitrógeno implicó la aplicación de componentes principales en R (R-Development-Core-Team, 2008). En base a dichos análisis se seleccionaron las variedades que tuvieron un mejor desempeño en los rendimientos de grano y de biomasa fresca. De acuerdo a la estrategia de selección del Cimmyt (Bänziger et al., 2000) que consiste en calificar a las líneas de maíz cultivadas en los ambientes sometidos a estrés, con los índices de componentes principales se listaron las mejores líneas; biomasa fresca, en el experimento de estrés hídrico: E-37, 147, 91, 99, 34, 104, 143, 117, 114, 87; en el experimento de bajo contenido de nitrógeno: G-3, 1, 7, 9, 6, 10, 11, 15, 8; grano: en el experimento de bajo contenido de nitrógeno, maíz amarillo: G-15, 18, 14, 17, 19, 16; maíz blanco: G-7, 5, 4, 9, 3, 10, 2, 11. Estas líneas mejoradas se recomendaron al grupo de trabajo en fisiología del Cimmyt para que las utilicen en sus siguientes ciclos de siembra. Es necesario enfatizar que para los programas de mejoramiento, lo más importante es la evaluación y la selección de las plantas con rasgos superiores. Los genotipos para ambos experimentos reportaron mayores rendimientos de biomasa y de grano, a partir de sus características metabólicas en el jugo del tallo. Cabe mencionar que el CIMMYT no se había hecho selección en base a la composición bioquímica del tallo, por lo que nuestra estrategia fue a la vez novedosa y útil para MasAgro y Sagarpa.

7.4 Ecuaciones lineales de predicción

El diseño de modelos de predicción (Acuña, 2011) del desempeño agronómico de plantas representa una oportunidad de comprobar el nivel de correlación e interacción, positiva o negativa, entre las características fenotípicas (altura de planta, peso de jilotes, peso de planta, etc.) con los metabolitos obtenidos con espectrometría de luz UV-VIS (glucosa, fructosa, sacarosa, almidón, aminoácidos), con cromatografía de gases (etanol, ácido láctico, ácido acético) u otra plataforma analítica de espectrometría de masas (García-Flores et al., 2012; Montero-Vargas et al., 2013). Del análisis de las tablas de correlación inferimos las relaciones entre los metabolitos y las características fenotípicas incluidas en el programa de mejora de la planta como son la fructosa y los aminoácidos con el peso total de la planta en el experimento de estrés hídrico o el peso total de la planta con la fructosa en el experimento de bajo contenido de nitrógeno. Combinando las características fenotípicas con los datos bioquímicos en un análisis de regresión múltiple, se requiere seleccionar en forma previa los factores

significativos que ajusten mejor en el modelo predictivo; lo anterior se alcanzó con la metodología estadística de selección de factores conocida como “paso a paso” (*step wise*). El programa genera una tabla de varianza con los valores de significancia ($\alpha = 0.05$) de los factores y con dicha lista se obtuvieron las ecuaciones de rendimiento de grano y de biomasa fresca. Las ecuaciones incluyen el valor de *lambda* de la variable de respuesta, los coeficientes de las variables y los signos de los términos del modelo. Toda vez que la validación mostró niveles de predicción del orden de 99 %, para los tratamientos individuales y el experimento global, podemos concluir que la metodología “paso a paso” fue eficiente para la selección de factores. Se puede modelar una ecuación de predicción para una variable de respuesta definida.

7.5 Experimento de estrés hídrico

La sequía es una de las principales causas de la pérdida de cultivos en todo el mundo, con una disminución drástica en los rendimientos (Bänziger et al., 2000). La mayoría de los efectos perjudiciales de la sequía están asociados a la reducción de la fotosíntesis en las hojas verdes y el crecimiento de tejidos reproductivos (alargamiento de estigmas y la aborción de embriones en el tallo floral) (Paul y Driscoll, 1997). Bajo sequía, algunas de las funciones fisiológicas del tallo vegetativo de la planta de maíz que son alteradas son: la fotosíntesis por la afectación de las enzimas por el estrés, la disponibilidad de oxígeno y el cambio en los metabolitos fermentativos (etanol y ácido láctico), el papel de acumulación de azúcares a partir de la floración, la regulación y partición de fotosintatos en las diferentes etapas de crecimiento de la planta incluyendo de manera importante al llenado del grano (Bänziger y Lafitte, 1997; Paul y Driscoll, 1997; Sherson et al., 2003). La variación en la intensidad de los iones responsables del ajuste osmótico podría explicarse como parte de un mecanismo de las células que facilitan la obtención constante de agua para mantener la turgencia y la expansión celular (Sherson et al., 2003). La identidad química de estos iones discriminantes todavía faltan por identificar con alguna de las técnicas avanzadas de espectrometría de masas con múltiple fragmentación o con el uso de equipos de RMN de alta resolución.

En experimentos metabolómicos dirigidos, algunos aminoácidos se acumularon en hojas (e. g. Trp, Asn, β -Ala, Phe, Thr, Ala, Pro, Val, Tyr, GABA, piro-Glu, Ile, Ser, Gly y homo-Ser); otros metabolitos decrecieron (e. g. eritritol, maltitol, maltosa, trehalosa, mioinositol malato, glicerato, treonato y fumarato) y algunos tienen una pequeña respuesta (e. g. xilosa, rafinosa y putrescina) (Obata et al., 2015). Además, el galactinol es un metabolito que por sus reducidos niveles reportados en algunas líneas de maíz se asocia a tolerancia a la sequía (Obata et al., 2015). También pudiese ser que los iones asociados a carbohidratos (e. g. sacarosa, glucosa y fructosa) aumentaran su intensidad por un proceso de acumulación temporal de metabolitos en el tallo como resultado de la degradación de almidón, para ser posteriormente transportado al tejido u órgano que lo metabolice (Paul y Foyer, 2001; Tiessen, 2009). Las variaciones en la intensidad de los iones podrían ser atribuidas a una nueva movilización de metabolitos en órganos sumidero (e. g. elotes), lo que significa que un órgano incrementa su fuerza sumidero (e. g. en la etapa de floración y llenado de grano) (Juárez, 2012). Esto apoya nuestra hipótesis que propone que el tallo tiene un metabolismo muy activo que puede cambiar dramáticamente en condiciones de sequía, y que estos cambios dependen del

genotipo y posiblemente también a la familia de genes de las invertasas (Pelleschi et al., 1997; Juárez, 2012).

7.6 Experimento de bajo contenido de nitrógeno

El estrés por nitrógeno disminuye la cantidad de clorofila y de proteína Rubisco y por consiguiente la fotosíntesis de las plantas. Reduce el desarrollo del área foliar y la tasa de fotosíntesis de la hoja acelerando la senescencia foliar. Alrededor del 50% de N de toda la planta está directamente involucrado en la fotosíntesis, ya sea como enzimas o como clorofila (Bänziger et al., 2000). En consecuencia, el crecimiento de la planta es menor ya que las funciones de reproducción celular y la síntesis de proteínas requieren en gran medida del N (Paul y Driscoll, 1997). Cuando el N disminuye sus niveles de disponibilidad en el suelo, las plantas remobilizan N a partir de tejido más viejo (hoja y/o tallo) a los tejidos más jóvenes (e. g., hojas en expansión), o reproductivos (e. g., granos), que conducen a la senescencia temprana del tejido más viejo y de las hojas más pequeñas. La degradación de enzimas y la interrupción de síntesis de proteínas aumentan la disponibilidad de moléculas nitrogenada(o)s. Juárez (2012), encontró en su experimento de estrés hídrico que la cantidad de aminoácidos libres en el jugo de tallo de maíz es mayor al igual que la acumulación de sacarosa, concluyendo que el estrés hídrico y el bajo contenido de nitrógeno son estreses que afectaron al desarrollo de la planta y que los procesos metabólicos están mutuamente relacionados, sobre todo en el ajuste osmótico de las células. La hipótesis es fortalecida que propone que el tallo tiene un metabolismo muy activo y que puede tener variaciones importantes en los niveles de intensidad de los iones en DIESI-MS en condiciones de bajo contenido nitrógeno. Se establece que dichas variaciones dependen del ambiente. Las características fisiológicas secundarias pueden aumentar la eficiencia de la selección para rendimiento de grano en los programas de mejoramiento de maíz dirigidos a ambientes de bajo contenido de nitrógeno (Bänziger y Lafitte, 1997).

7.7 Experimento fuente-sumidero

Los órganos fuente y sumidero, aunque estén alejados morfológicamente, tienen una relación de dependencia mutua. Pueden interactuar con el genotipo y el medio ambiente para definir qué nivel de metabolitos son sintetizados y remobilizados de una hoja fuente a un tejido sumidero, ajustando el nivel de partición entre los órganos productores y consumidores (Paul y Driscoll, 1997). Las alteraciones experimentales en las relaciones entre tejidos productores-consumidores afectan el nivel de rendimiento en biomasa seca y apoyaron la hipótesis sobre la funcionalidad del tallo como regulador de los metabolitos dirigidos hacia el tejido consumidor (Juárez, 2012). Con el enfoque no dirigido del estudio de metabolitos con DIESI-MS, se observan cambios en el nivel de intensidad de ciertos iones. Fue posible encontrar cambios en la relación fuente-sumidero dependientes del tiempo y del tratamiento (WE, WL, WM).

7.8 Experimento etapas de desarrollo

La finalidad de este conjunto de experimentos fue dilucidar la función de las hojas y de la vaina en su relación con el tallo. Los genotipos evaluados fueron los mismos en 2010 y 2011; el sitio de siembra fue dentro y fuera del invernadero, los genotipos, el tratamiento con o sin vaina, la repetición biológica y los días de colecta fueron los factores en el experimento (Bänziger et al., 2000; Juárez, 2012).

Con DIESI-MS (Montero-Vargas et al., 2013) se establecieron patrones y tendencias de variación en los iones de acuerdo a esos factores.

Del análisis de varianza en la metodología DIESI-MS se observó que el factor tratamiento con y sin vaina si fue significativo.

Los cambios más dramáticos en el contenido de carbohidratos no estructurales ocurren alrededor de la etapa de floración (Bänziger et al., 2000). Las hexosas presentaron un pico transitorio que coincidió con dicha etapa (Juárez, 2012). Algunos metabolitos tuvieron los más altos niveles en las primeras etapas, y más tarde se redujeron a niveles que eran 50% inferiores a los detectados en las etapas vegetativas anteriores. El metabolismo en el tallo de maíz es muy activo, pero el tallo también tiene un papel en el almacenamiento de carbohidratos como una fuente de esqueletos de carbono utilizados en etapas reproductivas de otros órganos o tejidos (Juárez, 2012). Con DIESI-MS se establecieron patrones de variación en los iones de acuerdo al genotipo, a la fecha de siembra y a la fecha de colecta. Del análisis de varianza en R observamos que la interacción de la fecha de siembra con la repetición biológica no son significativas por lo que se propone aumentar las fechas de siembra y las repeticiones biológicas en futuros experimentos. Estadísticamente no se debe construir un boxplot con dos datos sino al menos con cinco.

7.10 Oportunidades en fisiología de plantas

La combinación de datos fisiológicos y metabólicos abren nuevas oportunidades por la abundancia de la información disponible (García-Flores et al., 2012; Witt et al., 2012; Obata et al., 2015). El contraste de los resultados experimentales de las plantas de maíz bajo estrés hídrico y bajo contenido de N indican diferencias importantes en características fisiológicas como la altura y peso de la planta, peso de jilotes y peso de tallo. También hay diferencias en el rendimiento de grano y biomasa fresca y en el contenido de aminoácidos y carbohidratos no estructurales (Juárez, 2012; Witt et al., 2012; Obata et al., 2015). Las diferencias indican que las plantas en ambientes con bajo contenido de N, cultivadas en invernadero o en campo, se ven más afectadas en los rendimientos agronómicos. Las plantas con restricción de N detienen su crecimiento, se ponen amarillas y redirigen sus recursos energéticos a funciones reproductivas, que favorecieron la floración, redujeron la fotosíntesis y el contenido de clorofila en las hojas, produciendo la degradación enzimática en general y acelerando la senescencia de las plantas (Jones y Simmons, 1983; Bänziger et al., 2000). La correlación que se observó entre los metabolitos y las características fisiológicas de la planta nos dan idea de que las relaciones metabólicas entre tejidos y órganos son mutuamente dependientes (Paul y Driscoll, 1997). Su grado de dependencia nos hace preguntarnos sobre la naturaleza fisiológica que permite la afectación en el funcionamiento metabólico de un órgano por otro, en diferentes etapas y condiciones del ambiente.

Se ha propuesto la manipulación biotecnológica dirigida para producir el efecto favorable en el tejido potenciador de características deseables en las plantas (Jones y Simmons, 1983). Sin embargo, la biotecnología moderna es mucho más amplia que el solo uso de plantas transgénicas como lo proponen las empresas transnacionales. Se pueden lograr muchos avances genéticos con cruza sexuales y selección fenotípica y metabólica como lo demostramos en esta tesis. No es necesario recurrir a la transgénesis si se optimiza el manejo agronómico de forma

adecuada para la zona. Por ejemplo, tiene más efecto el usar fertiriego por goteo con soluciones nitrogenadas que el uso de maíz transgénico BT resistente a herbicidas.

Los transgénicos, más que ser útiles para el medio ambiente es una herramienta para el control comercial monopólico para las empresas extranjeras. De ahí que se utilicen patentes que originalmente era para inventos de objetos y máquinas y ahora los americanos esten patentando el maíz. Es cuestionable de que sea ético y moralmente correcto patentar seres vivos superiores. Para los gringos el maíz es un objeto, pero para los mexicanos el maíz es una planta mágica con vida y dignidad propia. Los transgénicos no incrementan el rendimiento de maíz ni tampoco disminuyen la aplicación de plaguicidas, al contrario, lo incrementan. Esto es aún más cierto para México donde somos centro de origen del maíz y el rendimiento de grano y de forraje depende más del ambiente que del genotipo. En nuestras gráficas *heatmap-bicluster*, demostramos por medio de los agrupamientos jerárquicos y los dendogramas que el ambiente tiene un efecto mucho más grande que las diferentes variedades de maíz utilizadas, ya sea de grano blanco y amarillo. De ahí radica la importancia del *fingerprinting* y de la metabolómica vegetal para la agricultura y el beneficio del campo, para la caracterización y la preservación de los maíces nativos mexicanos sin extinguirlos ni contaminarlos (con transgenes o pesticidas). De ahí también la utilidad del DIESI-MS para distinguir cultivos orgánicos de los demás.

