

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 434 737**

51 Int. Cl.:

**G01N 33/50** (2006.01)

**G01N 33/68** (2006.01)

**C12N 5/00** (2006.01)

**G06F 19/12** (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.10.2010 E 10014005 (2)**

97 Fecha y número de publicación de la concesión europea: **18.09.2013 EP 2447717**

54 Título: **Procedimiento rápido de selección de células (líneas celulares) dirigida**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**17.12.2013**

73 Titular/es:

**LONZA BIOLOGICS PLC. (100.0%)  
228-230 Bath Road  
Slough SL1 4DX, GB**

72 Inventor/es:

**LANG, DIETMAR;  
MARTIN, ELAINE B;  
MONTAGUE, GARY A;  
O'MALLEY, CHRISTOPHER J;  
ROOT, TRACY S;  
TRIM, CAROL M;  
POVEY, JANE F;  
SMALES, CHRISTOPHER M y  
RACHER, ANDREW J**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

ES 2 434 737 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

## DESCRIPCIÓN

Procedimiento rápido de selección de células (líneas celulares) dirigida

- 5 La presente invención se refiere a un proceso *in vitro* para la predicción de la productividad específica celular, del recuento integral de células viables o de los datos de concentración del producto celular de células de muestra, a un proceso *in vitro* para el aislamiento de dichas células y a un dispositivo para la predicción de la productividad celular específica, para el recuento integral de células viables o para los datos de concentración del producto celular de células de muestra.
- 10 Según la técnica anterior, los procedimientos para el aislamiento de líneas celulares recombinantes de mamífero con las propiedades de elaboración deseadas son ineficaces tanto desde el punto de vista de los recursos como por la capacidad para aislar las combinaciones específicas de las características deseadas.
- 15 Por ejemplo, Adrichem y col. (Anal. Chem., 1998, 70, 923 - 930) desvelan las investigaciones sobre patrones de proteínas y células de mamífero en sobrenadantes de cultivo mediante espectrometría de masas de desorción/ionización por láser asistida por matriz (espectrometría de masas MALDI). Dicha espectrometría de masas MALDI puede usarse para la caracterización de proteínas mediante la monitorización de las proteínas que han sido excretadas al medio o que están presentes en lisados celulares. El intervalo de masas detectable por la
- 20 espectrometría de masas MALDI es desde 16.000 hasta varios cientos de miles de Daltons. Los resultados así obtenidos son complementarios de la electroforesis de SDS-PAGE estándar. Por lo tanto, estos procedimientos pueden usarse, por ejemplo, para monitorizar el cultivo a gran escala de células de hibridoma que expresan un anticuerpo del tipo IgG.
- 25 Zhang y col. (J. Am. Soc. Mass Spectrom., 2006, 17, 490 - 499) desvelan la identificación de líneas celulares de mamífero usando espectrometría de masas MALDI - TOF y CL - ESI - EM / EM. Se establece que la espectrometría de masas MALDI - TOF es eficaz para la caracterización de péptidos directamente a partir de una única célula. El análisis mediante CL - ESI - EM / EM puede proporcionar una información útil de la secuencia después de una digestión triptica de las muestras celulares. Se averiguó que producía unos patrones de MALDI - EM únicos y
- 30 reproducibles, que pueden ser usados como huella molecular para identificar y distinguir las diferentes líneas celulares medidas. Sin embargo, los espectros de EM obtenidos se comparaban visualmente sobre la base de los claros picos visibles de EM para distinguir los tres diferentes tipos celulares de mamífero. Alternativamente, Zhang y col. demuestran que una técnica diferente - una combinación de cromatografía líquida seguida de una ionización por electronebulización y espectroscopía de masas en tándem (CL - ESI - EM / EM), para la que es necesario digerir las
- 35 muestras - también es útil para arrojar luz sobre las diferencias en los perfiles de los proteomas.
- Feng y col. (Rapid Commun. Mass Spectrom., 2010, 24, 1226 - 1230) desvelan una rápida caracterización de células CHO altas/bajas productoras usando una desorción/ionización por láser asistida por matriz con analizador del tiempo de vuelo (MALDI - TOF). El proceso desvelado en este documento es capaz de distinguir entre células
- 40 altas y bajas productoras cuando son producidas en el mismo cultivo a la misma escala mediante la aplicación de procedimientos estadísticos, a saber, el análisis de los componentes principales (PCA) y los mínimos cuadrados parciales lineales (PLS), para analizar los espectros de MALDI - TOF. Especialmente, el procedimiento según Feng y col. permite distinguir entre datos de productividad procedentes de diferentes líneas celulares que producen una proteína recombinante de IFN-gamma a la misma escala, es decir, crecidas a baja escala. Según Feng y col., esta metodología podría usarse posiblemente para predecir la productividad celular. La PLS lineal usada deriva su
- 45 utilidad de la capacidad para analizar datos con muchas variables, lo que es relevante para hallar subfamilias de líneas celulares a una escala dada.
- Según se mencionó anteriormente, los procedimientos según la técnica anterior, especialmente los de Feng y col., simplemente enseñan que los datos de MALDI - TOF que son analizados mediante programas estadísticos, a saber, PLS y PCA, pueden usarse para la diferenciación entre células altas y bajas productoras a una escala de cultivo
- 50 dada.
- Sin embargo, la técnica anterior no consigue enseñar un procedimiento que prediga las características celulares de una célula desconocida en una etapa posterior de la amplificación, en particular, en biorreactores de gran volumen, mientras estas células aún se están cultivando en un medio con un bajo volumen. Especialmente, los conocidos programas estadísticos de PCA y PLS producen unos datos que son insuficientes para ser usados como base para una predicción precisa y fiable de las características celulares en una etapa posterior de amplificación.
- 55
- 60 Una línea celular que expresa en la etapa temprana un balance específico de proteínas pueden mostrar, por ejemplo, una elevada productividad en esta etapa, pero después de la amplificación, en la etapa del biorreactor, esta productividad puede deteriorarse debido a diferentes factores que son, entre otros, fuerzas de cizallamiento, efectos del volumen, diferentes tipos o formatos de fermentador, parámetros de cultivo, por ejemplo, pH, y diferencias de densidad celular controladas por gas.
- 65
- Por lo tanto, existe una necesidad de un procedimiento con la capacidad de cribar rápidamente grandes cantidades

de líneas celulares tempranamente en el desarrollo de la línea celular y cultivadas únicamente en pequeños volúmenes, que sea capaz de predecir las características específicas de crecimiento y de producción, por ejemplo, unas productividades volumétricas aumentadas de dicha línea celular a una escala de cultivo dada y deseada, en particular a una escala de biorreactor, lo que significa a gran escala de volumen.

5 Por lo tanto, el problema técnico subyacente de la presente invención es proporcionar un procedimiento para superar los problemas identificados anteriormente, en particular, proporcionar un procedimiento para la predicción del cultivo celular, en particular de un biorreactor, del rendimiento de una célula con características celulares desconocidas, en particular ya en una etapa temprana de la amplificación, siendo adicionalmente ahorrador de tiempo y reductor de coste, pero con una elevada precisión de predicción para dicho rendimiento en un cultivo a una escala en gran volumen y/o en unas condiciones de alta productividad, en particular a una escala de un biorreactor.

Este problema es resuelto por las reivindicaciones independientes de la presente invención, que definen el procedimiento y el dispositivo en su sentido más amplio.

15 La presente invención proporciona por lo tanto un ventajoso proceso *in vitro* para la predicción precisa y fiable de la productividad celular específica, del recuento celular integral viable o de los datos de concentración del producto celular de al menos una célula de muestra con una productividad celular específica desconocida, un recuento integral de células viables o datos de concentración del producto celular, que permite el ahorro de tiempo y la reducción de costes en la predicción de la productividad celular específica en particular, del recuento integral de células viables o de los datos de concentración del producto celular. También puede prepararse una célula, y en particular aislarse, mediante dicho procedimiento, en el que dicha célula está caracterizada particularmente por un cultivo celular deseado, preferiblemente en biorreactor, unos datos de rendimiento, tales como una elevada productividad. Adicionalmente, la presente invención proporciona un dispositivo adaptado para proporcionar una predicción de la productividad celular específica, del recuento integral de células viables o de los datos de concentración del producto celular, cuando se suministra una muestra de al menos una célula de muestra que comprende: (a) un medio adaptado para someter una muestra de la al menos una célula de muestra a un análisis por EM (espectrometría de masas) para obtener unos datos de muestra en bruto de EM de la misma, (b) un medio adaptado para someter los datos en bruto de un estándar y en bruto de la muestra de EM a al menos un primer procedimiento de procesado de la señal de EM para obtener unos perfiles de EM pretratados del estándar y de la muestra, y (c) un medio adaptado para someter los datos de la productividad específica celular, del recuento integral de células viables o de la concentración del producto celular procedentes de una célula estándar, y los perfiles de EM pretratados de la muestra y del estándar, a un segundo procedimiento de procesado de la señal de EM que incluye un PLS - DA (un análisis discriminatorio por mínimos cuadrados parciales) basado en la evaluación comparativa, para predecir los datos de la productividad específica celular, del recuento integral de células viables o de la concentración del producto celular de la célula de muestra.

40 Al contrario que la técnica anterior, que habitualmente aumenta la escala desde una placa de 96 pocillos en primer lugar hasta una placa de 24 pocillos, después a un matraz con agitación y finalmente a escala de biorreactor, la presente invención obvia las etapas de cultivo en una placa de 24 pocillos y en un matraz con agitación. Directamente, después de la predicción de la productividad específica celular, el recuento integral de células viables o los datos de concentración del producto celular de al menos una célula de muestra cultivada en un volumen bajo, por ejemplo, en una placa de 96 pocillos, puede realizarse el cultivo en un biorreactor.

45 Según la presente invención, el PLS - DA permite la separación de unas clases muy específicas de observaciones sobre la base de una variable, de forma que mediante el uso del PLS - DA, se supera el problema de la clasificación de la productividad según las líneas celulares y puede predecirse su productividad específica celular, su recuento integral de células viables o la concentración de producto celular a diferentes escalas.

50 Por claridad, se entiende que el término "datos de rendimiento del biorreactor" se refiere a los datos sobre el comportamiento y las características de una célula, cuando dicha célula se cultiva o se reproduce en una condición de gran volumen y/o en unas condiciones de elevada productividad, en particular en un biorreactor. Los datos de rendimiento del biorreactor son preferiblemente datos sobre la productividad celular individual y específica para un producto celular específico, por ejemplo, una proteína, en particular un anticuerpo, un fragmento de anticuerpo o anticuerpo fusionado, antibióticos, necesidades del cultivo, crecimiento o esperanza de vida de dicha célula. En una forma de realización preferida, el producto celular es una proteína, en particular un anticuerpo, un péptido, un proteoglicano, una glucoproteína, un carbohidrato, un lípido, un antibiótico o una hormona.

60 El término "célula estándar" significa una célula con unos datos de rendimiento en cultivo celular conocidos, en particular unas características y un comportamiento conocidos a una escala dada, preferiblemente en una escala a gran volumen y/o unas condiciones de elevada productividad, en particular en una escala de un biorreactor. Estas características conocidas fueron medidas y analizadas mediante los procedimientos según al estado de la técnica. Por ejemplo, la productividad celular puede ser determinada mediante un ELISA (enzimoinmunoensayo de adsorción).

65 En el contexto de la presente invención, el término "célula de muestra" se refiere a una célula con al menos un dato

de rendimiento en cultivo celular desconocido, en particular, una característica celular. Preferiblemente, al menos una característica de dicha muestra es conocida.

5 En el contexto de la presente invención, el término "bajo volumen de un medio" significa que el medio tiene preferiblemente un volumen de desde preferiblemente desde 1  $\mu$ l hasta desde 100 l, preferiblemente desde 1  $\mu$ l hasta 90 l, preferiblemente desde 1  $\mu$ l hasta 80 l, preferiblemente desde 1  $\mu$ l hasta 70 l, preferiblemente desde 1  $\mu$ l hasta 60 l, preferiblemente desde 1  $\mu$ l hasta 50 l, preferiblemente desde 1  $\mu$ l hasta 40 l, preferiblemente desde 1  $\mu$ l hasta 30 l, preferiblemente desde 1  $\mu$ l hasta 20 l, preferiblemente desde 1  $\mu$ l hasta desde 10 l, preferiblemente desde 1  $\mu$ l hasta 5 l, preferiblemente desde 1  $\mu$ l hasta 4 l, preferiblemente desde 1  $\mu$ l hasta 3 l, preferiblemente desde 1  $\mu$ l hasta 2 l, preferiblemente desde 1  $\mu$ l hasta desde 1  $\mu$ l hasta desde 1 l, preferiblemente desde 1  $\mu$ l hasta 0.5 l, preferiblemente desde 1  $\mu$ l hasta 0,4 l, preferiblemente desde 1  $\mu$ l hasta 0,3 l, preferiblemente desde 1  $\mu$ l hasta 0,2 l, preferiblemente desde 1  $\mu$ l hasta 0,1 l, preferiblemente desde 1  $\mu$ l hasta 90 ml, preferiblemente desde 1  $\mu$ l hasta 80 ml, preferiblemente desde 1  $\mu$ l hasta 70 ml, preferiblemente desde 1  $\mu$ l hasta 60 ml, preferiblemente desde 1  $\mu$ l hasta 50 ml, preferiblemente desde 1  $\mu$ l hasta 40 ml, preferiblemente desde 1  $\mu$ l hasta 30 ml, preferiblemente desde 1  $\mu$ l hasta 20 ml, preferiblemente desde 1  $\mu$ l hasta desde 10 ml, preferiblemente desde 1  $\mu$ l hasta 5 ml, preferiblemente desde 1  $\mu$ l hasta 4 ml, preferiblemente desde 1  $\mu$ l hasta 3 ml, preferiblemente desde 1  $\mu$ l hasta 2 ml, preferiblemente desde 1  $\mu$ l hasta desde 1 ml, preferiblemente desde 1 hasta 999  $\mu$ l, preferiblemente desde 1  $\mu$ l hasta 0,5 ml, preferiblemente desde 1  $\mu$ l hasta 0,4 ml, preferiblemente desde 1  $\mu$ l hasta 0,3 ml, preferiblemente desde 1  $\mu$ l hasta 0,2 ml, preferiblemente desde 1  $\mu$ l hasta 0,1 ml, preferiblemente desde 1  $\mu$ l hasta 50  $\mu$ l, preferiblemente desde 1  $\mu$ l hasta 40  $\mu$ l, preferiblemente desde 1  $\mu$ l hasta 30  $\mu$ l, preferiblemente desde 1  $\mu$ l hasta 20  $\mu$ l, preferiblemente 1  $\mu$ l hasta 10  $\mu$ l, preferiblemente desde 20 hasta 90  $\mu$ l, preferiblemente 20 hasta 80  $\mu$ l, preferiblemente desde 20 hasta 70  $\mu$ l, preferiblemente desde 20 hasta 60  $\mu$ l, preferiblemente desde 20 hasta 50  $\mu$ l, preferiblemente desde 20 hasta 40  $\mu$ l, preferiblemente desde 20 hasta 30  $\mu$ l, preferiblemente desde 30 hasta 70  $\mu$ l, preferiblemente desde 30 hasta 60  $\mu$ l, preferiblemente desde 30 hasta 50  $\mu$ l, preferiblemente desde 30 hasta 40  $\mu$ l, preferiblemente desde 10 hasta 100  $\mu$ l, preferiblemente desde 10 hasta 90  $\mu$ l, preferiblemente desde 10 hasta 80  $\mu$ l, preferiblemente desde 10 hasta 70  $\mu$ l, preferiblemente desde 10 hasta 60  $\mu$ l, preferiblemente desde 10 hasta 50  $\mu$ l, preferiblemente desde 10 hasta 40  $\mu$ l, preferiblemente desde 10 hasta 30  $\mu$ l, preferiblemente desde 10 hasta 20  $\mu$ l, preferiblemente desde 40 hasta 60  $\mu$ l, preferiblemente desde 40 hasta 50  $\mu$ l.

30 En el contexto de la presente invención, el término "etapa temprana de amplificación" se entiende como un punto temporal en el que las células son cultivadas en un bajo volumen de un medio, según se ha definido anteriormente.

35 En el contexto de la presente invención, el término "biorreactor" significa un contenedor capaz de contener células para la producción de al menos un producto celular deseado, que preferiblemente permite una elevada productividad desde el punto de vista de velocidad de producción y/o de cantidad de dicho producto celular deseado. En una forma de realización particularmente preferida, el biorreactor es un dispositivo o un sistema que sustenta un entorno biológicamente activo. En una forma de realización particularmente preferida, un biorreactor de la presente invención es un contenedor adecuado para la producción industrial y comercial de dicho producto celular de interés. En una forma de realización particularmente preferida, dicho contenedor es capaz de crear unas condiciones de cultivo celular adecuadas para la producción del producto celular de interés con una elevada productividad. En una forma de realización particularmente preferida, el biorreactor contiene un volumen de medio de al menos 10 l, de al menos 20 l, de al menos 50 l, de al menos 100 l, de al menos 200 l, de al menos 300 l, de al menos 400 l, de al menos 500 l, de al menos 600 l, de al menos 700 l, de al menos 800 l, de al menos 900 l, de al menos 1.000 l, en particular de al menos 2.000 l, de al menos 3.000 l, o de al menos 4.000 l.

45 En el contexto de la presente invención, se entiende que la expresión "y/o" usada entre dos elementos designa que ambos elementos unidos por dicho término se denominan de una forma acumulativa o alternativa. Por lo tanto, la expresión "A y/o B" engloba los significados "A o B" y "A y B", lo que significa "cualquiera de A, B o ambos".

50 Por lo tanto, la presente invención proporciona un procedimiento *in vitro* que permite a una persona experta predecir las características de al menos una célula de muestra, especialmente de al menos una línea celular, proporcionando en una primera etapa a) una muestra de al menos una célula de muestra cultivada en un medio con un bajo volumen y una baja concentración celular en el mismo, preferiblemente de entre  $10^4$  y  $10^9$ , los datos de EM de la productividad específica celular, del recuento integral de células viables o de la concentración del producto celular de una célula estándar y en bruto estándar de la misma. Posteriormente, en una segunda etapa, la al menos una célula de muestra es analizada mediante un procedimiento de espectrometría de masas. En una tercera etapa, tanto los datos en bruto de la EM de la al menos una célula de muestra y la célula estándar se tratan mediante un procedimiento de procesamiento de la señal de EM para obtener unos perfiles pretratados de la EM de la al menos una célula de muestra y de la célula estándar. En una cuarta etapa se somete la productividad específica celular, el recuento integral de células viables o los datos de concentración del producto celular procedentes de la célula estándar y de los perfiles de EM pretratados de la muestra y del estándar a un procedimiento estadístico, a saber, PLS- DA. Este programa estadístico se emplea para comparar y evaluar los perfiles pretratados de la EM de la célula de muestra con los perfiles pretratados de la EM de una célula estándar.

En una forma de realización preferida de la presente invención, la concentración de las células es desde  $10^4$  hasta  $10^8$ , desde  $10^5$  hasta  $10^7$ , en particular desde  $10^6$  hasta  $10^7$  células viables/ml de medio de cultivo una vez finalizado el cultivo.

5 La presente invención permite la producción de la productividad específica celular, el recuento integral de células viables o los datos de concentración del producto celular, por ejemplo, la productividad, de la al menos una célula de muestra está siendo cultivada en una etapa temprana de la amplificación. El procedimiento según la presente invención permite el rápido desarrollo de, por ejemplo, unas líneas celulares altamente productivas, y reduce los costes de elaboración de proteínas terapéuticas y acelera el desarrollo de productos farmacéuticos.

10 El presente proceso proporciona una nueva herramienta de cribado para la identificación, preferiblemente temprana en el ciclo de desarrollo, de un producto, es decir, en una fase en la que se usa un medio de cultivo con un volumen de desde 1 hasta 999  $\mu$ l, de células, preferiblemente de líneas celulares, que tienen las propiedades deseadas, por ejemplo, una elevada productividad volumétrica, en particular en biorreactores a gran escala, en particular en biorreactores que contienen un medio de cultivo con un volumen superior a 10 litros. Además, el procedimiento según la presente invención mejora la probabilidad de encontrar una célula, especialmente una línea celular, con unas características celulares específicas, por ejemplo, que comprenda una elevada productividad del producto, preferiblemente de al menos 1 g/l, preferiblemente de al menos 5 g/l y muy preferiblemente de 10 g/l, tempranamente en el desarrollo. Además de la identificación de las células valiosas, preferiblemente de líneas celulares, el procedimiento también puede usarse preferiblemente para aislar nuevas células hospedadoras, preferiblemente líneas celulares, con unas propiedades mejoradas, por ejemplo, para la elaboración de proteínas terapéuticas, especialmente la elaboración de anticuerpos monoclonales. En particular, mediante la combinación del análisis por EM con el procedimiento estadístico de PLS - DA, preferiblemente mediante la combinación de PLS - DA con el PCA (análisis de componentes principales), el presente proceso permite la identificación de patrones a diferentes niveles de productividad del producto, para predecir las productividades a diferentes escalas de los mismos de una forma rápida y, si fuera aplicable, automática.

La elevada probabilidad de encontrar células altamente productoras, preferiblemente líneas celulares, tiene el potencial de reducir el número de células, preferiblemente de líneas celulares, que se necesita cribar antes de identificar la adecuada para la elaboración. Al reducir el número de células cribadas, preferiblemente de líneas celulares, se reducen los materiales requeridos. Consecuentemente se requerirán menos fuentes durante la construcción de la célula, preferiblemente de la línea celular, con la concomitante menor generación de materiales de desecho. Las células altamente productoras, preferiblemente líneas celulares, reducirán el número de cultivos en biorreactor requeridos para suministrar al mercado las necesidades del producto. El presente proceso es capaz de reducir los requisitos de los materiales de partida, en particular los costes, del proceso de producción, especialmente del agua que se usa tanto como material de partida como en la limpieza y la higienización del equipo.

Según la presente invención, los datos del rendimiento del biorreactor son la productividad específica celular, el recuento integral de células viables o la concentración del producto celular. En una forma de realización particularmente preferida, el dato de rendimiento del biorreactor es la productividad específica celular.

En una forma de realización preferida de la presente invención, la célula de muestra y/o la célula estándar se elige de entre el grupo que consiste en líneas celulares humanas, líneas celulares animales, líneas celulares vegetales, células de hongos, células de bacterias, células de levadura y células madre.

En una forma de realización preferida de la presente invención, tanto la célula de muestra como la célula estándar se eligen de entre el mismo tipo celular, especialmente la línea celular o la cepa.

En una forma de realización preferida de la presente invención, la célula de muestra es una línea celular CHO-K1 modificada. En una forma de realización preferida de la presente invención, la célula estándar es también una línea celular CHO-K1 modificada.

En una forma de realización preferida de la presente invención, el análisis por EM usado en la etapa b) se elige de entre el grupo que consiste en MALDI - TOF, CL - ESI - EM (espectrometría de masas con cromatografía líquida de ionización por electronebulización) y CL - ESI - EM / EM (cromatografía líquida acoplada en tándem a una espectrometría de masas con ionización por electronebulización).

En una forma de realización preferida de la presente invención, el análisis por EM usado en la etapa b) es MALDI - TOF. Las células de muestra sometidas a una MALDI - TOF no necesitan ser digeridas con, por ejemplo, tripsina, pero pueden ser incluidas en la matriz en su forma intacta mediante una preparación muy simple de las células.

En una forma de realización preferida de la presente invención, el análisis por EM usado en la etapa b) es CL - ESI - EM. El análisis mediante CL - ESI - EM proporciona una gran discriminación en particular entre las líneas celulares desde el punto de vista de la productividad, del crecimiento y de otras características deseables. El análisis mediante CL - ESI - EM proporciona una dimensión extra de información.

En una forma de realización preferida de la presente invención, los datos en bruto de la EM estándar se obtienen mediante un análisis por EM seleccionado de entre el grupo que consiste en MALDI - TOF, CL - ESI - EM y CL - ESI - EM / EM, preferiblemente MALDI - TOF.

5 En una forma de realización preferida de la presente invención, la ionización usada para la EM por MALDI - TOF o para la CL - ESI - EM se lleva a cabo en un modo de reflexión negativo o positivo o en un modo lineal positivo o negativo que sea óptimo según los parámetros específicos del instrumento, por ejemplo, siendo dependientes del instrumento, con o sin supresión de masa y extracción de ión pulsado.

