(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 686 300**

(51) Int. CI.:

C12Q 1/68 (2008.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **24.08.2011 PCT/US2011/048993**

(87) Fecha y número de publicación internacional: **01.03.2012 WO12027491**

(96) Fecha de presentación y número de la solicitud europea: **24.08.2011 E 11820610 (1)**

(97) Fecha y número de publicación de la concesión europea: **27.06.2018 EP 2609220**

(54) Título: **Asociación de variaciones genéticas raras recurrentes con el trastorno por déficit de atención e hiperactividad (TDAH) y procedimientos de uso de las mismas para el diagnóstico y tratamiento del mismo**

(30) Prioridad:

24.08.2010 US 376498 P
23.03.2011 US 201161466657 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
17.10.2018

(73) Titular/es:

THE CHILDREN'S HOSPITAL OF PHILADELPHIA (100.0%)
Research Institute Abramson Research Center Suite 140 Office Of Technology Transfer 3615 Civic Center Boulevard
Philadelphia, PA 19104, US

(72) Inventor/es:

GLESSNER, JOSEPH;
ELIA, JOSEPHINE y
HAKONARSON, HAKON

(74) Agente/Representante:

CARPINTERO LÓPEZ, Mario

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

Asociación de variaciones genéticas raras recurrentes con el trastorno por déficit de atención e hiperactividad (TDAH) y procedimientos de uso de las mismas para el diagnóstico y tratamiento del mismo

La presente solicitud reivindica la prioridad de las Solicitudes Provisionales de los Estados Unidos n.º 61/376.498 y 61/466.657 presentadas el 24 de agosto de 2010 y el 23 de marzo de 2011, respectivamente, cuyos contenidos se incorporan por referencia como si se expusiesen en su totalidad.

Campo de la invención

La presente invención se refiere a los campos de la genética y al diagnóstico del trastorno por déficit de atención e hiperactividad (TDAH). De manera más específica, la invención proporciona composiciones y procedimientos útiles para el diagnóstico y tratamiento del TDAH.

Antecedentes de la invención

A lo largo de la memoria descriptiva se citan varias publicaciones y documentos de patente para describir el estado de la técnica al que pertenece la presente invención. Cada una de las citas se incorpora al presente documento por referencia como si se expusiese en su totalidad.

El trastorno por déficit de atención e hiperactividad (TDAH) es un trastorno neuropsiquiátrico común con una estimación de heredabilidad en el intervalo del 30 al 90% (Derks y col. 2008; Wood, y col. 2008; Haberstic, y col. 2008). La mayoría de los trastornos del neurodesarrollo han sido resistentes a la estrategia de asociación por todo el genoma (GWA, por sus siglas en inglés), aunque se han hechos avances recientes en el autismo (Glessner y col. 2009; Derks, y col. 2008; Wod, y col. 2008; Haberstic, y col. 2008; Wang, y col. 2009). Se han comunicado estudios GWA de TDAH utilizando una cohorte de 958 tríos de progenitores-niño reclutados a través del estudio *International Multicentre ADHD Genetics* (IMAGE, por sus siglas en inglés). Los resultados de estos estudios no aportaron asociación alguna con un nivel de significación por todo el genoma (Franke y col. 2009; Neale, y col. 2008). Usando mediciones cuantitativas de TDAH, Lasky-Su y colaboradores comunicaron recientemente una evidencia nominal a partir de un análisis PBAT de marcaje de SNP ubicados en *CDH13* (rs6565113) y *GFOD1* (rs552655) (Lasky-Su, y col. 2008). También se informó de un SNP en fuerte desequilibrio de unión con rs6565113 que tiene impacto en *CDH13* en un estudio GWA de una muestra independiente de adultos con TDAH (Lesch, y col. 2008). Los presentes inventores informaron acerca de loci de variación de número de copias (CNV) observados en los primeros 335 casos de TDAH que reclutaron (Elia y col. 2009). Aunque ninguno de los loci de CNV detectados en ese estudio satisfizo los criterios de significación, resulta destacable que en una familia se observó que tenían una eliminación de *GRM5* que tenía impacto en los tres niños afectados, heredada de su padre afectado. Además, en una familia con TDAH se detectó una eliminación de *GRM7* (Elia y col. 2009). Los CNV de los receptores de glutamato metabotrópicos (mGluR) además del descubrimiento del gen NK3 en el TDAH han sugerido nuevas estrategias terapéuticas para el tratamiento del TDAH.

Es altamente deseable el desarrollo de pruebas diagnósticas precisas mejoradas para este trastorno basándose en alteraciones genéticas asociadas. Dichas pruebas facilitarían un diagnóstico definitivo y proporcionarían vías para el desarrollo de agentes terapéuticos eficaces para el tratamiento del TDAH.

Sumario de la invención

De acuerdo con la presente invención, se proporcionan procedimientos para el diagnóstico y tratamiento del TDAH. Un procedimiento ilustrativo implica detectar la presencia de al menos un CNV en un polinucleótido diana en el que en caso de que dicho(s) CNV esté(n) presente(s), dicho paciente tiene un riesgo aumentado de desarrollar TDAH.

En un aspecto de la presente invención, se proporciona un procedimiento para detectar la propensión a desarrollar trastorno por déficit de atención e hiperactividad (TDAH) en un paciente que lo necesite. Un procedimiento ilustrativo implica detectar la presencia de al menos un ácido nucleico que contiene SNP en un polinucleótido diana, siendo dicho SNP informativo de la presencia de una variación del número de copias (CNV) asociada con el TDAH, en el que en caso de que esté presente dicho SNP, dicho paciente tiene un riesgo aumentado de desarrollar TDAH, en el que dicho ácido nucleico que contiene SNP se proporciona en la tabla 13.

En otra realización de la invención, se proporciona un procedimiento para identificar agentes que alteran la señalización y/o la morfología neuronal. Dicho procedimiento comprende proporcionar células que expresan al menos un ácido nucleico que comprende los CNV asociados con el TDAH de la invención, (etapa a); proporcionar células que expresan las secuencias de tipo silvestre afines que carecen del CNV (etapa b); poner en contacto las células de cada muestra con un agente de ensayo y analizar si dicho agente altera la señalización y/o morfología neuronal de las células de la etapa a) en relación a las de la etapa b), identificando de este modo agentes que alteran la señalización y morfología neuronal. En una realización preferida, el agente de ensayo modula la actividad génica del receptor de glutamato metabotrópico (mGluR). En otra realización, el agente de ensayo se selecciona entre el grupo que consiste en moduladores alostéricos positivos (PAM) de mGluR (por ejemplo, PAM de mGluR5, PAM de mGluR7), un modulador alostérico negativo (NAM) de mGluR (por ejemplo, NAM de mGluR2/3) y un

antagonista del receptor de taquicinina-3/neurocinina-3 (TAC3/NK3R). En otra realización, el agente de ensayo se selecciona entre el grupo que consiste en ADX63365, ADX50938, ADX71149, ADX48621, AMN082, derivados de 1-(hetero)aril-3-amino-pirrolidina (por ejemplo, los proporcionados en la Publicación de Solicitud de Patente de los Estados Unidos n.º 2008/0300266), LY341495, GSK1144814 y SB223412. La presente invención también abarca procedimientos para tratar a pacientes con TDAH mediante la administración de agentes de ensayo identificados usando los procedimientos descritos en el presente documento en pacientes que lo necesiten. La invención también proporciona al menos un ácido nucleico que contiene SNP relacionado con el TDAH seleccionado entre el grupo listado en la **tabla 13**. En una realización, se proporciona un panel de SNP multiplexado que contiene todos los SNP informativos de la **tabla 13**. Dichos ácidos nucleicos que contienen SNP que indican la presencia de CNV asociados con el TDAH pueden estar contenidos opcionalmente en un vector de expresión adecuado para su expresión en células neuronales. Como alternativa, pueden estar inmovilizados sobre un soporte sólido. En otra alternativa más, el panel puede proporcionarse *in silico*.

De acuerdo con otro aspecto más de la presente invención, se proporciona un procedimiento para tratar el TDAH en un paciente en el que se determina que tiene al menos un polimorfismo de un solo nucleótido prescrito indicativo de la presencia de una variación en el número de copias asociada con el TDAH, como se describe más adelante en el presente documento, mediante la administración al paciente de una cantidad terapéuticamente eficaz de al menos un miembro de la familia de agentes nootrópicos de piracetam. Este procedimiento proporciona una prueba y un paradigma de tratamiento, mediante el cual se usa el perfil genético de un paciente para personalizar el tratamiento con agentes terapéuticos dirigidos contra defectos neurofisiológicos específicos hallados en individuos que presentan TDAH. Dicha prueba y modelo de tratamiento pueden beneficiar a hasta un 50% de los pacientes con TDAH con mayor eficacia y menos efectos secundarios que el tratamiento no personalizado. Por lo tanto, puede evaluarse la presencia de dicha alteración genética en cualquiera de los pacientes que presenten una alteración en la señalización glutaminérgica y después, tratarse con el producto farmacéutico adecuado, tal como los agentes listados anteriormente.