Las plantas en sus etapas de desarrollo siempre experimentan algún tipo de estrés por lo que es bueno considerar el diseño de experimentos en el que se apliquen diferentes tipos de estrés como sequía y calor porque las plantas lo reciben de manera simultánea (Bänziger et al., 2000).

Las herramientas bioinformáticas y estadísticas permiten dilucidar los efectos de los factores que se incluyen en un experimento. Es necesario determinar si los efectos son aditivos o multiplicativos para identificar los metabolitos que puedan ser utilizados como marcadores metabólicos de un estado fisiológico de las plantas con fines de hibridación (Roessner et al., 2001; Leflon et al., 2005; R-Development-Core-Team, 2008). La regresión lineal múltiple permite encontrar las relaciones significativas entre factores metabólicos y los fisiológicos, en un mismo tratamiento o entre ellos, con objeto de predecir rendimientos de grano y biomasa (Castaño y Domínguez, 2003; Montgomery, 2008). Lo ideal es obtener biomarcadores que puedan ser usados para selección. Con los datos DIESI-MS se podrán evaluar ecuaciones de regresión múltiple para modelar los eventos fisiológicos en otras plantas modelo (Tiessen, 2009). La metodología DIESI-MS es muy prometedora y vale la pena seguir ahondando en su perfeccionamiento para aplicaciones agrícolas dentro de la biotecnología vegetal, para beneficio del ser humano y del medio ambiente, en atención a problemas nacionales y expandiendo nuestras fronteras de conocimientos.

8. CONCLUSIONES

Los resultados de la presente tesis permiten derivar las siguientes conclusiones principales:

- a) La metabolómica vegetal no se logra con una sola técnica analítica, sino que surge de la combinación de varias plataformas y métodos, incluyendo UV-VIS, HPLC, GC-MS, LC-MS y RMN. La técnica DIESI-MS tiene una utilidad para el fingerprinting metabólico. Puede ser útil para la biotecnología agrícola ya que permite una clasificación rápida y automática de muestras usando criterios estadísticos (Garcia-Flores et al., 2015).
- b) El protocolo optimizado DIESI-MS permite obtener datos de más de 400 iones diferentes en el rango de 50 a 1200 m/z (Garcia-Flores et al., 2012). Los iones positivos tienen una intensidad más abundante que los iones negativos.
- c) El solvente que se utiliza y la dilución de la muestra determina el *fingerprint* iónico que se observa en DIESI-MS. No siempre hay una relación lineal entre el metabolito en la muestra y la intensidad del ión (no publicado).
- d) Las figuras Heatmap-Bicluster hechas con scripts optimizados de R son muy útiles para organizar y visualizar la gran cantidad de datos DIESI-MS (Garcia-Flores et al., 2012; Montero-Vargas et al., 2012; Garcia-Flores et al., 2015).
- e) El perfil metabolómico obtenido con DIESI-MS de las muestras de jugo del tallo de maíz en floración permitieron evaluar el estado fisiológico de las plantas (Garcia-Flores et al., 2012).
- f) Los factores agronómicos y bioquímicos medidos en el jugo del tallo de maíz en la etapa temprana de floración permiten predecir el rendimiento final de forraje y de grano (Garcia-Flores et al., 2015).
- g) La metodología de fenotipo bioquímico usando DIESI-MS genera una gran cantidad de información, la cual tiene mejor relación de costo/ beneficio que otras técnicas analíticas basadas en cromatografía (Garcia-Flores et al., 2015).
- h) Los datos generados por DIESI-MS, como el *fingerprinting*, permiten fraccionar la varianza fenotípica observada de la intensidad de cada ion m/z

en efectos genéticos (G), efectos ambientales (E) o su interacción (G x E) (García-Flores et al., 2015). A pesar de la complejidad de la muestra, la supresión de iones y la formación de aductos y conglomerados moleculares, la información fenotípica que se obtiene con DIESI-MS es útil para el mejoramiento y el manejo agronómico.

- i) Una limitante del método DIESI-MS es que no es posible asignarle un nombre bioquímico a un ion específico, por lo que hemos decidido nombrar el ion solamente con un valor numérico de su masa carga (m/z). El hecho de que un ion tenga un valor de m/z que corresponda al valor teórico de un compuesto de interés, no significa que realmente sea ese compuesto. Para identificar los iones se requiere un trabajo de doctorado extra. Para determinar la naturaleza química de un ion discriminante se debe aplicar una batería de análisis bioquímicos adicionales basados en ensayos enzimáticos específicos.

PERSPECTIVAS DE LA TESIS:

- a) La gran cantidad de información cuantitativa que se genera por medio de la metodología optimizada de DIESI-MS para el fenotipo bioquímico permitió plantear un posible proyecto de aplicación a la industria agroalimentaria relacionado con la distinción entre tomates orgánicos de tomates no orgánicos (Tesis doctoral Adrián García, 2017). Esta tecnología permitiría una certificación de alimentos en base a análisis de producto final y no sólo basada en proceso, visitas de campo y registros de bitácora, como es común hasta ahora. Un grupo de compañeros (Adrián García, Josaphat Montero y Martín García supervisados por el Dr Robert Winkler y el Dr Axel Tiessen) decidimos someter la idea de proyecto en 2014 al concurso de Cinvestav ID4Biz. Después de ser evaluada por un panel de expertos de diferentes áreas, la idea recibió el premio principal para ser apoyada con un monto de 500 mil pesos. Derivado de eso, los participantes recibimos asesoría, capacitación y entrenamiento en diferentes temas de los bionegocios, emprendimiento, innovación y leyes. La propiedad intelectual de la tecnología DIESI-MS quedo en manos de Cinvestav y de los profesores responsables. Cabe mencionar que la capacitación que recibieron mis compañeros en Bionegocios sirvió entonces de base para que fundarán una nueva bio-empresa con el nombre comercial de INFOOD®.
- b) El presente proyecto de doctorado demuestra que los algoritmos optimizados para generar *heatmaps* con escala de grises y con agrupamiento jerárquico a partir de los datos DIESI-MS son muy útiles para clasificar las muestras según el estado fisiológico. Esto permite usar los *heatmaps biclusters* para programas de mejoramiento genético de especies vegetales como el maíz, en el caso del proyecto Vitamaíz (Tiessen et al., 2017). Sin embargo, hay que cuidar los parámetros que se utilizan para construir los dendogramas, ya sea en base a distancia euclídeana, valores de correlación R o R² (Tiessen et al., 2017).
- c) Los experimentos control presentados en la tesis demuestran que el comportamiento de los iones en la fuente de ionización ESI es en muchos casos no lineal. Aún no se sabe la razón de esta complejidad. Es posible que no sólo sean iones desnudos los que transitan en el espectrómetro de masas, sino que sean más bien aductos y conglomerados de moléculas los que permanecen juntos en el alto vacío y se aceleran por el campo electromagnético hasta llegar al detector. Una perspectiva importante de la presente tesis es estudiar más a fondo los fenómenos de ionización en modalidad DIESI-MS usando fragmentación MSⁿ con el triple cuádrupolo (TSQ2).
- d) Otra perspectiva de la tesis es usar otras tecnologías de metabolómica como resonancia magnética nuclear (RMN) para comparar los resultados de metabolitos y de agrupamientos jerárquicos con DIESI-MS y con RMN.

- e) Otra perspectiva es usar espectrómetros con más precisión de masa como el Orbitrap, para tratar de identificar molecularmente los iones más discriminantes entre el tratamiento de sequía (WS) o bajo contenido de nitrógeno (LN), y con eso obtener un panorama más amplio que permita explorar las rutas metabólicas y pasos enzimáticos que están regulados en maíz en dependencia del medio ambiente (estrés abiótico).

- f) Otra perspectiva global, es promover el uso de la metabolómica para aplicaciones agrícolas dentro del Laboratorio Nacional PlanTECC. El fenotipo bioquímico usando espectrometría de masas y DIESI-MS necesita demostrar con casos concretos (ciencia aplicada) de que es útil y confiable para las empresas. Se requiere convencer a autoridades federales (SAGARPA, SNITT, COFUPRO, CESAVER, COFEPRIS, CONACYT) de que la metabolómica es importante para el desarrollo agrícola sustentable de nuestro país. Es por ello que sería deseable que más estudiantes, técnicos agrónomos, maestros, licenciados y políticos pudieran tener el privilegio de formarse académicamente dentro de CINVESTAV Irapuato, para que en un futuro tengamos más emprendedores y candidatos diputados que entiendan los alcances de la ciencias ómicas para el beneficio del ser humano y del ambiente. La perspectiva por lo tanto es lograr que la metabolómica y la biotecnología vegetal de CINVESTAV Irapuato tengan un impacto más allá de las revistas científicas en el desarrollo del país.

9. BIBLIOGRAFÍA

- Acuña E** (2011) Análisis de regresión. Departamento de Matemáticas, Universidad de Puerto Rico
- Adams WW, 3rd, Muller O, Cohu CM, Demmig-Adams B** (2013) May photoinhibition be a consequence, rather than a cause, of limited plant productivity? *Photosynthesis Research* **117**: 31-44
- Agilent T** (2008) Inyectores automáticos preparativos y estándar Agilent serie 1200. *In* A Technologies., ed, Manual del usuario, Alemania
- Angeles-Nunez JG, Tiessen A** (2010) Arabidopsis sucrose synthase 2 and 3 modulate metabolic homeostasis and direct carbon towards starch synthesis in developing seeds. *Planta* **232**: 701-718
- Araujo WL, Nunes-Nesi A, Nikoloski Z, Sweetlove LJ, Fernie AR** (2012) Metabolic control and regulation of the tricarboxylic acid cycle in photosynthetic and heterotrophic plant tissues. *Plant, Cell and Environment* **35**: 1-21
- Araus Ortega JL, Serret Molins MD, Edmeades GO** (2012) Phenotyping maize for adaptation to drought. *Frontiers in Physiology* **305**:1-20
- Banerjee S, Prakash H, Mazumdar S** (2011) Evidence of molecular fragmentation inside the charged droplets produced by electrospray process. *Journal of the American Society for Mass Spectrometry* **22**: 1707-1717
- Bänziger M, Betran F, Lafitte H** (1997) Efficiency of high-nitrogen selection environments for improving maize for low-nitrogen target environments. *Crop Science* **37**: 1103-1109
- Bänziger M, Edmeades G, Beck D, Bellon M** (2000) Breeding for drought and nitrogen stress tolerance in maize: from theory to practice. *In*, Vol 1. International Maize and Wheat Improvement Center (CIMMYT), México, p 68
- Bänziger M, Edmeades G, Lafitte H** (1999) Selection for drought tolerance increases maize yields across a range of nitrogen levels. *Crop Science* **39**: 1035-1040
- Bänziger M, Edmeades G, Lafitte H** (2002) Physiological mechanisms contributing to the increased N stress tolerance of tropical maize selected for drought tolerance. *Field Crops Research* **75**: 223-233
- Bänziger M, Lafitte H** (1997) Efficiency of secondary traits for improving maize for low-nitrogen target environments. *Crop Science* **37**: 1110-1117
- Beecher WWC** (2003) The Human Metabolome. *In*: GGG Harrigan, Royston (Eds.), ed, *Metabolic Profiling: Its Role in Biomarker Discovery and Gene Function Analysis*. Kluwer Academic Publishers, Norwell, Mass
- Butelli E, Titta L, Giorgio M, Mock H-P, Matros A, Peterrek S, Schijlen EGWM, Hall RD, Bovy AG, Luo J, Martin C** (2008) Enrichment of tomato fruit with health-promoting anthocyanins by expression of select transcription factors. *Nature Biotechnology* **26**: 1301-1308
- Castaño TE, Domínguez DJ** (2003) Experimentos: Estrategia y análisis en ciencia y tecnología, Vol 1. Centro de Investigaciones en Matemáticas, México
- Chai W, Piskarev V, Lawson AM** (2002) Branching pattern and sequence analysis of underivatized oligosaccharides by combined MS/MS of singly and doubly charged molecular ions in negative-ion electrospray mass spectrometry. *Journal of the American Society for Mass Spectrometry* **13**: 670-679

- Douglas D** (2009) Linear quadrupoles in mass spectrometry. *Mass spectrometry reviews* **28**: 937-960
- Dunn WB** (2008) Current trends and future requirements for the mass spectrometric investigation of microbial, mammalian and plant metabolomes. *Physical Biology* **5**: 011001
- Enfissi EMA, Fraser PD, Lois LM, Boronat A, Schuch W, Bramley PM** (2005) Metabolic engineering of the mevalonate and non-mevalonate isopentenyl diphosphate-forming pathways for the production of health-promoting isoprenoids in tomato. *Plant Biotechnology Journal* **3**: 17-27
- Fernie AR, Schauer N** (2009) Metabolomics-assisted breeding: a viable option for crop improvement? *Trends in Genetics* **25**: 39-48
- Fiehn O** (2001) Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and Functional Genomics* **2**: 155-168
- Fraser PD, Romer S, Shipton CA, Mills PB, Kiano JW, Misawa N, Drake RG, Schuch W, Bramley PM** (2002) Evaluation of transgenic tomato plants expressing an additional phytoene synthase in a fruit-specific manner. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 1092-1097
- Garcia-Flores M, Juarez-Colunga S Fau - Garcia-Casarrubias A, Garcia-Casarrubias A Fau - Trachsel S, Trachsel S Fau - Winkler R, Winkler R Fau - Tiessen A, Tiessen A** (2015) Metabolic profiling of plant extracts using direct-injection electrospray ionization mass spectrometry allows for high-throughput phenotypic characterization according to genetic and environmental effects. *Journal of Agricultural and Food Chemistry* **3**: 1042-1052
- Garcia-Flores M, Juarez-Colunga S, Montero-Vargas JM, Lopez-Arciniega JA, Chagolla A, Tiessen A, Winkler R** (2012) Evaluating the physiological state of maize (*Zea mays* L.) plants by direct-injection electrospray mass spectrometry (DIESI-MS). *Molecular BioSystems* **8**: 1658-1660
- Geigenberger P, Stitt M, Fernie A** (2004) Metabolic control analysis and regulation of the conversion of sucrose to starch in growing potato tubers. *Plant, Cell and Environment* **27**: 655-673
- Geigenberger P, Stitt M, Fernie AR** (2004) Metabolic control analysis and regulation of the conversion of sucrose to starch in growing potato tubers. *Plant Cell and Environment* **27**: 655-673
- Go EP** (2010) Database Resources in Metabolomics: An Overview. *Journal of Neuroimmune Pharmacology* **5**: 18-30
- Hall RD** (2011) Plant Metabolomics in a Nutshell: Potential and Future Challenges. *In* BP Ltd., ed, *Annual Plant Reviews. Biology of plants Metabolomics*, Vol 43, pp 1-24
- Jones R, Simmons S** (1983) Effect of altered source-sink ratio on growth of maize kernels. *Crop Science* **23**: 129-134
- Juárez CS** (2012) El tallo de maíz: ¿un simple soporte conductor, o un reservorio dinámico de carbohidratos? *Cinvestav, Irapuato, México*
- Junker J, Bielow C, Bertsch A, Sturm M, Reinert K, Kohlbacher O** (2012) TOPPAS: a graphical workflow editor for the analysis of high-throughput proteomics data. *Journal of Proteome Research* **11**: 3914-3920
- Knoblauch M, Peters WS** (2013) Long-distance translocation of photosynthates: a primer. *Photosynthesis Research* **117**: 189-196