10 En una forma de realización preferida de la presente invención, se usan los siguientes parámetros del instrumento espectrómetro de masas de la MALDI - TOF para el procedimiento según la presente invención:

15	Polaridad:	+ve
	Supresión a:	1.000 Da (Daltons)
	Intervalo:	desde 1.000 hasta 50.000 Da

20 En una forma de realización preferida de la presente invención, la supresión de masa durante el análisis mediante EM de la MALDI - TOF y de la CL - ESI - EM está por debajo de 500 Da, preferiblemente por debajo de 1.000 Da y lo más preferido por debajo de 1.500 Da.

En una forma de realización preferida de la presente invención, la supresión de masa durante el análisis mediante EM de la MALDI - TOF y de la CL - ESI - EM está por encima de 500 Da, preferiblemente por encima de 1.000 Da y lo más preferido por encima de 1.500 Da.

25 En una forma de realización preferida de la presente invención, el intervalo detectado durante el análisis mediante EM de la MALDI - TOF y de la CL - ESI - EM es desde 200 hasta 100.000 Da, preferiblemente desde 500 hasta 50.000 Da, preferiblemente desde 1.000 hasta 100.000 Da, 1.000 hasta 18.000 Da, preferiblemente desde 500 hasta 10.000 Da y lo más preferiblemente desde 200 hasta 8000 Da.

30 En una forma de realización preferida de la presente invención, la al menos una célula de muestra se lava, preferiblemente con una disolución tampón, antes de ser sometida al análisis mediante una EM de la MALDI - TOF.

35 En una forma de realización preferida de la presente invención, la al menos una célula de muestra se lava únicamente con disolución salina tamponada con fosfato (PBS) o seguido de una etapa de lavado con una disolución acuosa de sacarosa, en particular con desde 0,2 hasta 0,7 M, preferiblemente desde 0,3 hasta 0,5 M, preferiblemente desde 0,35 M de sacarosa, antes de ser sometida al análisis mediante una EM de la MALDI - TOF.

40 En una forma de realización preferida de la presente invención, la matriz usada para el análisis de MALDI - TOF es ácido sinapínico (SA). Según la presente invención, el uso del ácido sinapínico (SA) como matriz para el análisis de MALDI - TOF mediante EM proporciona espectros ventajosos con un particular amplio intervalo de picos, preferiblemente de hasta 70 kDa, y estando particularmente bien resueltos En una forma de realización adicional de la presente invención, también puede usarse ácido 2,5-dihidroxibenzoico (DHB) como matriz.

45 En una forma de realización preferida de la presente invención, la muestra de la al menos una célula de muestra sometida en la etapa b) comprende no más de  $1 \times 10^6$  células, preferiblemente entre  $0,015 \times 10^6$  y  $0,0625 \times 10^6$  células, en particular  $0,03 \times 10^6$  células.

50 En una forma de realización preferida de la presente invención, los perfiles de la EM se toman después de entre 1 y 5 horas, preferiblemente entre 1 y 4 horas, en particular entre 1 y 3 horas de aclimatación a baja temperatura, por ejemplo, a entre 0 y 10 °C, preferiblemente a entre 2 y 8 °C, en particular a 4 °C, dando como resultado la mejor reproductibilidad y unos espectros de masa que contienen una menor proporción entre señal y ruido.

55 En una forma de realización preferida de la presente invención, las células de muestra se someten al análisis mediante EM en un momento específico del crecimiento. Los momentos de toma de muestra preferidos son en la fase logarítmica media y/o final del crecimiento celular.

60 En una forma de realización preferida de la presente invención, la señal de los datos en bruto de la muestra de la EM obtenidos en la etapa b) y/o de los datos en bruto de la muestra estándar EM proporcionados en la etapa a) es procesada mediante una operación elegida del grupo que consiste en corrección de la línea de base, normalización, alineación, filtrado y recorte.

65 Los perfiles de los datos de la EM muestran normalmente una línea de base variada debido a aspectos tales como el ruido químico en la matriz de MALDI y la sobrecarga de iones. Esto es indeseable cuando se usan técnicas de análisis de datos para comparar los perfiles de la EM como matriz de distancia utilizada para medir la similitud entre perfiles. Por lo tanto, en una forma de realización preferida de la presente invención, los datos en bruto de la señal de EM de la muestra y/o del estándar son procesados mediante la operación de corrección de la línea de base.

Otro fenómeno observado habitualmente en los perfiles de la EM es la variación en la amplitud de las intensidades iónicas. Esto puede ser provocado por varios factores, tales como variaciones en la preparación de la muestra o cambios en la sensibilidad del instrumento. Por lo tanto, en una forma de realización preferida de la presente invención, los datos en bruto de la señal de la EM de la muestra y/o del estándar son procesados mediante la operación de normalización.

La alineación de los picos se usa para corregir la variación entre los valores observados de M/Z y el tiempo de vuelo real. Estos errores se producen habitualmente como resultado de errores de calibración y pueden observarse como un desplazamiento sistemático entre picos. Por lo tanto, en una forma de realización preferida de la presente invención, los datos en bruto de la señal de la EM de la muestra y/o del estándar son procesados mediante la operación de alineación de los picos.

El filtrado de los perfiles de la EM se realiza suavizando la señal, preferiblemente mediante un filtro de Savitzky-Golay. Por lo tanto, en una forma de realización preferida de la presente invención, los datos en bruto de la señal de EM de la muestra y/o del estándar son procesados mediante la operación de filtrado.

El recorte de los perfiles de la EM se realiza para eliminar partes de la señal que contienen poca o ninguna información. Preferiblemente, se elimina de los espectros de EM el intervalo de desde 0 hasta 500 unidades m/z. Por lo tanto, en una forma de realización preferida de la presente invención, los datos en bruto de la señal de EM de la muestra y/o del estándar son procesados mediante la operación de recorte.

En una forma de realización preferida de la presente invención, la muestra de la al menos una célula de muestra es re-muestreada. Esta etapa específica preferida del proceso permite muestrear por aumento y muestrear por disminución la señal original, mientras se conserva la información contenida en el espectro, es decir, alterando la cantidad de canales de los diferentes puntos de datos medidos. Normalmente, el re-muestreado se utiliza en situaciones en las que la alta resolución de la señal original de EM sería considerada poco práctica para trabajar con ella debido a restricciones computacionales tales como la carencia de memoria en el ordenador. El re-muestreado también puede usarse para crear un intervalo coherente de m/z, que facilite la alineación de múltiples espectros.

En una forma de realización preferida de la presente invención, los datos en bruto de la EM del estándar y de la muestra son analizados visualmente.

En una forma de realización preferida de la presente invención, los perfiles pretratados de la EM del estándar y de la muestra obtenidos en la etapa c) son analizados ópticamente.

En una forma de realización preferida de la presente invención, la célula de muestra con los datos predichos de la productividad específica celular, el recuento integral de células viables o la concentración del producto celular integral viable evaluados en la etapa d) se cultiva en un cultivo celular, preferiblemente en un biorreactor, de forma que se verifiquen sus datos de productividad específica celular, de recuento integral de células viables o de concentración del producto celular.

En una forma de realización preferida de la presente invención, los perfiles en bruto de la muestra de la EM obtenidos en la etapa b) y los datos verificados de la productividad específica celular, del recuento integral de células viables o de la concentración del producto celular de la célula de muestra se usan en la etapa a), los datos estándar de la EM y los datos de la productividad específica celular, el recuento integral de células viables o la concentración del producto celular integral viable procedentes de la célula estándar. Al proporcionar un elevado número de perfiles de la EM a partir de las células estándar con sus datos conocidos de la productividad específica celular, del recuento integral de células viables o de la concentración del producto celular integral viable, la probabilidad de la predicción es más precisa y fiable.

En una forma de realización preferida de la presente invención, el modelo de PLS - DA según la etapa (d) requiere dos conjuntos diferentes de información, a saber, el bloque x y el bloque y. En una forma de realización preferida de la presente invención, el bloque x contiene la información procedente de los datos de la EM de la muestra pretratada generados en la etapa del proceso de la 96 DWP (placa de pocillos profundos). En una forma de realización preferida de la presente invención, cada dato de la EM de la muestra pretratada se trata como una muestra, con las intensidades de la señal registradas en un intervalo específico de valores de m/z que se tratan como las variables. En una forma de realización preferida de la presente invención, el bloque y contiene la información que asigna cada dato de la EM del estándar pretratado a una variable de clase. En una forma de realización preferida de la presente invención, el bloque y contiene una información relacionada preferiblemente con las mediciones específicas de productividad de una línea celular a escala de biorreactor, es decir, la concentración del producto, la productividad específica o el recuento integral de células viables.

- En una forma de realización preferida de la presente invención, usando los datos del bloque x se realiza un cartografiado de PLS de las variables originales en el espacio variable latente. Esto tiene el efecto de reducir la dimensionalidad del problema a la vez que describe tanta variabilidad como sea posible en los datos originales. En una forma de realización preferida de la presente invención, el algoritmo de PLS - DA utiliza la información del
- 5 bloque y para ajustar los límites de discriminación lineal que separan mejor los datos del bloque x basándose en la clase de información almacenada en el bloque y. Si sólo hay dos clases descritas en el bloque y, es suficiente un único límite de discriminación. En los casos en los que hay presentes tres o más clases, las muestras de dentro de las clases deberían compararse con las muestras de fuera de las clases para cada clase disponible.
- 10 Adicionalmente, la presente invención proporciona un proceso *in vitro* para la preparación, en particular el aislamiento, de una célula con unos datos deseados de productividad específica celular, de recuento integral de células viables o de concentración del producto celular, en el que se realiza un proceso para la predicción de los datos de la productividad específica celular, del recuento integral de células viables o de la concentración del producto celular de al menos una célula de muestra, y se prepara la al menos una célula deseada, preferiblemente
- 15 se aísla.
- Puede obtenerse una célula, en particular aislarse, mediante un proceso según la presente invención.
- 20 La célula aislada puede estar caracterizada por una productividad de proteínas de al menos 1 g/l, preferiblemente de al menos 5 g/l, preferiblemente de al menos 6, 7, 8, 9 o preferiblemente de al menos 10 g/l.
- La presente invención resuelve su problema subyacente también mediante un dispositivo adaptado para proporcionar una predicción de los datos de la productividad específica celular, el recuento integral de células viables o la concentración del producto celular cuando se suministra una muestra de al menos una célula de muestra que comprende: a) un medio adaptado para someter una muestra de la célula de muestra a un análisis por EM
- 25 (espectrometría de masas) para obtener unos datos de muestra en bruto de la EM de la misma, (b) un medio adaptado para someter los datos de la EM en bruto de un estándar y en bruto de la muestra a al menos un primer procedimiento de procesado de la señal de la EM para obtener unos perfiles de la EM pretratados del estándar y de la muestra, y (c) un medio adaptado para someter los datos de la productividad específica celular, del recuento integral de células viables o de la concentración del producto celular procedentes de una célula estándar y de los
- 30 perfiles de EM pretratados de la muestra y del estándar, a un segundo procedimiento de procesado de la señal de la EM que incluye un PLS - DA (un análisis discriminatorio por mínimos cuadrados parciales) basado en la evaluación comparativa, para predecir los datos de la productividad específica celular, del recuento integral de células viables o de la concentración del producto celular de la célula de muestra. El dispositivo se define en la reivindicación 11.
- 35 Algunas formas de realización adicionales preferidas son el aspecto en cuestión de las subreivindicaciones.
- La presente invención se ilustra adicionalmente mediante los siguientes ejemplos y las correspondientes figuras.
- 40 Las Figuras 1a y 1b muestran un perfil típico de la EM antes y después del re-muestreado de 1.000 puntos de datos.
- Las Figuras 2a y 2b muestran los perfiles típicos de la EM antes y después de la corrección de la línea de base.
- 45 Las Figuras 3a y 3b muestran los perfiles típicos de la EM antes y después de la normalización.
- Las Figuras 4a y 4b muestran los perfiles típicos de la EM antes y después del suavizado de Savitzky-Golay.
- 50 Las Figuras 5a y 5b muestran los perfiles típicos de la EM antes y después del recorte de las regiones de los espectros con poca o ninguna información.
- Las Figuras 6a y 6b muestran un gráfico de puntos de PC1 frente a PC2 usando los perfiles en bruto de la EM (6a) y con un perfil típico asociado de la EM (6b).
- 55 Las Figuras 7a y 7b muestran un gráfico de puntos de PC1 frente a PC2 usando los perfiles de la EM sometidos a una corrección de la línea de base (7a) con un perfil típico asociado de la EM (7b).
- Las Figura 8a y 8b muestran un gráfico de puntos de PC1 frente a PC2 usando los perfiles de la EM sometidos a una corrección de la línea de base y a una normalización (8a) con un perfil típico asociado de la EM (8b).
- 60 Las Figuras 9a y 9b muestran un gráfico de puntos de PC1 frente a PC2 usando los perfiles de la EM sometidos a una corrección de la línea de base, a una normalización y a un recorte (9a) con un perfil típico asociado de la EM (9b).
- 65 La Figura 10 muestra el diagrama de flujo del análisis de PLSDA en el PLS\_Toolbox.

Las Figuras 11a y 11b muestran el cuadro de diálogo importado de PLSDA del PLS\_Toolbox.

La Figura 12 muestra el cuadro de diálogo de selección del grupo de clases de PLSDA del PLS\_Toolbox.

5 La Figura 13 muestra las opciones de re-procesado de datos de PLSDA del PLS\_Toolbox.

La Figura 14 muestra el cuadro de diálogo de validación cruzada de PLSDA del PLS\_Toolbox.

10 Las Figuras 15a a 15d muestran varios gráficos de puntos de PLSDA: (15a) gráfico de puntos bivariado, (15b) gráfico T2 de Hotelling, (15c) representación de las predicciones del modelo y (15d) representación de la probabilidad de las predicciones del modelo.

15 Las Figuras 16a y 16b muestran los gráficos de carga de PLSDA: (16a) variable latente 1 y (16b) variable latente 2.

La Figura 17 muestra el gráfico de los límites de decisión de PLSDA para la Clase 1 Y predicha.

La Figura 18 muestra una gráfica de probabilidad de PLSDA para la Clase 1 Y predicha.

20 La Figura 19 muestra un análisis de PLSDA que compara los datos de CL - ESI - EM procedentes de las líneas celulares CHO productoras de anticuerpos 2, 42 y 52.

25 La Figura 20 muestra un análisis de PLSDA que compara los datos de LC-ESI-MS procedentes de las siete líneas celulares CHO productoras de anticuerpos (2, 42, 52, 75, 106, 144 y 164) con las muestras agrupadas en menor de 2 mg/l y mayor de 2 mg/l.

Ejemplo 1 - Generación de la línea celular (según el estado de la técnica)

30 Se usó un vector de expresión GS (Lonza) que contiene los genes optimizados de la cadena pesada y los genes de la cadena ligera para la expresión de un anticuerpo quimérico modelo IgG4 o IgG1 de ratón-humano (Kalwy y col., Mol. Biotechnol 2006, 34, 151 - 156) para generar líneas celulares recombinantes GS-CHO que expresan anticuerpos. El vector se introdujo en la línea celular hospedadora, CHOK1SV (un derivado de CHO-K1; Lonza), usando procedimientos de electroporación estándar, y la mezcla de transfección se distribuyó entre ochenta placas de 96 pocillos. Las placas se incubaron a 37 °C en una atmósfera de aire humidificado con un 10 % de CO<sub>2</sub>. Al día siguiente se añadió medio nuevo a la suspensión celular de las placas. La concentración de EMX (metionina sulfoxamina) en el medio era tal que la concentración final de EMX en cada pocillo era de 50 µM. En primer lugar las placas se cribaron para buscar las transfectantes independientes de glutamina aproximadamente 3 semanas después de la transfección. Las colonias transfectantes aisladas (cada una identificada como procedente de un pocillo con una única colonia) se hicieron progresar a través de todas las etapas de evaluación de una estrategia de construcción de una línea celular típica.

45 La concentración celular de los cultivos se determinó mediante el uso de un analizador de viabilidad celular automatizado Vi-CELL™ (Beckman Coulter). Los cultivos se establecieron en matraces con agitación de 125 ml con una concentración celular dirigida de  $2,0 \times 10^5$  células viables/ml y un volumen final normalmente de 30 ml. Las líneas celulares se subcultivaron sucesivamente en un régimen de 4 días. Una vez alcanzadas unas concentraciones celulares aceptables en el subcultivo, y cuando cesó cualquier gran fluctuación en la concentración celular viable entre los subcultivos, comenzaron las etapas de evaluación realizadas en el cultivo en suspensión. La evaluación del 'lote alimentado' se realizó después de que las líneas celulares habían sido clasificadas tras una primera evaluación de la suspensión ('lote'). Para la evaluación del lote alimentado se determinó la concentración celular de los cultivos en los días 7 y 14 usando un analizador de viabilidad celular automatizado Vi-CELL™. Se realizó la adición de un bolo de alimento A el día 3 y se elaboraron bolos adicionales de alimento B los días 8 y 11. Las muestras del sobrenadante de cultivo se tomaron en días diferentes para la determinación de la concentración anticuerpos. El análisis de viabilidad celular podría realizarse alternativamente con un analizador MACSQuant®.

55 Ejemplo 2 - Preparación de las células de muestra para el análisis de MALDI-TOF

60 Salvo que se especifique de otro modo, todos los experimentos se han realizado en las mismas condiciones de cultivo a las descritas en el ejemplo 1 (primer párrafo). Antes de someter las sondas de muestra al análisis por EM, las células se contaron y se calculó el volumen requerido de cultivo para proporcionar el número apropiado de células. Las células se extrajeron de la estufa de incubación (placa de 96 pocillos) inmediatamente antes de su tratamiento. Se transfirió el volumen requerido de cada muestra a un tubo de Eppendorf, se centrifugó durante 5 minutos a 960 rcf (3.000 rpm) en una micrófuga Eppendorf (modelo 5417c, rotor F-45-30-11) y se eliminó el sobrenadante. Después las células se lavaron con 1 ml de PBS (disolución salina tamponada con fosfato) pipetando suavemente arriba y abajo y después se centrifugó como anteriormente. Cuando se indicó, las células se lavaron subsiguientemente con 1 ml de sacarosa 0,35 M y el sobrenadante se eliminó después de la centrifugación según se describió anteriormente. Todos esos agregados celulares puntuales pudieron ser almacenados (-80 °C) para su

manipulación adicional en el futuro o procesados inmediatamente para su análisis por EM. En caso de haber sido almacenados, los agregados celulares congelados necesitaban ser equilibrados hasta la temperatura ambiente antes de ser usados tras la descongelación. Se preparó una disolución de 20 g/ml de ácido sinapínico en tampón de matriz (40 % de acetonitrilo, 60 % de TFA al 0,1 %) que da como resultado una disolución saturada. La disolución de ácido sinapínico se colocó después en un baño de agua con ultrasonidos durante 15 minutos antes de su centrifugación a 17.900 rcf (13.000 rpm) durante 5 minutos en una micrófuga Eppendorf (modelo 5417c, rotor F-45-30-11).

Después se añadió disolución de matriz (50  $\mu$ l) a cada muestra y las células se resuspendieron pipeteando manualmente la disolución arriba y abajo. Después de su resuspensión, las células se colocaron a 4 °C durante hasta varias horas. Al extraerlas de los 4 °C, las células se resuspendieron golpeando suavemente el tubo y después se añadió 1  $\mu$ l de cada muestra en una rejilla de acero de 384 MTP de una placa de MALDI TOF (Bruker). Las muestras se dejaron secar al aire antes de poner la placa en la máquina de MALDI TOF (Bruker Ultraflex), y se analizaron las muestras.

### Ejemplo 3 - Preparación de las células de muestra para su análisis mediante CL - ESI - EM usando tampón de lisis de Dunn

Recolección de muestras: se hizo crecer un conjunto de líneas celulares CHO en matraces de cultivo celular en suspensión de 250 ml. Las células se contaron usando un Vi- CELL™ y se pipeteó el número requerido de células (entre  $1 \times 10^6$  y  $0,015625 \times 10^6$ ) en tubos de Eppendorf de 1,5 ml y se centrifugaron a 960 rcf en una micrófuga Eppendorf (modelo 5417c, rotor F-45-30-11) durante 5 min y se eliminó del sobrenadante. Los agregados se almacenaron a -80 °C hasta su uso.

Lisis celular: los agregados se descongelaron y se resuspendieron en 400 ml de tampón de lisis de Dunn (urea ultra pura 9,5 M, CHAPS al 2 %, DTT al 1 %) se agitaron vorticialmente cuidadosamente y se incubaron a temperatura ambiente (TA) durante 1 h con una breve agitación vorticial 30 min después del inicio de la incubación. Después las muestras se centrifugaron a 985,6 g durante 1 min para eliminar los desechos celulares, y el sobrenadante se pipeteó en Eppendorfs de 2 ml. Después se usó una muestra de 50  $\mu$ l para su precipitación con acetona.

Precipitación con acetona: se incubó una dilución 4:1 de acetona al 100 % enhiada en hielo con la muestra durante 1 h a -20 °C. La muestra diluida se centrifugó después a 8870,4 g durante 10 min, se eliminó al sobrenadante y el agregado se dejó secar brevemente al aire (no más de 5 min).

Digestión trípica en disolución: el agregado se resuspendió en 50  $\mu$ l de urea 8 M, bicarbonato amónico 0,4 M ( $\text{NH}_4\text{HCO}_3$ ) pipeteando la muestra arriba y abajo para desplazar inicialmente el agregado, seguido de una breve agitación vorticial. La muestra se redujo químicamente mediante la adición de 2,5  $\mu$ l de ditiotreitól (DTT) 100 mM en  $\text{NH}_4\text{HCO}_3$  50 mM durante 1 h en una estufa de incubación a 37 °C. La muestra se alquiló entonces mediante la adición de 5  $\mu$ l de yodoacetamida 100 mM en  $\text{NH}_4\text{HCO}_3$  50 mM durante 15 min a TA en la oscuridad. La concentración de urea se diluyó hasta < 2 M mediante la adición de 192,5  $\mu$ l de agua de calidad HPLC seguido de la adición de 10  $\mu$ l de 0,25  $\mu$ g/ $\mu$ l de tripsina modificada (Promega). Entonces se dejó proceder la digestión trípica durante una noche en una estufa de incubación a 37 °C. La muestra se secó después usando una centrífuga Savant speed vac (SC110A) en una configuración baja y se resuspendió en 20  $\mu$ l de ácido fórmico al 0,1 %, se centrifugó a 8870,4 g durante 1 min, se eliminó el sobrenadante y cualquier agregado resuspendido y se centrifugó de nuevo a 8870,4 g durante 1 min después se pipeteó en viales con tapón de rosca con insertos y se congeló a -80 °C hasta su análisis mediante CL - ESI - EM.

### Ejemplo 4 – Análisis mediante CL - ESI - EM

En la Tabla 1 se muestra un procedimiento de HPLC usado para el análisis mediante CL - ESI - EM (ESI - EM (Bruker o Waters) acoplado a una HPLC (Donex o Agilent)) que dio como resultado unos espectros de EM apropiados.

Los archivos producidos por la CL - ESI - EM se convirtieron después desde el formato de archivo patentado (Bruker o Waters) a un estándar universal (mzXML) (es decir, mediante el uso de CompassExport) y los archivos y los datos resultantes se sometieron a un procedimiento de empaquetado. La metodología de empaquetado, que es estándar para el análisis de este tipo de datos de EM, permite la comparación de múltiples conjuntos de datos de ESI - EM procedentes de diferentes muestras (o de las mismas) alineándolas, e implica a dividir el tiempo de retención (tiempo de elución desde el sistema de CL) y el intervalo de m/z (proporción entre masa y carga de los iones detectados en la ESI - EM) en intervalos separados equidistantemente, por ejemplo, usando un intervalo del tiempo de retención de 60 segundos y un intervalo entre masa y carga de 1 unidad m/z por intervalo.

Tabla 1  
Ejemplo de ejecución de una HPLC en gradiente. Un gradiente de 35 min con un caudal de 0,3 ml/min durante la ejecución, usando un gradiente multietapa según se muestra a continuación, con un tampón A que comprende un 0,1 % de ácido fórmico y un tampón B que comprende un 80 % de acetonitrilo (ACN) y un 0,1 % de ácido fórmico,

Tiempo	% de tampón B	Caudal (μl/min)
0	4	0,3
0	4	0,3
10	55	0,3
11	90	0,3
16	90	0,3
17	4	0,3
35	4	0,3

#### Ejemplo 5 – Procedimiento de procesado de la señal 1

##### 5.1 Procesado de los datos y desarrollo del programa informático

5 En el presente ejemplo, para el cribado y la generación de una línea celular basados en la EM se usó como herramienta un programa informático – ejecutado a través de una interfaz de Windows - que permite una predicción rápida y una capacidad de amplificación de la productividad de la línea celular. Se compila con MATLAB (versión 2008b, referencia) usando el conjunto de herramientas MATLAB Bioinformatics y Statistics así como la PLS\_Toolbox (www.eigenvector.com).