Breve descripción de los dibujos

Figura 1. Una distribución gráfica de las llamadas de CNV por casos individuales (panel superior) en comparación con los controles (panel inferior).

Figura 2. Una presentación gráfica de los datos de nivel de SNP normalizados de Perlegen 600K. El eje X muestra la posición del par de bases en megabases en el cromosoma 11. Datos de nivel de SNP en bruto que muestran la eliminación de GRM5 en cinco muestras de los datos de IMAGE Perlegen 600K normalizados mediante el protocolo adaptado de PennCNV-Affy. Se representan los datos de genotipado denominados frecuencia del alelo B (BAF) y los datos de intensidad denominados Relación de Log R (LRR).

Figura 3. Gráficas de todos los datos a nivel de SNP: A) Datos de Perlegen 600K normalizados, B datos de Illumina 1M PUWMa normalizados y C) datos de Affymetrix 5.0 IMAGE II normalizados.

Figura 4. Una presentación gráfica de los datos de validación independientes de IMAGE Perlegen 600K. Se diseñaron ensayos de qPCR a base de sondas fluorescentes usando sondas universales de Roche para validar cada CNV candidato con una prueba completamente independiente (se encontraban disponibles para su validación 11 de las 14 muestras de IMAGE con llamadas de CNV replicantes para los loci indicados y todas se validaron en comparación con pares de control; los otros 3 loci se validaron visualmente). Las barras de error indican la desviación estándar de ejecuciones por cuadruplicado. Del, eliminación; Dup, duplicación.

Figura 5. Una ilustración del análisis de componentes principales de Eigenstrat. Se analizaron simultáneamente los casos y los controles para minimizar la subestructura de la población en la asociación de CNV de control de casos. Se retiraron las muestras que se desviaban de la agrupación caucásica mostrada. El factor de inflación genómica (GIF, por sus siglas en inglés) dentro de Plink se encontraba a un nivel aceptable (GIF=1,02409). También se comprobó la concordancia de población emparejada para comprobar y filtrar los grados de relación crípticos que podrían dar lugar a CNV raros específicos para subpoblaciones ultra-estratificadas de Europa.

Figura 6. Un ejemplo de la estadística basada en SNP aplicada y de la región de más alga significación llamada. Se muestran ejemplos del crom. 3: A) 1.327.963-2.376.095 y B) 1.847.000-1.862.261. El solapamiento de CNV complejos se simplifica produciendo estadísticas basadas en SNP. Tal como se observa en las gráficas para casos eliminados y controles eliminados, cada SNP tiene un número específico de CNV. Los casos y controles se comparan con una prueba exacta de Fisher y en la tercera gráfica se muestra el valor de log p negativo. Se comunican regiones de significación dentro de una potencia de diez y se comunica la región de significación más elevada (valor de p mínimo local) dentro de 1MB. La gráfica de casos de IMAGE eliminados muestra únicamente un caso, la muestra n.º 11939, ya que el resto de regiones en rojo 3' son progenitores.

Figura 7. Validación independiente de Illumina Human Hap550 del CHOP usando qPCR. Se diseñaron ensayos qPCR basados en sondas fluorescentes usando sonda universal de Roche para validar cada CNV candidato con una prueba completamente independiente (se muestran series representativas para cada locus en los pares de caso y control). Las barras de error indican la desviación estándar de ejecuciones por cuadruplicado. Del,

eliminación; Dup, duplicación.

Figura 8. Ejemplos de observación de CNV basándose en la frecuencia de alelo B (BAF) y la relación de log R (LRR).

Figura 9. Una ilustración de la eliminación que tiene impacto directo en *GRM5,* exclusiva para casos de TDAH y replicada en IMAGE y PUWMa. Se encontraron en cuatro casos de TDAH del CHOP eliminaciones hemicigóticas en *GRM5* replicadas por 2 eliminaciones y 3 eliminaciones mayores en IMAGE y 1 eliminación de PUWMa. La cobertura de SNP de las matrices Illumina 550k, Perlegen 600k, Illumina 1M y Affymetrix 5.0 se muestran como líneas azules verticales.

Figura 10. Una presentación de redes de interacción génica del receptor GRM afectadas en el TDAH. Los genes receptores GRM se muestran como nódulos grandes con forma de rombo mientras que otros genes que interactúan con 2 grados de interacción se muestran como nódulos circulares más pequeños. Los nódulos se colorean para representar el enriquecimiento de CNV: las eliminaciones enriquecidas en los casos en rojo oscuro, las eliminaciones enriquecidas en los controles en rojo claro, las duplicaciones enriquecidas en los casos en verde oscuro, las duplicaciones enriquecidas en los controles en verde claro y los diploides y desprovistos de CNV en gris. Las líneas discontinuas gruesas azules destacan los bordes conectados a al menos un gen GRM mientras que las líneas estrechas de puntos representan todas las demás interacciones génicas. Los módulos enriquecidos respecto de anotaciones funcionales significativas se resaltan por elipses sombreadas en azul.

Figura 11. Una vista esquemática que muestra la interacción de los receptores GRM afectados en el TDAH con módulos de genes enriquecidos respecto de su significación funcional. Los genes de receptor GRM se muestran como rombos de color verde o rojo para representar duplicaciones y eliminaciones enriquecidas respectivamente en los casos. Los recuadros destacan módulos funcionales definidos por la red de genes que interactúan que están enriquecidos significativamente para las anotaciones GO. Los módulos funcionales describen anotaciones funcionales significativas y se marcan con el nombre agrupado y el número de genes componentes entre paréntesis. Se muestran en negrita las anotaciones funcionales que pueden ser particularmente pertinentes para la fisiopatología subyacente al TDAH. Los vértices de la red conectan genes de receptor GRM a los módulos funcionales: las líneas sólidas indican la pertenencia del gen que interactúa con GRM en el módulo funcional y las líneas de puntos indican una relación de primer grado entre los genes receptores GRM y al menos un gen componente de un módulo funcional.

Figura 12. Un ejemplo de asociación falsa positiva de península de CNV. Se muestra un ejemplo del crom. 2 (ubicación 51.777.616-51.784.033). Todos los CNVR significativos se revisan para las penínsulas de CNV que indican incertidumbre en el truncamiento de los límites.

Descripción detallada de la invención

El trastorno por déficit de atención e hiperactividad (TDAH) es un trastorno neuropsiquiátrico común hereditario de etiología desconocida. Recientemente, los presentes inventores han comunicado un enriquecimiento de variantes raras en genes implicados en el aprendizaje, comportamiento, la transmisión sináptica y el desarrollo del sistema nervioso central en el autismo[1], lo que sugiere que en la etiología del TDAH, un trastorno neuropsiquiátrico relacionado, también podrían desempeñar un papel raras variantes estructurales heredadas.

Para continuar estos estudios, se llevó a cabo un estudio de CNV en todo el genoma en una cohorte de 1.013 casos de TDAH y 4.105 niños sanos con antepasados europeos en los que se genotiparon 550.000 marcadores de SNP. Los hallazgos positivos se evaluaron en múltiples cohortes independientes, totalizando 2.493 casos de TDAH y 9.222 controles con antepasados europeos, con las respectivas cohortes de control de caso genotipadas en plataformas emparejadas.

Los presentes resultados identificaron varios CNV que tienen impacto en los genes de receptor de glutamato metabotrópico que se encontraban significativamente enriquecidos en todas las cohortes independientes ($P= 2{,}1x10^{-9}$). Entre ellos, se produjeron eliminaciones en *GRM5* (receptor de glutamato metabotrópico 5) en diez casos en tres cohortes independientes y en solo un sujeto de control ($P=1{,}36x10^{-6}$). Además, se produjeron eliminaciones en *GRM7* en seis casos y de *GRM8* en ocho casos, ambos con una frecuencia de control de cero. *GRM1* se duplicó en ocho casos, una frecuencia notablemente enriquecida por encima de los controles. Las variantes observadas se validaron experimentalmente usando PCR cuantitativa. El posterior análisis de red génica demostró que los genes que interactúan con receptores GRM están significativamente enriquecidos para CNV en los casos en comparación con los controles ($P=4{,}38x10^{-10}$), con un impacto colectivo del -10% de los casos de TDAH en este estudio. Además, se observó que los GRM sirven como nódulos críticos que coordinan módulos altamente conectados de genes que interactúan, muchos de los cuales portan CNV que están enriquecidos para funciones biológicas sinápticas y neuronales.