- Koek MM, Jellema RH, van der Greef J, Tas AC, Hankemeier T** (2011) Quantitative metabolomics based on gas chromatography mass spectrometry: status and perspectives. *Metabolomics* **7**: 307-328
- Lafitte H, Edmeades G, Taba S** (1997) Adaptive strategies identified among tropical maize landraces for nitrogen-limited environments. *Field Crops Research* **49**: 187-204
- Leflon M, Lecomte C, Barbottin A, Jeuffroy M, Robert N, Brancourt-Hulmel M** (2005) Characterization of Environments and Genotypes for Analyzing Genotypex Environment Interaction: Some Recent Advances in Winter Wheat and Prospects for QTL Detection. *Journal of Crop Improvement* **14**: 249-298
- Liu Z, Rochfort S** (2014) Recent progress in polar metabolite quantification in plants using liquid chromatography–mass spectrometry. *Journal of Integrative Plant Biology* **56**: 816-825
- Lopes MS, Araus JL, Van Heerden PD, Foyer CH** (2011) Enhancing drought tolerance in C4 crops. *Journal of Experimental Botany* **62**: 3135-3153
- Madalinski G, Godat E, Alves S, Lesage D, Genin E, Levi P, Labarre J, Tabet JC, Ezan E, Junot C** (2008) Direct introduction of biological samples into a LTQ-Orbitrap hybrid mass spectrometer as a tool for fast metabolome analysis. *Analytical Chemistry* **80**: 3291-3303
- Martín GMCaBGM** (2012) Espectrometría de masas y análisis de biomarcadores. *Real Academia Nacional de Farmacia* **1**:113-168
- Memelink J, Gantet P** (2007) Transcription factors involved in terpenoid indole alkaloid biosynthesis in *Catharanthus roseus*. *Phytochemistry Reviews* **6**: 353-362
- Monneveux P, Sanchez C, Tiessen A** (2008) Future progress in drought tolerance in maize needs new secondary traits and cross combinations. *The Journal of Agricultural Science* **146**: 287-300
- Montero-Vargas JM, González-González LH, Gálvez-Ponce E, Ramírez-Chávez E, Molina-Torres J, Chagolla A, Montagnon C, Winkler R** (2013) Metabolic phenotyping for the classification of coffee trees and the exploration of selection markers. *Molecular BioSystems* **9**: 693-699
- Montgomery D** (2009) *Design and Analysis of Experiments*, Ed 7th Vol 1. John Wiley & sons, Inc
- Naqvi S, Zhu C, Farre G, Ramessar K, Bassie L, Breitenbach J, Perez Conesa D, Ros G, Sandmann G, Capell T, Christou P** (2009) Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 7762-7767
- Nielsen J** (2003) It is all about metabolic fluxes. *Journal of Bacteriology* **185**: 7031-7035
- Obata T, Witt S, Lisek J, Palacios-Rojas N, Florez-Sarasa I, Yousfi S, Araus JL, Cairns JE, Fernie AR** (2015) Metabolite Profiles of Maize Leaves in Drought, Heat, and Combined Stress Field Trials Reveal the Relationship between Metabolism and Grain Yield. *Plant Physiology* **169**: 2665-2683
- Paponov IA, Sambo P, Erley GS, Presterl T, Geiger HH, Engels C** (2005) Grain yield and kernel weight of two maize genotypes differing in nitrogen use efficiency at various levels of nitrogen and carbohydrate availability during flowering and grain filling. *Plant and Soil* **272**: 111-123

- Paul MJ, Driscoll SP** (1997) Sugar repression of photosynthesis: the role of carbohydrates in signalling nitrogen deficiency through source:sink imbalance. *Plant, Cell and Environment* **20**: 110-116
- Paul MJ, Foyer CH** (2001) Sink regulation of photosynthesis. *Journal of experimental botany* **52**: 1383-1400
- Pelleschi S, ROCHER JP, PRIOUL JL** (1997) Effect of water restriction on carbohydrate metabolism and photosynthesis in mature maize leaves. *Plant, Cell and Environment* **20**: 493-503
- R-Development-Core-Team** (2008) R: A Language and Environment for Statistical Computing;. *In. R Foundation for Statistical Computing Vienna, Austria.* <https://www.r-project.org/>
- Raines CA** (2003) The Calvin cycle revisited. *Photosynthesis Research* **75**: 1-10
- Rivoal J, Hanson AD** (1994) Metabolic Control of Anaerobic Glycolysis (Overexpression of Lactate Dehydrogenase in Transgenic Tomato Roots Supports the Davies-Roberts Hypothesis and Points to a Critical Role for Lactate Secretion. *Plant Physiology* **106**: 1179-1185
- Roessner U, Luedemann A, Brust D, Fiehn O, Linke T, Willmitzer L, Fernie AR** (2001) Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *The Plant Cell* **13**: 11-29
- Rolland F, Moore B, Sheen J** (2002) Sugar sensing and signaling in plants. *The Plant Cell* **14**: 185-205
- Saglio PH, Raymond P, Pradet A** (1983) Oxygen Transport and Root Respiration of Maize Seedlings A Quantitative Approach Using the Correlation between ATP/ADP and the Respiration Rate Controlled by Oxygen Tension. *Plant Physiology* **72**: 1035-1039
- Schwinn K, Venail J, Shang YJ, Mackay S, Alm V, Butelli E, Oyama R, Bailey P, Davies K, Martin C** (2006) A small family of MYB-regulatory genes controls floral pigmentation intensity and patterning in the genus *Antirrhinum*. *Plant Cell* **18**: 831-851
- Sherson SM, Alford HL, Forbes SM, Wallace G, Smith SM** (2003) Roles of cell-wall invertases and monosaccharide transporters in the growth and development of Arabidopsis. *Journal of Experimental Botany* **54**: 525-531
- Slewiniski TL, Meeley R, Braun DM** (2009) Sucrose transporter1 functions in phloem loading in maize leaves. *Journal of Experimental Botany* **60**: 881-892
- Smith AM, Stitt M** (2007) Coordination of carbon supply and plant growth. *Plant Cell and Environment* **30**: 1126-1149
- Spielbauer G, Armstrong P, Baier JW, Allen WB, Richardson K, Shen B, Settles AM** (2009) High-throughput near-infrared reflectance spectroscopy for predicting quantitative and qualitative composition phenotypes of individual maize kernels. *Cereal chemistry* **86**: 556-564
- Stitt M, Sulpice R, Keurentjes J** (2010) Metabolic Networks: How to Identify Key Components in the Regulation of Metabolism and Growth. *Plant Physiology* **152**: 428-444
- Sturm M, Bertsch A, Groepl C, Hildebrandt A, Hussong R, Lange E, Pfeifer N, Schulz-Trieglaff O, Zerck A, Reinert K, Kohlbacher O** (2008) OpenMS-An open-source software framework for mass spectrometry. *Bmc Bioinformatics* **9**: 1-11
- Sweetlove LJ, R.G.** (2011) Flux-balance modeling of plant metabolism. *Frontiers in Plant Science* **2**: 1-10

- Tiessen A** (2009) Fundamentos y metodologías innovadoras para el mejoramiento genético del maíz. Fundación ciencia activa, México. https://www.google.com.mx/?gfe_rd=cr&ei=NWhHV4-9G4XYqAPWm6nQDg#q=Fundamentos+y+metodolog%C3%ADas+innovadoras+para+el+mejoramiento+gen%C3%A9tico+del+ma%C3%ADz
- Tiessen A, Hendriks JH, Stitt M, Branscheid A, Gibon Y, Farre EM, Geigenberger P** (2002) Starch synthesis in potato tubers is regulated by post-translational redox modification of ADP-glucose pyrophosphorylase: a novel regulatory mechanism linking starch synthesis to the sucrose supply. *Plant Cell* **14**: 2191-2213
- Wait R** (1993) Introduction to Mass Spectrometry. *In* BM Christopher Jones, Adrian H. Thomas, ed, *Spectroscopic Methods and Analyses*, Vol 1. Springer, Totowa New Jersey, pp 191-213
- Winkler R** (2010) ESIprot: a universal tool for charge state determination and molecular weight calculation of proteins from electrospray ionization mass spectrometry data. *Rapid Communications in Mass Spectrometry* **24**: 285-294
- Witt S, Galicia L, Lisek J, Cairns J, Tiessen A, Araus JL, Palacios-Rojas N, Fernie AR** (2012) Metabolic and phenotypic responses of greenhouse-grown maize hybrids to experimentally controlled drought stress. *Molecular Plant* **5**: 401-417
- Worku M, Bänziger M, Friesen D, Horst WJ** (2007) Nitrogen uptake and utilization in contrasting nitrogen efficient tropical maize hybrids. *Crop Science* **47**: 519-528

Evaluating the physiological state of maize (*Zea mays* L.) plants by direct-injection electrospray mass spectrometry (DIESI-MS)[†]

Martín García-Flores,^a Sheila Juárez-Colunga,^a Josaphat Miguel Montero-Vargas,^b Janet Ana Isabel López-Arciniega,^b Alicia Chagolla,^b Axel Tiessen^a and Robert Winkler^{*b}

Received 17th February 2012, Accepted 31st March 2012

DOI: 10.1039/c2mb25056j

Climatic change is an increasing challenge for agriculture that is driving the development of suitable crops in order to ensure supply for both human nutrition and animal feed. In this context, it is increasingly important to understand the biochemical responses of cells to environmental cues at the whole system level, an aim that is being brought closer by advances in high throughput, cost-efficient plant metabolomics. To support molecular breeding activities, we have assessed the economic, technical and statistical feasibility of using direct mass spectrometry methods to evaluate the physiological state of maize (*Zea mays* L.) plants grown under different stress conditions.

Plants react to external signals by changes of their gene expression, changes in proteomic profiles and adjustments of metabolic fluxes and steady state levels. The metabolic ability of sessile plants to adapt to a dynamic environment is crucial for their survival. The drastic climatic change caused by human activities is creating ever more extreme agro-ecosystems: in vulnerable sub-tropical countries such as Mexico, the productivity of food crops such as maize (*Zea mays* L.) is threatened by rising temperatures, lixiviation, lack of water and extended drought periods. Small-scale farmers, in particular, are projected to experience significant losses, since they can neither afford fertilizers nor irrigation systems.¹ For the successful breeding of stress tolerant maize varieties it will become more important not only to consider the genomic sequence of an organism, but also its actual ability to respond adequately and rapidly to hostile conditions.^{2,3}

For the discovery of molecules which might indicate abnormal physiological processes, so-called biomarkers, mass spectrometry is widely applied in human medicine.⁴ Usually blood is chosen as sample material from patients, because of its wide dispersion in the body, its physiological importance and availability. Analogously, in plants, the composition of stem fluids should permit monitoring of their physiological state.

^a Department of Genetic Engineering, CINVESTAV Unidad Irapuato, Km. 9.6 Libramiento Norte Carr. Irapuato-León, 36821 Irapuato Gto., México

^b Department of Biotechnology and Biochemistry, CINVESTAV Unidad Irapuato, Km. 9.6 Libramiento Norte Carr. Irapuato-León, 36821 Irapuato Gto., México.
E-mail: robert.winkler@ira.cinvestav.mx; Fax: +52 462 624 58 46; Tel: +52 462 6239 635

[†] Electronic supplementary information (ESI) available: supplemental methods, mass spectrometry data and the TOPPAS and R scripts. See DOI: 10.1039/c2mb25056j

In maize, several analytical approaches have been implemented for rapid metabolic profiling. Near-infrared reflectance spectroscopy (NIRS) is employed to evaluate protein, tryptophan and lysine content in quality maize breeding programs.⁵ However, this approach requires the collection of reference data from several hundred samples and the reliability for following analyses is limited to previously validated models. Therefore, for a profiling study -which aims to provide quantitative information about a broad range of individual metabolites- neither nuclear magnetic resonance (NMR) nor NIRS is suitable.

High pressure liquid chromatography (HPLC) enables the quantification of compounds and their unequivocal assignment based on the comparison with standards. However, HPLC methods are optimized for certain classes of compounds and therefore limited to a relatively small number of metabolites, such as amino acids, sugars or carotenoids.⁶ Additionally, HPLC runs take at least 30 min per sample, thus impeding the processing of high sample numbers.

Gas chromatography (GC) coupled to mass spectrometry (MS) generates complex data sets which permit both the quantification and identification of many relevant metabolites.³ The maximal throughput of GC-MS is limited due to several reasons. Firstly, the metabolites have to be extracted and derivatized, in order to include non-volatile compounds, which otherwise would not be accessible by gas chromatography. Further, the chromatography step typically requires an hour, and the processing and interpretation of GC-MS data is intellectually demanding and not easily automatizable.

Since a high through-put is desirable, the pre-treatment of samples should be reduced to a minimum. Further, fractionation and separation of the compound mixture should be omitted. Therefore, we examined the applicability of direct MS based fingerprinting methods on maize stem fluids and the possible classification of plant stress responses based on the automatically processed raw data. For this strategy, the focus is the detection of significant differences among MS profiles more than the identification of compounds themselves.