10 La aplicación del programa informático comienza con la disponibilidad de los perfiles de EM procedentes de las líneas celulares de muestra y estándar que se han hecho crecer en diferentes condiciones y escalas de cultivo. Las herramientas de procesado de la señal se han aplicado a los perfiles de la EM para extraer unos patrones de datos únicos de la EM indicativos de los diferentes niveles de líneas celulares productoras del producto.

##### 5.2 Remuestreo de los perfiles de EM

20 El remuestreo de los perfiles de la EM se realiza usando la función 'msresample' del conjunto de herramientas MATLAB Bioinformatics ([http:// tinyurl.com/ msresample](http://tinyurl.com/msresample)). Esto permite el muestreo por arriba y por debajo de la señal original, mientras se conserva la información contenida dentro de los espectros. Las Figuras 1a y 1b muestran un espectro típico de 96DWP antes y después de la aplicación del remuestreo.

25 Normalmente el remuestreo se utiliza en aquellas situaciones en las que la alta resolución de la señal original de la EM sería considerada poco práctica para trabajar con ella debido a restricciones computacionales tales como la carencia de memoria en el ordenador. El re-muestreo también puede usarse para crear un intervalo coherente de m/z, que facilite la alineación de múltiples espectros. Debe tenerse cuidado al remuestrear los perfiles de la EM para que el número de unidades remuestreadas no sea demasiado bajo. Esto provocará que las señales pierdan resolución y puede dar como resultado una pérdida de características.

##### 5.3 Corrección de la línea de base de los perfiles de la EM

35 Los perfiles de datos de la EM muestran normalmente una línea de base variada debido a aspectos tales como el ruido químico de la matriz de MALDI y la sobrecarga de iones. Esto puede ser indeseable cuando se usan técnicas de análisis de datos para comparar los perfiles de EM, ya que utilizan criterios de distancia para medir la similitud entre perfiles. Por lo tanto, se prefiere eliminar estos efectos antes de cualquier de forma de análisis comparativo de las señales. Esto se realiza usando la función 'msbackadj' del conjunto de herramientas MATLAB Bioinformatics (<http://tinyurl.com/msbackadj>). Las Figuras 2a y 2b muestran una selección de espectros típicos de 96DWP antes y después de la aplicación de la corrección de la línea de base.

40 Cuando se aplican varios pretratamientos espectrales en serie, debería usarse la corrección de la línea de base después del muestreo por debajo y antes de corregir la calibración, ya que el ruido presente afectará al resultado.

##### 5.4 Normalización

45 Otro fenómeno observado habitualmente en los perfiles de EM es una variación en la amplitud de las intensidades iónicas. Esto puede estar provocado por varios factores, tales como la variación en la preparación de la muestra o

cambios en la sensibilidad del instrumento. El procedimiento estándar para contrarrestar esta variación es normalizar el área bajo la curva de la EM a la del promedio del grupo (normalmente se usa la media o la mediana). Esto se realiza usando la función 'msnorm' del conjunto de herramientas MATLAB Bioinformatics (<http://tinyurl.com/msnormal>). Las Figuras 3a y 3b muestran una selección de espectros típicos de 96DWP antes y después de la aplicación de la normalización del área bajo la curva.

Quando se aplican varios pretratamientos espectrales en serie, debería usarse la normalización de las muestras después de restar la línea de base, ya que el elemento de ruido introducido por la matriz de cristalización puede afectar a los resultados.

### 5.5 Alineación de la EM

La alineación de los picos se usa para corregir la variación entre el valor observado de  $m/z$  y el tiempo de vuelo real. Estos errores se producen habitualmente como resultado de errores de calibración y pueden observarse como un desplazamiento sistemático entre los picos. La corrección de estas incoherencias puede realizarse usando la función 'msalign' del conjunto de herramientas MATLAB Bioinformatics (<http://tinyurl.com/msalign>).

Un procedimiento de alineación de espectros es añadir a las muestras una sustancia con un perfil espectral conocido, y alinear las muestras basándose en esta. Sin embargo, en situaciones en las que a las muestras no se les ha añadido nada, las muestras pueden ser alineadas con respecto a un espectro de referencia tal como el perfil medio.

### 5.6 Filtrado de los perfiles de la EM

Un perfil típico de EM contiene una mezcla de señal y ruido, el suavizado de la señal mediante el uso de un filtro de Savitzky-Golay puede ayudar a reducir el impacto del componente de ruido de la señal durante el subsiguiente procesado. Los filtros de Savitzky-Golay se aplican normalmente a señales de EM, ya que usan polinomios de orden alto para ajustar las curvas. Esto da como resultado una mayor conservación de las características de la señal, tal como la altura de los picos. Este proceso se realiza usando la función 'mssgolay' del conjunto de herramientas MATLAB Bioinformatics (<http://tinyurl.com/mssgolay>). Las Figuras 4a y 4b muestran una selección de espectros típicos de 96DWP antes y después de la aplicación del filtrado de Savitzky-Golay.

### 5.7 Recorte de los perfiles de la EM

El recorte de los perfiles de la EM se realiza para eliminar partes de la señal que contienen poca o ninguna información. También permite que los espectros sean divididos en subsecciones. Esto permite el análisis de regiones específicas de los perfiles de la EM en lugar de los espectros completos. Las Figuras 5a y 5b muestran una selección de espectros típicos de 96DWP antes y después del recorte de la señal en el intervalo de desde 0 hasta 500  $m/z$ .

### 5.8 Comparación de los efectos del pretratamiento

Para demostrar el efecto de la aplicación de técnicas de procesado de la señal a los datos de la EM, se analizó un grupo de 118 líneas celulares (medidas por duplicado) usando un análisis de los componentes principales (PCA). Las Figuras 6a y 6b muestran un gráfico de las puntuaciones de los componentes principales de PC1 frente a PC2. Este gráfico describe dos fuentes principales de variación en el conjunto de datos.

La aplicación de la corrección de la línea de base a los perfiles en bruto de la EM da como resultado una reducción de la cantidad de dispersión observada en el primer y el segundo componente principal. Esto puede observarse en las Figuras 7a y 7b.

La corrección de la línea de base sólo contrarresta el ruido en la señal debido a la matriz de MALDI. Se prefiere que la variación en las amplitudes de la señal sea eliminada usando una normalización. Las Figuras 8a y 8b muestran el efecto de la aplicación de la normalización al grupo de espectros. Se muestra claramente una reducción en la variación observada en PC1, la principal fuente de variación.

La etapa final de procesado de la señal se realiza para eliminar las partes de la señal que se sabe que no contienen información útil. Las Figuras 9a y 9b muestran el efecto de la eliminación de los puntos de datos en el perfil de EM en el intervalo de desde 0 hasta 500 unidades  $m/z$ , ya que el MALDI se ajustó para no registrar intensidades en este intervalo. A partir de los gráficos de puntos está claro que no se observa ningún efecto al eliminar estos datos, ya que el gráfico de puntos es idéntico al observado en las Figuras 8a y 8b.

## Ejemplo 6 - Procedimiento de procesado de la señal 2

### 6.1 Modelado por PLSDA de los criterios de productividad

La PLSDA es una aplicación de modelado por mínimos cuadrados multivariados formulada específicamente para la clasificación predictiva. La metodología desarrollada de huella molecular de EM utilizada en el ejemplo emplea la implementación del PLS\_Toolbox de la PLSDA, publicada por Eigenvector Research, Inc. (EVRI) ([www.eigenvector.com](http://www.eigenvector.com)).

5

## 6.2 Prueba de un modelo de PLSDA

Para probar un modelo de PLSDA se requieren dos conjuntos diferentes de información; el bloque x y el bloque y. En la metodología esquematizada para realizar la construcción de una nueva línea celular, el bloque x contiene la información de los perfiles espectrales generados en la etapa del proceso de 96DWP. Cada perfil se trata como una muestra, registrándose las intensidades de la señal en un intervalo específico de valores de m/z que se tratan como las variables. El bloque y contiene la información que asigna cada una de las muestras de prueba a una variable de clase. En este ejemplo, el bloque y contiene la información relativa a los datos específicos del cultivo celular sobre productividad de una línea celular a una escala de biorreactor, es decir, la concentración de producto, la productividad específica o el recuento integral de células viables.

15

Usando los datos del bloque x se realiza un cartografiado de PLS de las variables originales en el espacio de variable latente. Esto tiene el efecto de reducir la dimensionalidad del problema, a la vez que describe todo lo posible la variabilidad en los datos originales. Entonces el algoritmo de PLSDA utiliza la información del bloque y para ajustar los límites de discriminación lineal que separan mejor los datos del bloque x basándose en la clase de información almacenada en el bloque y. Si sólo hay dos clases descritas en el bloque y, es suficiente un único límite de discriminación; en los casos en los que hay presentes tres o más clases, las muestras de dentro de las clases deberían compararse con las muestras de fuera de las clases para cada clase disponible.

20

## 6.3 Diagrama de flujo del análisis

La Figura 10 muestra el diagrama de flujo de las operaciones requeridas para construir un modelo de PLSDA usando la implementación gráfica del algoritmo encontrado en MATLAB PLS\_Toolbox.

30 1. Cargar los datos X – Este botón indica al usuario que importe los datos del bloque x al programa informático. El programa informático de huella molecular, ms\_preproc, diseñado para que actúe como una interfaz entre el procesamiento de la señal y las técnicas de análisis, convierte automáticamente los datos de la EM en el formato requerido para trabajar con el conjunto de herramientas de PLS y guarda la variable del espacio de trabajo de MATLAB como la variable 'Xblock' (Figuras 11a y 11b).

35

2. Cargar las clases (opcional) – Este botón se usa para importar los datos del bloque y al programa informático. Sin embargo, la información de la clase también puede almacenarse en la variable 'Xblock' (para más detalles, véase [http://wiki.eigenvector.com/index.php?title=DataSet\\_Object](http://wiki.eigenvector.com/index.php?title=DataSet_Object)). El programa informático ms\_preproc incorpora la información de la clase en la variable 'Xblock'; por lo tanto, esta etapa es opcional.

40

3. Seleccionar los grupos de clase – Este botón presenta al usuario la opción de seleccionar los grupos de clase con los que construir el modelo. La Figura 12 muestra un ejemplo típico con tres clases (alta, media, baja). El modelo calculará los límites de decisión para las clases añadidas a la columna de la derecha.

45

4. Elegir el preprocesado – Este botón puede usarse para aplicar varias técnicas de preprocesado a los archivos espectrales (Figura 13). Si esta etapa se realiza usando las herramientas de preprocesado espectral del conjunto de herramientas bioinformáticas, la mayoría de esta etapa es opcional. Sin embargo, también relacionado con el preprocesado de los datos, está el aspecto de las probabilidades previas asignadas durante la prueba. Por defecto, el algoritmo asume una probabilidad igual de que cada clase sea seleccionada. Algunas veces no es el caso, y los valores deben ser ajustados para que reflejen las probabilidades reales. Esto se realiza alterando las 'Opciones del Procedimiento (PLSDA)', que están en la sección de 'Opciones' del menú 'Editar'.

50

5. Elegir la validación cruzada – Este botón permite al usuario realizar una validación cruzada durante el proceso de prueba. Este proceso se usa a menudo para proporcionar un grado de confianza mejorado de un resultado y sirve como comprobación cruzada del rendimiento del clasificador. La metodología estándar es reservar una porción de los datos de prueba y usar esto para probar el rendimiento del clasificador. Normalmente, este proceso se repite después con una porción diferente de los datos reservados. La Figura 14 muestra los procedimientos disponibles para el racionamiento de los datos usando la PLS\_toolbox ([http://wiki.eigenvector.com/index.php?title=Using\\_Cross-Validation](http://wiki.eigenvector.com/index.php?title=Using_Cross-Validation)).

55

60

6. Construir el modelo – Este botón calcula las variables latentes de PLS y coloca la posición óptima de los límites de discriminación para maximizar el número de clasificaciones correctas en el conjunto de datos de prueba.

65

7. Elegir los componentes – Este botón está disponible una vez que se ha calculado el modelo de PLSDA. Produce gráficos que ayudan al usuario en la selección del número de Variables Latentes a conservar en el

modelo de PLSDA. Otro procedimiento para conseguir esto es asumir que la variación explicada relativa al bloque y debería estar en la región del 70 - 80 % para no sobrejustar el modelo.

5 8. Revisar las puntuaciones – Este botón permite el acceso a varias gráficas relacionadas con las puntuaciones de la variable latente. Esto puede usarse para analizar el rendimiento del modelo. Las Figuras 15a a 15d muestran algunos de los gráficos más útiles; (a) gráfico bivariado de las cargas, (b) gráfico T2 de Hotelling, (c) gráfico de predicciones del modelo y (d) probabilidades asociadas para el gráfico de predicciones del modelo.

10 9. Revisar las cargas – Este botón permite el acceso a varias gráficas relacionadas con las cargas de la variable latente. Esto puede usarse para identificar las variables que tienen la influencia más significativa sobre cada variable latente. Esto puede ser útil para identificar las áreas de la señal espectral que tienen la mayor probabilidad de influir en la discriminación.

15 10. Cargar los datos de ensayo – Una vez que se ha construido el modelo, la siguiente etapa es utilizar los modelos para elaborar predicciones sobre los datos no observados. Al hacer clic, este botón presenta al usuario el mismo cuadro de diálogo que en las Figuras 11a/b. Los datos de ensayo también pueden cargarse con el bloque x de datos original en la etapa 1 si no se asigna una variable de clase a las muestras. Cualquier muestra no asignada a una clase por nuestro programa informático ms\_preproc durante la etapa 2 se considera automáticamente como dato de ensayo. Estas pueden considerarse como las muestras clasificadas como desconocidas en las Figuras 16a y 16b.

11. Aplicar el modelo – Este botón ajusta los datos de ensayo al modelo de PLSDA probado. Permite al usuario determinar la clase más probable a la que puedan pertenecer las muestras desconocidas.

25 Ejemplo 7 - Perfiles de EM sometidos a MALDI-TOF y su modelado estadístico

El párrafo ejemplificado se centra en los resultados obtenidos a partir del modelado de los datos del análisis por EM durante el proceso de generación de la nueva línea celular. Las Figuras 17 y 18 muestran los resultados obtenidos para el modelo BNCD (los datos están corregidos en la línea de base, normalizados y recortados con muestras por duplicado incluidas). Los triángulos muestran los datos de prueba para la clase "Alta" (> 4.000 mg/l), los asteriscos muestran los datos de prueba para la clase "Baja" (≤ 3.999 mg/l), los puntos negros sin un número de identificación de la línea celular representan las muestras para las que los datos de clase son desconocidos, y los puntos con el número de identificación de la línea celular muestran las muestras que están por encima de los límites de clasificación (línea punteada gris superior).

35 Usando la información procesada a partir del proceso de construcción de la línea celular (i) podría construirse un modelo de predicción que incluyera cientos de datos de EM generados durante el proceso de generación de la línea celular. Basándose en el modelo, se recopiló una lista de las líneas celulares que se esperaba que produjeran diferentes cantidades de MAb (> 4.000 mg/l; ≤ 3.999 mg/l). La Tabla 2 destaca las diversas líneas celulares que pueden ser identificadas en las figuras 17 y 18. Las líneas celulares se hicieron crecer con sus valores de título registrados para medir el rendimiento del procedimiento de predicción (Tabla 2).

Tabla 2 Líneas celulares altas/bajas productoras predichas frente a la productividad observada

ID de la línea celular	Predicción	Biorreactor de 10 l
262B7	> 4.000 mg/l	6.524 mg/l
281D8	> 4.000 mg/l	4.555 mg/l
241B6	≤ 3.999 mg/l	1.219 mg/l
243D11	≤ 3.999 mg/l	663 mg/l
246F9	≤ 3.999 mg/l	964 mg/l

45 Los resultados de la ejecución de la validación verificaron con éxito la aplicación de la selección predictiva de las líneas celulares en el proceso de generación de la línea celular, de forma que, en una etapa de desarrollo temprana, los perfiles recopilados de las EM de las líneas celulares individuales reflejan su comportamiento en la escala de elaboración posterior.

50 Ejemplo 8 – Los resultados posteriores al modelado estadístico de los perfiles de la EM sometidos a una CL - ESI - EM

55 Se procesó la señal de los perfiles en bruto de la EM obtenidos mediante CL - ESI - EM mediante el procedimiento de procesado de señales I (ejemplo 5) y el procedimiento de procesado de señales II (ejemplo 6). Los resultados se

mostraron en las figuras 19a/b y 20.

Las Figuras 19a/b muestran la separación de tres líneas celulares CHO recombinantes diferentes usando un análisis mediante PLSDA (las Figuras 19a/b muestran LV1 v LV3) y la Figura 20 muestra la separación de siete líneas celulares CHO recombinantes diferentes tras un análisis mediante PLSDA cuando las muestras se agrupaban en < 2 g/l y > 2 g/l. Los datos muestran dos muestras de cada grupo de línea celular recombinante, y que el algoritmo de PLSDA es capaz de discriminar entre líneas celulares pertenecientes a grupos diferentes. La metodología es adecuada para la identificación de la huella molecular de líneas celulares recombinantes sobre la base de características deseables (por ejemplo, productividad).

La Tabla 3 muestra la concentración del producto de las líneas celulares CHO 2, 42, 52, 75, 106, 144 y 164 cultivadas en una placa de 24 pocillos, por lotes, por lote alimentado y en un biorreactor. Especialmente, la concentración de producto en la escala de biorreactor fue correctamente predicha mediante el análisis por PLS - DA usando los datos de la CL - ESI - EM (Figura 20).

Tabla 3

Línea celular CHO	Concentración del producto en una placa de 24 pocillos (mg/l)	Concentración del producto por lotes (mg/l)	Concentración del producto por lote alimentado (mg/l)	Concentración del producto a escala de biorreactor (mg/l)	Agrupamiento según se muestra en el análisis PLS - DA de los datos de la CL - ESI - EM
42	230	538	2.404	3.220,00	> 2 g/l
52	31,5	31,5	101	24,00	< 2 g/l
2	236	480	1.680,5	2.594,00	> 2 g/l
144	175	391	1.592	1.816,00	< 2 g/l
75	241	606	1.001	1.826,00	< 2 g/l
106	221	766	969,5	2.325,00	> 2 g/l
164	202	534	881,5	2.307,00	> 2 g/l

**REIVINDICACIONES**

1. Un proceso *in vitro* para la predicción de los datos de productividad específica celular, de recuento integral de células viables o de concentración del producto celular de al menos una célula de muestra en una etapa tardía de la amplificación, en particular en biorreactores de gran volumen, mientras la al menos una célula de muestra es cultivada en un medio con un bajo volumen que comprende las etapas:
  - (a) proporcionar (i) una muestra de la al menos una célula de muestra, (ii) datos de productividad específica celular, recuento integral de células viables o concentración del producto celular a partir de una célula estándar y (iii) datos en bruto estándar de EM (espectrometría de masas) a partir de la célula estándar,
  - (b) someter la muestra de la al menos una célula de muestra a un análisis por EM para obtener los datos de muestra en bruto de la EM de la misma,
  - (c) someter los datos de la EM en bruto del estándar y en bruto de la muestra a al menos un primer procedimiento de procesado de la señal de la EM para obtener unos perfiles de la EM pretratados del estándar y de la muestra, y
  - (d) someter los datos de la productividad específica celular, de recuento integral de células viables o de concentración del producto celular procedentes de la célula estándar de la etapa (a) y los perfiles de la EM pretratados de la muestra y del estándar obtenidos en la etapa (c) a un segundo procedimiento de procesado de la señal de la EM que incluye un PLS - DA (un análisis discriminatorio por mínimos cuadrados parciales) basado en la evaluación comparativa, para predecir los datos de productividad específica celular, de recuento integral de células viables o de concentración del producto celular de la al menos una célula de muestra.
2. El proceso según la reivindicación 1, en el que la célula se elige de entre el grupo que consiste en líneas celulares humanas, líneas celulares animales, líneas celulares vegetales, células de hongos, células de bacterias, células de levadura y células madre.
3. El proceso según una cualquiera de las reivindicaciones anteriores, en el que la célula es una línea celular CHO.
4. El proceso según una cualquiera de las reivindicaciones anteriores, en el que el análisis por EM de la etapa (b) es MALDI -TOF.
5. El proceso según una cualquiera de las reivindicaciones anteriores, en el que la muestra de las células de muestra sometida a la etapa (b) comprende entre  $0,015 \times 10^6$  y  $0,0625 \times 10^6$  células.
6. El proceso según una cualquiera de las reivindicaciones anteriores, en el que la señal de los datos de la EM en bruto de la muestra obtenidos en la etapa (b) y los datos de la EM en bruto del estándar proporcionados en la etapa (a) se procesa mediante una operación elegida de entre el grupo que consiste en corrección de la línea de base, normalización, alineación, filtrado y recorte.
7. El proceso según una cualquiera de las reivindicaciones anteriores, en el que los perfiles de la EM pretratados del estándar y de la muestra obtenidos en la etapa (c) son analizados ópticamente.
8. El proceso según una cualquiera de las reivindicaciones anteriores, en el que la célula de muestra, con los datos predichos de productividad específica celular, de recuento integral de células viables o de concentración del producto celular evaluados en la etapa (d) se cultiva en un cultivo celular para verificar sus datos de productividad específica celular, de recuento integral de células viables o de concentración del producto celular.
9. El proceso según la reivindicación 8, en el que los perfiles en bruto de la EM de la muestra obtenidos en la etapa (b) y los datos verificados de productividad específica celular, de recuento integral de células viables o de concentración del producto celular de la célula de muestra, se usan en la etapa (a) como datos estándar de la EM y como datos de productividad específica celular, de recuento integral de células viables o de concentración del producto celular a partir de una célula estándar.
10. Un proceso *in vitro* para el aislamiento de una célula con unos datos deseados de productividad específica celular, de recuento integral de células viables o de concentración del producto celular, en el que se realiza un proceso según una cualquiera de las reivindicaciones anteriores y se aísla la célula deseada.
11. Un dispositivo adaptado para proporcionar una predicción de los datos de productividad específica celular, de recuento integral de células viables o de concentración del producto celular de al menos una célula de muestra en una etapa tardía de la amplificación, en particular, en biorreactores de gran volumen, cuando al dispositivo se le suministra una muestra de al menos una célula de muestra cultivada en un medio con un bajo volumen que comprende:
  - (a) un medio adaptado para someter una muestra de la al menos una célula de muestra a un análisis por EM (espectrometría de masas) para obtener unos datos en bruto de la EM de la muestra de la misma,
  - (b) un medio adaptado para someter los datos de la EM en bruto de un estándar y en bruto de la muestra a al

menos un primer procedimiento de procesado de la señal de la EM para obtener unos perfiles de la EM pretratados del estándar y de la muestra, y

- 5 (c) un medio adaptado para someter los datos de productividad específica celular, de recuento integral de células viables o de concentración del producto celular a partir de una célula estándar y los perfiles de la EM pretratados de la muestra y del estándar, a un segundo procedimiento de procesado de la señal de la EM que incluye un PLS - DA (un análisis discriminatorio por mínimos cuadrados parciales) basado en la evaluación comparativa, para predecir los datos de productividad específica celular, de recuento integral de células viables o de concentración del producto celular de la célula de muestra.