Las siguientes definiciones se proporcionan para facilitar la comprensión de la presente invención.

I. Definiciones:

A efectos de la presente invención, "un" o "una" entidad se refiere a una o más de esa entidad; por ejemplo, "un ADNc se refiere a uno o más ADNc o a al menos un ADNc. Como tales, los términos "un" o "una", "uno o más" y "al menos uno" pueden usarse de manera indistinta en el presente documento. También cabe destacar que los términos "comprende", "incluye" y "tiene" pueden usarse de manera intercambiable. Además, un compuesto "seleccionado entre el grupo que consiste en" se refiere a uno o más de los compuestos en la lista a continuación, incluyendo mezclas (es decir, combinaciones) de dos o más de los compuestos. De acuerdo con la presente invención, una molécula aislada o biológicamente pura es un compuesto que se ha extraído de su medio natural. Como tal, "aislado" y "biológicamente puro" no reflejan necesariamente el alcance hasta el que se ha purificado el compuesto. Un compuesto aislado de la presente invención puede obtenerse de su fuente natural, puede producirse usando técnicas sintéticas de laboratorio o puede producirse por cualquier ruta sintética química similar.

La expresión "alteración genética", tal como se usa en el presente documento, se refiere a un cambio respecto de la secuencia de tipo silvestre o de referencia de una o más moléculas de ácido nucleico. Las alteraciones genéticas incluyen, sin limitación, sustituciones, adiciones y eliminaciones de pares de bases de al menos un nucleótido de una molécula de ácido nucleico de secuencia conocida.

Un "polimorfismo de un solo nucleótido (SNP)" se refiere a un cambio en el que una sola base en el ADN difiere de la base habitual en esa posición. Estos cambios de una sola base se denominan SNP o "snips". Se han catalogado millones de SNP en el genoma humano. Algunos SNP, tales como los que causan las células falciformes, son responsables de enfermedades. Otros SNP son variaciones normales en el genoma.

Una "variación en el número de copias (CNV)" se refiere al número de copias de un gen particular o un segmento del mismo en el genoma de un individuo. Las CNV representan un componente genético importante de la diversidad fenotípica humana. Se sabe que la susceptibilidad a trastornos genéticos está asociada no solo con polimorfismos de un solo nucleótido (SNP), sino también con variaciones estructurales y otras variaciones genéticas, incluyendo las CNV. Una CNV representa un cambio en el número de copias que implica un fragmento de ADN que es de ~1 kilobases (kb) o mayor (Freuk y col. 2006a). Las CNV descritas en el presente documento no incluyen aquellas variantes que surgen de la inserción/eliminación de elementos transpondibles (por ejemplo, repeticiones de KpnI de ~6 kb) para minimizar la complejidad de futuros análisis de CNV. Por lo tanto, el término CNV abarca términos anteriormente presentados, tales como variantes de número de copias a gran escala (LCV; Iafrate y col. 2004), polimorfismos de número de copias (CNP; Sebat y col. 2004) y variantes de tamaño intermedio (ISV; Tuzun y col. 2005), pero no inserciones de retroposón. La expresión "CNV que contiene duplicación" también se usa más adelante en el presente documento de manera coherente con la definición de CNV proporcionada.

Un "SNP asociado con el TDAH" o un "marcador específico asociado con el TDAH" o una "molécula de secuencia informativa asociada con el TDAH" es un SNP o una secuencia marcadora que está asociada con un riesgo aumentado o reducido de desarrollar TDAH no encontrada en pacientes normales que no tienen esta enfermedad. Dichos marcadores pueden incluir, pero sin limitación, ácidos nucleicos, proteínas codificadas por los mismos u otras moléculas pequeñas. Por lo tanto, la expresión "ácido nucleico que contiene un SNP asociado con el TDAH" queda abarcada por la descripción anterior.

La expresión "matriz sólida", tal como se usa en el presente documento, se refiere a cualquier formato, tal como perlas, micropartículas, una micromatriz, la superficie de un pocillo de microtitulación o un tubo de ensayo, una varilla graduada o un filtro. El material de la matriz puede ser poliestireno, celulosa, látex, nitrocelulosa, nilón, poliacrilamida, dextrano o agarosa.

La expresión "que consiste esencialmente en", cuando se refiere a un nucleótido o aminoácido particular, se refiere a una secuencia que tiene las propiedades de una SEQ ID NO dada. Por ejemplo, cuando se usa en referencia a una secuencia de aminoácidos, la frase incluye la secuencia en sí y modificaciones moleculares que podrían no afectar a las características funcionales y novedosas de la secuencia.

La expresión "CNV informativa parcial" se usa en el presente documento para referirse a un ácido nucleico que hibrida con secuencias que comprenden una duplicación en un cromosoma; sin embargo, la CNV informativa parcial puede no ser idéntica a la duplicación, en su lugar, la CNV puede corresponder a solo una parte de la duplicación, pero seguir siendo informativa de la misma.

"Ácido nucleico diana", tal como se usa en el presente documento, se refiere a una región previamente definida de un ácido nucleico presente en una mezcla compleja de ácido nucleico, en la que la región de tipo silvestre definida contiene al menos una variación de nucleótidos conocida que puede asociarse o no con el TDAH pero es informativa del riesgo de TDAH. La molécula de ácido nucleico puede aislarse de una fuente natural mediante clonación de ADNc o hibridación sustractiva o sintetizarse manualmente. La molécula de ácido nucleico puede sintetizarse manualmente mediante el procedimiento sintético de triéster o usando un sintetizador de ADN automatizado.

En referencia a los ácidos nucleicos usados en la invención, en ocasiones se emplea la expresión "ácido nucleico aislado". Esta expresión, cuando se aplica a ADN, se refiere a una molécula de ADN que se separa de secuencias con las que se encuentra inmediatamente contigua (en las direcciones 5' y 3') en el genoma de origen natural del

organismo del que procede. Por ejemplo, el "ácido nucleico aislado" puede comprender una molécula de ADN insertada en un vector, tal como un plásmido o un vector vírico o estar integrado en el ADN genómico de un procariota o eucariota. Una "molécula de ácido nucleico aislada" también puede comprender una molécula de ADNc. Una molécula de ácido nucleico aislada insertada en un vector también se cita en ocasiones en el presente documento como una molécula de ácido nucleico recombinante.

Con respecto a moléculas de ARN, la expresión "ácido nucleico aislado" se refiere principalmente a una molécula de ARN codificada por una molécula de ADN aislada como se ha definido anteriormente. Como alternativa, la expresión puede referirse a una molécula de ARN que se ha separado lo suficientemente de moléculas de ARN con las que se asociaría en su estado natural (es decir, en células o tejidos), de tal forma que existe en una forma "sustancialmente pura".

Mediante el uso del término "enriquecido", en referencia a un ácido nucleico, se entiende que la secuencia de ADN o ARN constituye una fracción significativamente mayor (2-5 veces) del ADN o ARN total presente en las células o la solución de interés que en células normales o en las células de las que se extrajo la secuencia. Esto podría estar causado por una persona mediante una reducción preferencial en la cantidad de otro ADN o ARN presente o mediante un aumento preferencial en la cantidad de la secuencia de ADN o ARN específica o mediante una combinación ambos. Sin embargo, cabe destacar que "enriquecido" no implica que no haya presencia de otras secuencias de ADN o ARN, solo que la cantidad relativa de la secuencia de interés ha aumentado significativamente.

También es ventajoso para ciertos fines que una secuencia de nucleótidos se encuentre en forma purificada. El término "purificado", en referencia a un ácido nucleico, no requiere una pureza absoluta (tal como una preparación homogénea); en cambio, representa una indicación de que la secuencia es relativamente más pura que en el ambiente natural (en comparación con el nivel natural, este nivel debe ser al menos 2-5 veces mayor, por ejemplo, en términos de mg/ml). Los clones individuales aislados de una biblioteca de ADNc pueden purificarse hasta su homogeneidad por electroforesis. Las moléculas de ADN reivindicadas obtenidas de estos clones pueden obtenerse directamente de ADN total o de ARN total. Los clones de ADNc no son de origen natural, sino que se obtienen preferentemente mediante manipulación de una sustancia de origen natural parcialmente purificada (ARN mensajero). La construcción de una biblioteca de ADNc a partir de ARNm implica la creación de una sustancia sintética (ADNc) y pueden aislarse clones de ADNc individuales puros a partir de la biblioteca sintética mediante selección clonal de las células que portan la biblioteca de ADNc. Por lo tanto, el procedimiento que incluye la construcción de una biblioteca de ADNc a partir de ARNm y el aislamiento de distintos clones de ADNc proporciona una purificación de aproximadamente 10^{-6} veces del mensaje natural. Por lo tanto, se contempla expresamente la purificación de al menos un orden de magnitud, preferentemente, de dos o tres órdenes de magnitud y más preferentemente, de cuatro o cinco órdenes de magnitud. Por lo tanto, la expresión "sustancialmente pura" se refiere a una composición que comprende al menos un 50-60% en peso del compuesto de interés (por ejemplo, ácido nucleico, oligonucleótido, etc.). Más preferentemente, la preparación comprende al menos un 75% en peso y más preferentemente, un 90-99% en peso, del compuesto de interés. La pureza se mide por procedimientos adecuados para el compuesto de interés.