First, we evaluated the feasibility of analysing protein profiles. Changes in proteins indicate physiological states and are an important tool in modern human medicine.⁴ Further, protein profiles measured by matrix assisted laser desorption/ionization (MALDI) – time of flight (ToF) mass spectrometry have been used successfully in, for example, the fast and secure identification

of clinically relevant pathogens.⁷ For this so-called ‘biotyping’ approach, microbial proteins are extracted and precipitated by ethanol. Applying a similar strategy to maize stem juices we observed protein losses of 40–84% (protein quantification with a ninhydrin-based assay^{8,9}). Even after dialysis, the analysis of the protein fractions by gel electrophoresis exhibited protein smears rather than defined bands, indicating non-proteic contaminants (data not shown).

The application of those protein fractions from ethanolic precipitation resulted in sticky spots which did not yield MALDI-ToF spectra. Therefore we tried a more harsh Trichloroacetic Acid/Acetone precipitation, which is nearly quantitative.

Gel electrophoresis, however, still displayed smearing. After additional protein purification with a C18-ZipTip and using sinapinic acid as matrix, we could obtain protein mass profiles for the range of 4–20 kDa, although the signal intensity was low, necessitating excessive raw spectra processing. Consequently, a relatively low number of peaks could be extracted with satisfactory signal-to-noise ratio, as demonstrated in Fig. 1. Taking into account the preparation effort, through-put and sample cost, we concluded that protein profiles are not the first choice for molecular phenotyping in maize breeding programs, and therefore we moved on to metabolic profiling.

Just a few years ago, the use of nano-assisted laser desorption/ionization (NALDI) targets was introduced for the high through-put analysis of small organic molecules.^{10,11} Contrary to MALDI methods, the sample preparation does not require the addition of matrix compounds. However, the direct application of maize juices resulted in dense plaques, from which no ionization could be achieved. Taking the supernatant after ethanolic protein precipitation, signals could be measured. However, the number of detected molecules and the signal intensities were relatively low (data not shown).

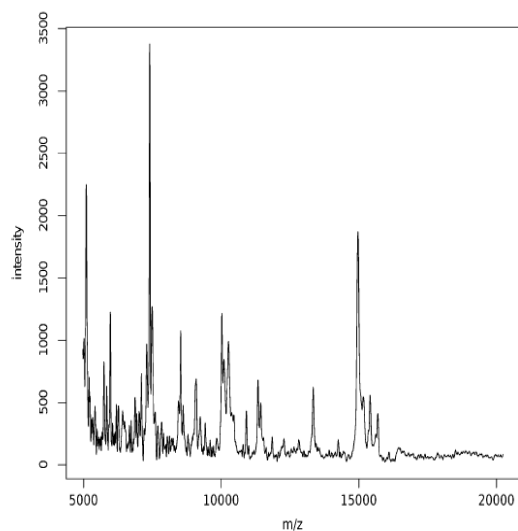


Fig. 1 Protein profile of maize stem juice as obtained by MALDI-ToF. Several sample preparation steps are necessary to extract and purify the protein fraction. The spectra are of low intensity and resolution and therefore require significant data processing. About 40 protein peaks can be assigned in a mass range for proteins of 4–20 kDa.

We then implemented direct injection electrospray mass spectrometry (DIESI-MS) for metabolic phenotyping. In contrast to GC-MS, HPLC and NALDI, the sample preparation effort for DIESI-MS is minimal. The liquid sample has only to be diluted, filtered and acidified.

For testing the potential of the direct-injection strategy, we optimized the protocol in order to analyse stem juices of maize plants suffering either drought stress or nitrogen deficiency alongside their respective control samples.

Since a quadrupole analyser was employed for the study, both qualitative and quantitative MS data could be obtained. Measurement in positive ion mode and data collection for 1 min (averaging of 6 scans of 10 s each) produced satisfactory spectral quality. A typical mass profile is shown in Fig. 2. In order to automatize the data evaluation (spectral smooting, peak picking and data exporting), we employed a OpenMS/ TOPPAS^{12,13} pipeline and wrote an R script for automatic data processing and hierarchical clustering. In contrast to principal component analysis (PCA), our script took all metabolite signals with a defined characteristic into consideration. For our pilot experiment we selected the 100 most abundant ions. With this procedure, all samples (biological triplicates) could be classified correctly according to the physiological treatment (control, low-nitrogen and drought stress), as illustrated in Fig. 3. Additionally the different genotypes can be distinguished by their metabolic fingerprint. The metabolic heat map visually displayed a large amount of quantitative data and further allowed the extraction of individual signals, which have the potential to be used as metabolic markers, thus complementing studies based on molecular¹⁴ and protein¹⁵ markers.

Direct mass spectrometry, *i.e.* without coupling to a separation technique such as gas chromatography reduces the total analysis time and data complexity for subsequent analyses. Eliminating the derivatization step and the reduction of manual handling

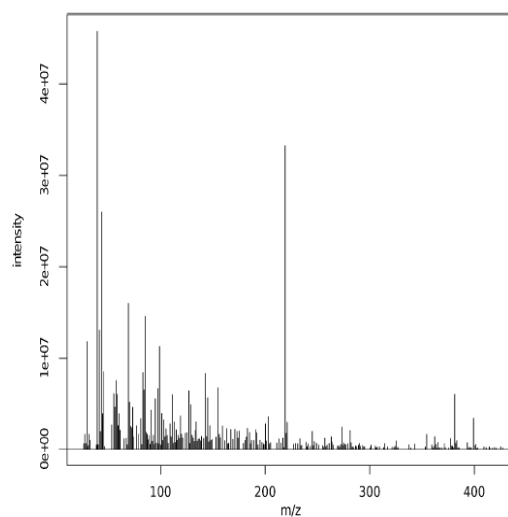


Fig. 2 Direct-injection electrospray mass spectrometry (DIESI-MS) of maize juices. Besides removal of particles and acidification, no further sample processing is required. The monoisotopic resolution of the signals facilitates subsequent data operations. More than 300 peaks were detected in a mass range for metabolites of 15–2000 Da.

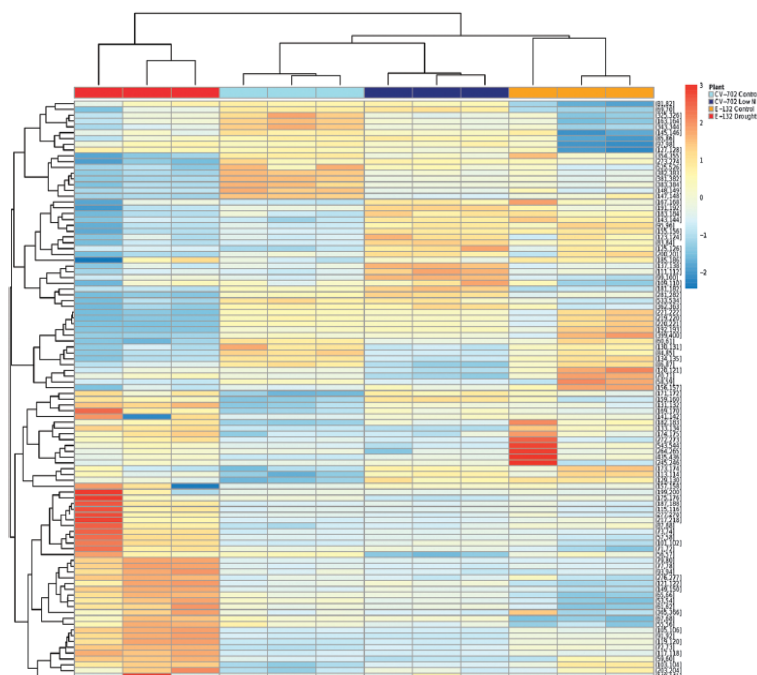


Fig. 3 Metabolic heat map based on DIESI-MS data. The hierarchical clustering results in correct assignment of the plants grown under drought stress or nitrogen limitation. From the differential signals, metabolic markers for breeding can be extracted.

minimizes the influences of sample manipulation on the results and brings down costs.

Our DIESI-MS methodology enabled the collection of large amount of quantitative metabolic data. The rapidity of the analyses permits the 'bed-side' monitoring of the physiological state of plants. In our case, DIESI-MS of maize stem fluids yielded metabolic information which enabled the evaluation of the physiological state of the plants (see Fig. 3).

The mass/charge ratio and the isotopic pattern of the metabolites can be extracted from the raw spectra, leading to the molecular weight and possible composition of the compounds. Certainly, exact metabolite identification is limited by the device accuracy (~ 0.3 Da) and the lack of fragmentation data. Nevertheless, identification of compounds could be realized by subsequent analyses on high-resolution devices and fragmentation experiments. Taking into account *a priori* knowledge about participating metabolites in maize stem fluids, sufficient information can be extracted from the spectra for follow-up studies.

The application of the DIESI-MS method was rapid and cost-efficient, therefore we decided to apply it to a large scale breeding project characterizing the metabolic stress responses of various *Zea mays* L. genotypes. Derived metabolic markers can complement the DNA-based markers for ambitious cereal breeding projects such as IMIC-MasAgro in Latin America or HarvestPlus and DTMA in Africa funded by the Bill Gates Foundation.

Acknowledgements

We thank Prof Dr Ruairidh J. H. Sawers for proof reading the manuscript and the PANGAEA team for preparation and

archiving of our data. This work was partially supported by the initiative MasAgro funded by SAGARPA through grant number IMIC-2011-011. AT and RW also acknowledge basic science funding by Consejo Nacional de Ciencia y Tecnología (CONACYT México). JAILA, JMMV, MGF and SJC acknowledge postgraduate fellowships by CONACYT.

Notes and references

- 1 M. Bellon, D. Hodson and J. Hellin, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 13432.
- 2 P. Monneveux, C. Sanchez and A. Tiessen, *J. Agric. Sci.*, 2008, **146**, 287.
- 3 S. Witt, L. Galicia, J. Lisee, J. Cairns, A. Tiessen, J. L. Araus, N. Palacios-Rojas and A. Fernie, *Molecular plant*, 2012, **5**, 401.
- 4 M. Kiehntopf, D. Schmerler, F. M. Brunkhorst, R. Winkler, K. Ludwig, D. Osterloh, F. Bloos, K. Reinhart and T. Deufel, *Shock*, 2011, **36**, 560.
- 5 A. Rosales, L. Galicia, E. Oviedo, C. Islas and N. Palacios-Rojas, *J. Agric. Food Chem.*, 2011, **59**, 10781.
- 6 M. Kimura, C. Kobori, D. Rodriguez-Amaya and P. Nestel, *Food Chem.*, 2007, **100**, 1734.
- 7 G. Marklein, M. Josten, U. Klanke, E. Müller, R. Horré, T. Maier, T. Wenzel, M. Kostrzewa, G. Bierbaum, A. Hoerauf and H. Sahl, *J. Clin. Microbiol.*, 2009, **47**, 2912.
- 8 B. Starcher, *Anal. Biochem.*, 2001, **292**, 125.
- 9 M. González-González, K. Mayolo-Deloisa, M. Rito-Palomares and R. Winkler, *Process Biochem.*, 2011, **46**, 413.
- 10 R. Daniels, S. Dikler, E. Li and C. Stacey, *J. Assoc. Lab. Autom.*, 2008, **13**, 314.
- 11 E. Guénin, M. Lecouvey and J. Hardouin, *Rapid Commun. Mass Spectrom.*, 2009, **23**, 1395.
- 12 O. Kohlbacher, K. Reinert, C. Gropl, E. Lange, N. Pfeifer, O. Schulz-Trieglaff and M. Sturm, *Bioinformatics*, 2007, **23**, e191.
- 13 M. Sturm, A. Bertsch, C. Groepf, A. Hildebrandt, R. Hussong, E. Lange, N. Pfeifer, O. Trieglaff, A. Zerck, K. Reinert and O. Kohlbacher, *BMC Bioinformatics*, 2008, **9**, 1.
- 14 J. Ribaut and M. Ragot, *J. Exp. Bot.*, 2007, **58**, 351.
- 15 R. Winkler and S. García-Lara, *Mol. BioSyst.*, 2010, **6**, 1810.

Metabolic Profiling of Plant Extracts Using Direct-Injection Electrospray Ionization Mass Spectrometry Allows for High-Throughput Phenotypic Characterization According to Genetic and Environmental Effects

Martín García-Flores,[†] Sheila Juárez-Colunga,[†] Adrián García-Casarrubias,[†] Samuel Trachsel,[‡] Robert Winkler,[§] and Axel Tiessen^{*,†}

[†]Departamento de Ingeniería Genética, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV), Unidad Irapuato, Km 9.6 Libramiento Norte, 36821 Irapuato, Guanajuato, Mexico

[‡]Global Maize Program, Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), Km 45 Carretera Mexico-Veracruz, 56130 El Batán, Texcoco, State of Mexico, Mexico

[§]Departamento de Biotecnología y Bioquímica de Plantas, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (CINVESTAV), Unidad Irapuato, Km 9.6 Libramiento Norte, 36821 Irapuato, Guanajuato, Mexico

Supporting Information

ABSTRACT: In comparison to the exponential increase of genotyping methods, phenotyping strategies are lagging behind in agricultural sciences. Genetic improvement depends upon the abundance of quantitative phenotypic data and the statistical partitioning of variance into environmental, genetic, and random effects. A metabolic phenotyping strategy was adapted to increase sample throughput while saving reagents, reducing cost, and simplifying data analysis. The chemical profiles of stem extracts from maize plants grown under low nitrogen (LN) or control trial (CT) were analyzed using optimized protocols for direct-injection electrospray ionization mass spectrometry (DIESI-MS). Specific ions significantly decreased or increased because of environmental (LN versus CT) or genotypic effects. Biochemical profiling with DIESI-MS had a superior cost-benefit compared to other standard analytical technologies (e.g., ultraviolet, near-infrared reflectance spectroscopy, high-performance liquid chromatography, and gas chromatography with flame ionization detection) routinely used for plant breeding. The method can be successfully applied in maize, strawberry, coffee, and other crop species.