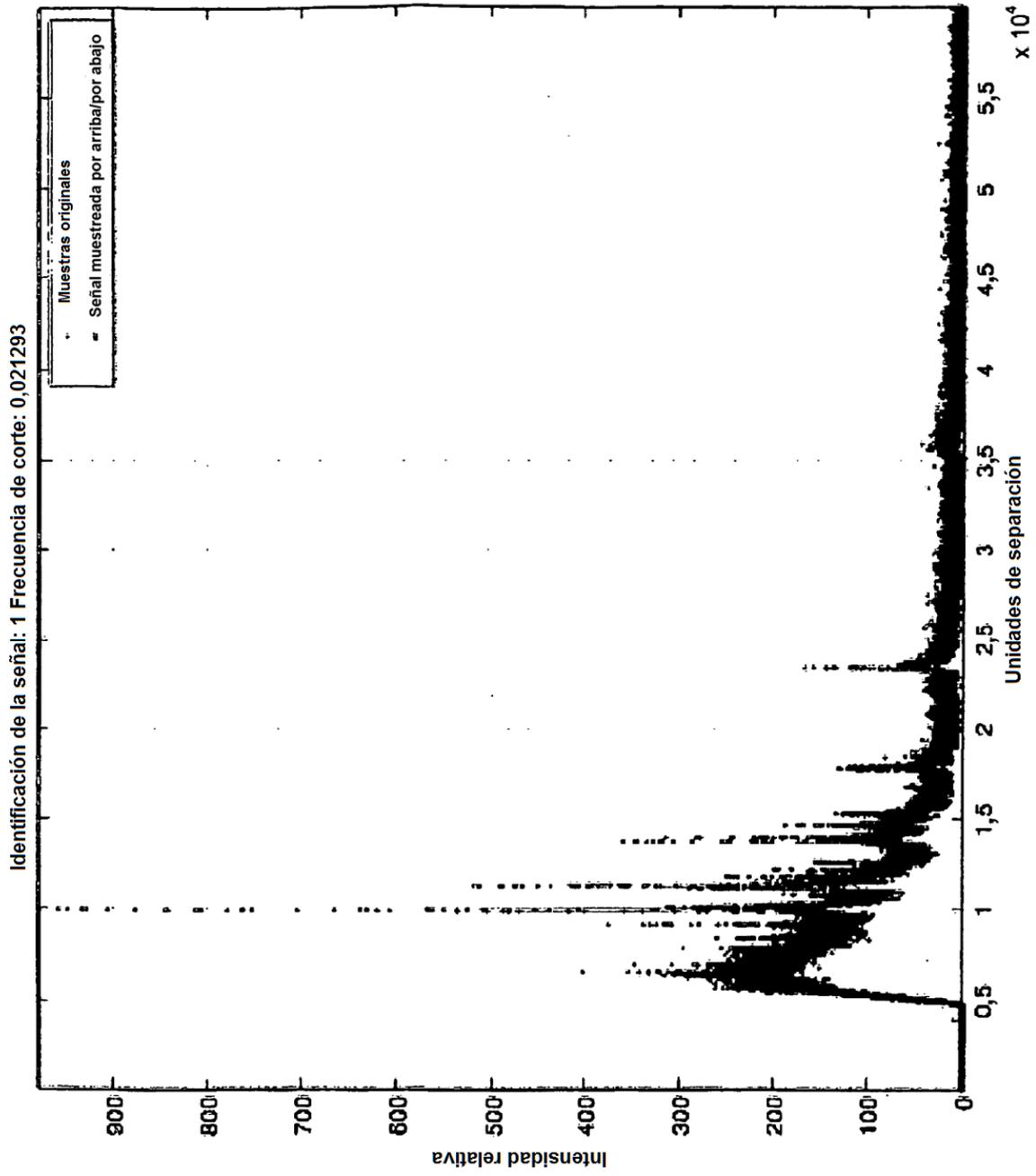


Fig. 1a

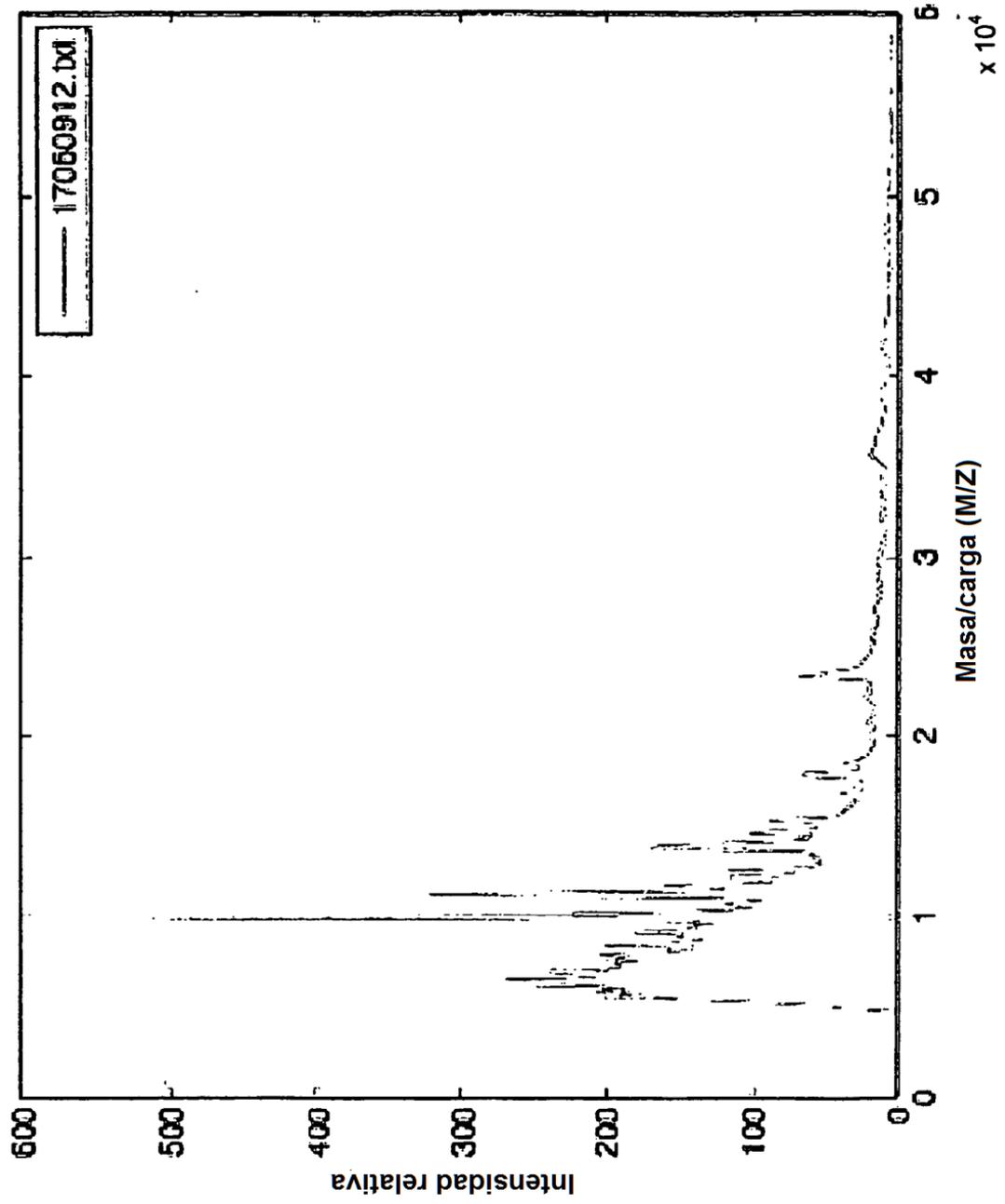


Fig. 1b

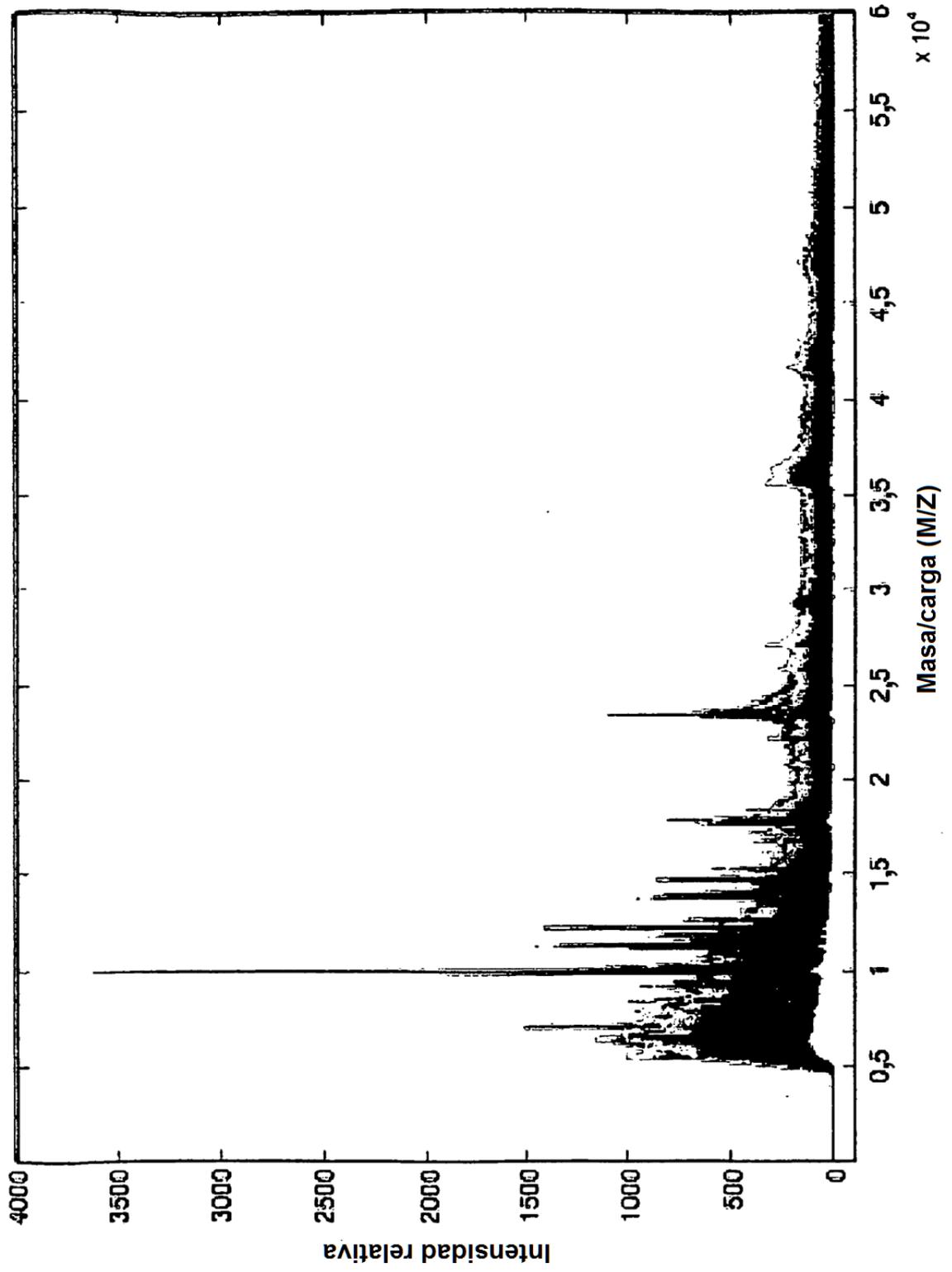


Fig. 2a

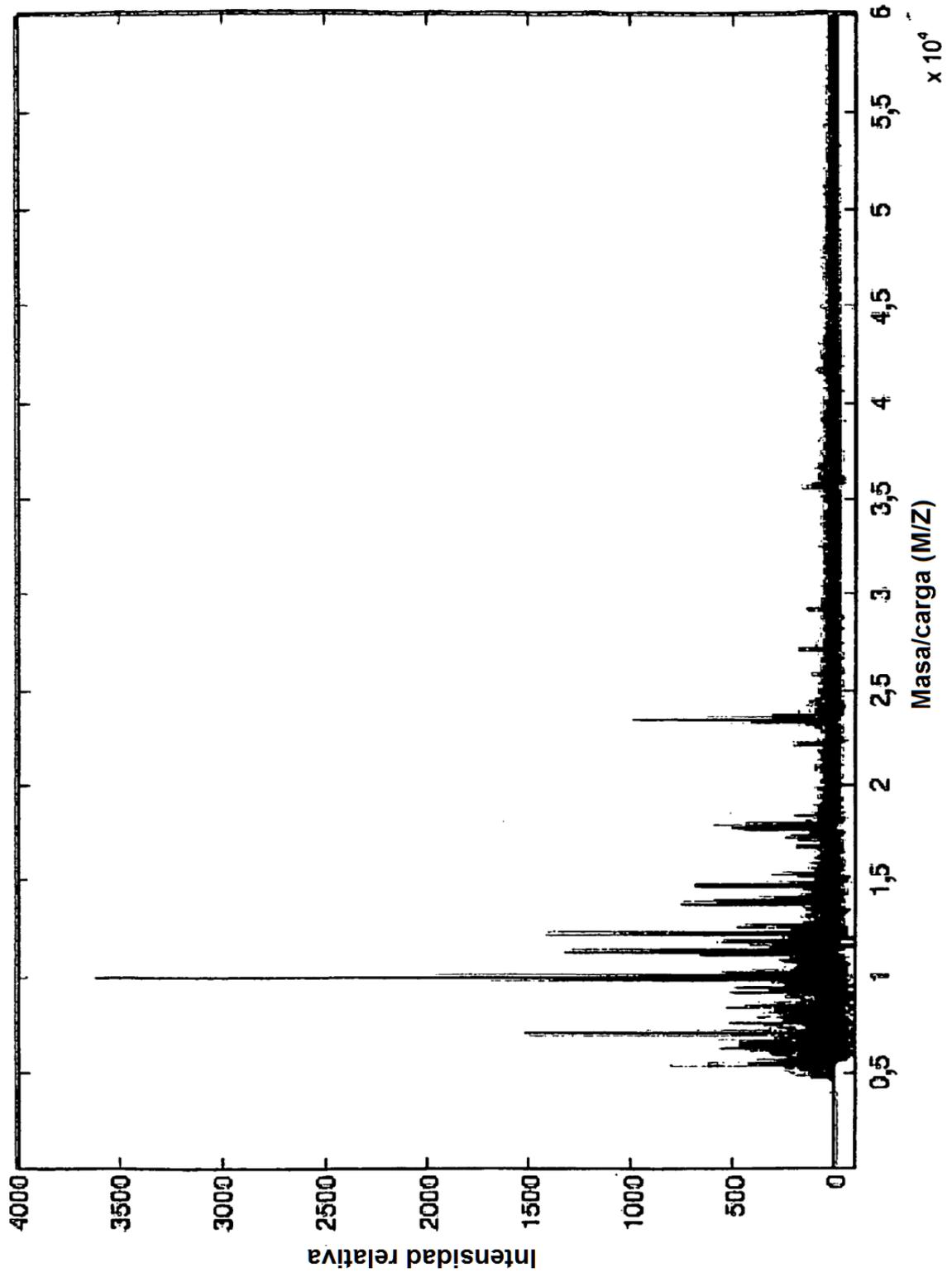


Fig. 2b

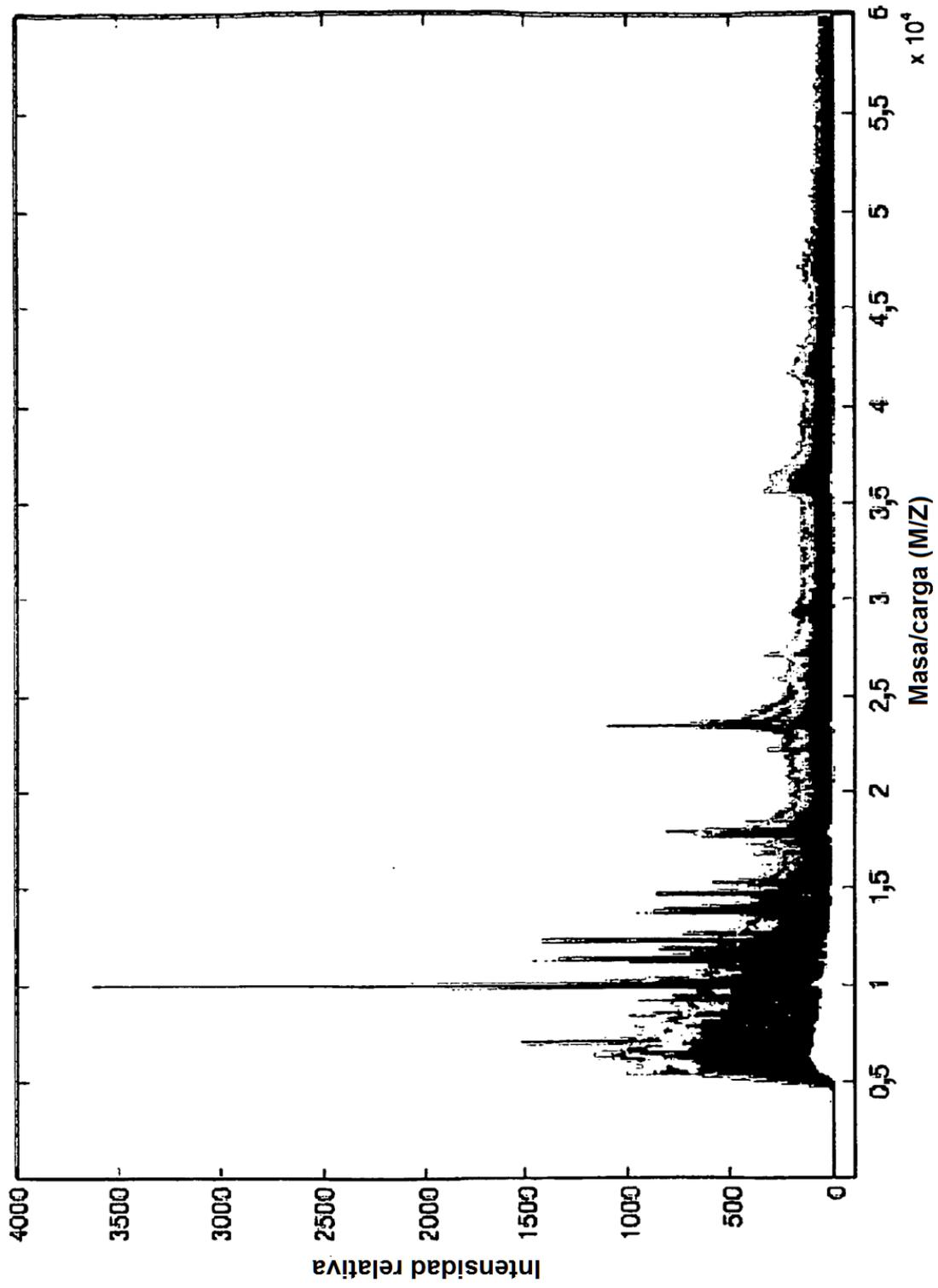


Fig. 3a

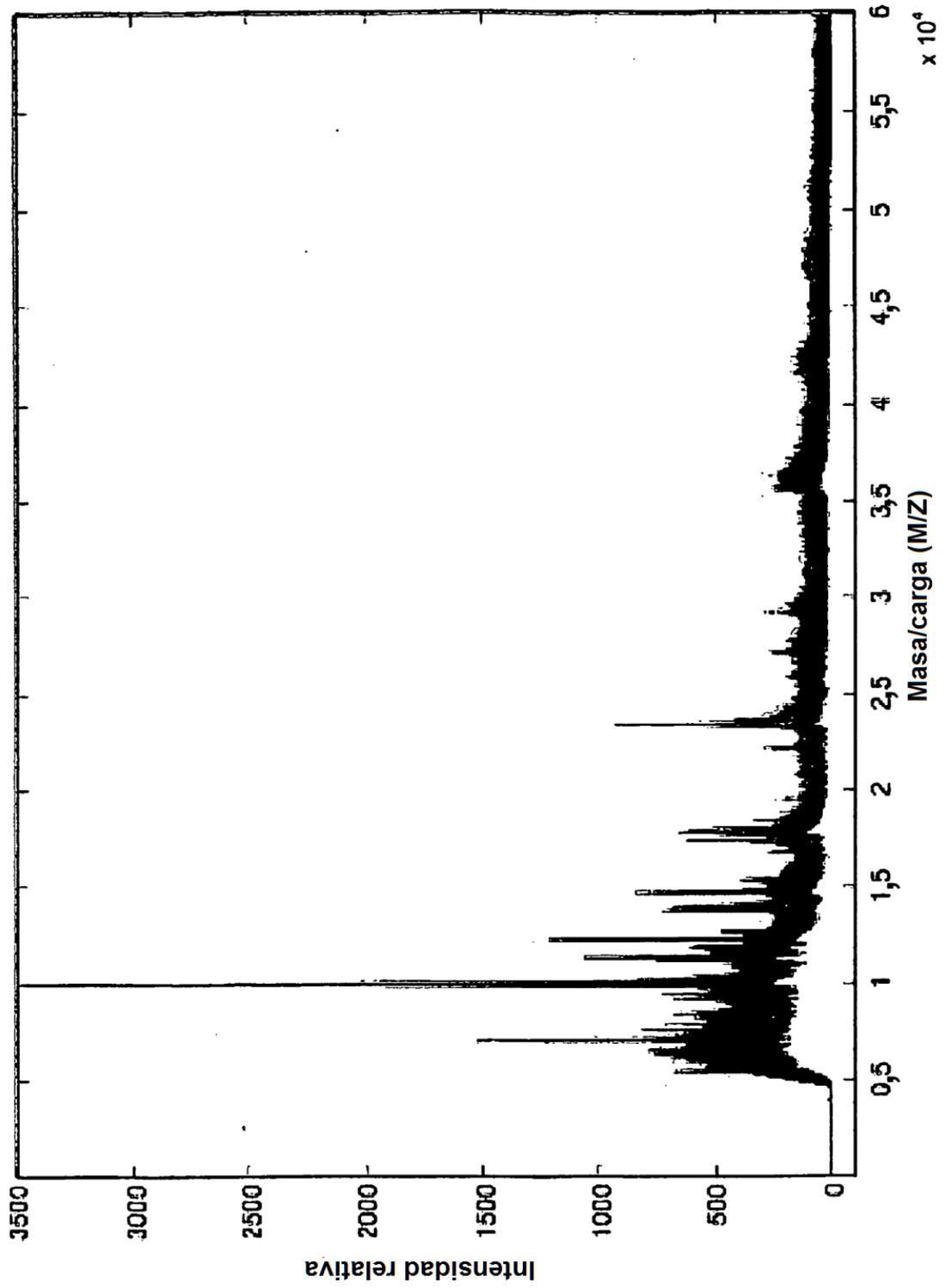


Fig. 3b

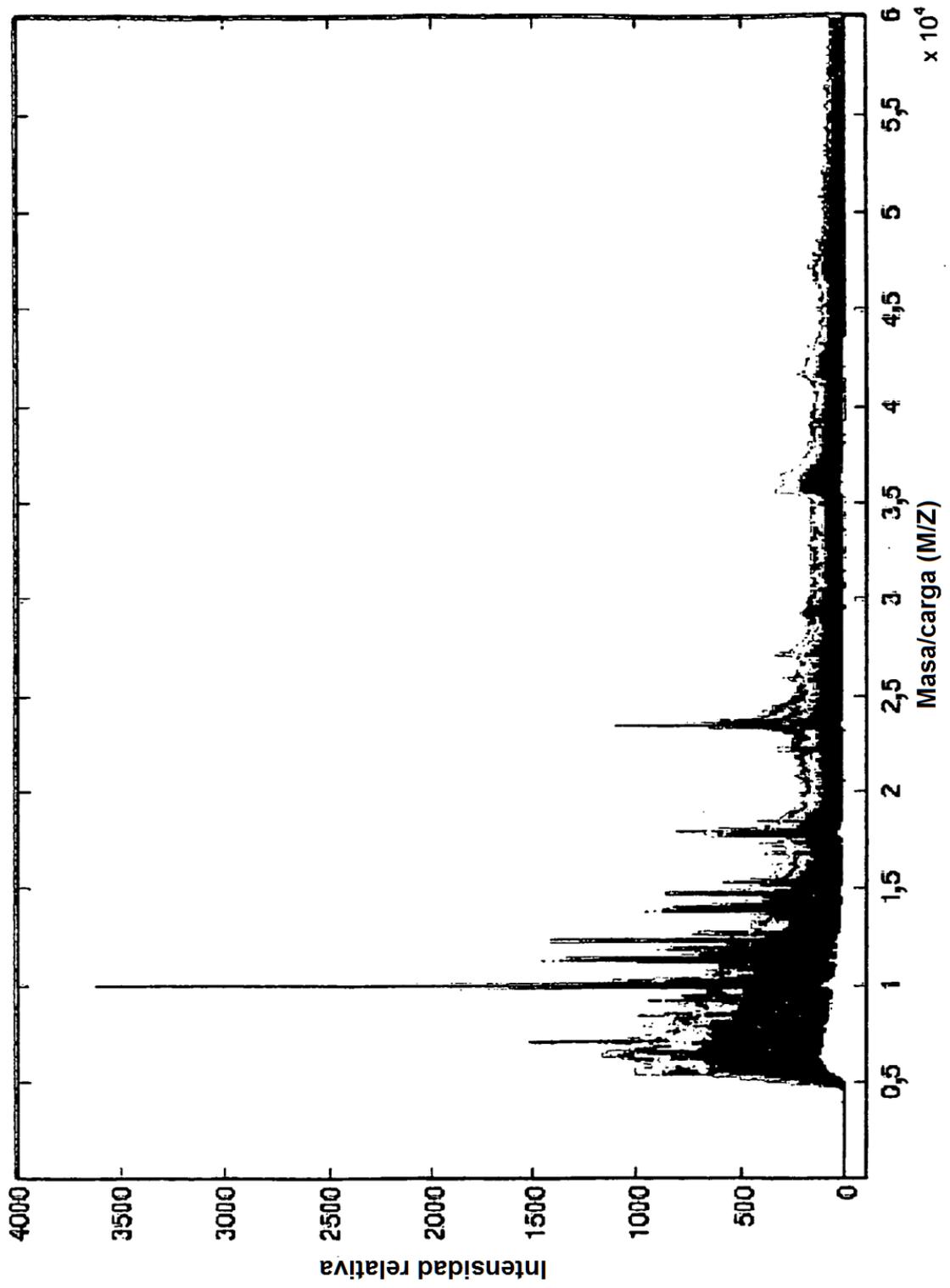


Fig. 4a

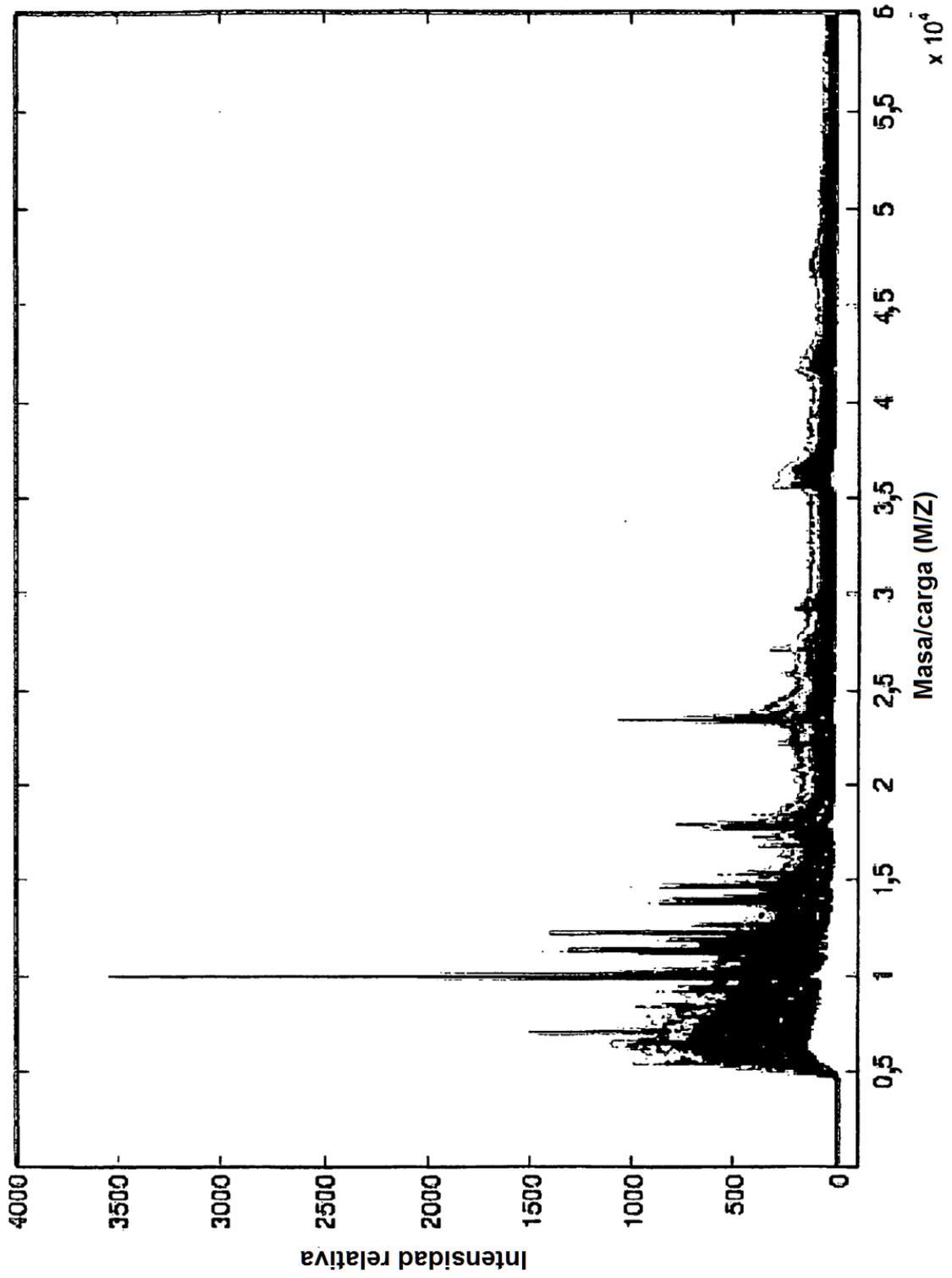