El término "complementario" describe dos nucleótidos que pueden formar múltiples interacciones favorables entre sí. Por ejemplo, la adenina es complementaria a la timina, ya que pueden formar dos enlaces de hidrógeno. De forma similar, la guanina y la citosina son complementarias, ya que pueden formar tres enlaces de hidrógeno. Por lo tanto, si una secuencia de ácido nucleico contiene la siguiente secuencia de bases, timina, adenina, guanina y citosina, un "complemento" de esta molécula de ácido nucleico sería una molécula que contiene adenina en el lugar de la timina, timina en el lugar de la adenina, citosina en el lugar de la guanina y guanina en el lugar de la citosina. Ya que el complemento puede contener una secuencia de ácido nucleico que forma interacciones óptimas con la molécula de ácido nucleico parental, dicho complemento puede unirse con alta afinidad a su molécula parental.

Con respecto a los ácidos nucleicos monocatenarios, en particular, oligonucleótidos, la expresión "que hibrida específicamente" se refiere a la asociación entre dos moléculas de nucleótidos monocatenarias con una secuencia lo suficientemente complementaria como para permitir dicha hibridación en condiciones predeterminadas usadas generalmente en la técnica (en ocasiones denominadas "sustancialmente complementarias"). En particular, el término se refiere a la hibridación de un oligonucleótido con una secuencia sustancialmente complementaria contenida en una molécula de ADN o ARN monocatenaria de la invención, hasta la exclusión sustancial de la hibridación del oligonucleótido con ácidos nucleicos monocatenarios de secuencia no complementaria. Por ejemplo, la hibridación específica puede referirse a una secuencia que hibrida con un gen marcador o ácido nucleico específico de TDAH, pero no hibrida con otros nucleótidos. Asimismo, un polinucleótido que "hibrida específicamente" puede hibridar únicamente a un marcador neuroespecífico concreto, tal como un marcador específico de TDAH mostrado en las tablas contenidas en el presente documento. Se conocen bien en la técnica condiciones adecuadas que permiten la hibridación específica de moléculas de ácido nucleico monocatenarias de diverso grado de complementariedad.

Por ejemplo, a continuación se expone una fórmula común para calcular las condiciones de rigurosidad necesarias para lograr la hibridación entre moléculas de ácido nucleico con una homología de secuencia específica (Sambrook y col., Molecular Cloning, Cold Spring Harbor Laboratory (1989):

$$T_m = 81,5°C + 16,6Log\ [Na+] + 0,41(\%\ G+C) - 0,63\ (\%\ formamida) - 600/\#pb\ en\ dúplex$$

Como ilustración de la fórmula anterior, usando [Na+] = [0,368] y formamida al 50%, con un contenido de GC del 42% y un tamaño medio de la sonda de 200 bases, la T_m es de 57°C. La T_m de un dúplex de ADN se reduce en 1 - 1,5°C con cada reducción de un 1% de homología. Por lo tanto, podrían observarse dianas con más de aproximadamente un 75% de identidad de secuencia usando una temperatura de hibridación de 42°C.

La rigurosidad de la hibridación y el lavado dependen principalmente de la concentración de sal y la temperatura de las soluciones. En general, para maximizar la tasa de hibridación de la sonda con su diana, la hibridación se lleva normalmente a cabo en condiciones de salinidad y temperatura que son 20-25°C inferiores a la T_m del híbrido. Las condiciones de lavado han de ser tan rigurosas como sea posible para el grado de identidad de la sonda por la diana. En general, las condiciones de lavado se seleccionan para que sean aproximadamente 12-20°C menores a la T_m del híbrido. Con respecto a los ácidos nucleicos de la presente invención, la hibridación de rigurosidad moderada se define como una hibridación en SSC 6X, solución de Denhardt 5X, SDS al 0,5% y 100 µg/ml de ADN de esperma de salmón desnaturalizado a 42°C y lavado en SSC 2X y SDS al 0,5% a 55°C durante 15 minutos. Una hibridación de rigurosidad elevada se define como hibridación en SSC 6X, solución de Denhardt 5X, SDS al 0,5% y 100 µg/ml de ADN de esperma de salmón desnaturalizado a 42°C y lavado en SSC 1X y SDS al 0,5% a 65°C durante 15 minutos. Una hibridación de rigurosidad muy elevada se define como hibridación en SSC 6X, solución de Denhardt 5X, SDS al 0,5% y 100 µg/ml de ADN de esperma de salmón desnaturalizado a 42°C y lavado en SSC 0.1X y SDS al 0,5% a 65°C durante 15 minutos.

El término "oligonucleótido", tal como se usa en el presente documento, se define como una molécula de ácido nucleico formada por dos o más ribo o desoxirribonucleótidos, preferentemente más de tres. El tamaño exacto del oligonucleótido dependerá de diversos factores y de la aplicación y el uso concreto del oligonucleótido. Los oligonucleótidos, que incluyen sondas y cebadores, pueden ser de cualquier longitud, desde 3 nucleótidos hasta la longitud completa de la molécula de ácido nucleico e incluyen explícitamente cualquier número posible de ácidos nucleicos desde 3 hasta la longitud completa del polinucleótido. Preferentemente, los oligonucleótidos son de al menos aproximadamente 10 nucleótidos de longitud, más preferentemente, de al menos 15 nucleótidos de longitud, más preferentemente, de al menos aproximadamente 20 nucleótidos de longitud.

El término "sonda", tal como se usa en el presente documento, se refiere a un oligonucleótido, polinucleótido o ácido nucleico, ya sea de ARN o ADN, tanto de origen natural como en una digestión con enzimas de restricción como producido de manera sintética, que es capaz de hibridar o de hibridar específicamente con un ácido nucleico con secuencias complementarias a la sonda. Una sonda puede ser monocatenaria o bicatenaria. La longitud exacta de la sonda dependerá de muchos factores, incluyendo la temperatura, la fuente de la sonda y el uso del procedimiento. Por ejemplo, para aplicaciones diagnósticas, dependiendo de la complejidad de la secuencia diana, la sonda de oligonucleótidos contiene normalmente 15-25 o más nucleótidos, aunque puede contener menos nucleótidos. Las sondas en el presente documento se seleccionan para que sean complementarias a diferentes hebras de una secuencia de ácido nucleico diana particular. Esto significa que las sondas han de ser lo suficientemente complementarias para que sean capaces de "hibridar específicamente" o ligarse con sus hebras respectivas en una serie de condiciones predeterminadas. Por lo tanto, la secuencia de la sonda no refleja necesariamente la secuencia complementaria exacta de la diana. Por ejemplo, puede unirse un fragmento de nucleótido no complementario al extremo 5' o 3' de la sonda, siendo el resto de la secuencia de la sonda complementaria a la hebra diana. Como alternativa, pueden intercalarse en la sonda bases no complementarias o secuencias más largas, a condición de que la secuencia de la sonda tenga una complementariedad suficiente con la secuencia del ácido nucleico para hibridarse de manera específica con el mismo.

El término "cebador", tal como se usa en el presente documento, se refiere a un oligonucleótido, ya sea de ARN o ADN, ya sea monocatenario o bicatenario, ya proceda de un sistema biológico, generado por digestión con enzimas de restricción o producido sintéticamente, que, cuando se coloca en el ambiente adecuado, es capaz de actuar funcionalmente como iniciador de la síntesis de ácido nucleico dependiente de molde. Cuando se presenta con un molde de ácido nucleico adecuado, precursores de nucleótido trifosfato adecuados de los ácidos nucleicos, una enzima polimerasa, cofactores adecuados y condiciones, tales como una temperatura y pH adecuados, el cebador puede extenderse en su extremo 3' mediante la adición de nucleótidos gracias a la acción de una polimerasa o una actividad similar para proporcionar un producto de extensión del cebador. La longitud del cebador puede variar dependiendo de las condiciones particulares y los requisitos de la aplicación. Por ejemplo, en aplicaciones diagnósticas, el cebador de oligonucleótidos tiene normalmente una longitud de 15-25 o más nucleótidos. El cebador ha de ser lo suficientemente complementario al molde deseado para cebar la síntesis del producto de extensión deseado, es decir, ser capaz de hibridar con la hebra molde deseada de una manera suficiente para proporcionar el resto hidroxilo 3' del cebador en yuxtaposición adecuada para su uso en el inicio de la síntesis mediante una polimerasa o una enzima similar. No es necesario que la secuencia de cebador represente un complemento exacto del molde deseado. Por ejemplo, puede unirse una secuencia de nucleótidos no complementaria al extremo 5' de un cebador por lo demás complementario. Como alternativa, pueden intercalarse bases no complementarias dentro de la secuencia de cebador de oligonucleótido, a condición de que la secuencia de cebador tenga una complementariedad suficiente con la secuencia de la hebra molde deseada para proporcionar funcionalmente un complejo molde-cebador para la síntesis del producto de extensión.