KEYWORDS: *biochemical phenotyping, crop physiology, fingerprinting, maize genetic improvement, low nitrogen stress*

■ INTRODUCTION

The progress in deoxyribonucleic acid (DNA) technologies allow for the sequencing of whole eukaryotic genomes, but the technical advances in metabolic profiling strategies have lagged behind, thus increasing the gap of knowledge on the functional link between genotypes and phenotypes¹ (see also the Table of Content graphic). Genetic improvement relies strongly on quantitative phenotypic data.² Molecular breeding based on DNA markers has led to some advances^{3,4} but not as anticipated initially because target traits, such as grain yield (GY) and nutritional quality (e.g., content of carotenoids, amino acids, etc.), depend upon multiple factors in addition to genetics. The environment can modulate the expression of a mendelian trait (gene) at multiple stages, such as affecting DNA structure, the transcription of genes, the translation of proteins, the activity of enzymes, carbon fixation, and the translocation of metabolites. The model shown in Figure 1 illustrates the reciprocal effects at various levels. At the top, environmental factors can cause epigenetic modification (e.g., DNA methylation), affecting transcriptional activity and downstream molecular processes (Figure 1). At the bottom, the availability of light, water, and nitrogen determines photosynthesis, growth, development, partitioning, biomass accumulation, and yield (Figure 1). According to this model

(Figure 1), a metabolite is much closer related to the agronomic phenotype than any chosen fragment of nucleic acid.^{1,5,6}

If that is true, why do plant breeding initiatives employ molecular markers based on DNA sequences more often than metabolomics? One possible answer is as follows: technical progress has made sequencing of DNA cheaper than biochemical profiling and even more affordable than multi-location multi-year field trials. Metabolic phenotyping with high-performance liquid chromatography (HPLC) or ultra performance liquid chromatography–mass spectrometry (UPLC–MS) is too slow and expensive to be used routinely for plant breeding, whereas near-infrared reflectance spectroscopy (NIRS) provides quantitative data about only too few metabolites grouped in families of compounds. Therefore, in agricultural sciences, there is an urgent need to develop novel biochemical profiling strategies particularly adapted to widen metabolic coverage, increase information density, and sample throughput. To be effectively implemented for breeding,

Received: October 7, 2014

Revised: December 9, 2014

Accepted: December 22, 2014

Published: January 14, 2015

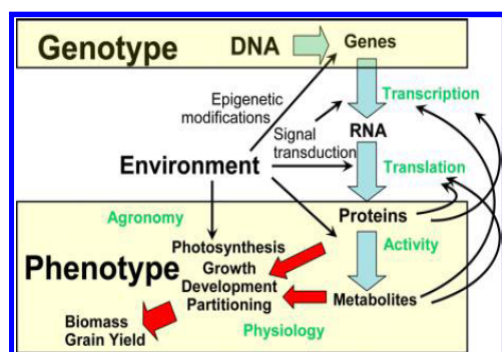


Figure 1. Model showing the large gap between genotype and phenotype. The functional link between DNA and yield is complex because the environment acts at many different levels. Genes can be modified epigenetically by methylation and histone acetylation, which are both affected by the environment. Some epigenetic states can be stably inherited and are therefore relevant for breeding. Multiple feed-forward and feedback signaling loops between metabolites, proteins, and small ribonucleic acids (RNAs) complicate the specific pattern of gene expression. Genes represent “potential”, whereas transcripts reveal the short-term “strategy”. Proteins and enzymes do the “work”, whereas the final “results” are the metabolites that are accumulated in a certain storage organ (e.g., starch in seeds). From this model, it is apparent that metabolites are much closer to grain yield than any DNA sequence, and therefore, selection for biomarker metabolites should be more relevant for breeding than selection for molecular markers [e.g., microsatellites, single nucleotide polymorphisms (SNPs), quantitative trait loci (QTL), etc.].

analytical methods should be retailored to be faster, save reagents, reduce cost, and simplify/automatize data analysis.⁶

Mass spectrometry (MS) allows for the analysis of a large number of chemical compounds and is a promising tool for agricultural research.^{5–7} Traditionally, qualitative separation and quantitative detection are performed in two steps: upstream separation, such as liquid chromatography (LC) and gas chromatography (GC) first filters molecules because of physicochemical differences (affinity to a stationary or mobile phase), and then detection of positive or negative ions is performed according to the mass-to-charge ratio (m/z). In GC–MS, post-chromatographic molecular fragmentation is performed by electron impact (EI), a hard ionization method that results in the fragmentation of the molecules.⁸ MS methods combining different chromatographic and ionization technologies have been used for phenotyping purposes, such as liquid chromatography–electrospray ionization–mass spectrometry (LC–ESI–MS) and gas chromatography–electron impact–mass spectrometry (GC–EI–MS).^{9–12} Metabolic profiling based on GC–EI–MS is the most widely used in plant biology, e.g., for the phenotypic characterization of mutants, transgenic plants, and recombinant inbred line (RIL) populations.^{6,13–16} It has also been used to characterize the metabolic response of maize hybrids subjected to drought.¹²

UPLC–MS is too expensive to be implemented for routine breeding in agricultural institutions of the subtropics. GC–EI–MS is more affordable, but unfortunately, it has several technical limitations. Each single chromatographic run requires 30–60 min, and therefore, less than 24–48 samples can be analyzed daily. Besides low throughput, GC–MS-based metabolic profiling depends upon volatility and ionizability of molecules, with both aims being somehow in contrast. On the one hand, ionizable molecules have less propensity to escape

from the liquid phase the more charged that they are. On the other hand, uncharged molecules are volatile but do not ionize efficiently. Therefore, GC–EI–MS samples need to be previously derivatized to increase chemical coverage. The most widely used procedure is the addition of trimethylsilyl (TMS) groups,¹⁴ which makes highly hydrophilic compounds, such as amino acids and carboxylic acids, volatile in the gas phase. Unfortunately, biomolecules are derivatized to several possible TMS compounds, each having different retention times (RTs) in the GC column. One single metabolite, such as citric acid, can generate more than 10 different RT peaks in a GC–EI–MS chromatogram. The difficulty is to decide which chromatographic peak should be used to quantify citric acid in the original sample. It becomes even more difficult because GC–EI–MS data are split in two dimensions (RT and m/z). Therefore, sophisticated data processing workflows are needed for data integration and statistical interpretation.¹² Manual processing and intuitive decisions have to be made relying on recommendations of experienced users.¹² Each group may follow particular philosophies, and therefore, GC–EI–MS and LC–ESI–MS data analysis can be considered both mathematics and art.^{12,17} The bottleneck of LC–ESI–MS and GC–EI–MS profiling is the manual processing and individual interpretation of chromatographic data by the final user. Typically, a student requires several months for training, and even when it learns to perform GC–EI–MS by himself, a single user cannot fully process and interpret more than 24 samples per day, much less produce publication-ready graphics in less than 1 week (final throughput is very low).

Given the previously mentioned considerations, we decided to implement a metabolic profiling strategy that overcomes the following problems: (1) time-consuming chromatographic separation, (2) limited compound volatility and inefficient ionization of hydrophilic metabolites, (3) generation of multiple peaks because of chemical derivatization (data redundancy), (4) reliable, sensitive, and quantitative detection, (5) manual handling of chromatograms, and (6) complex data workflow and statistical processing (e.g., automatic data reduction and visualization). Inspired by the requirements of plant breeders, we aimed to develop an analytical platform that allows for a high sample throughput, produces a vast amount of quantitative chemical information, and is at the same time cheap, robust, and easy to handle bioinformatically.

Therefore, we compared several analytical methodologies, with particular focus on direct-injection electrospray ionization mass spectrometry (DIESI–MS).¹⁸ We documented the genotypic and environmental effects on the metabolic profile of maize stem extracts. Our leading questions were as follows: Is DIESI–MS data quantitative and repeatable? Can we detect specific ions that have larger variance because of the environment (E) or the genotype (G)? How large is the throughput and how expensive is DIESI–MS profiling compared to other methods? Does this metabolic phenotyping strategy work for other plant samples in addition to maize?

■ MATERIALS AND METHODS

Plant Materials and Growth Conditions. *Maize Experiments under Nitrogen Stress.* The low nitrogen (LN) and control trial (CT) were performed during the 2011 summer season in El Batán, Texcoco, Mexico [19° 30' N latitude, 98° 53' O longitude, 2247 m above sea level (masl) altitude]. It was performed according to the field management standards established by the International Maize and Wheat Improvement Center (CIMMYT).² The LN treatment consisted of suppressing nitrogen fertilization during the previous

~4 years (8 cycles), whereas the control field was fertilized chemically with urea at the usual frequency (3 times per growth cycle) and dose (~400 kg/ha). Planting density was 67 000 plants/ha in two rows of 5 m distributed in randomized blocks (α lattice).² Six plants were sampled per genotype for each treatment (LN and CT). The CIMMYT field trial HL2011-11-01-PVB-2 was harvested ~10 days after silking (~85 days after sowing), while the field trial HL2011-11-01-PVB-1 was harvested ~12 days after silking (~85 days after sowing). DIESI-MS results are shown for 684 maize samples: 19 genotypes, 2 conditions (LN and CT), and 18 replicates (6 maize plants extracted 3 times each). Figures show the number of n behind each boxplot. Each of the 684 DIESI-MS runs lasted for 60 s, and all MS spectra acquired were averaged to increase accuracy and precision. In total, 4104 individual MS spectra were processed for the maize experiments.

Maize Stem Juice Collection. Representative healthy plants from each genotype and treatment were selected in the field and processed within ~15 min after harvest. Morphological and physiological data were rapidly registered: total plant weight, plant height (up to the flag leaf), number of ears, and ear weight. The tassel and the leaf blades were removed, and the supporting stem (including leaf sheath) was weighted and cut into segments (~30 cm) for tissue homogenization. The stem juice (phloem sap + xylem sap + cellular extracts) was obtained using a mechanical extractor (industrial-scale juice extractor model EXS, International Co., Mexico). The stem juice (~80 mL) was collected in a plastic beaker, weighted accurately (± 0.2 g), aliquoted (1 mL), and frozen in dry ice immediately (<30 s). Collected samples were stored at -80 °C for ~4 months in sealed deep well plates (96 well format). For analysis, the stem juice was unfrozen and kept on ice for short time periods (<1 h). Samples were first centrifuged (10 min at 5000g and 4 °C), and the supernatant was used directly for metabolite analysis (see below).

Other Plant Samples. Leaves from different coffee and tea species were obtained as described previously.^{18,19} Strawberry fruits were divided in three levels of ripeness (see Figure S1 of the Supporting Information). The level of ripeness was determined according to the following parameters: fruit turgor, fruit color, and peduncle color (see Table S1 of the Supporting Information).

Metabolite Analysis. Analytical methods were evaluated considering costs, sample throughput, and the requirements of chromatographic separation and chemical derivatization. We also evaluated the required workflow for computational and statistical analyses of the generated data. We excluded NIRS because it is well-known that it can simultaneously detect/discriminate only 1–3 different compounds reliably,²⁰ whereas MS-based methods produce information on more than 300 different metabolites. We also excluded high-end UPLC-MS and nuclear magnetic resonance (NMR) because machine costs are prohibitive for routine plant breeding.

Enzyme-Coupled Spectrophotometry Assays. Soluble sugars were measured in microplate format.^{21,22} The insoluble starch pellet was dissolved in 0.5 mL of 10 mM KOH at 99 °C and autoclaved for 30 min. Starch was hydrolyzed in 50 mM *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES) (pH 7.5) at 37 °C overnight by the addition of 10 units of α -amylase (EC 3.2.1.1, Roche) and 10 units of amyloglucosidase (EC 3.2.1.3, Roche). Samples were centrifuged (13000g for 5 min); the resulting supernatant was stored at 4 °C; and the pellet was hydrolyzed again for 30 min at 37 °C. Both supernatants were combined, and an aliquot was enzymatically assayed for glucose, as described above.

DIESI-MS. The DIESI-MS assay was performed as described previously,^{18,19} with some adjustments (see below). The stem juice supernatant was filtered through a 0.45 μ m polyvinylidene difluoride (PVDF) filter and diluted 1:100 with deionized water to avoid saturation of the capillary and cone soiling. To improve ionization, to 475 μ L of diluted sample and 25 μ L pure formic acid were added [final concentration of 5% (v/v) formic acid]. Control experiments were carried out to determine optimal parameters of sample pretreatment and dilution (data not shown). The diluted and acidified samples were analyzed by DIESI-MS in positive mode, employing a Micromass ZQ 2000 with a quadrupole analyzer (Waters) and MassLynx 4.0 as

control software.¹⁸ A constant volumetric flow rate (10 μ L/min) was achieved using a syringe pump loaded with 100 μ L of sample. Capillary voltage was set to 3 kV; the cone voltage was set to 60 V; the extractor voltage was set to 3 V; and the radio frequency (RF) lens was left at 0.5 V. The source temperature (80 °C) and desolvation temperature (150 °C) were used with a nitrogen gas flow of 250 L/h and a cone gas flow of 50 L/h. In the analyzer section, low-mass (LM) and high-mass (HM) resolution of 15.0 and ion energy of 0.5 were set. The multiplier was adjusted to a value of 650. Continuous spectra were collected in a range of m/z 15–2000, with a total run duration of 1 min, a scan time of 10 s, and an interscan time of 0.1 s, producing six spectra per sample.

Data Analysis and Script Optimization. The raw MS spectra files were converted to *.mzXML using masswolf, version 1.4. Analysis of mass spectra was performed with the OpenMS/TOPP suite, version 1.8.0, and later further processed with the free statistical software R (<http://www.r-project.org>).²³ A TOPPAS pipeline was written executing the following sequential tasks: file conversion to mzML, spectra merger with a block method (six replicate spectra), NoiseFilter sgolay with a frame length of 21 and polynomial order of 4, followed by a PeakPicker with a signal-to-noise ratio (S/N) of 1 and a peak width of 0.15. Finally, data tables were converted into TOPPAS out files using raw text format, containing tab-separated columns for genotype names, treatment factors, m/z values, and corresponding signal intensities of the individual ions. Individual data files were then automatically imported with a tailored script into R.²³

Besides sample harvesting and DIESI-MS workflow optimization, other advantages/novelities presented in this paper are the improved algorithms for data processing. The new R script imported all text files from a given directory and allowed for the definition of analysis of variation (ANOVA) experimental factors based on the original file name. Variable assignment was performed inside the R environment and did not require additional software for file merging or renaming. The R script was also optimized to process the data of hundreds of samples in $1/10$ of the time of previous scripts.^{18,19}

Instead of naming peaks by interval boundaries,^{18,19} m/z peaks were named by their center of maximal intensities. Instead of defining mass bins across the entire m/z range,^{18,19} only real occurring m/z peaks were transferred to the rectangular result matrix used for the heatmap graphics. Instead of defining bin boundaries centered at 0, a mass shift was introduced to reduce bin jumping. The m/z bin boundaries were adjusted according to the observed m/z pattern (metabolic space) in various extracts: in the mass range of 50–400 Da, m/z peaks were centered at +0.2 Da from the unitary mass. In the mass range of 400–700 Da, a mass shift of +0.3 Da was more accurate. Ions were less frequent in the mass range of 800–2000 and had more variability in mass shift, so that it was more convenient to define bin centers based on the occurrence of specific ion peaks (metabolic space).