Fig. 4b

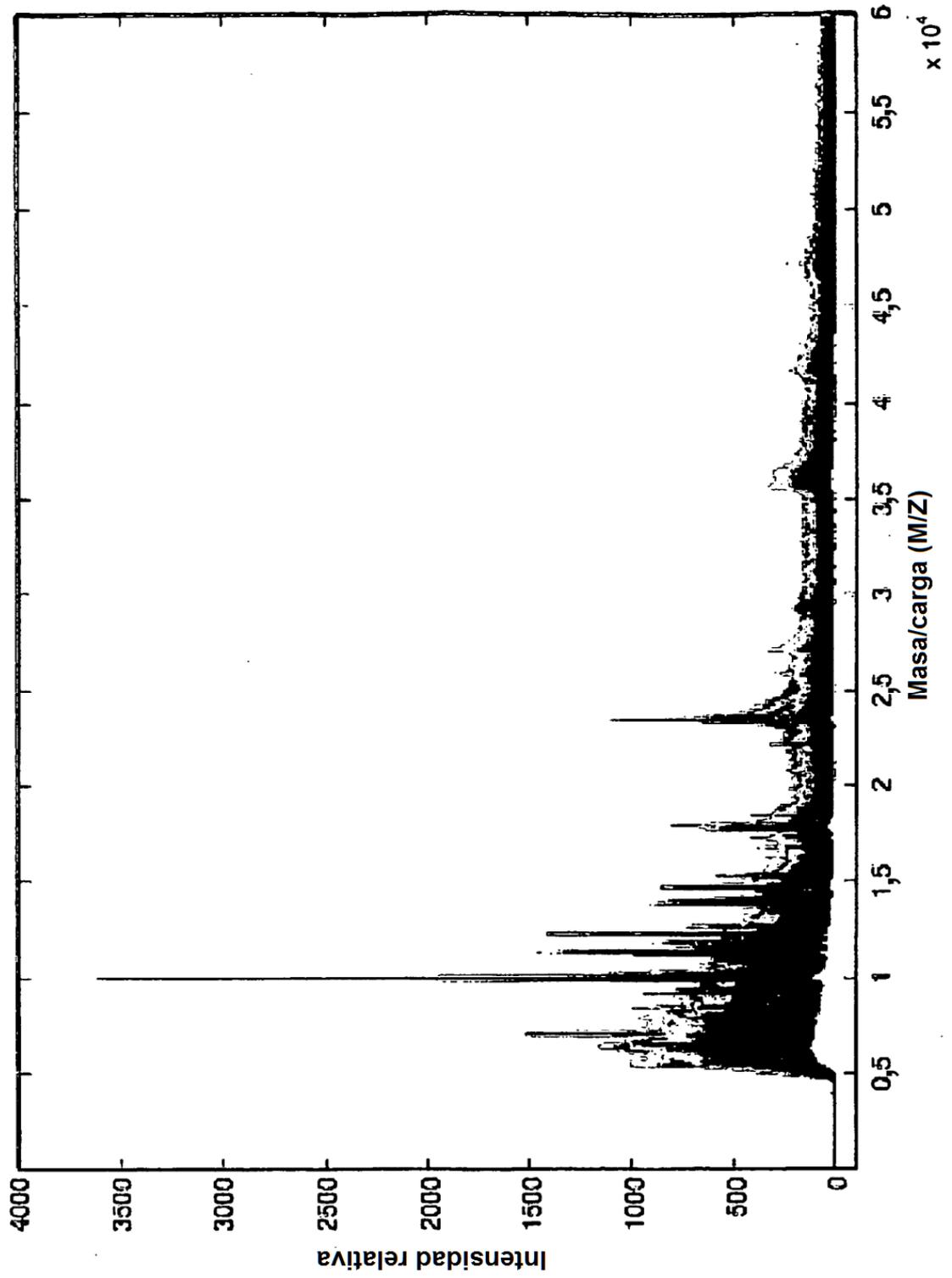


Fig. 5a

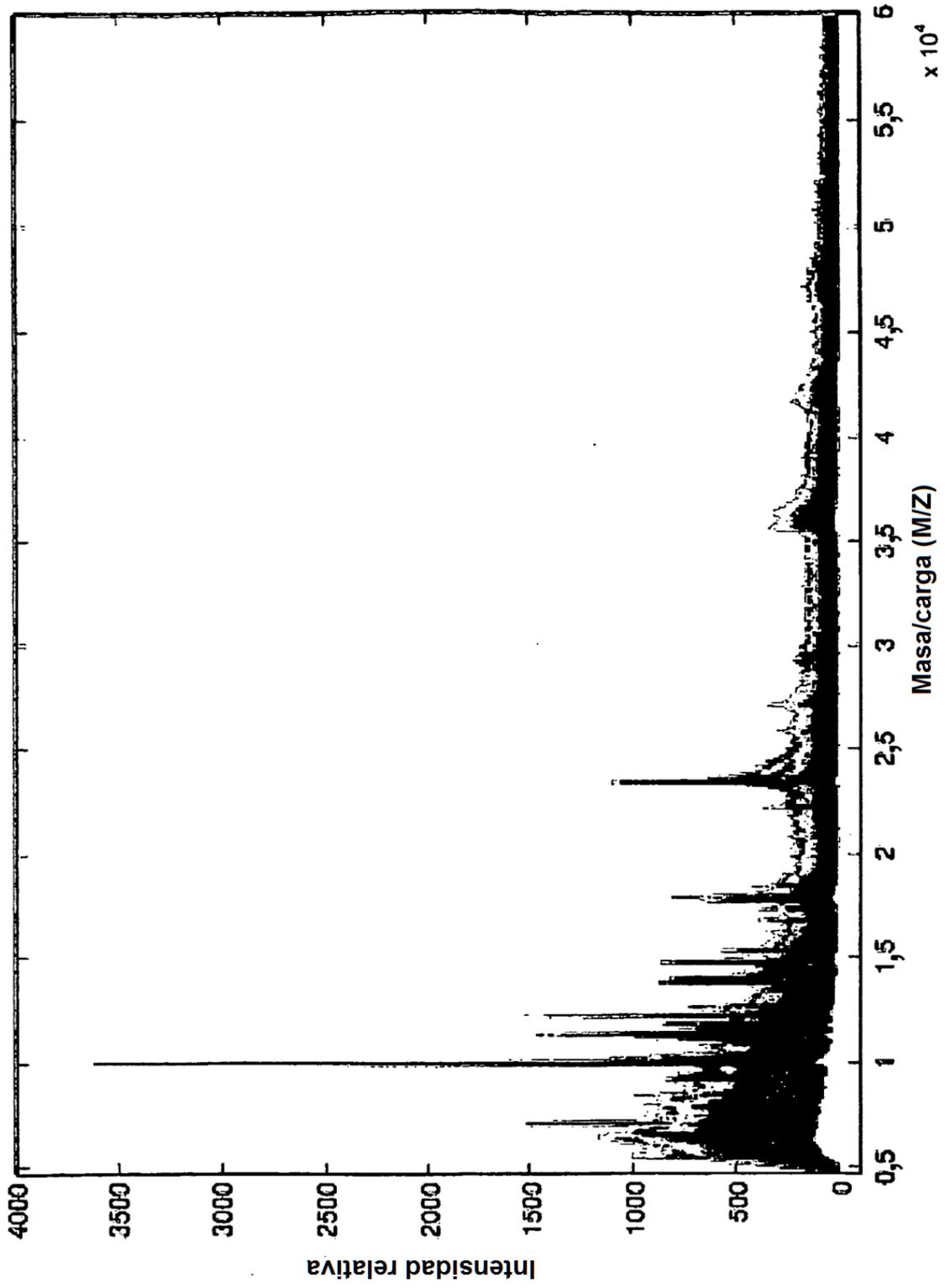


Fig. 5b

Gráfico de muestras/puntuaciones

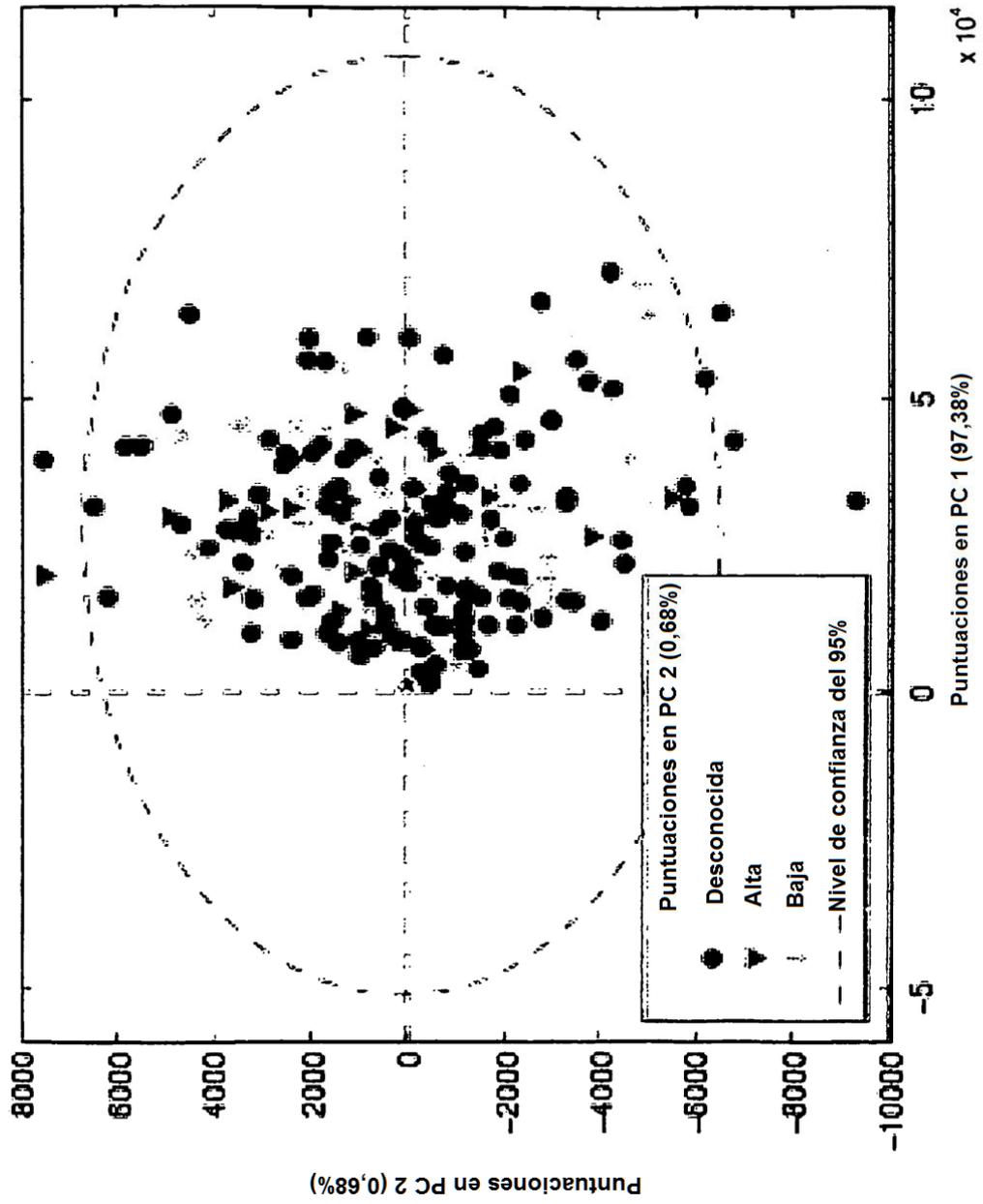


Fig. 6a

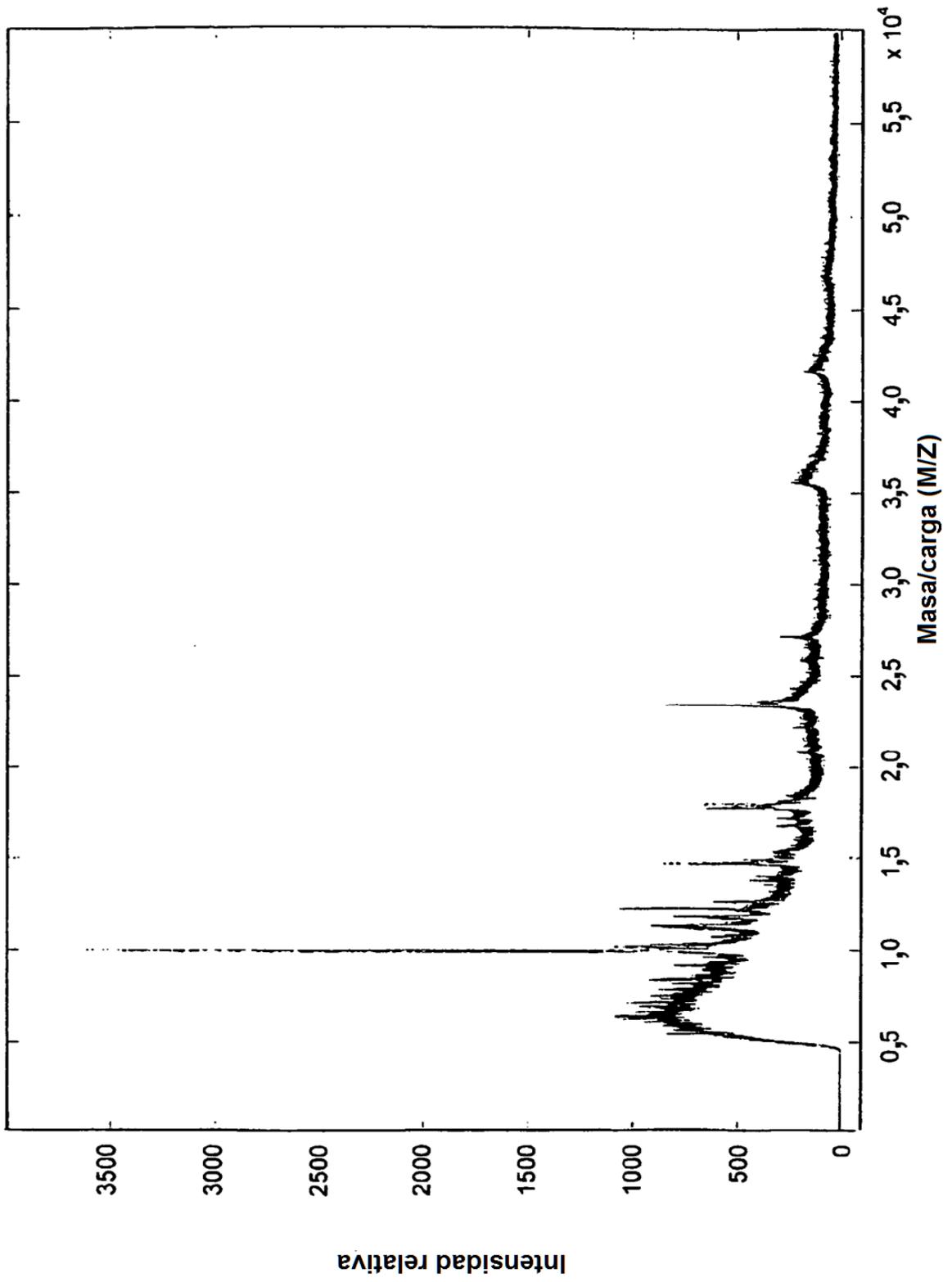


Fig. 6b

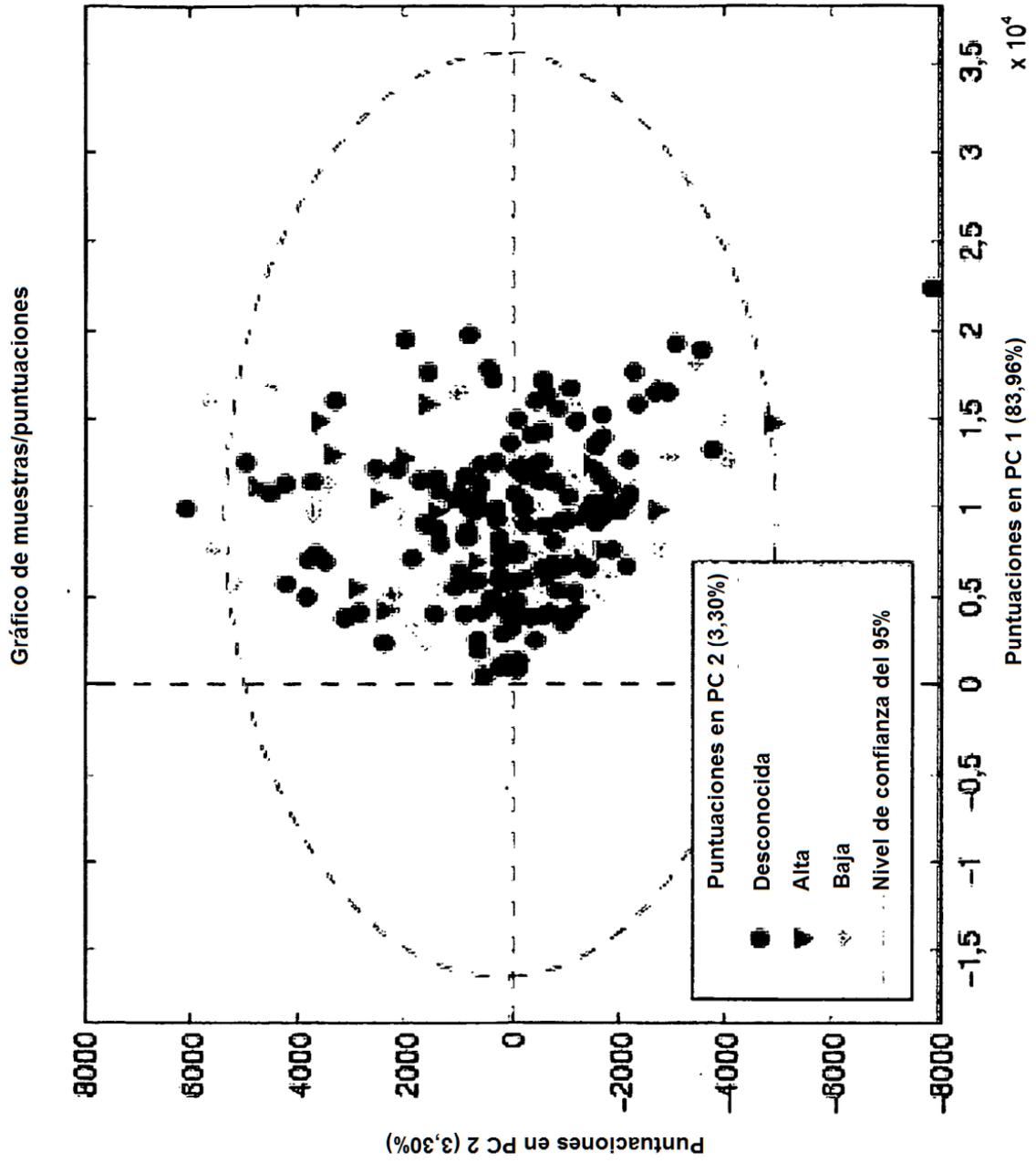


Fig. 7a

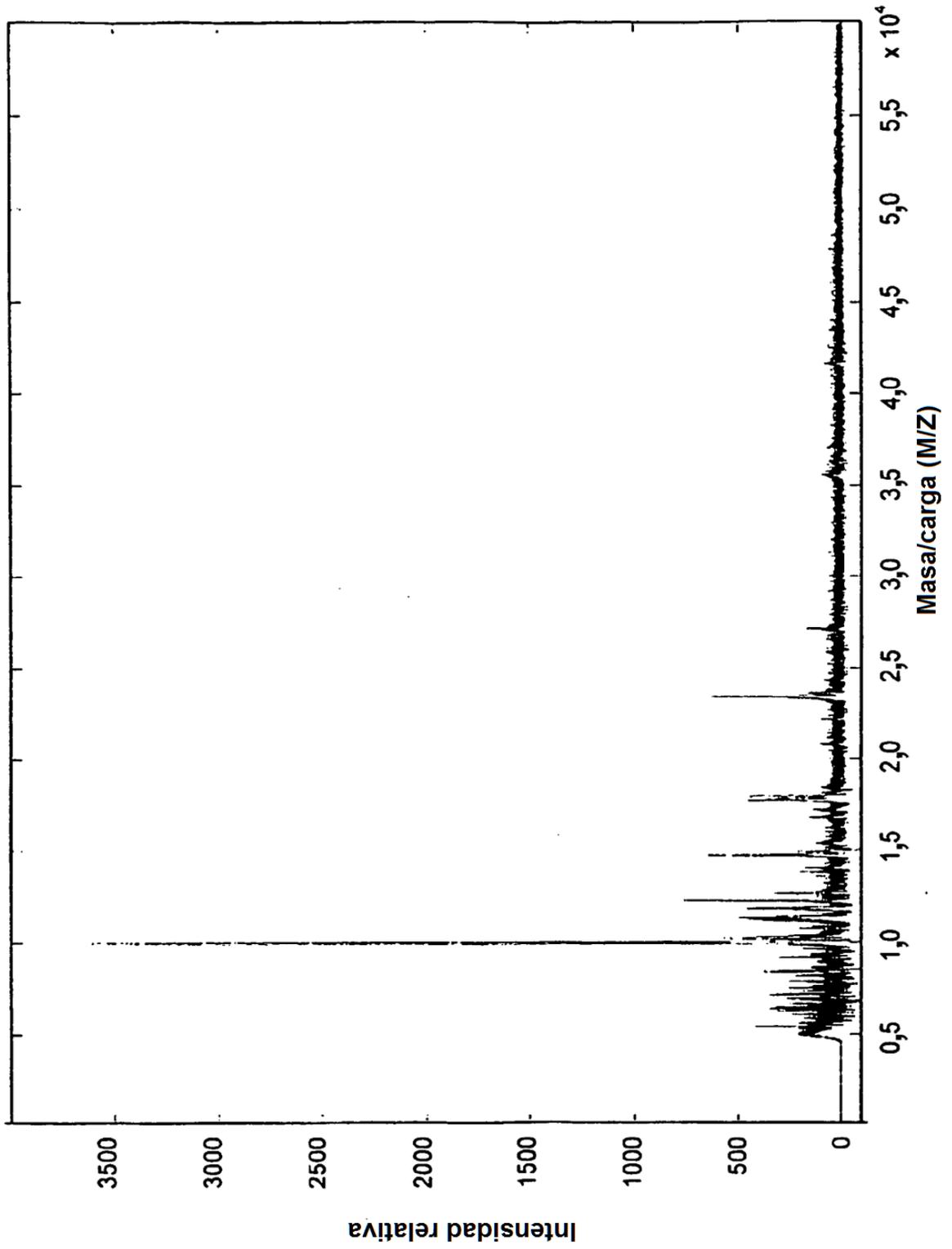


Fig. 7b