La reacción en cadena de la polimerasa (PCR) se ha descrito en las Patentes de los Estados Unidos 4.683.195, 4.800.195 y 4.965.188, cuyas divulgaciones se incorporan al presente documento por referencia.

El término "vector" se refiere a una molécula de ácido nucleico circular mono o bicatenaria que puede usarse para infectar, transfectar o transformar células y replicarse de manera independiente o dentro del genoma de la célula hospedadora. Una molécula de ácido nucleico bicatenaria circular puede cortarse y de este modo linealizarse tras su tratamiento con enzimas de restricción. Se encuentra fácilmente disponible para los expertos en la materia una serie de vectores, enzimas de restricción y el conocimiento de secuencias de nucleótidos sobre las que actúan de manera específica enzimas de restricción e incluyen cualquier replicón, tal como un plásmido, cósmido, bácmido, fago o virus, al que puede unirse otra secuencia o elemento genético (ya sea de ADN o ARN), de tal forma que se produce la replicación de la secuencia o elemento unido. Puede insertarse una molécula de ácido nucleico de la invención en un vector cortando el vector con enzimas de restricción y ligando las dos piezas entre sí.

Los expertos en la materia tienen a su disposición diversas técnicas para facilitar la transformación, la transfección o la transducción de la construcción de expresión en un organismo procariota o eucariota. Los términos "transformación", "transfección" y "transducción" se refieren a procedimientos para insertar un ácido nucleico y/o una construcción de expresión en una célula u organismo hospedador. Estos procedimientos implican una serie de técnicas, tales como el tratamiento de las células con altas concentraciones de sal, un campo eléctrico o detergente, para hacer que la membrana externa o la pared celular de la célula hospedadora sea permeable a las moléculas de ácido nucleico de interés, microinyección, fusión con PEG y similares.

La expresión "elemento promotor" describe una secuencia de nucleótidos que se incorpora en un vector que, una vez se encuentra dentro de una célula adecuada, puede facilitar la unión a factores de transcripción y/o a polimerasas y la posterior transcripción de porciones del ADN del vector en ARNm. En una realización, el elemento promotor de la presente invención precede al extremo 5' de la molécula de ácido nucleico marcador específica de TDAH, de tal forma que esta última se transcribe en ARNm. Después, la maquinaria de la célula hospedadora traduce el ARNm en un polipéptido.

Los expertos en la materia reconocerán que un vector de ácido nucleico puede contener elementos de ácido nucleico distintos al elemento promotor y la molécula de ácido nucleico marcador específica de TDAH. Estos otros elementos de ácido nucleico incluyen, pero sin limitación, orígenes de replicación, sitios de unión a ribosomas, secuencias de ácido nucleico que codifican enzimas de resistencia a fármacos o enzimas metabólicas de aminoácidos y secuencias de ácido nucleico que codifican señales de secreción, señales de localización o señales útiles para la purificación de polipéptidos.

Un "replicón" es cualquier elemento genético, por ejemplo, un plásmido, cósmido, bácmido, plástido, fago o virus, que tiene capacidad de replicación en gran medida bajo su propio control. Un replicón puede ser de ARN o ADN y puede ser mono o bicatenario.

Un "operón de expresión" se refiere a un segmento de ácido nucleico que puede poseer secuencias de control de la transcripción y la traducción, tales como promotores, potenciadores, señales de inicio de la traducción (por ejemplo, codones ATG o AUG), señales de poliadenilación, terminadores y similares y que facilitan la expresión de una secuencia codificante de polipéptidos en una célula u organismo hospedador.

Tal como se usa en el presente documento, los términos y expresiones "indicador", "sistema indicador", "gen indicador" o "producto génico indicador" han de entenderse como un sistema genético operativo en el que un ácido nucleico comprende un gen que codifica un producto que, cuando se expresa, produce una señal indicadora que es fácilmente medible, por ejemplo, mediante ensayo biológico, inmunoensayo, radioinmunoensayo o por procedimientos colorimétricos, fluorogénicos, quimioluminiscentes u otros. El ácido nucleico puede ser ARN o ADN, lineal o circular, mono o bicatenario, con polaridad antisentido o con sentido y está unido operativamente a los elementos de control necesarios para la expresión del producto génico indicador. Los elementos de control necesarios variarán dependiendo de la naturaleza del sistema indicador y de si el gen indicador se encuentra en forma de ADN o ARN, aunque puede incluir, pero sin limitación, elementos tales como promotores, potenciadores, secuencias de control de la traducción, señales de adición de poli A, señales de terminación de la transcripción y similares.

El ácido nucleico introducido puede integrarse o no (unirse covalentemente) en el ácido nucleico de la célula u organismo receptor. En células bacterianas, de levadura, vegetales y de mamífero, por ejemplo, puede mantenerse el ácido nucleico introducido como un elemento episómico o como un replicón independiente, tal como un plásmido. Como alternativa, el ácido nucleico introducido puede quedar integrado en el ácido nucleico de la célula u organismo receptor y mantenerse de manera estable en esa célula u organismo y posteriormente pasarse o heredarse por las células u organismos descendientes de la célula u organismo receptor. Finalmente, el ácido nucleico introducido puede existir en la célula u organismo receptor solo de manera transitoria.

La expresión "gen marcador de selección" se refiere a un gen que cuando se expresa, confiere un fenotipo seleccionable, tal como resistencia a los antibióticos, en una célula transformada.

La expresión "unido operativamente" significa que las secuencias reguladoras necesarias para la expresión de la

secuencia codificante se colocan en la molécula de ADN en las posiciones adecuadas en relación con la secuencia codificante para efectuar la expresión de la secuencia codificante. En ocasiones esta misma definición se aplica a la disposición de unidades de transcripción y otros elementos de control de la transcripción (por ejemplo, potenciadores) en un vector de expresión.

Las expresiones "organismo recombinante" u "organismo transgénico" se refieren a organismos que tienen una nueva combinación de genes o moléculas de ácido nucleico. Puede introducirse una nueva combinación de genes o moléculas de ácido nucleico en un organismo usando una gran variedad de técnicas de manipulación de ácido nucleico disponibles para los expertos en la materia. El término "organismo" se refiere a cualquier ser vivo formado por al menos una célula. Un organismo puede ser tan simple como una célula eucariota o tan complejo como un mamífero. Por lo tanto, la expresión "un organismo recombinante" abarca una célula recombinante, así como organismos eucariotas y procariotas.

En ocasiones se usa en el presente documento la expresión "proteína aislada" o "proteína aislada y purificada". Esta expresión se refiere principalmente a una proteína producida mediante la expresión de una molécula de ácido nucleico aislada de la invención. Como alternativa, esta expresión puede referirse a una proteína que se ha separado suficientemente de otras proteínas con las que se asociaría de manera natural, de tal forma que existe en una forma "sustancialmente pura". Por "aislado" no se pretende excluir mezclas artificiales o sintéticas con otros compuestos o materiales o la presencia de impurezas que no interfieren con la actividad fundamental y que pueden estar presentes, por ejemplo, debido a una purificación incompleta, a la adición de estabilizantes o a su incorporación en, por ejemplo, preparaciones inmunogénicas o preparaciones farmacéuticamente aceptables.

Un "par de unión específico" comprende un miembro de unión específico (mue) y un compañero de unión (cu) que tienen una especificidad particular entre sí y que en condiciones normales, se unen entre sí de manera preferente frente a otras moléculas. Los ejemplos de pares de unión específicos son antígenos y anticuerpos, ligandos y receptores y secuencias de nucleótidos complementarias. El experto en la materia conocerá muchos ejemplos diferentes. Además, la expresión "par de unión específico" también puede aplicarse a casos en los que uno de los dos o ambos del miembro de unión específico y el compañero de unión comprenden una parte de una molécula grande. En las realizaciones en las que el par de unión específico comprende secuencias de ácido nucleico, pueden ser de cualquier longitud para hibridar entre sí en las condiciones de ensayo, preferentemente de más de 10 nucleótidos de longitud, más preferentemente de más de 15 o 20 nucleótidos de longitud.