The improved R script also allowed us to define variably the tolerance of m/z peak assignment (bin size). However, for our plant samples, there was little gain of information by defining mass bins smaller than ± 0.25 Da, and therefore, bin size was conveniently defined in 1 Da intervals (± 0.5 Da from the center) across the entire m/z range (30–2000 Da).

When double peaks appeared in the same bin (e.g., isotopic peaks), intensities were summed up and not averaged or deleted. The optimized R script produced separate tables documenting the number of peaks per bin, intensity of single peaks, and summed intensity of multiple peaks. Automatic quality control was carried out to detect double-peak occurrence in individual mass bins.

In addition to presenting metabolic data as colored heatmaps, the optimized R script produced gray heatmap biclusters and various boxplot figures showing statistics and quality controls. For example, in addition to reporting the intensity values for given m/z ion bins, the script also plotted the exact m/z value for each sample group. The m/z residual (difference between the real observed mass and the unitary mass) was calculated for each peak and sample and plotted in parallel to the intensity boxplots (e.g., panels D–F of Figure 8). From this approach, users can decide if a given m/z value in one plant sample corresponds to the same metabolite in another plant sample. Ideally,

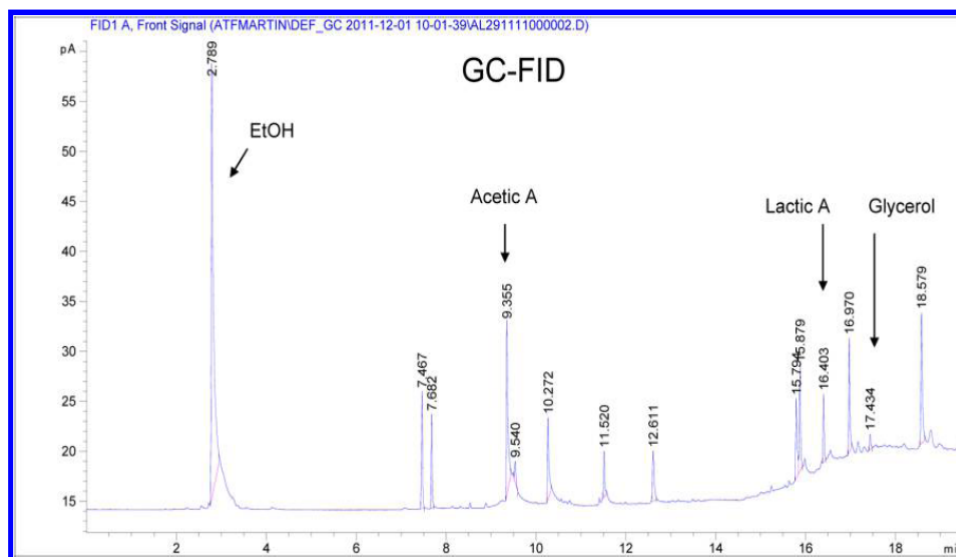


Figure 2. Typical GC–FID chromatogram of juice extracts of maize stem samples. Results correspond to the maize genotype CLQ-RCWQ83 harvested under well-watered conditions. RT's of standards were as follows: methanol, 2.53 min; ethanol, 2.78 min; acetic acid, 9.19 min; lactic acid, 16.38 min; and glycerol, 17.31 min.

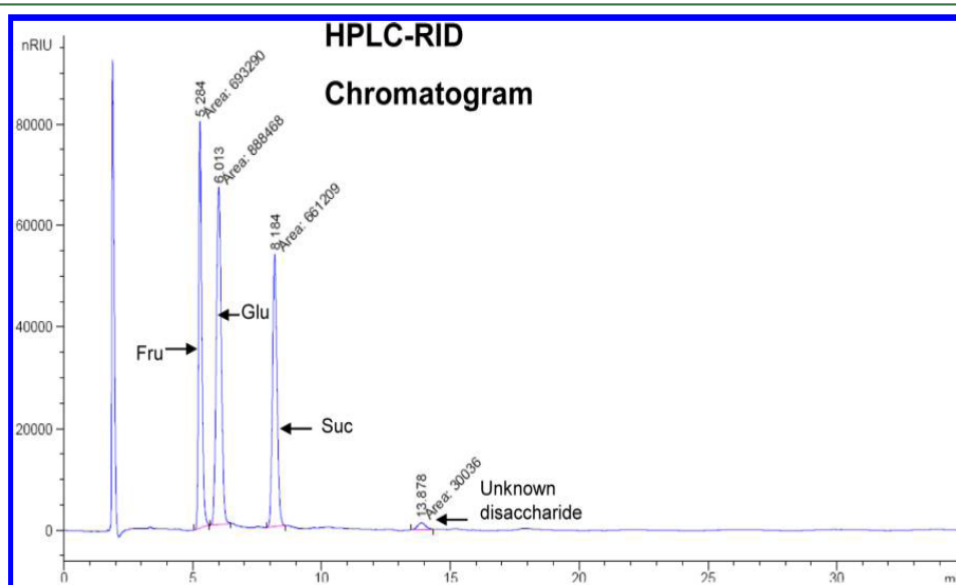


Figure 3. Chromatogram of juice extracts of maize stem samples acquired with HPLC–RID. The three main sugars were fructose, glucose, and sucrose. A fourth minor disaccharide of unknown nature was also observed in maize extracts. Other compounds rather than sugars are not detectable with this method.

m/z values of the same compound across samples should be similar, but in practice, they did vary ± 0.2 Da with our quadrupole MS equipment.

Overall, the optimized algorithm permitted automatic import of thousands of files, allowed for more rapid data processing, and also significantly increased reliability by diminishing “peak jumping” and eliminating the requirement of manual post-processing of data.

Statistical Procedures. All experiments were conducted using randomized complete block designs with an adequate number of controls and replicas (see the legends). ANOVAs and Tukey's least significant difference (LSD) test were applied with $p \leq 0.05$ ($\alpha = 0.05$). Mixed effect linear models were used for the analysis of data, with separate factors for the effect of the G, E, development stage

(DS), and their respective interactions. Bonferroni correction was used to address the false discovery rate. Statistical analysis was performed with R, version 2.10.²³

RESULTS AND DISCUSSION

Comparison of Information Density Generated by Different Analytical Methods. Various samples were analyzed with the particular interest of determining the information density of different phenotyping strategies. Maize was grown under different environmental conditions under water and nitrogen stress.^{12,18} Plants were harvested at the post-flowering stage, and stem juice aliquots were collected and

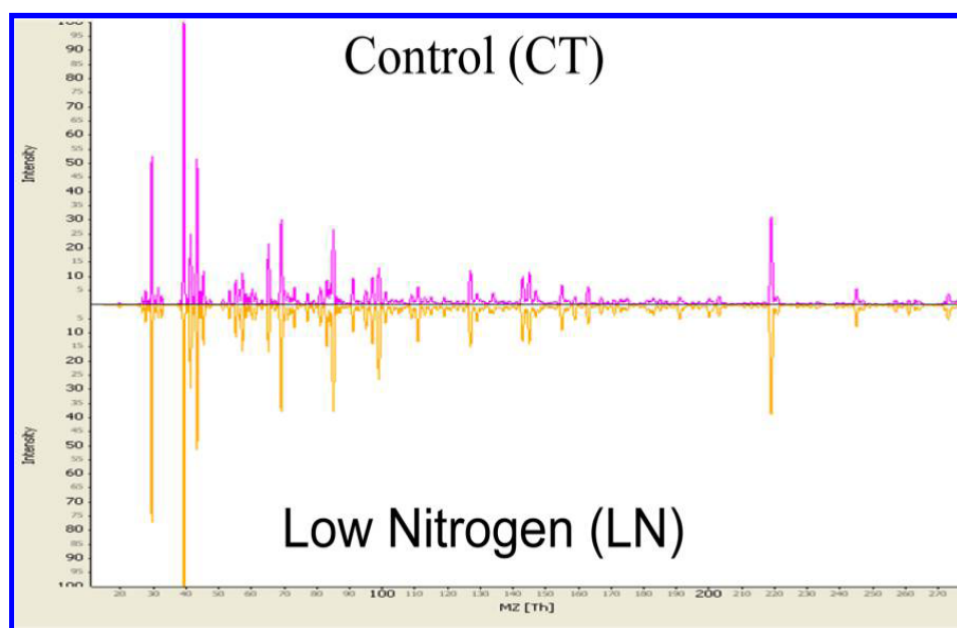


Figure 4. Mass spectrum of stem juice extracts of maize acquired with DIESI–MS showing ~200 compounds. Maize were grown under normal CT and LN stress, and the stem was harvested ~12 days after female flowering. Stem juice was extracted mechanically, and aliquots were frozen in liquid nitrogen until analysis. Data from a representative sample (genotype CV-702) are shown.

Table 1. Comparison of Analytical Platforms for Metabolic Phenotyping and Breeding Purposes with Respect to Throughput, Information Density, and Cost^a

method/analytical platform	sample throughput per day technical/statistical	number of metabolites per assay	cost per sample (\$ U.S./sample)	cost per metabolite data point (\$ U.S./metabolite)
UV–vis enzymatic	384/384	1–3	~10.00	3.33–10.00
UV–vis colorimetric	384/500	1	~0.50	0.50
GC–FID	30/10	5–20	~50.00	2.50–10.00
GC–EI–MS	24/2	50–100	~80.00	0.80–1.60
UPLC–MS	48/2	200–1000	~200.00	0.20–1.00
HPLC–DAD	24/10	3–15	~80.00	5.30–26.67
NIRS	80/500	1–4	~2.00	2.00–0.50
DIESI–MS	80/500	100–500	~10.00	0.02–0.10

^aValues were estimated from our experiences across several research projects at CINVESTAV and other institutes. Throughput is separated in technical (amount of samples that a laboratory can process) and statistical [amount of data files that an undergraduate student can analyze considering worktime for chromatogram integration, metabolite assignment, statistical analysis, etc. (~human worktime)] throughput.

frozen for the simultaneous analysis with various analytical procedures [GC, HPLC, ultraviolet–visible (UV–vis), nano-assisted laser desorption ionization time-of-flight mass spectrometry (NALDI–TOF–MS), and DIESI–MS]. The gas chromatography with flame ionization detection (GC–FID) method allowed for the detection of ~10 different peaks with satisfactory results (Figure 2). The chromatographic peaks had RTs of known alcohols and organic acids that are typically present in plant extracts (e.g., ethanol, acetic acid, lactic acid, glycerol, etc.) (Figure 2).

Using high-performance liquid chromatography with refractive index detection (HPLC–RID), ~4 different peaks were detected with satisfactory results (Figure 3). The HPLC chromatographic peaks had retention times corresponding to glucose (glu), fructose (fru), and sucrose (suc) (Figure 3). Enzymatic assays coupled to NADPH using a UV–vis spectrophotometer (microplate reader) specifically detected the three main sugars (glu, fru, and suc) that are typically present in plant samples. From this targeted method, the

presence and amount of all sugars were confirmed in the maize stem extracts. The sugar ratio was dependent upon the G but also strongly dependent upon the DS and E conditions in which the maize plants were grown. Using ANOVA, the effects because of E, G, DS, and their interactions were discriminated. The variance effects because of G were in the range of 3–20% of the total phenotypic variance, which is excellent for breeding purposes in maize, because selection is only effective for the genetic variance.

We also employed analytical methods based on MS. NALDI–TOF–MS has the advantage to be relatively cheap if it is implemented routinely and further requires no previous chromatographic separation. With NALDI–TOF–MS we found ~20 different peaks with a S/N of >10 (data not shown; see also ref 18). The information density generated with NALDI–TOF–MS did not surpass the previous analytical methods, and therefore, we switched to another MS method that neither employs chromatography nor chemical derivatization: DIESI–MS. Using DIESI–MS in positive-ion mode, more

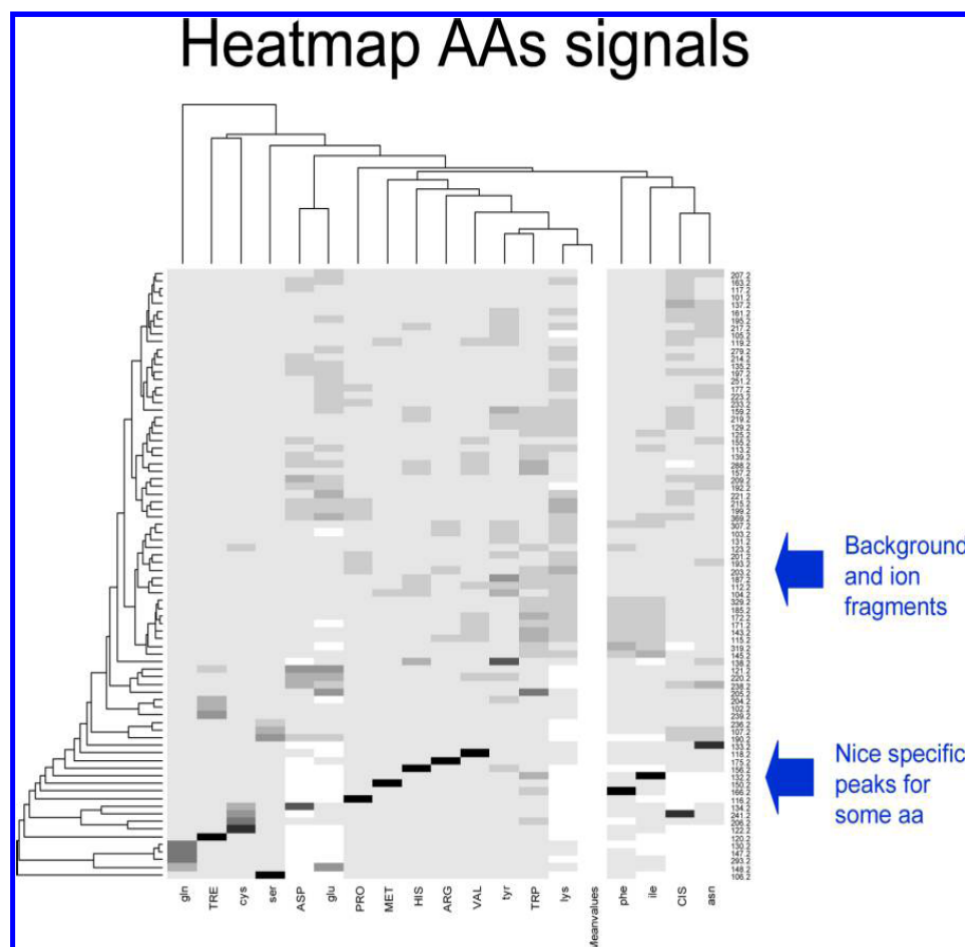


Figure 5. Heatmap profile of amino acid standards. Hierarchical clustering is presented for the ion fragments (horizontal rows) and amino acid standards (vertical rows).

than 200 different peaks were detected with an excellent S/N of >10 (Figure 4). Up to ~600 other ions could also be quantified with acceptable results ($S/N > 3$). DIESI-MS yielded 1–2 orders of magnitude more biochemical information at a fraction of the cost compared to other methods (Table 1).