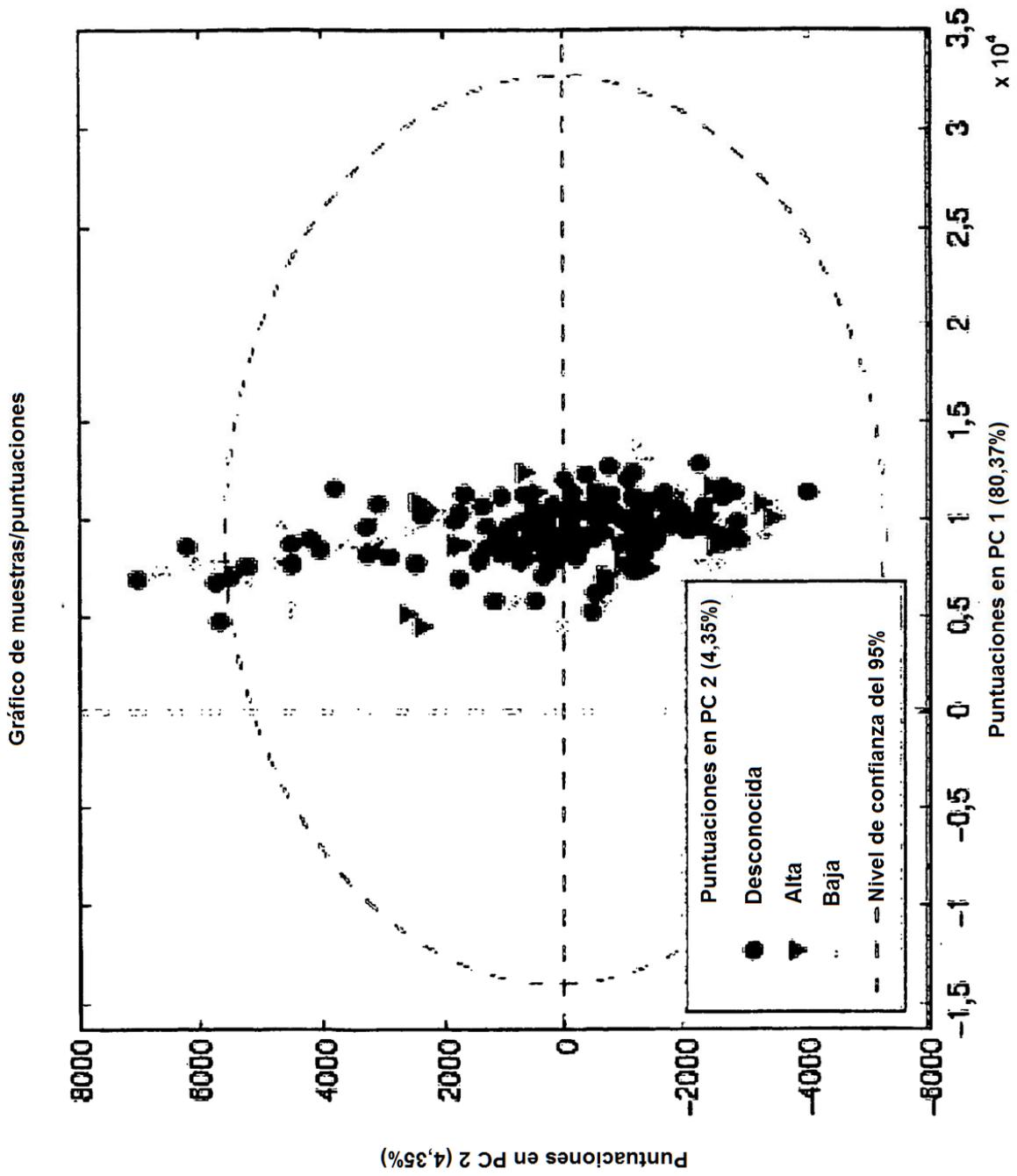


Fig. 8a

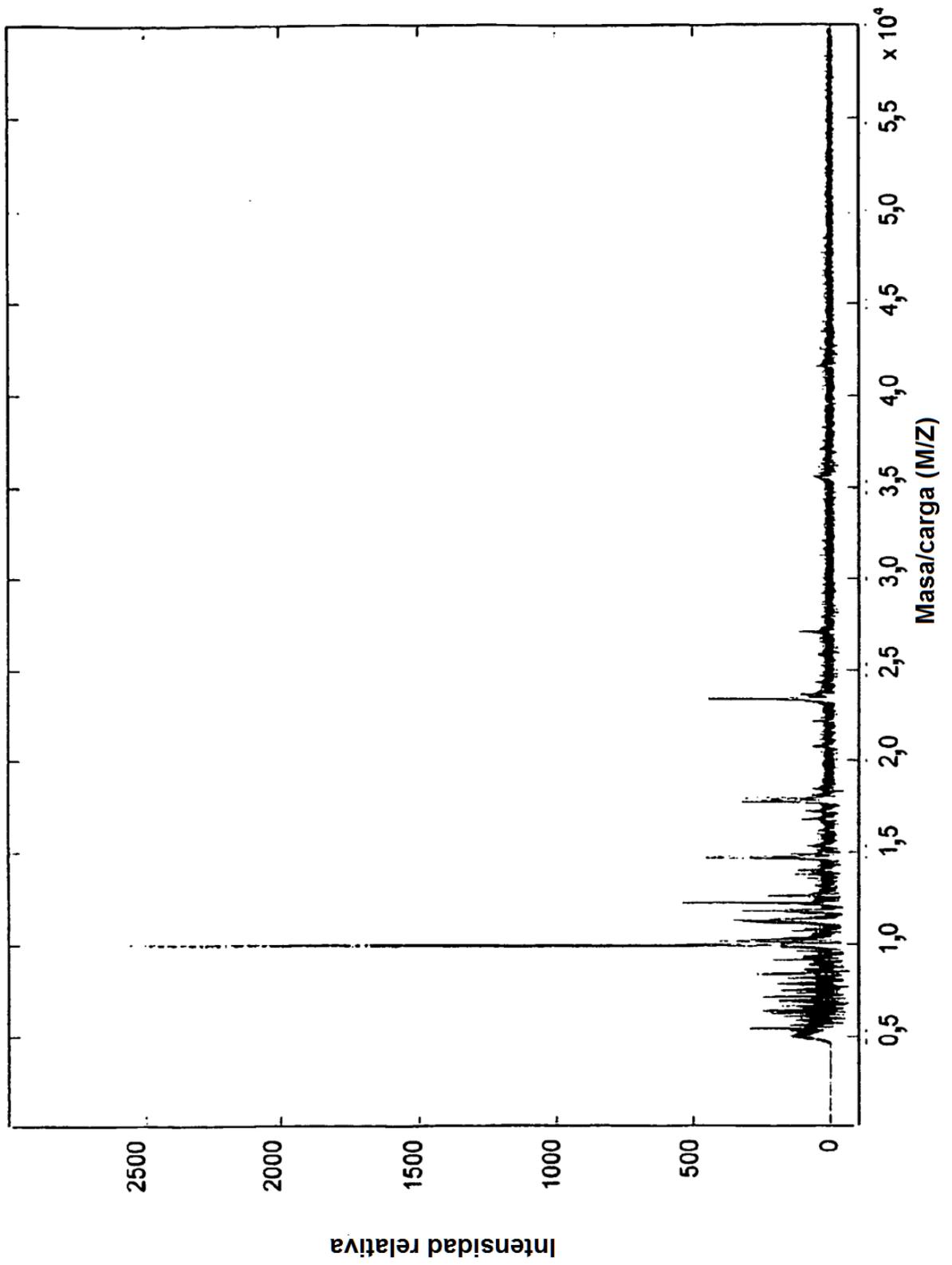


Fig. 8b

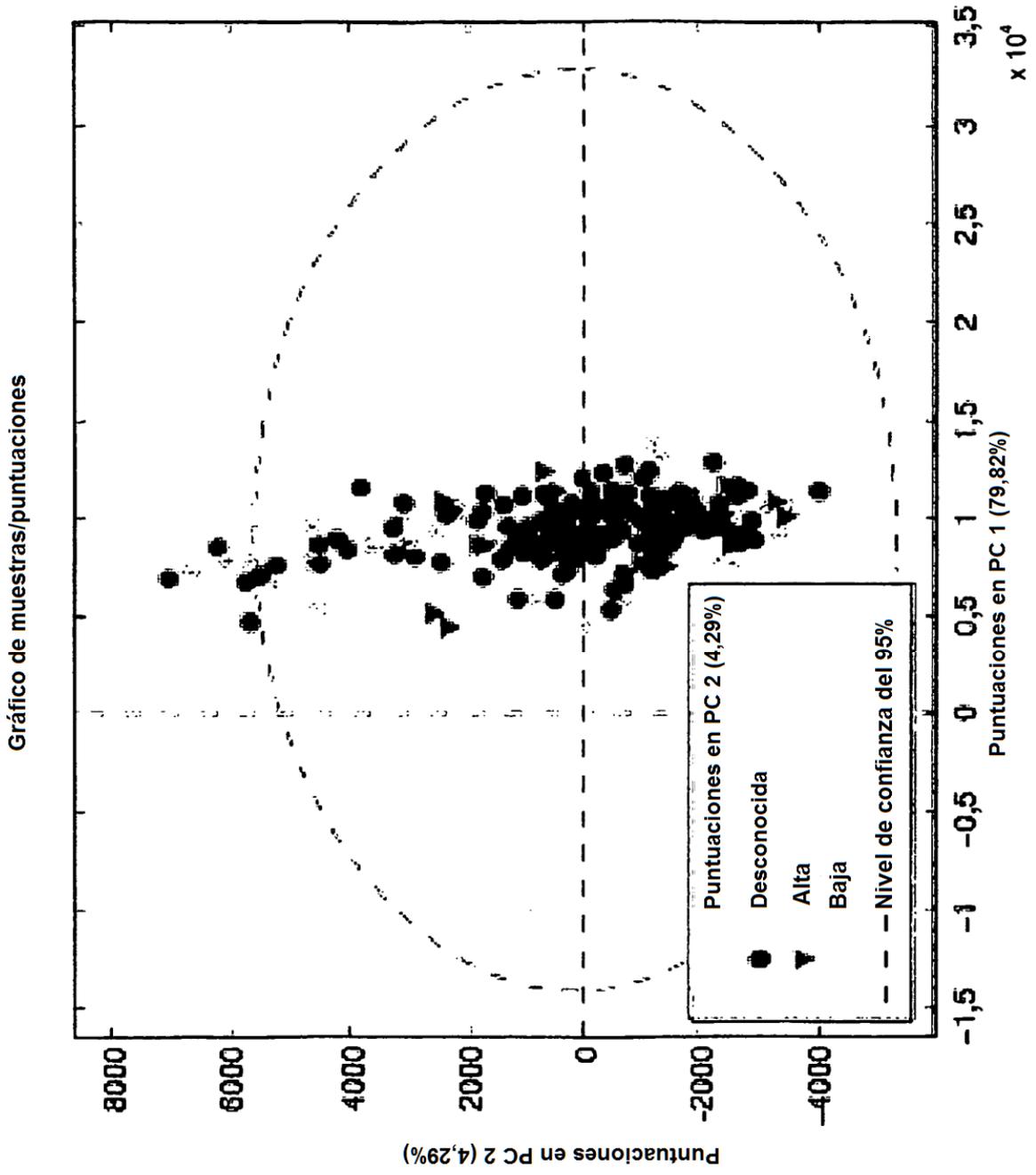


Fig. 9a

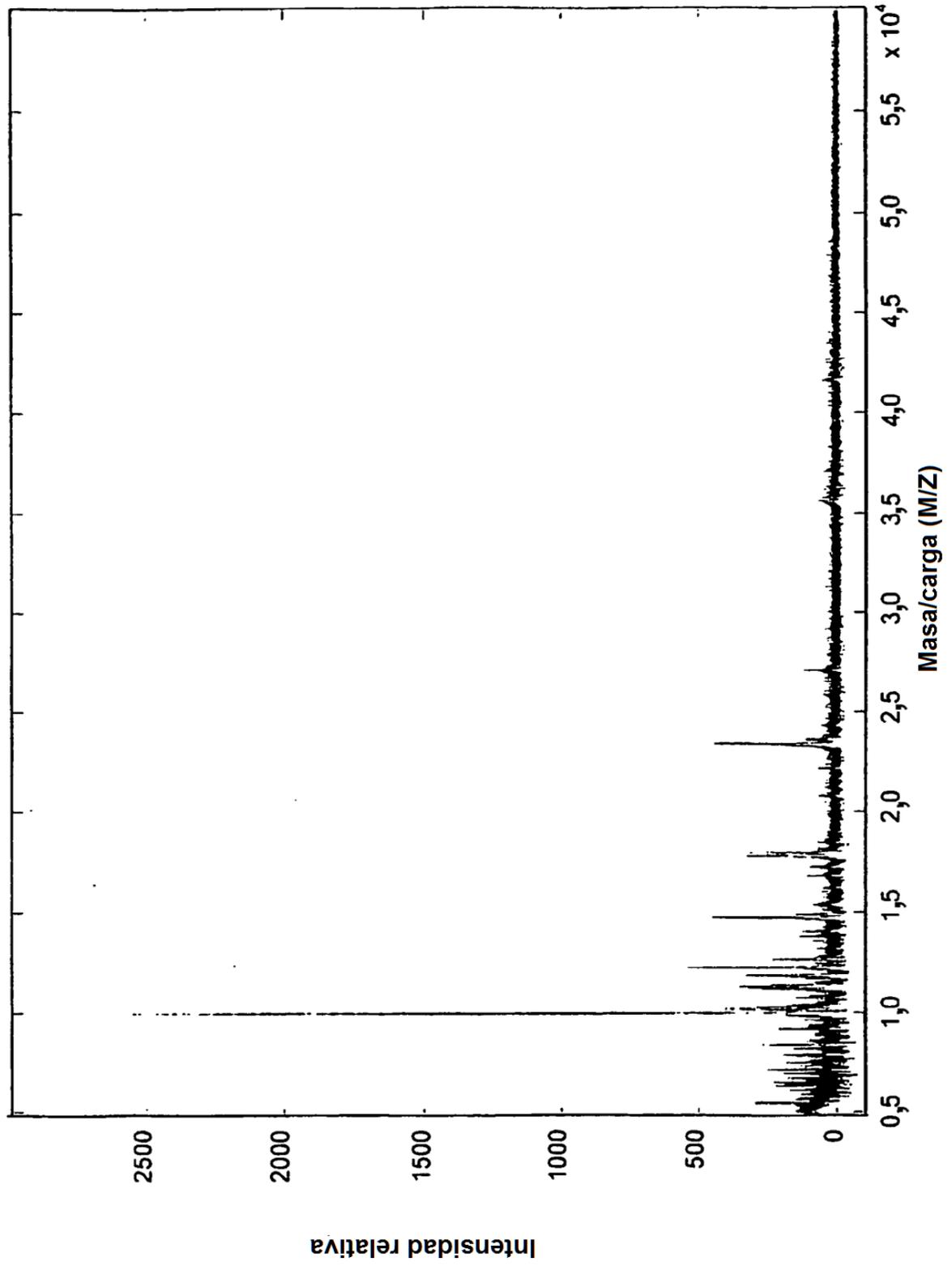


Fig. 9b

Fig. 10

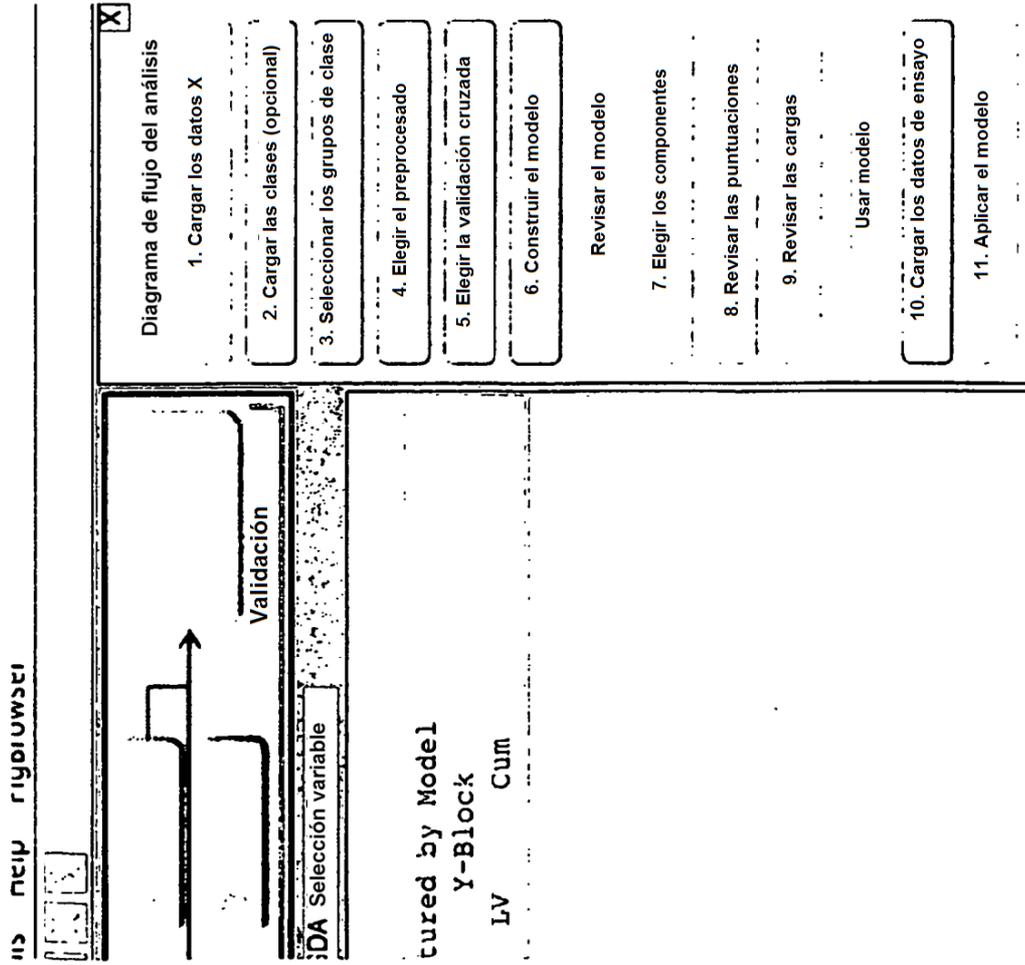


Fig. 11a

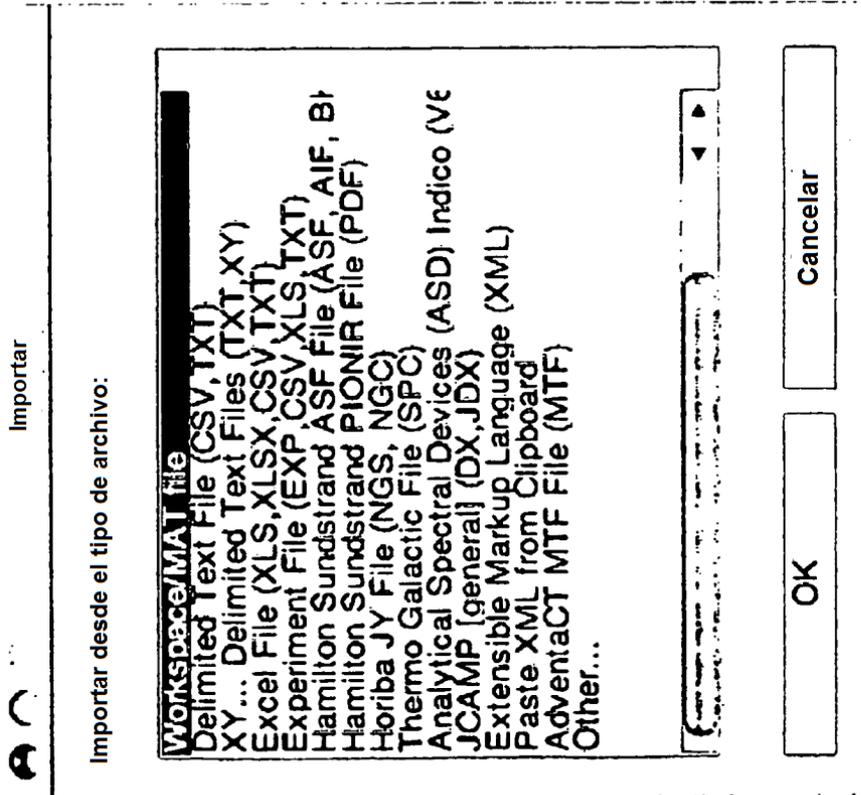


Fig. 11b

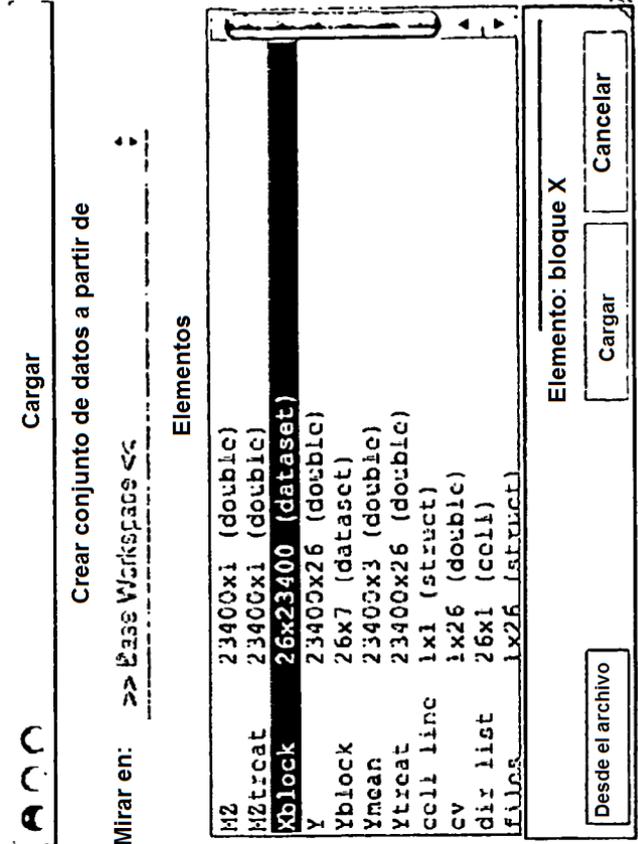
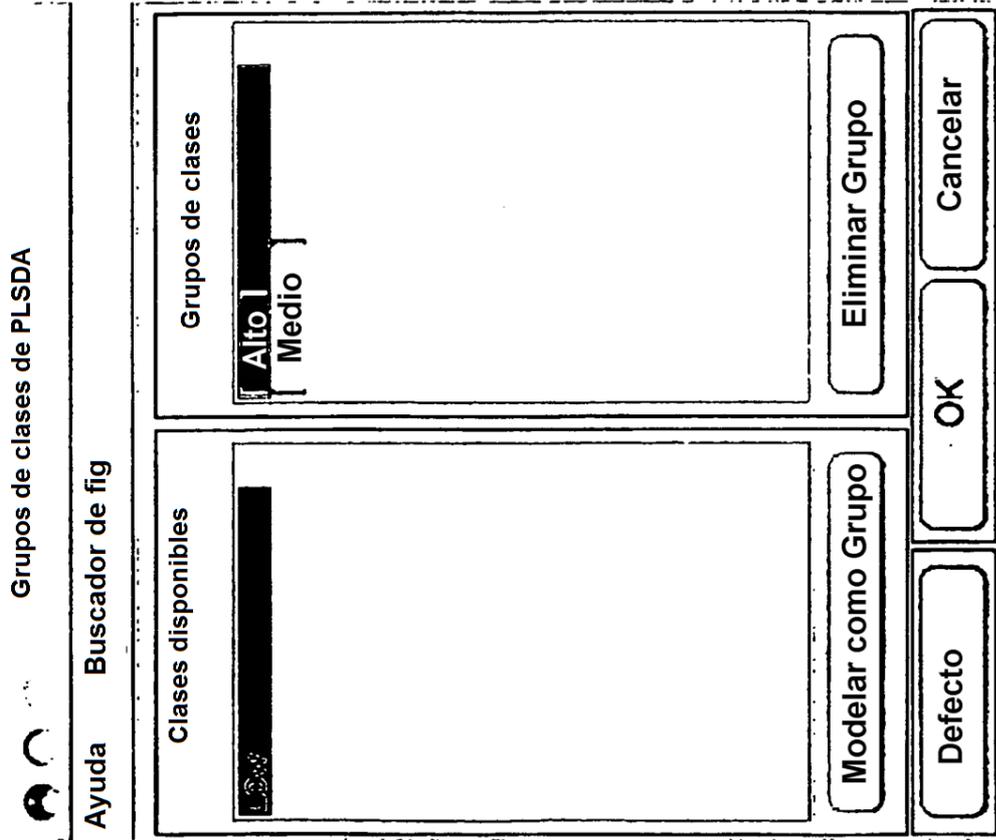