"Muestra" o "muestra de paciente" o "muestra biológica" se refiere en general a una muestra que puede evaluarse respecto de una molécula particular, preferentemente, una molécula marcadora específica de TDAH, tal como un marcador descrito más adelante en el presente documento. Las muestras pueden incluir, pero sin limitación, células, fluidos corporales, incluyendo sangre, suero, plasma, fluido cefalorraquídeo, orina, saliva, lágrimas, fluido pleural y similares.

Los términos "agente" y "compuesto" se usan de manera indistinta en el presente documento e indican un compuesto químico, una mezcla de compuestos químicos, una molécula biológica o un extracto obtenido de materiales biológicos, tales como células o tejidos bacterianos, vegetales, fúngicos o animales (en particular de mamífero). Las macromoléculas biológicas incluyen ARNpi, ARNhc, oligonucleótidos antisentido, péptidos, complejos de péptido/ADN y cualquier molécula a base de ácidos nucleicos que muestre la capacidad de modular la actividad de los ácidos nucleicos que contienen CNV o SNP descritos en el presente documento o sus proteínas codificadas. Los agentes y compuestos también pueden citarse como "agentes de ensayo" o "compuestos de ensayo", que se evalúan respecto de su potencial actividad biológica mediante su inclusión en los ensayos de exploración descritos más adelante en el presente documento.

El término "modular", tal como se usa en el presente documento, se refiere a aumentar/promover o reducir/inhibir una función biológica o de señalización celular particular asociada con las actividades normales de las moléculas que contienen CNV descritas en el presente documento o las proteínas codificadas por las mismas. Por ejemplo, el término modular se refiere a la capacidad de un compuesto de ensayo o de un agente de ensayo para interferir con la señalización o actividad de un gen o proteína de la presente invención.

II. Procedimientos para usar CNV y/o SNP asociados con el TDAH para diagnosticar la propensión a desarrollar TDAH

La presente invención proporciona procedimientos para diagnosticar el TDAH en un paciente o procedimientos para identificar a un paciente que tiene un riesgo aumentado de desarrollar TDAH. Diagnóstico, tal como se usa en el presente documento, incluye no solo la identificación inicial de TDAH asociado con las alteraciones genéticas descritas en el presente documento en un paciente, sino pruebas confirmatorias o exploración en pacientes que hayan sido identificados previamente como afectados o probablemente afectados de TDAH. Los procedimientos incluyen las etapas de proporcionar una muestra biológica del paciente, medir la cantidad de conjuntos particulares o cualquiera de todos los marcadores asociados con el TDAH (tabla 13) presentes en la muestra biológica, preferentemente, una muestra de tejido y/o plasma sanguíneo y determinar si el paciente tiene una probabilidad mayor de TDAH basándose en la cantidad y/o el tipo de nivel de expresión del marcador de TDAH en relación con dichos niveles de expresión identificados en cohortes de pacientes de resultado conocido. Un paciente tiene

mayores probabilidades de tener TDAH cuando la muestra tiene un perfil de expresión de marcador de CNV asociado con pacientes diagnosticados previamente de TDAH. Las composiciones y procedimientos de la invención son útiles para el pronóstico, el diagnóstico y el tratamiento del TDAH.

En otro aspecto, la muestra del paciente puede haberse genotipado previamente y por lo tanto, puede estar disponible para el médico el perfil de expresión genética en la muestra. Por consiguiente, el procedimiento puede implicar almacenar información de secuencias de marcadores asociados con el TDAH de referencia en una base de datos, es decir, aquellas CNV estadísticamente asociadas con un pronóstico más favorable o menos favorable, tal como se describe en las tablas en el presente documento y llevar a cabo un análisis genético comparativo en el ordenador, identificando de este modo a aquellos pacientes que tengan un riesgo aumentado de TDAH.

Los ácidos nucleicos que contienen CNV o SNP relacionados con el TDAH, incluyendo, pero sin limitación, los listados más adelante pueden usarse para diversos fines de acuerdo con la presente invención. Puede usarse ADN, ARN o fragmentos de los mismos que contienen CNV o SNP asociados con el TDAH como sondas para detectar la presencia y/o la expresión de marcadores específicos de TDAH. Los procedimientos en los que pueden utilizarse ácidos nucleicos marcadores específicos de TDAH como sondas para dichos ensayos incluyen, pero sin limitación: (1) hibridación *in situ*; (2) hibridación de Southern (3) hibridación de northern; y (4) reacciones de amplificación variadas, tales como reacciones en cadena de la polimerasa (PCR).

Además, pueden llevarse a cabo ensayos para detectar CNV o SNP asociados con el TDAH en cualquier tipo de muestra biológica, incluyendo, pero sin limitación, fluidos corporales (incluyendo sangre, orina, suero, lavado gástrico, fluido cefalorraquídeo), cualquier tipo de célula (tal como células cerebrales, glóbulos blancos sanguíneos, células mononucleadas, células fetales en la circulación materna) o tejidos corporales.

Claramente, Los ácidos nucleicos que contienen CNV o SNP asociados con el TDAH, los vectores que expresan los mismos, las proteínas marcadoras que contienen CNV o SNP de TDAH y los anticuerpos marcadores específicos anti-TDAH de la invención pueden usarse para detectar CNV o SNP asociados con el TDAH en tejidos corporales, células o fluidos y alterar la expresión de la proteína marcadora que contiene CNV o SNP de TDAH a fin de evaluar las interacciones genéticas y de proteína implicadas en el desarrollo del TDAH.

En la mayoría de realizaciones para la exploración de CNV o SNP asociados con el TDAH, el ácido nucleico que contiene CNV o SNP asociados con el TDAH en la muestra se amplificará inicialmente, por ejemplo, usando PCR, para aumentar la cantidad de los moldes en comparación con otras secuencias presentes en la muestra. Esto permite la detección de las secuencias diana con un alto grado de sensibilidad si están presentes en la muestra. Puede evitarse esta etapa inicial usando técnicas de matriz altamente sensibles que son importantes en la materia.

Como alternativa, las nuevas tecnologías de detección pueden superar esta limitación y permitir el análisis de pequeñas muestras que contienen nada más que 1 µg de ARN total. Usando la tecnología de dispersión de luz de resonancia (RLS), a diferencia de las técnicas de fluorescencia tradicionales, múltiples lecturas pueden detectar bajas cantidades de ARNm usando dianas hibridadas marcadas con biotina y anticuerpos anti-biotina. Otra alternativa a la amplificación mediante la PCR implica la tecnología guía de onda planar (PWG, por sus siglas en inglés) para aumentar las relaciones de señal a ruido y reducir la interferencia de fondo. Ambas técnicas se encuentran disponibles comercialmente de Qiagen Inc. (EE. UU.).

Puede usarse cualquiera de las técnicas anteriormente mencionadas para detectar o cuantificar la expresión de marcadores de CNV o SNP asociados con el TDAH y por consiguiente, diagnosticar el TDAH.

III. Kits y artículos de fabricación

Puede incorporarse cualquiera de los productos anteriormente mencionados en un kit que puede contener un polinucleótido marcador específico de CNV o SNP asociado con el TDAH o uno o más de dichos marcadores inmovilizados sobre un chip génico, un oligonucleótido, un polipéptido, un péptido, un anticuerpo, una etiqueta, marcador, indicador, un vehículo farmacéuticamente aceptable, un vehículo fisiológicamente aceptable, instrucciones para su uso, un recipiente, un vial para administración, un sustrato de ensayo o cualquier combinación de los mismos.

IV. Procedimientos para usar CNV y/o SNP asociados con el TDAH para el desarrollo de agentes terapéuticos

Debido a que las CNV y los SNP identificados en el presente documento se han asociado con la etiología del TDAH, los procedimientos para identificar agentes que modulan la actividad de los genes y sus productos codificados que contienen dichas CNV y/o SNP deberían dar como resultado la generación de agentes terapéuticos eficaces para el tratamiento de este trastorno.

Varias regiones del genoma humano proporcionan dianas adecuadas para el diseño racional de agentes terapéuticos. Pueden aprovecharse moléculas de ácido nucleico pequeñas o moléculas de péptido correspondientes a estas regiones en el diseño de agentes terapéuticos que modulan de manera eficaz la actividad de las proteínas codificadas.

El modelado molecular debería facilitar la identificación de moléculas orgánicas específicas con capacidad para unirse al sitio activo de las proteínas codificadas por los ácidos nucleicos que contienen CNV o SNP basándose en la conformación o en restos de aminoácido clave necesarios para la función. Se usará un enfoque de química combinatoria para identificar moléculas con mayor actividad y después, se generarán iteraciones de estas moléculas para ciclos posteriores de selección. Varias de las moléculas disponibles en este ensayo de selección, aunque no limitan el procedimiento, incluyen moduladores alostéricos positivos (PAM) del receptor de glutamato metabotrópico (mGluR), moduladores alostéricos negativos (NAM) y antagonistas del receptor de taquicinina-3/neurocinina-3 (TACR-3/NK3R). Una lista específica incluye ADX63365, ADX50938, ADX71149, ADX48621, AMN082, derivados de 1-(hetero)aril-3-amino-pirrolidina (por ejemplo, los proporcionados en la Publicación de Solicitud de Patente de los Estados Unidos n.º 2008/0300266), LY341495, GSK1144814 y SB223412 (**tabla 1**).