DIESI-MS Generates a Metabolic Fingerprint Revealing Specific Ions Differing between Treatments. When samples from maize plants grown under different environmental conditions (normal CT and LN) are compared, many peak intensities were similar, the mass spectra appeared symmetrical, but detailed inspection revealed that there were also specific ions that increased or decreased (Figure 4). A similar response was also observed for maize plants subjected to drought stress and for plants grown at different localities and conditions (data not shown). Overall, the results confirmed our previous report on the possible application of DIESI-MS for the physiological evaluation of maize plants using stem extracts collected at the flowering time.¹⁸

Method Development and Optimization. The promising results motivated us to further optimize the DIESI-MS platform for high-throughput metabolic phenotyping of maize and other species. We simplified the extraction and sample preparation procedure, but more particularly, we improved the scripts, file handling, statistical procedures, and boxplot figure

output compared to previous reports.¹⁸ The statistical analysis of the DIESI-MS data using loop programming and the R scripts were optimized as described extensively in the methods. We also optimized sample treatment and handling, allowing for the realistic measurement of ~80 samples per day (Table 1).

Control Experiments: DIESI-MS Ion Pattern of Standards. We then asked: How repeatable and redundant is the DIESI-MS data? Do standard compounds ionize with the theoretical positive mass, or does our DIESI-MS method produce ions fragmenting other non-standard m/z peaks? To answer those questions, additional control experiments were performed. Candidate primary metabolites that might change under different stress conditions were evaluated. We focused on amino acids, because it could be expected that CT and LN treatments severely affects the nitrogen status of plants. Theoretically, they should all ionize well in positive-ion mode; therefore, the 20 standard amino acids and cystine were injected into the MS quadrupole, and the occurrence and the magnitude of ion fragmentation with electrospray ionization (ESI) was documented using gray heatmaps (Figure 5).

DIESI-MS produced an ion pattern (mass spectrum) that was highly specific for each amino acid (Figure 5). The most intense peaks (black bars) had indeed the expected m/z according to the theoretical mass of non-fragmented proto-

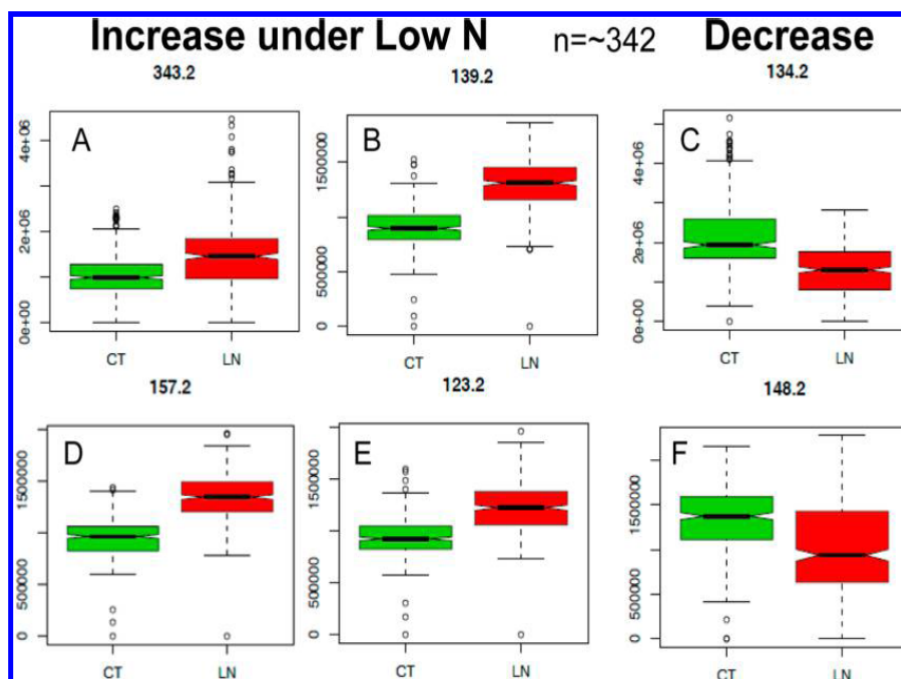


Figure 6. Intensity data of selected m/z ions quantified with DIESI–MS. Maize plants were grown under either normal CT (green) or LN (red) conditions, and stem juice extracts were sampled for DIESI–MS metabolic profiling. Boxplots summarize data from 19 genotypes, 6 biological, and 3 technical replicates ($n = 342$). Boxplot whiskers are shown for significant differences. (A, B, D, and E) Data from four representative ions that increase under LN conditions. (C and F) Data from two representative ions that decrease under LN conditions.

nated amino acids (AA + 1H) (Figure 5). Most amino acids ionized very well with our DIESI–MS protocol (e.g., Gly, Ser, Thr, Gln, Leu, and Ala). Unfortunately, other amino acids did not ionize equally well (e.g., His, Met, Arg, Asn, Tyr, Trp, Phe, Glu, and Asp), and the reason for this remained unknown. The atomic events that lead to ionization in ESI are complex, so that efficient ionization cannot be theoretically predicted nor explained by the chemical formula but must rather be confirmed by a trial and error approach.

In addition to the expected m/z peaks (main ion fragments), other minor ions of lower intensity also appeared in the DIESI–MS spectra (Figure 5). Minor accompanying peaks were both smaller and larger than the expected molecular mass of the unitary ion. These peaks could be ion fragments (smaller mass) or ion aggregates (larger mass) of the original molecules. Nevertheless, ion fragmentation with ESI was much less than that with EI, and the avoidance of chemical derivatization reduced the problem of data redundancy (e.g., with GC–TOF, citric acid and some amino acids generate more than six RT peaks across several fragmented masses).

Taken together, the control experiments indicated that DIESI–MS efficiently discriminates the chemical profile of samples and that data are repeatable and fairly non-redundant. However, because of fragmentation, aggregation, and matrix interaction, the assignment of a given m/z peak to a given biomolecule should be performed with extreme care. Even if a DIESI–MS peak has a m/z value that is identical to the theoretical mass of a standard known compound, we did not assume that the intensity value corresponded to the concentration of the known compound in the plant sample. In fact, direct-injection spectra provide a numeric phenotype of complex ion mixtures that reveals the physiological status of plants,¹⁸ but it is not the best choice for metabolite

identification, chemical discovery, and/or structural assignment as with other high methods, such as UPLC–MS–Orbitrap or NMR. For example, the DIESI–MS peak m/z 106.2 in theory corresponds to positively ionized serine, but other compounds could also generate signals or suppress ionization in that m/z range. A search in chemical databases yielded dozens of chemical compounds for that given m/z value. If the mass tolerance was increased to ± 0.5 Da, then hundreds of theoretical compounds were listed. Many molecules are certainly absent in plants or living cells, but unequivocal assignment of m/z values to metabolites requires many control experiments that are out of the focus of the present paper. In the midterm, we will start building a biochemical database of DIESI–MS peaks that are frequently found in various plant extracts (we call it “metabolic space” of each species and tissue) (see also ref 19).

Given the previous limitations, we avoided assigning chemical names to DIESI–MS peaks but only named them by the numerical m/z value (center of maximal intensity of the positive ion). From the practical perspective of metabolic phenotyping for breeding purposes, it does not matter the exact identity or name of the compound but rather the following questions: Is it related to yield? Can we measure it cheaply? Is it quantitative? How much variance depends upon the E, and how much depends upon the G? Will it be useful for plant breeding? Which ions increase or decrease significantly under stress? Which peaks should be used as biomarkers for selection and/or screening?

Discrimination of Phenotypic Variance Because of E and G Effects. To answer above questions, we performed field trials using 18 maize hybrids grown under CT and LN conditions (see the Materials and Methods). The optimized DIESI–MS data analysis pipeline and the novel R scripts

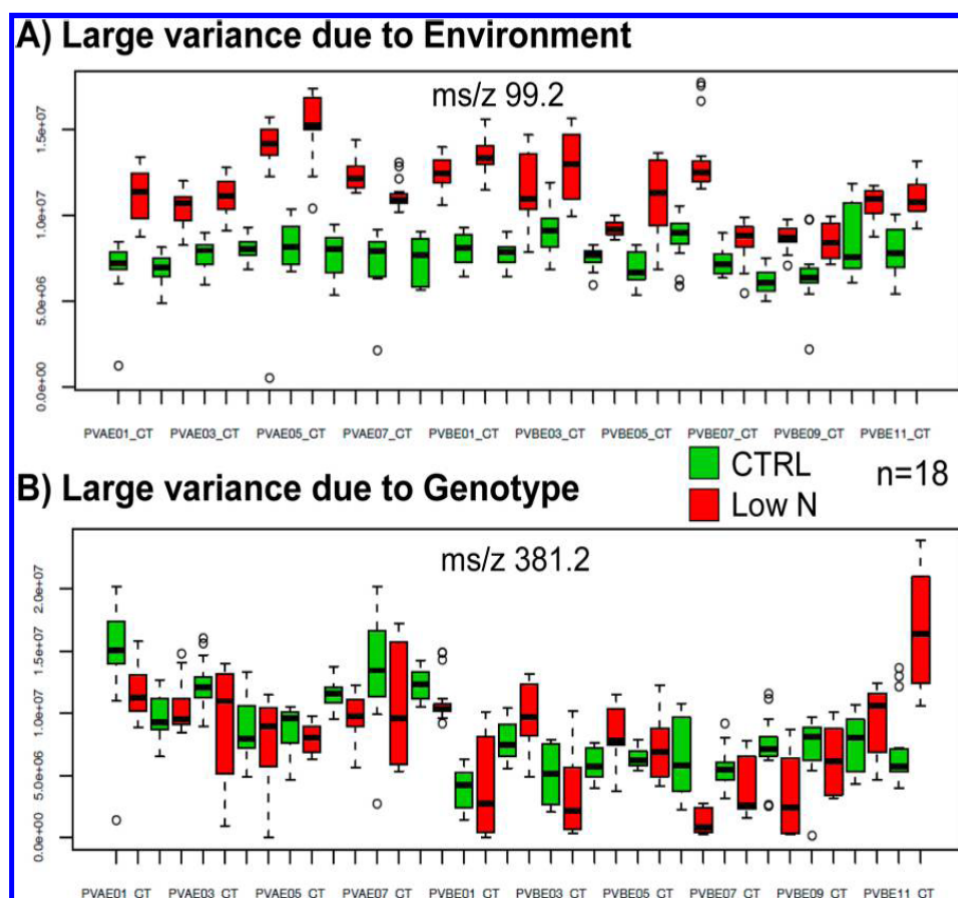


Figure 7. Evaluation of signals that primarily depend upon environmental or genetic factors. Intensity data of selected m/z ions were quantified with DIESI–MS. A total of 19 maize genotypes were grown under either normal CT (green) or LN (red) conditions, and stem juice extracts were sampled for DIESI–MS metabolic profiling. Boxplots summarize data from 6 biological and 3 technical replicates ($n = 18$). ANOVA was used to discriminate environmental and genotypic variance. (A) Representative m/z ion 99.2 with a large environmental effect. (B) Representative m/z ion 381.2 with a lower environmental effect but larger genetic effect.

allowed us to represent the data of given m/z peaks automatically as boxplots (Figure 6). The pooled data per environmental condition revealed several m/z peaks that significantly increased (panels A, B, D, and E of Figure 6) or decreased (panels C and F of Figure 6) under LN treatment.

When the data for each genotype were analyzed separately (Figure 7), several m/z signals, such as 99.2, displayed a strong dependence upon the environment (Figure 7A), while others, such as 381.2, were less dependent upon the growth conditions but rather presented large variance because of the genotype (Figure 7B). Such information is required for phenotyping purposes because it is important to identify candidate m/z peaks to be used as biomarkers for a given physiological condition or stress. Overall, DIESI–MS allowed for the partitioning of variance of individual peaks according to E and G effects and can be combined with standard agronomic measurements within yield trials to predict hybrid performance under variable conditions.

Comparison of the Cost–Benefit Analysis of Different Metabolic Profiling Strategies. For research purposes, it is attractive to find reliable quantitative biomarkers for genotypic selection, but a maize breeder with limited resources would ask the following questions: How expensive is DIESI–MS

profiling? How much training and manual worktime is required? How many samples can a user process per week or season?

Systematic comparison of several analytical platforms revealed that DIESI–MS allowed for a high throughput of samples, required little time for statistical analysis, and was the cheapest method per metabolic data point (Table 1). The cost per metabolic data point was in the range of 2–10 cents in U.S. dollars. Among all metabolic profiling strategies, DIESI–MS had the greatest potential to be extensively used for plant phenotyping.