Fig. 12



Preprocesado del bloque X

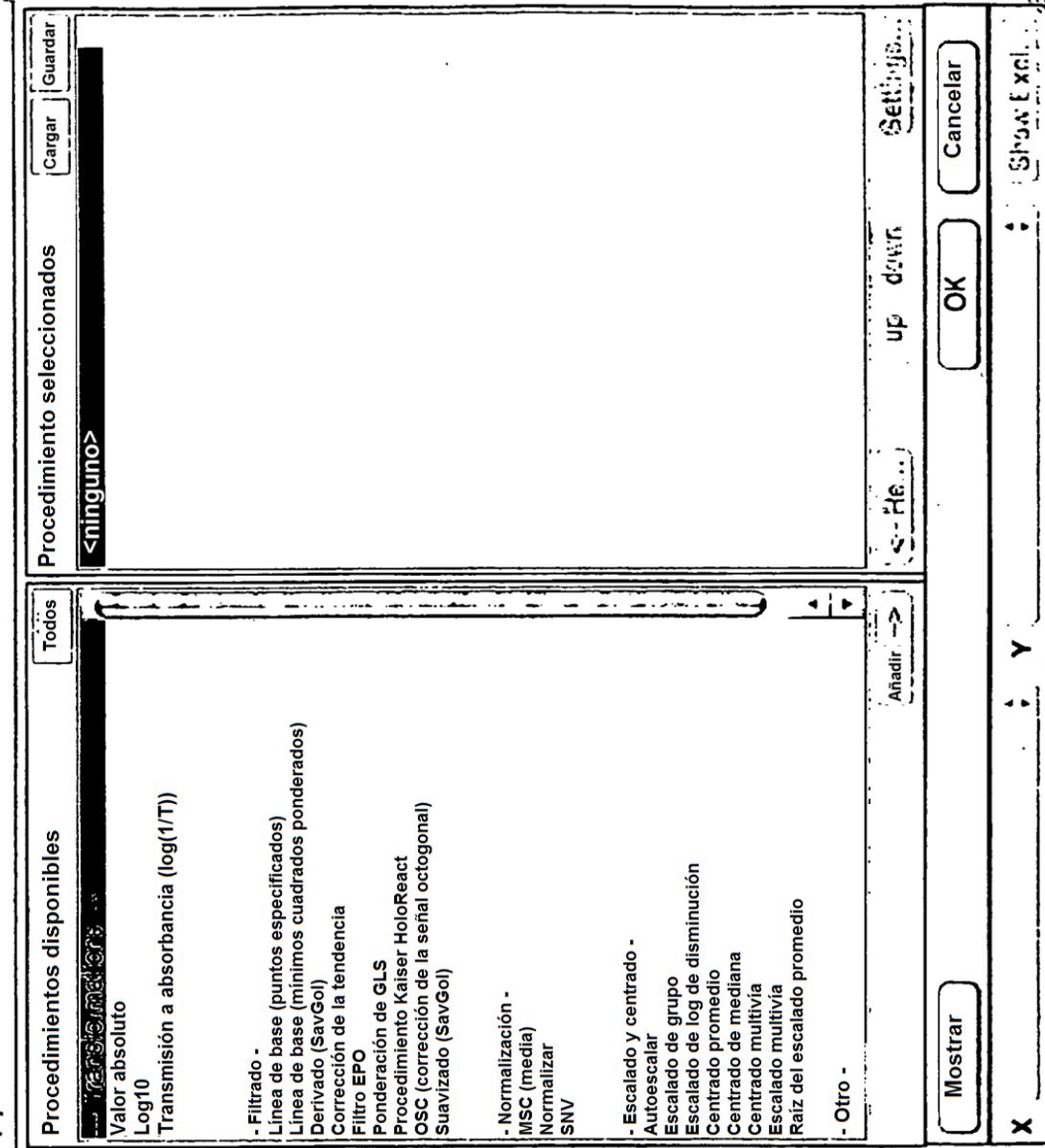


Fig. 13

Fig. 14

**Validación cruzada**

Procedimiento de validación cruzada

- ninguno
- Dejar uno fuera
- Persiana veneciana
- Bloque contiguo
- Subconjuntos aleatorios
- Personalizado