Tabla 1. Moléculas y agentes terapéuticos disponibles para un enfoque de química combinatoria.

Nombre del producto	Nombre de la compañía	Indicación	Mecanismo de acción
ADX63365	Addex Pharmaceuticals	Esquizofrenia	Modulador alostérico positivo del receptor de glutamato metabotrópico 5 (GRM5)
ADX63365	Merck & Co Inc	Esquizofrenia	Modulador alostérico positivo del receptor de glutamato metabotrópico 5 (GRM5)
ADX50938	Addex Pharmaceuticals	Esquizofrenia	Modulador alostérico positivo del receptor de glutamato metabotrópico 5 (GRM5)
ADX50938	Addex Pharmaceuticals	Enfermedad de Alzheimer	Modulador alostérico positivo del receptor de glutamato metabotrópico 5 (GRM5)
ADX71149	Addex Pharmaceuticals	Esquizofrenia	Modulador alostérico positivo del receptor de glutamato metabotrópico 2 (GRM2)
AMN082	-	Esquizofrenia, Depresión, Enfermedad de Alzheimer	Modulador alostérico positivo del receptor de glutamato metabotrópico 7 (GRM7)
derivados de 1-(hetero)aril-3-amino-pirrolidina	Eli Lilly & Co	Migraña	Antagonista del receptor de glutamato metabotrópico 3 (GRM3)
LY341495	Eli Lilly & Co	Trastornos del sistema nervioso central	Antagonista del receptor de glutamato metabotrópico 2 (GRM2); Antagonista del receptor de glutamato metabotrópico 3 (GRM3)
Addex Pharmaceuticals	ADX48621	Enfermedad de Parkinson	Modulador alostérico negativo del receptor de glutamato metabotrópico 5 (GRM5)
GSK1144814	Glaxo Smith Kline	Esquizofrenia	Antagonista de receptores de neurocinina-3
SB223412 (Talnetant)	Glaxo Smith Kline	Esquizofrenia	Antagonista de receptores de neurocinina-3

Los polipéptidos o fragmentos empleados en los ensayos de exploración de fármacos pueden encontrar o bien libres en solución, fijados a un soporte sólido o dentro de una célula. Un procedimiento de selección de fármacos utiliza células eucariotas o procariotas que se transforman de manera estable con polinucleótidos recombinantes que expresan el polipéptido o el fragmento, preferentemente en ensayos de unión competitiva. Dichas células, ya se encuentren en forma viable o fija, pueden usarse para ensayos de unión convencionales. Se puede determinar, por ejemplo, la formación de complejos entre el polipéptido o el fragmento y el agente que se está ensayando o examinar el grado en que se ve afectada la formación de un complejo entre el polipéptido o el fragmento y un sustrato conocido por el agente que se esté ensayando.

Otra técnica para la exploración de fármacos proporciona exploración de alto rendimiento para compuestos que tienen una afinidad de unión adecuada para los polipéptidos codificados y se describe en detalle en Geysen, Solicitud Publicada PCT WO 84/03564, publicada el 13 de septiembre de 1984. Dicho brevemente, se sintetiza sobre un sustrato sólido, tal como puntas de plástico u otra superficie diferente, grandes números de compuestos de ensayo peptídicos pequeños, tales como los descritos anteriormente. Los compuestos de ensayo peptídicos se

hacen reaccionar con el polipéptido diana y se lavan. Después, se detecta el polipéptido unido mediante procedimientos bien conocidos en la técnica.

Una técnica adicional para la exploración de fármacos implica el uso de líneas celulares hospedadoras eucariotas o células (tales como las descritas anteriormente) que tienen un gen asociado con el TDAH no funcional o alterado. Estas líneas celulares o células hospedadoras son defectuosas a nivel de polipéptido. Las líneas celulares o células hospedadoras se cultivan en presencia de compuesto farmacológico. La función glutaminérgica alterada de las células hospedadoras se mide para determinar si el compuesto tiene la capacidad de regular esta función en las células defectuosas. Las células hospedadoras contempladas para su uso en la presente invención incluyen, pero sin limitación, células bacterianas, células fúngicas, células de insecto, células de mamífero y células vegetales. Sin embargo, se prefieren células de mamífero, en particular, células neuronales. Las moléculas de ADN que codifican CNV o SNP asociados con el TDAH pueden introducirse de manera individual en dichas células hospedadoras o en combinación para determinar el fenotipo de las células conferido por dicha expresión. Los procedimientos para introducir moléculas de ADN también son de sobra conocidos por los expertos habituales en la materia. Dichos procedimientos se exponen en Ausubel y col. eds., Current Protocols in Molecular Biology, John Wiley & Sons, NY, N.Y. 1995, cuya divulgación se incorpora al presente documento por referencia.

Hay disponible una gran variedad de vectores de expresión que pueden modificarse para expresar las nuevas secuencias de ADN de la presente invención. Los vectores específicos presentados en el presente documento son únicamente ilustrativos y no pretenden limitar el ámbito de la invención. Se describen procedimientos de expresión por Sambrook y col. Molecular Cloning: A Laboratory Manual o Current Protocols in Molecular Biology 16.3-17.44 (1989). También se describen procedimientos de expresión en *Saccharomyces* en Current Protocols in Molecular Biology (1989).

Los vectores adecuados para su uso en la práctica de la invención incluyen vectores procariotas, tales como los vectores pNH (Stratagene Inc., 11099 N. Torrey Pines Rd., La Jolla, Calif. 92037), vectores pET (Novogen Inc., 565 Science Dr., Madison, Wis. 53711) y los vectores pGEX (Pharmacia LKB Biotechnology Inc., Piscataway, N.J. 08854). Los ejemplos de vectores eucariotas útiles en la práctica de la presente invención incluyen los vectores pRc/CMV, pRc/RSV y pREP (Invitrogen, 11588 Sorrento Valley Rd., San Diego, Calif. 92121); pcDNA3.1/V5&His (Invitrogen); vectores baculovíricos, tales como pVL1392, pVL1393 o pAC360 (Invitrogen); y vectores de levadura, tales como YRP17, YIP5 y YEP24 (New England Biolabs, Beverly, Mass.), así como pRS403 y pRS413 (Stratagene Inc.); vectores para *Picchia* tales como pHIL-DI (Phillips Petroleum Co., Bartlesville, Okla. 74004); vectores retrovíricos, tales como PLNCX y pLPCX (Clontech); y vectores adenovíricos y víricos adenoasociados.

Los promotores para su uso en los vectores de expresión de la presente invención incluyen promotores que son operativos en células procariotas o eucariotas. Los promotores que son operativos en células procariotas incluyen elementos de control de lactosa (lac), elementos de control de bacteriófago lambda (pL), elementos de control de arabinosa, elementos de control de triptófano (trp), elementos de control de bacteriófago T7 e híbridos de los mismos. Los promotores que son operativos en células eucariotas incluyen promotores del virus de Epstein Barr, promotores de adenovirus, promotores de SV40, promotores del virus del sarcoma de Rous, promotores de citomegalovirus (CMV), promotores de baculovirus, tales como el promotor de polihedrina de AcMNPV, promotores de *Picchia*, tales como el promotor de alcohol oxidasa y promotores de *Saccharomyces*, tales como el promotor inducible por gal4 y el promotor constitutivo PGK, así como el promotor de factor de crecimiento derivado de plaquetas (PDGF) específico de neuronas, el promotor Thy-1, el promotor de prión de hámster y ratón (MoPrP) y la proteína ácida fibrilar glial (GFAP) para la expresión de transgenes en células gliales.

Además, un vector de la presente invención puede contener uno cualquiera de un número de marcadores diversos que facilitan la selección de una célula hospedadora transformada. Dichos marcadores incluyen genes asociados con la sensibilidad a la temperatura, la resistencia a los fármacos o enzimas asociadas con características fenotípicas de los organismos hospedadores.

Las células hospedadoras que expresan las CNV y/o los SNP asociados con el TDAH de la presente invención o fragmentos funcionales de los mismos proporcionan un sistema en el que explorar compuestos o agentes potenciales respecto de su capacidad para modular el desarrollo del TDAH. Por lo tanto, en una realización, las moléculas de ácido nucleico de la invención pueden usarse para crear líneas celulares recombinantes para su uso en ensayos para identificar agentes que modulan aspectos del metabolismo celular asociados con el TDAH y la función glutaminérgica aberrante. En el presente documento también se proporcionan procedimientos para explorar compuestos capaces de modular la función de proteínas codificadas por ácidos nucleicos que contienen CNV y SNP.