DIESI–MS Results from Various Plant Extracts. Because DIESI–MS was generating promising results, several colleges asked: does it work for other samples in addition to maize stems? To answer their questions, we combined the DIESI–MS data from different plant samples (Figure 8). We included previously obtained spectra from coffee leaves (*Coffea arabica* and *Coffea canephora*), red and black tea leaves (*Camellia sinensis*), and coca leaves (*Erythroxylum coca*).¹⁸ Novel data were added from maize stem juice extracts (grown under CT and LN conditions) and also strawberry (*Fragaria vesca*) samples of different maturity stages. A pseudo-extract was also included as a blank control (Figure 8).

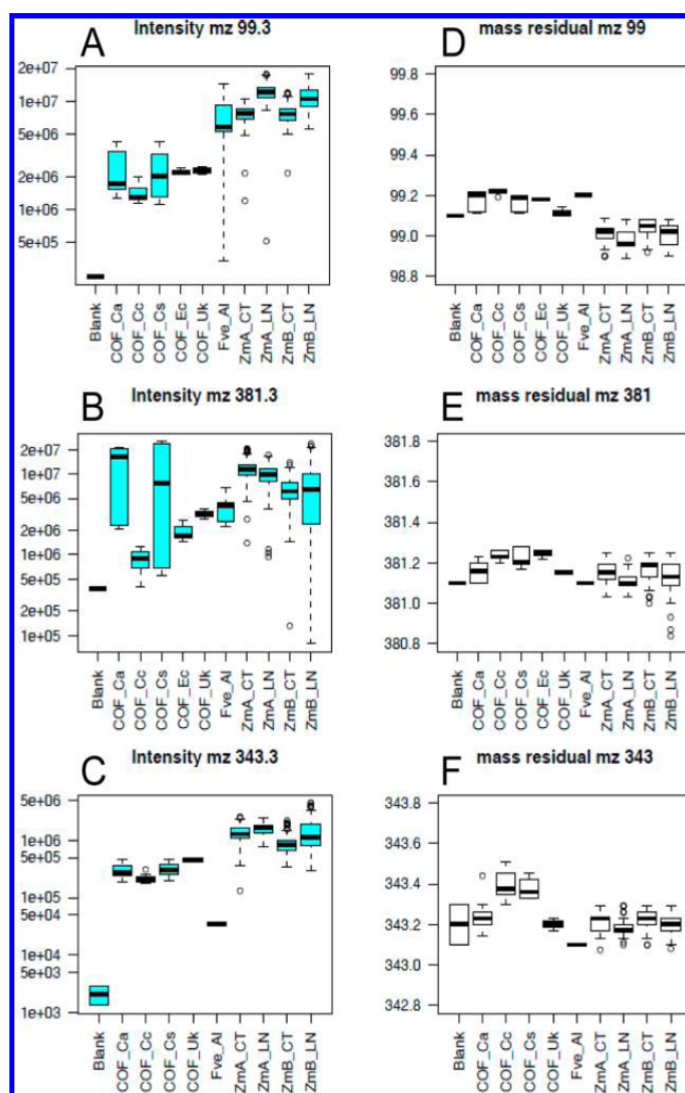


Figure 8. Ion intensity data of nine different plant samples. (A–C) Representative m/z ions. (D–F) Corresponding m/z values. (A) Ion 99.3 was low in leaves of tea samples but was equally abundant in strawberry fruits and maize stem. It increased under LN stress compared to CT but was similar among yellow (ZmA) and white (ZmB) maize hybrids. (B) Ion 381.3 did not change much because of environmental conditions but varied more according to genotype. (C) Ion 343.3 increased slightly because of LN stress in maize but was very low in strawberry fruits. *Zea mays* samples were pooled into white grain genotypes (ZmA) or yellow grain genotypes (ZmB) and grown under either normal CT or LN conditions. Boxplots summarize data from each species or treatment ($n = 6–500$). Data were from coffee leaves [*C. arabica* (COF-Ca) and *C. canephora* (COF-Cc)], tea leaves [*C. sinensis* (COF-Cs)], and coca leaves [*E. coca* (COF-Ec)]¹⁸ as well as strawberry (*F. vesca*) fruit samples of different maturity stages (Fve-Al). A pseudo-extract was used as a blank.

In addition to reporting the intensity values for given m/z ion bins (panels A–C of Figure 8), we also found it relevant to plot the exact m/z values (panels D–F of Figure 8). Peak m/z 99.3 was highest in all maize samples and strawberry fruits, whereas it was lower in all leaf samples (Figure 8A). In maize stem extracts, peak m/z 99.3 increased under LN compared to CT treatment (Figure 8A). The m/z values were similar across leaf samples (Figure 8D). There was a slight tendency that the m/z value was lower in maize, indicating that the metabolite is not necessarily the same in tea leaves or maize stems (Figure 8D).

Peak m/z 381.3 was high in maize stems, intermediate in strawberry fruits, and highly variable in leaves of different

species (Figure 8B). It was higher in *C. arabica* than in *C. canephora* (Figure 8B). In maize stem extracts, peak m/z 381.3 did not change significantly under LN or CT conditions (Figure 8B). The m/z values were very similar across all species and treatments (Figure 8E).

Peak m/z 343.3 was high in maize, intermediate in leaves, and rather low in strawberry fruits (Figure 8B). This metabolite was high in maize stem extracts, particularly low in fruits, and totally absent in coca leaves (Figure 8C). No difference was found between LN or CT conditions for peak m/z 343.3 (Figure 8C). The m/z values varied across plant species, so that most likely intensity values correspond to different metabolites in different samples (Figure 8F).

The metabolite detected in *C. arabica* samples probably corresponds to the metabolite detected in maize and strawberry, but it might be different from the metabolite detected in *C. canephora* and *C. sinensis* (Figure 8F).

We conclude from these analyses that the boxplot figures provide useful quality controls for the DIESI-MS results. Our protocol worked for many different plant samples, and it even allowed for the study of phylogenetic relationships. Some specific ions allowed for the identification of plant organs or species. Other ions allowed for the specific identification of environmental variance or the selection according to genetic variance.

Potential Applications of DIESI-MS for Plant Breeding. It can be concluded that, among all metabolic profiling strategies, DIESI-MS generates a vast amount of information at the cheapest price (Table 1) and, therefore, has the greatest potential to be routinely implemented for biochemical phenotyping of plant cultivars. Because metabolites are closer related to GY than DNA sequences (Figure 1), we postulate that metabolic phenotyping using DIESI-MS complements genome-wide selection (GWS) strategies.

In summary, DIESI-MS is a great phenotyping tool, which will assist plant research and breeding. We expect it will become common in the near future, because there is still large room for methodological improvements to increase information density, bring down cost, and further simplify procedures. Success will depend upon the establishment of strategic alliances between plant physiologists, biochemical researchers, biostatisticians, breeders, and agronomists.

■ ASSOCIATED CONTENT

📄 Supporting Information

Strawberry batch divided in different groups according to their different levels of ripeness, with stage 1 representing the less ripe and stage 3 representing the most ripened fruit (Figure S1) and parameters used to identify the different ripeness stages in a strawberry (*F. vesca*) batch (Table S1). This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: atiessen@ira.cinvestav.mx.

Funding

This work was partially supported by the initiative MASAGRO funded by SAGARPA through Grant IMIC-2012-016 to Axel Tiessen. Axel Tiessen and Robert Winkler acknowledge basic science funding by Consejo Nacional de Ciencia y Tecnología (CONACYT, Mexico). Martín García-Flores, Sheila Juárez-Colunga, and Adrián García-Casarrubias obtained a postgraduate fellowship by CONACYT.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank J. P. Délano and R. Sawers for proofreading the manuscript.

■ ABBREVIATIONS USED

DNA, deoxyribonucleic acid; DIESI-MS, direct-injection electrospray ionization mass spectrometry; GC-FID, gas chromatography with flame ionization detection; GY, grain yield; QTL, quantitative trait loci; *m/z*, mass-to-charge ratio;

NALDI-TOF-MS, nanoassisted laser desorption ionization time-of-flight mass spectrometry; NIRS, near-infrared reflectance spectroscopy; RNA, ribonucleic acid; UPLC, ultra performance liquid chromatography; UV, ultraviolet

■ REFERENCES

- (1) Fiehn, O. Metabolomics—The link between genotypes and phenotypes. *Plant Mol. Biol.* **2002**, *48* (1–2), 155–171.
- (2) Monneveux, P.; Sanchez, C.; Tiessen, A. Future progress in drought tolerance in maize needs new secondary traits and cross combinations. *J. Agric. Sci.* **2008**, *146*, 287–300.
- (3) Moose, S. P.; Mumm, R. H. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiol.* **2008**, *147* (3), 969–977.
- (4) Ribaut, J. M.; de Vicente, M. C.; Delannay, X. Molecular breeding in developing countries: Challenges and perspectives. *Curr. Opin. Plant Biol.* **2010**, *13* (2), 213–218.
- (5) Schauer, N.; Fernie, A. R. Plant metabolomics: Towards biological function and mechanism. *Trends Plant Sci.* **2006**, *11* (10), 508–516.
- (6) Fernie, A. R.; Schauer, N. Metabolomics-assisted breeding: A viable option for crop improvement? *Trends Genet.* **2009**, *25* (1), 39–48.
- (7) Stitt, M.; Fernie, A. R. From measurements of metabolites to metabolomics: An 'on the fly' perspective illustrated by recent studies of carbon–nitrogen interactions. *Curr. Opin. Biotechnol.* **2003**, *14* (2), 136–144.
- (8) Cappiello, A.; Famigliini, G.; Palma, P.; Pierini, E.; Termopoli, V.; Trufelli, H. Direct-EL in LC-MS: Towards a universal detector for small-molecule applications. *Mass Spectrom. Rev.* **2011**, *30* (6), 1242–1255.
- (9) Geigenberger, P.; Tiessen, A.; Meurer, J. Use of non-aqueous fractionation and metabolomics to study chloroplast function in *Arabidopsis*. *Methods Mol. Biol.* **2011**, *775*, 135–160.
- (10) Zwiener, C.; Frimmel, F. H. LC-MS analysis in the aquatic environment and in water treatment—A critical review—Part I: Instrumentation and general aspects of analysis and detection. *Anal. Bioanal. Chem.* **2004**, *378* (4), 851–861.
- (11) Bino, R. J.; Hall, R. D.; Fiehn, O.; Kopka, J.; Saito, K.; Draper, J.; Nikolau, B. J.; Mendes, P.; Roessner-Tunali, U.; Beale, M. H.; Trethewey, R. N.; Lange, B. M.; Wurtele, E. S.; Sumner, L. W. Potential of metabolomics as a functional genomics tool. *Trends Plant Sci.* **2004**, *9* (9), 418–425.
- (12) Witt, S.; Galicia, L.; Liseic, J.; Cairns, J.; Tiessen, A.; Araus, J. L.; Palacios-Rojas, N.; Fernie, A. R. Metabolic and phenotypic responses of greenhouse-grown maize hybrids to experimentally controlled drought stress. *Mol. Plant* **2012**, *5* (2), 401–417.
- (13) Roessner, U.; Luedemann, A.; Brust, D.; Fiehn, O.; Linke, T.; Willmitzer, L.; Fernie, A. R. Metabolic profiling allows comprehensive phenotyping of genetically or environmentally modified plant systems. *Plant Cell* **2001**, *13* (1), 11–29.
- (14) Fiehn, O. Metabolite profiling in *Arabidopsis*. *Methods Mol. Biol.* **2006**, *323*, 439–447.
- (15) Meyer, R. C.; Steinfath, M.; Liseic, J.; Becher, M.; Witucka-Wall, H.; Torjek, O.; Fiehn, O.; Eckardt, A.; Willmitzer, L.; Selbig, J.; Altmann, T. The metabolic signature related to high plant growth rate in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104* (11), 4759–4764.
- (16) Liseic, J.; Meyer, R. C.; Steinfath, M.; Redestig, H.; Becher, M.; Witucka-Wall, H.; Fiehn, O.; Torjek, O.; Selbig, J.; Altmann, T.; Willmitzer, L. Identification of metabolic and biomass QTL in *Arabidopsis thaliana* in a parallel analysis of RIL and IL populations. *Plant J.* **2008**, *53* (6), 960–972.
- (17) Kind, T.; Fiehn, O. Metabolomic database annotations via query of elemental compositions: Mass accuracy is insufficient even at less than 1 ppm. *BMC Bioinf.* **2006**, *7*, 234–243.
- (18) Garcia-Flores, M.; Juarez-Colunga, S.; Montero-Vargas, J. M.; Lopez-Arciniega, J. A. I.; Chagolla, A.; Tiessen, A.; Winkler, R.

Evaluating the physiological state of maize (*Zea mays* L.) plants by direct-injection electrospray mass spectrometry (DIESI-MS). *Mol. BioSyst.* 2012, 8 (6), 1658–1660.

(19) Montero-Vargas, J. M.; Gonzalez-Gonzalez, L. H.; Galvez-Ponce, E.; Ramirez-Chavez, E.; Molina-Torres, J.; Chagolla, A.; Montagnon, C.; Winkler, R. Metabolic phenotyping for the classification of coffee trees and the exploration of selection markers. *Mol. BioSyst.* 2013, 9 (4), 693–699.

(20) Spielbauer, G.; Armstrong, P.; Baier, J. W.; Allen, W. B.; Richardson, K.; Shen, B.; Settles, A. M. High-throughput near-infrared reflectance spectroscopy for predicting quantitative and qualitative composition phenotypes of individual maize kernels. *Cereal Chem.* 2009, 86 (5), 556–564.

(21) Tiessen, A.; Hendriks, J. H. M.; Stitt, M.; Branscheid, A.; Gibon, Y.; Farre, E. M.; Geigenberger, P. Starch synthesis in potato tubers is regulated by post-translational redox modification of ADP-glucose pyrophosphorylase: A novel regulatory mechanism linking starch synthesis to the sucrose supply. *Plant Cell* 2002, 14 (9), 2191–2213.

(22) Angeles-Núñez, J.; Tiessen, A. *Arabidopsis* sucrose synthase 2 and 3 modulate metabolic homeostasis and direct carbon towards starch synthesis in developing seeds. *Planta* 2010, 232 (3), 701–718.

(23) R Development Core Team, R. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2008; <http://www.R-project.org>.

10.3 Anexos