Número máximo de LVs

Número de división de datos

Número de iteraciones

Aplicar Restablecer OK Cerrar

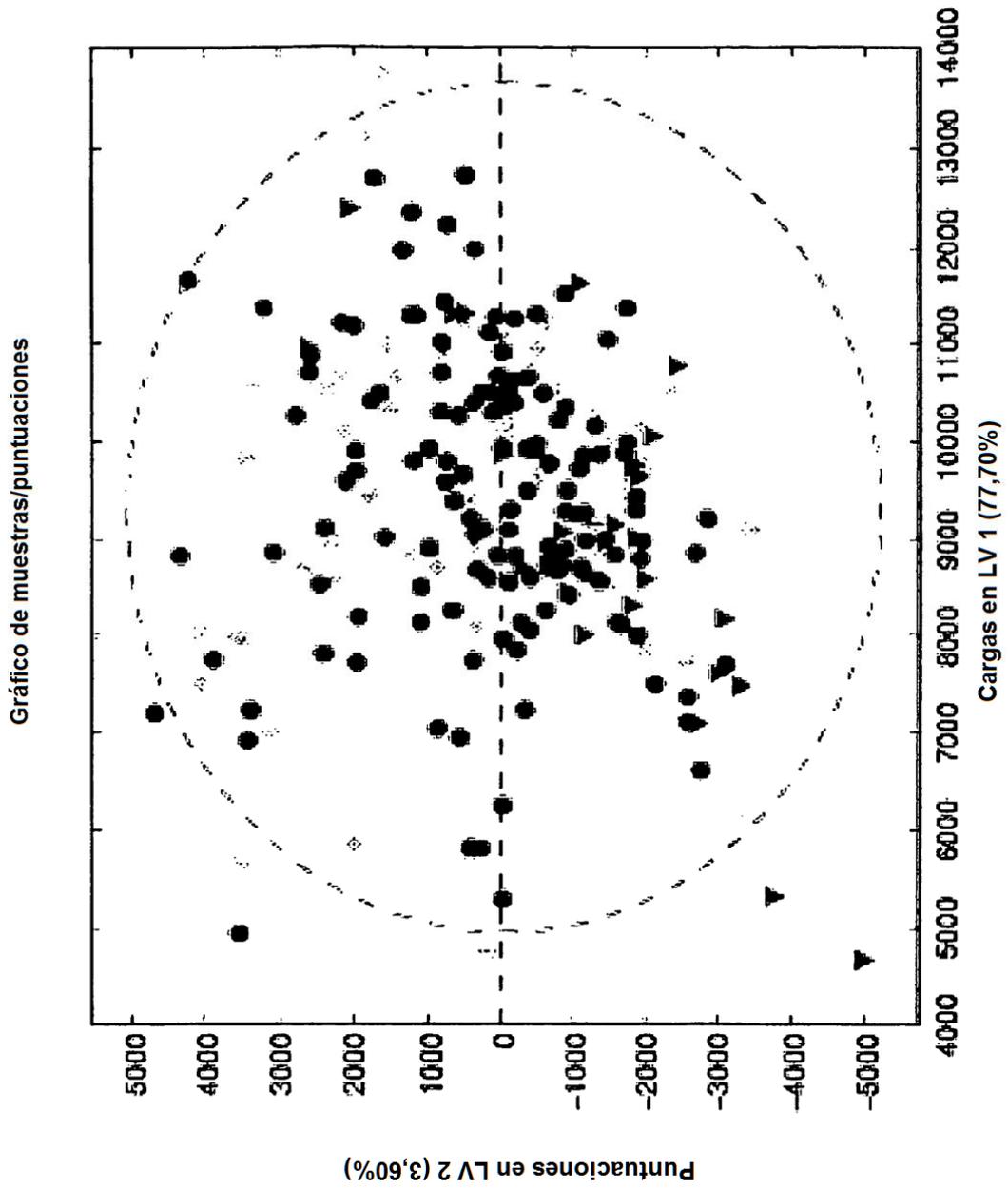


Fig. 15a

Gráfico de muestras/puntuaciones

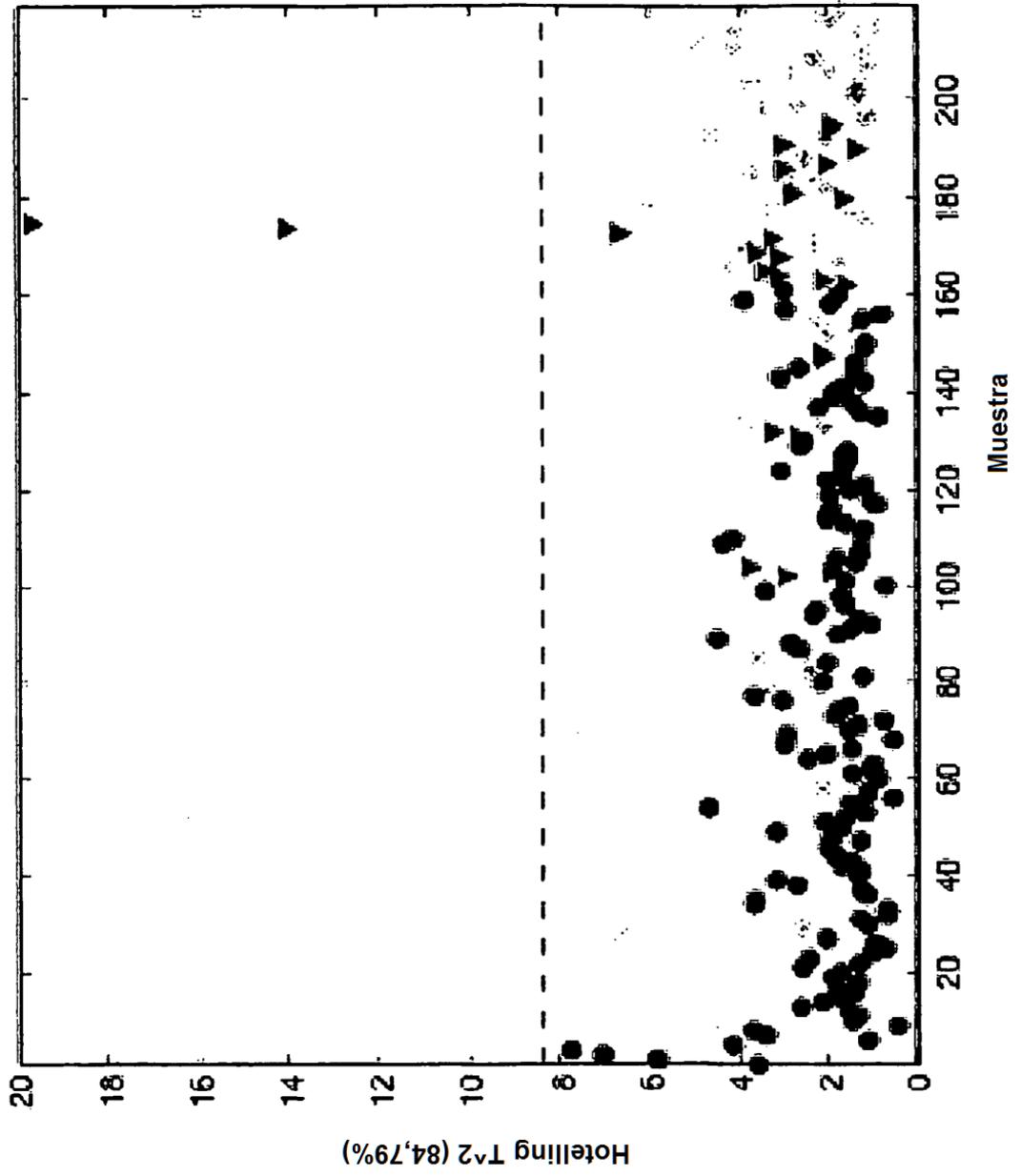


Fig. 15b

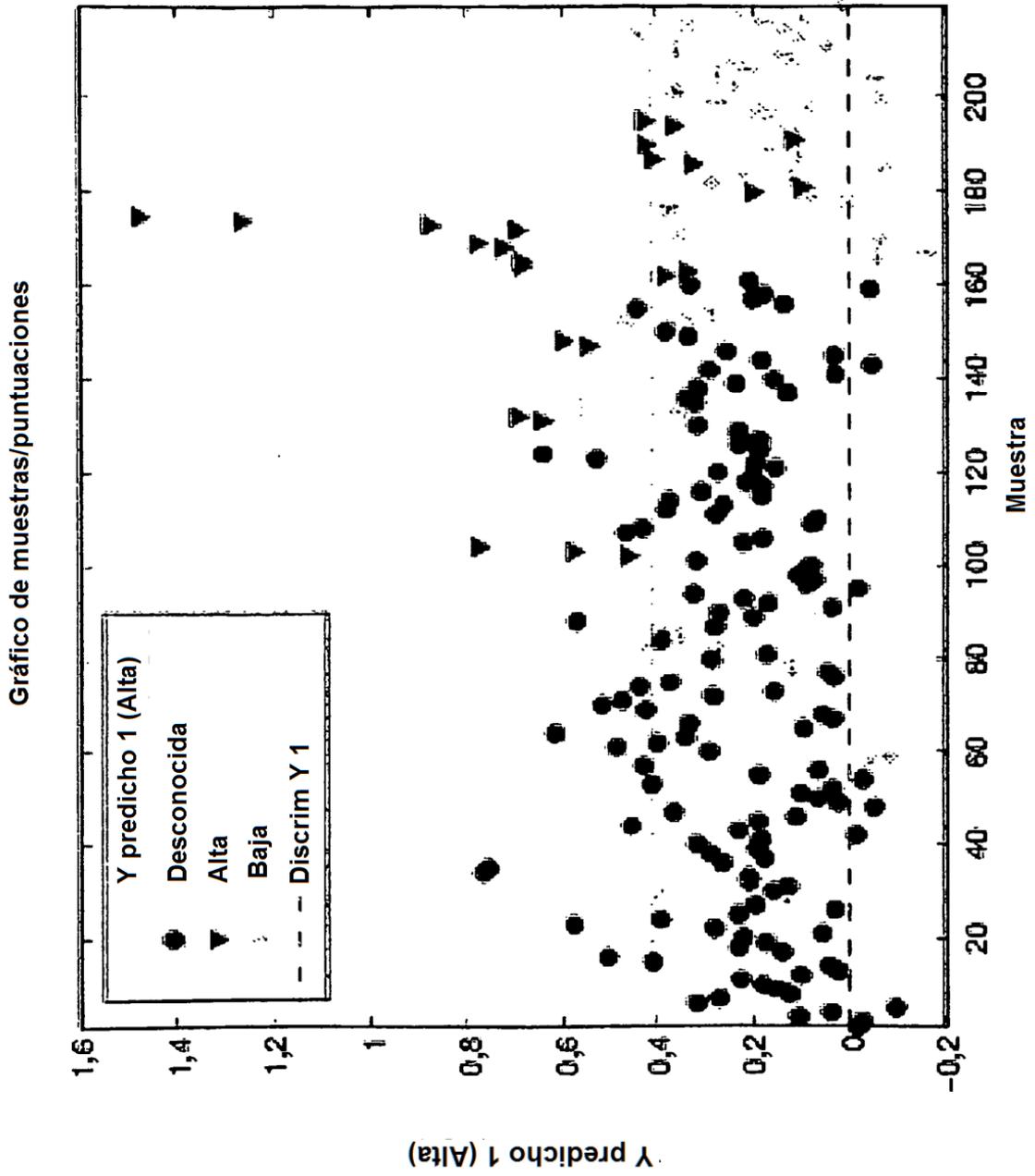


Fig. 15c

Gráfico de muestras/puntuaciones

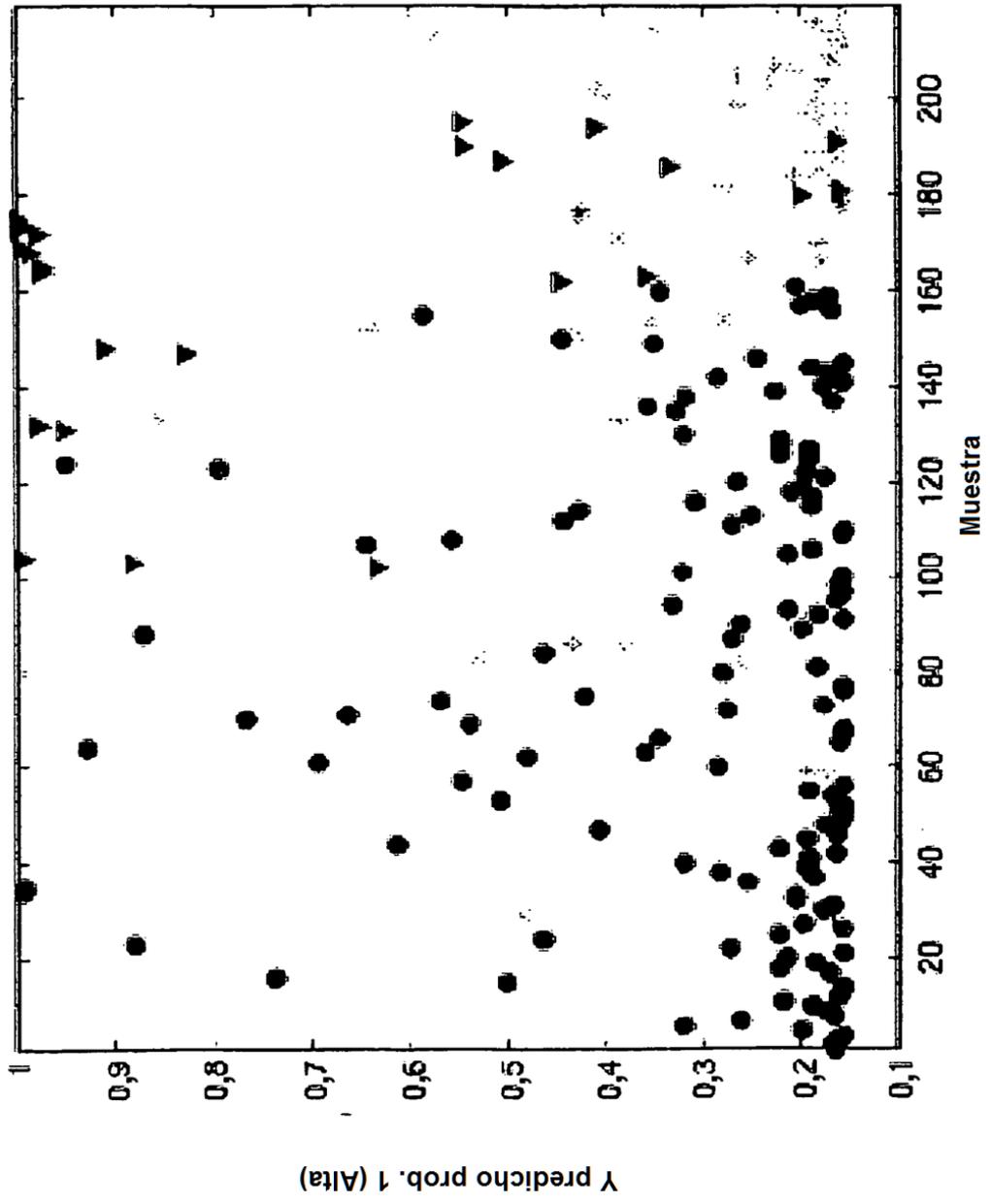


Fig. 15d

Gráfico de variables/cargas

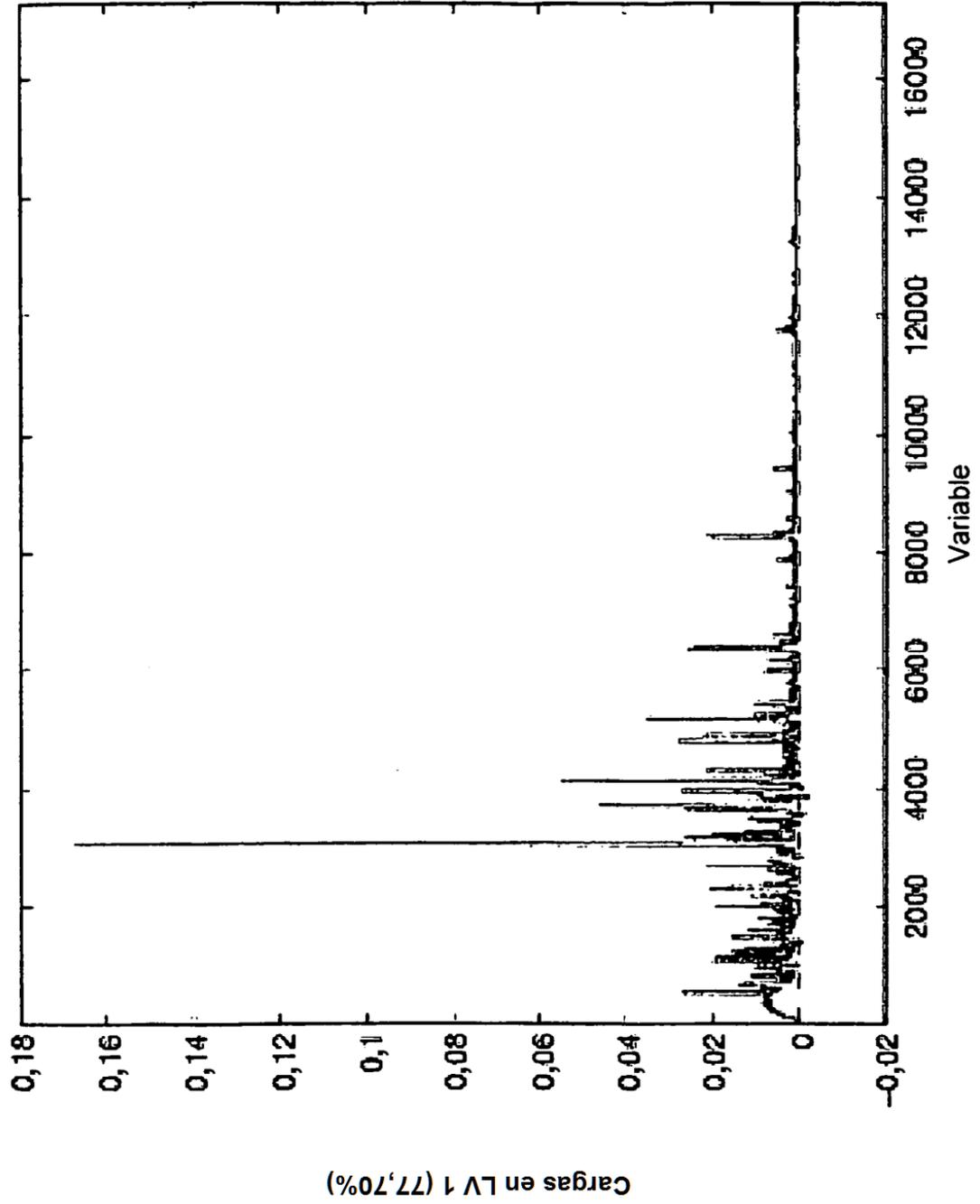


Fig. 16a

Gráfico de variables/cargas

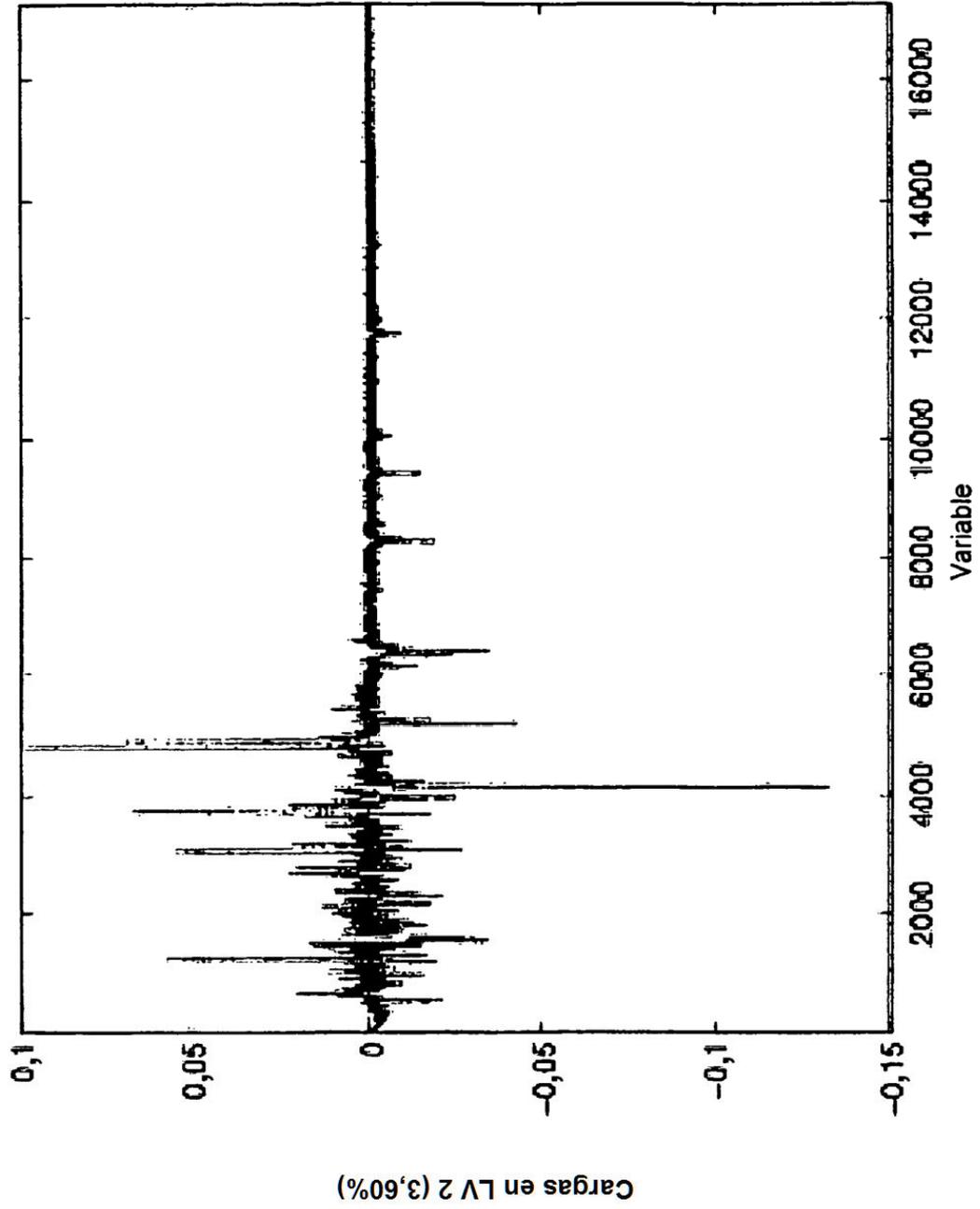


Fig. 16b

Figura 1. Muestras/Puntuaciones - datos de PLSDA S LVs

Archivo Editar Ver Insertar Herram. Escritorio Ventana Ayuda Buscador de fig. Gráfico CUI

Fig. 17

Gráfico de muestras/puntuaciones

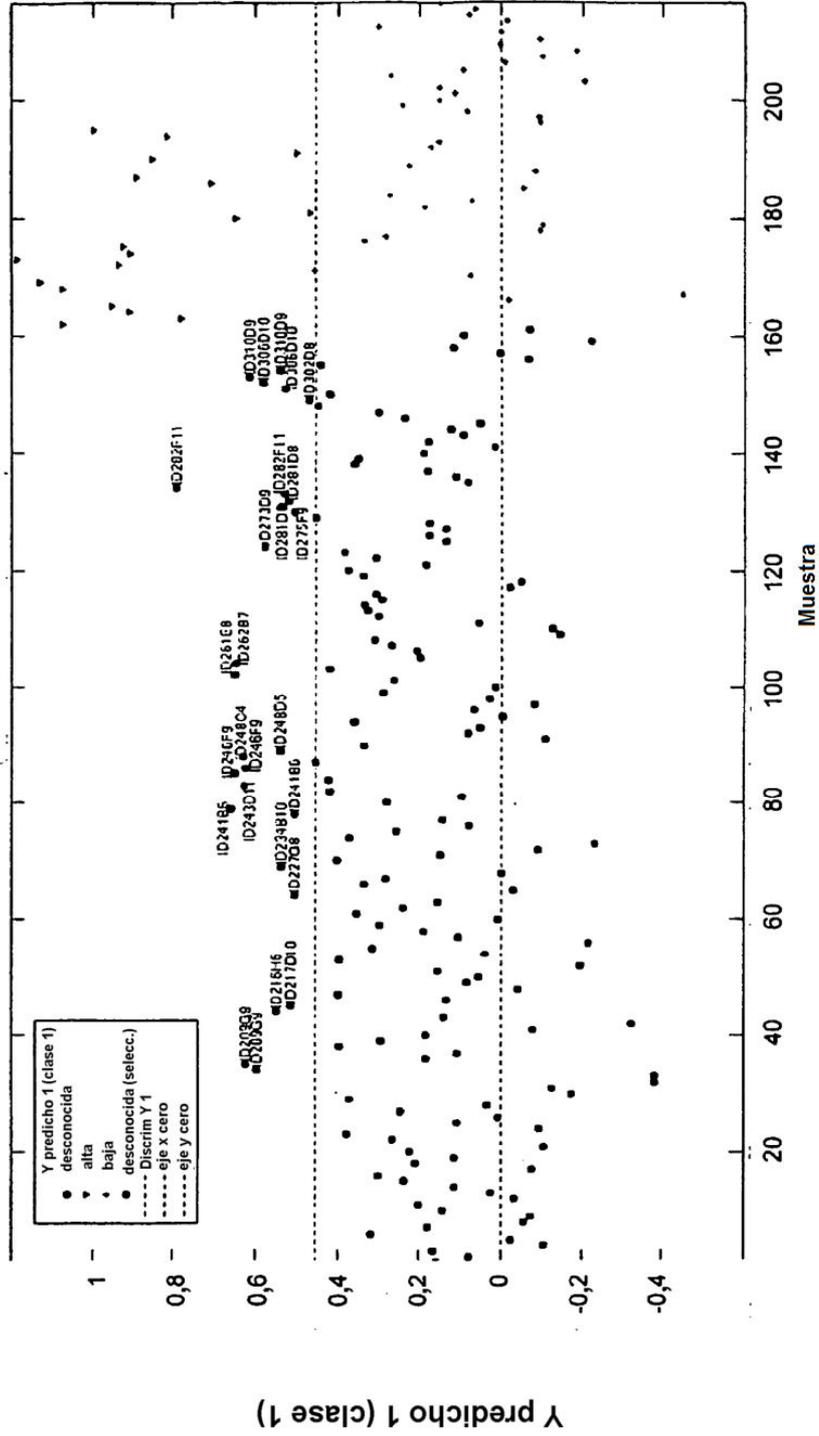


Figura 1. Muestras/Puntuaciones - datos de PLSDA S LVs

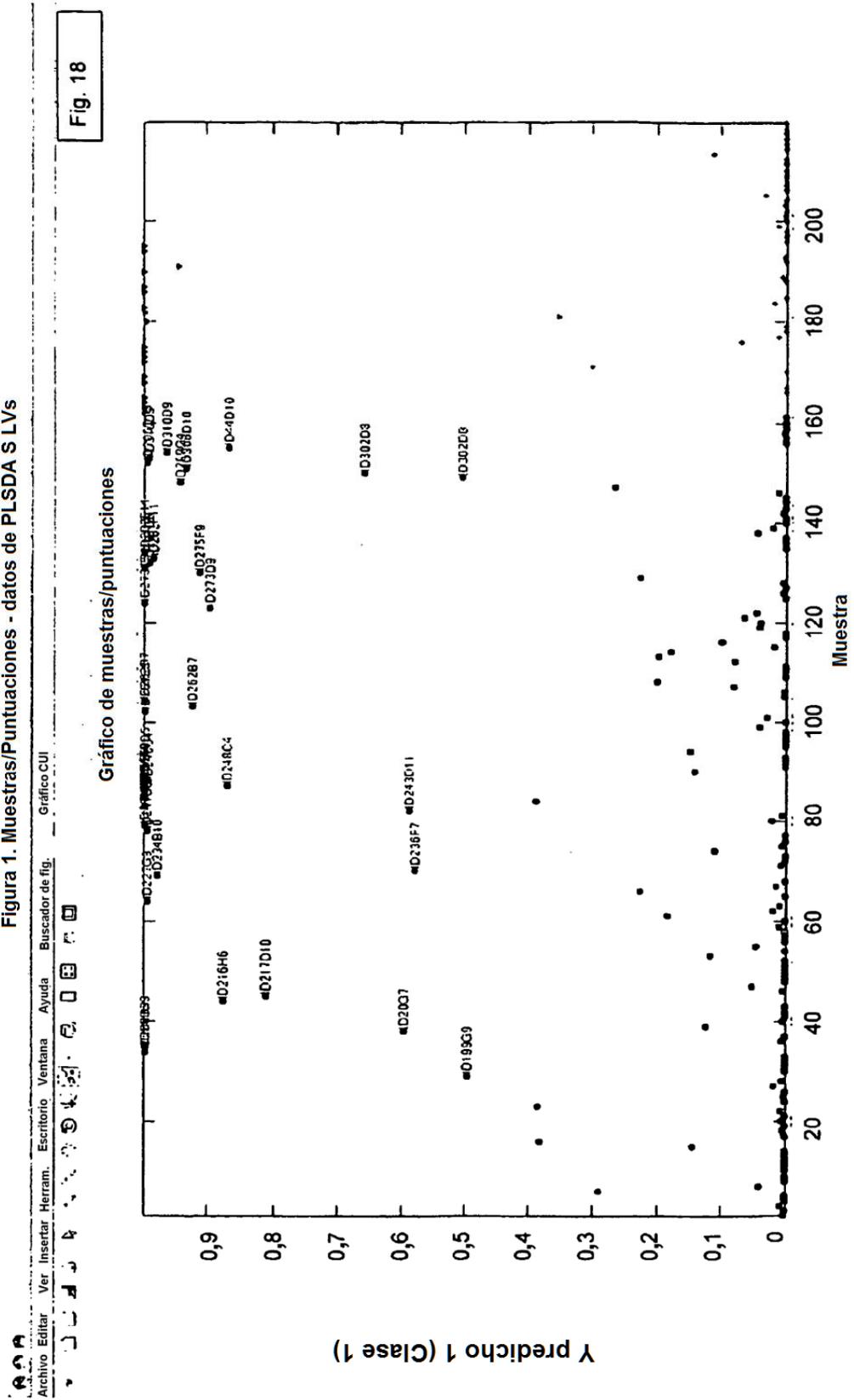


Fig. 18

Archivo Editar Ver Insertar Herram. Escritorio Ventana Ayuda Buscador de fig. Gráfico CUI

Fig. 19a

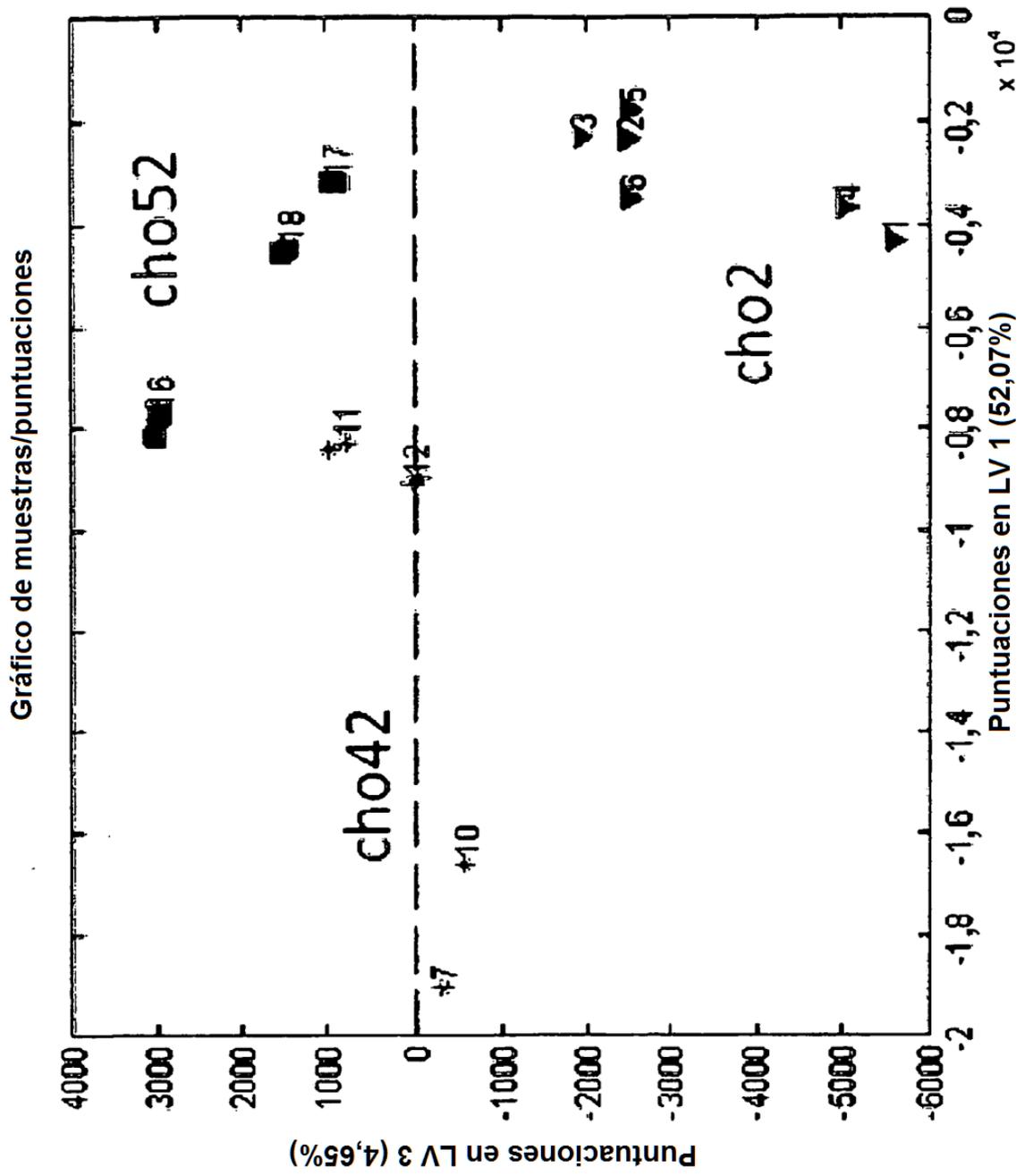
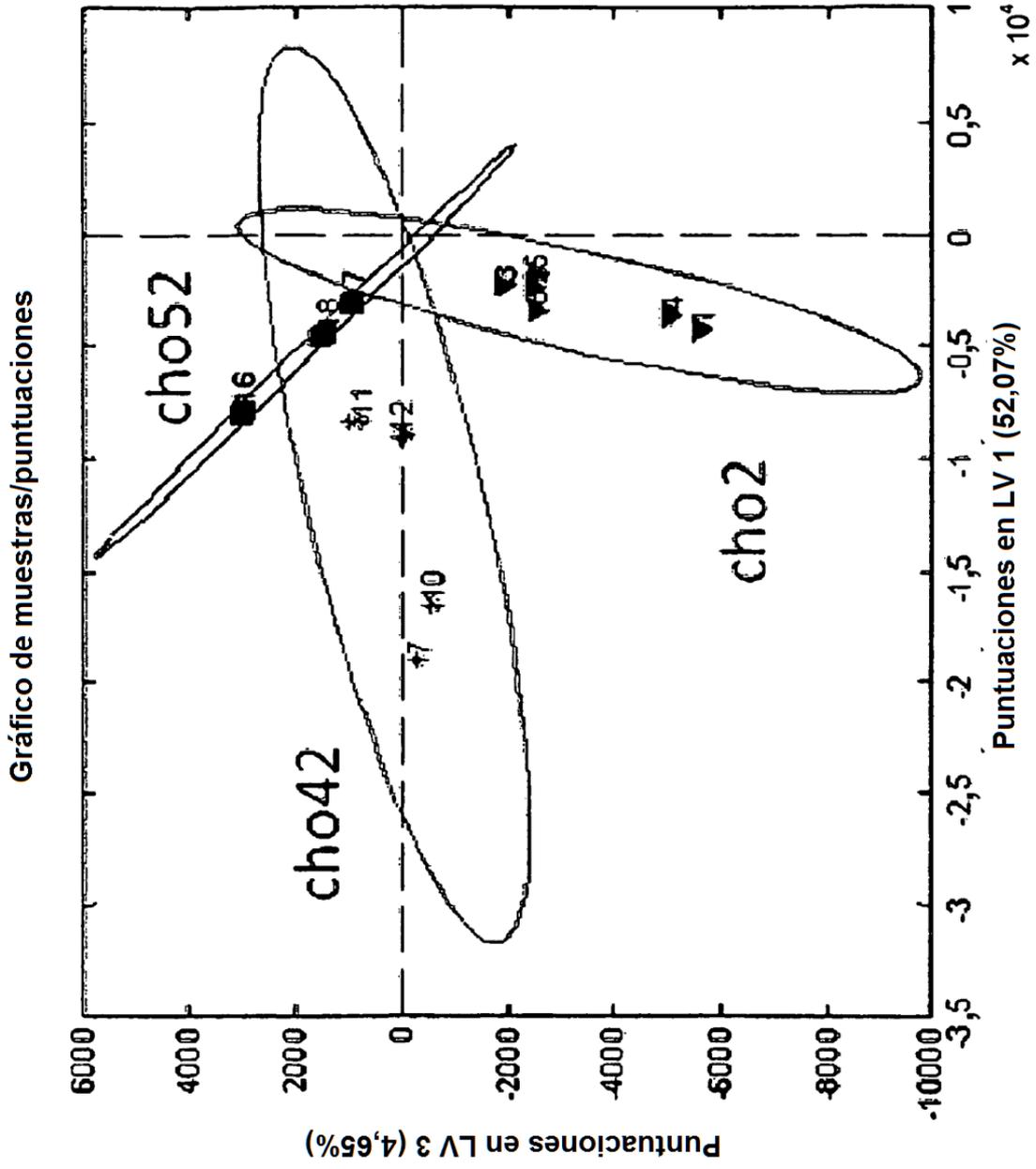


Fig. 19b



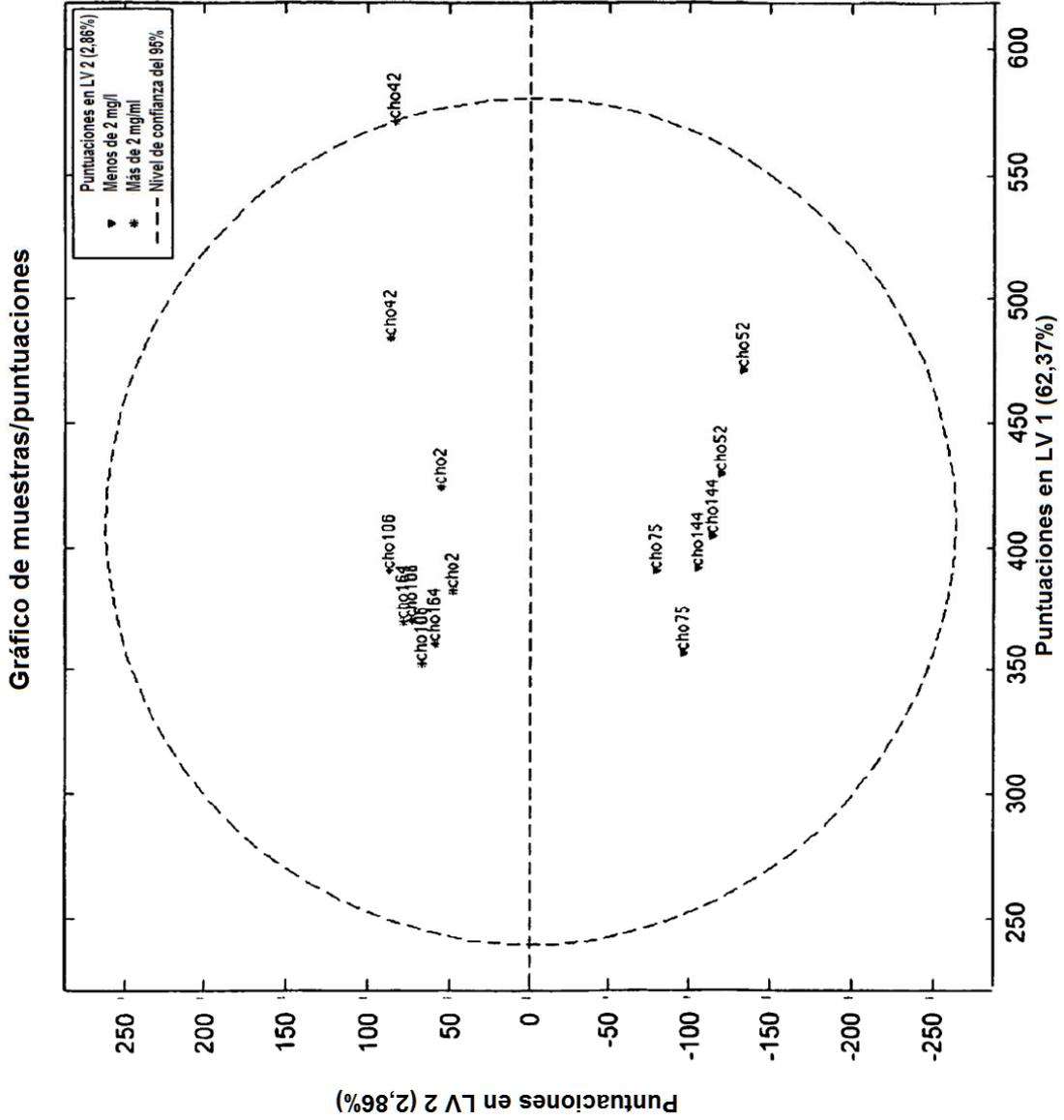


Fig. 20