Otra estrategia implica el uso de bibliotecas de presentación en fagos diseñadas para expresar fragmentos de los polipéptidos codificados por los ácidos nucleicos que contienen CNV o SNP en la superficie del fago. Dichas bibliotecas se ponen después en contacto con una biblioteca química combinatoria en condiciones en las que puede detectarse la afinidad de unión entre el péptido expresado y los componentes de la biblioteca química. Las Patentes de los Estados Unidos 6.057.098 y 5.965.456 proporcionan procedimientos y aparatos para llevar a cabo dichos ensayos.

El objetivo del diseño racional de fármacos es producir análogos estructurales de polipéptidos biológicamente activo de interés o de moléculas pequeñas con las que interactúan (por ejemplo, agonistas, antagonistas, inhibidores) para diseñar fármacos que son, por ejemplo, formas más activas o estables del polipéptido o que, por ejemplo, mejoran o interfieren con la función de un polipéptido *in vivo*. Véase, por ejemplo, Hodgson, (1991) Bio/Technology 9:19-21. En un enfoque, analizado anteriormente, se resuelve la estructura tridimensional de una proteína de interés o, por ejemplo, del complejo proteína-sustrato, mediante cristalografía de rayos X, mediante resonancia magnética nuclear, mediante modelado computarizado o más típicamente, mediante una combinación de enfoques. Con menos frecuencia, ha de obtenerse información útil relativa a la estructura de un polipéptido mediante modelado basándose en la estructura de proteínas homólogas. Un ejemplo de diseño racional de fármacos es el desarrollo de inhibidores de proteasa del VIH (Erickson y col., (1990) Science 249:527-533). Además, los péptidos pueden analizarse mediante barrido de alanina (Wells, (1991) Meth. Enzym. 202:390-411). En esta técnica, se reemplaza un resto de aminoácido por Ala y se determina su efecto en la actividad del péptido. Se analiza de este modo cada uno de los restos de aminoácido del péptido para determinar las regiones importantes del péptido.

También es posible aislar un anticuerpo específico de diana, seleccionado mediante un ensayo funcional y después, resolver su estructura cristalina. En principio, esta estrategia proporciona un farmacóforo en el que puede basarse el posterior diseño del fármaco.

Se puede evitar completamente la cristalografía de las proteínas generando anticuerpos antidiotípicos (anti-id) para un anticuerpo funcional, farmacológicamente activo. Como imagen especular de una imagen especular, podría esperarse que el sitio de unión de los anti-id sea un análogo de la molécula original. Después, podría usarse el anti-id para identificar y aislar péptidos de bancos de péptidos producidos química o biológicamente. Después, los péptidos seleccionados podrían actuar como farmacóforo.

Por lo tanto, se pueden diseñar fármacos que tengan, por ejemplo, actividad o estabilidad de polipéptido mejorada o que actúen como inhibidores, agonistas, antagonistas, etc. de la actividad polipeptídica. Gracias a la disponibilidad de secuencias de ácido nucleico que contienen CNV o SNP descritas en el presente documento, pueden producirse cantidades suficientes del polipéptido codificado para llevar a cabo dichos estudios analíticos, tales como cristalografía de rayos X. Además, el conocimiento de la secuencia de proteína proporcionada en el presente documento guiará a los que empleen técnicas de modelado asistido por ordenador en lugar de o además de la cristalografía de rayos X.

En otra realización, la disponibilidad de ácidos nucleicos que contienen CNV o SNP asociados con el TDAH permite la producción de estirpes de ratones de laboratorio que portan los SNP o CNV asociados con el TDAH de la invención. Los ratones transgénicos que expresan las CNV o los SNP asociados con el TDAH de la invención proporcionan un sistema modelo en el que examinar el papel de la proteína codificada por el ácido nucleico que contiene CNV o SNP en el desarrollo y la progresión hacia el TDAH. Los expertos en la materia conocen procedimientos para introducir transgenes en ratones de laboratorio. Tres procedimientos comunes incluyen: 1. integración de vectores retrovíricos que codifican el gen exógeno de interés en un embrión temprano; 2. inyección de ADN en el pronúcleo de un óvulo recién fertilizado; y 3. la incorporación de células madre embrionarias genéticamente manipuladas en un embrión temprano. La producción de los ratones transgénicos descritos anteriormente facilitará la dilucidación molecular del papel que desempeña una proteína diana en diversos procesos metabólicos celulares, incluyendo: función glutaminérgica aberrante, señalización de receptor de ligando neuroactivo alterada y neurotransmisión aberrante o morfología neuronal alterada y recrecimiento de neuritas. Dichos ratones proporcionan una herramienta de exploración *in vivo* para estudiar posibles fármacos terapéuticos en un modelo de animal completo y están abarcados por la presente invención.

El término "animal" se usa en el presente documento para incluir todos los animales vertebrados, a excepción de seres humanos. También incluye un animal individual en todas las etapas de desarrollo, incluyendo las etapas embrionarias y fetales. Un "animal transgénico" es cualquier animal que contiene una o más células que portan información genética alterada o recibida, directa o indirectamente, por manipulación genética deliberada a nivel subcelular, tal como mediante recombinación dirigida o microinyección o infección con virus recombinantes. La expresión "animal transgénico" no pretende abarcar el cruzamiento clásico o la fertilización *in vitro*, sino que pretende abarcar animales en los que se alteran una o más células o reciben una molécula de ADN recombinante. Esta molécula puede dirigirse específicamente a un locus genético definido, integrarse aleatoriamente dentro de un cromosoma o puede ser ADN de replicación extracromosómica. La expresión "animal transgénico de línea germinal celular" se refiere a un animal transgénico en el que la alteración genética o la información genética se introdujo en una célula de línea germinal, confiriendo de este modo la capacidad de transferir la información genética a la descendencia. En caso de que dicha descendencia, de hecho, posea parte o la totalidad de esa alteración o información genética, también se les considera animales transgénicos.

La alteración de información genética puede ser exógena para la especie animal a la que pertenece el receptor o exógeno solo para el animal receptor o puede ser información genética que el receptor ya posee. En el último caso, el gen alterado o introducido puede expresarse de una manera diferente al gen nativo. Dicha información genética alterada o exógena podría abarcar la introducción de secuencias de nucleótidos que contienen CNV o SNP asociados con el TDAH.

El ADN usado para alterar un gen diana puede obtenerse mediante una gran variedad de técnicas que incluyen, pero sin limitación, aislamiento de fuentes genéticas, preparación de ADNc a partir de moldes de ARNm aislados, síntesis directa o una combinación de las mismas.

Un tipo preferido de célula diana para la introducción de transgenes es las células madre embrionarias (ES). Las células ES pueden obtenerse de embriones antes de la implantación cultivados *in vitro* (Evans y col., (1981) Nature 292:154-156; Bradley y col., (1984) Nature 309:255-258; Gossler y col., (1986) Proc. Natl. Acad. Sci. 83:9065-9069). Los transgenes pueden introducirse de manera eficaz en las células ES mediante técnicas convencionales, tales como transfección de ADN o mediante transducción mediada por retrovirus. Las células ES transformadas resultantes pueden combinarse posteriormente con blastocistos de un animal no humano. Las células ES introducidas colonizan posteriormente el embrión y contribuyen a la línea germinal del animal quimérico resultante.

Una estrategia para el problema de determinar las contribuciones de genes individuales y sus productos de expresión es usar genes con CNV o SNP asociados con el TDAH como casetes de inserción para inactivar de manera selectiva un gen de tipo silvestre en células ES totipotentes (tales como las descritas anteriormente) y después generar ratones transgénicos. Se ha descrito el uso de células ES con genes dirigidos en la generación de ratones transgénicos con genes dirigidos y se ha revisado en otras partes (Frohman y col., (1989) Cell 56:145-147; Bradley y col., (1992) Bio/Technology 10:534-539).

Hay técnicas disponibles para inactivar o alterar cualquier región genética para una mutación deseada usando recombinación de homólogos dirigida para insertar cambios específicos en alelos cromosómicos. Sin embargo, en comparación con la recombinación de homólogos extracromosómica, que se produce con una frecuencia próxima al 100%, se ha comunicado que la recombinación de plásmido homólogo-cromosoma se detecta con frecuencias de entre 10^{-6} y 10^{-3}. Es más frecuente que las interacciones de plásmido no homólogo-cromosoma se produzcan a niveles de 10^5 veces a 10^2 veces mayores que la inserción de homólogos comparables.

Para superar esta baja proporción de recombinación dirigida en células ES murinas, se han desarrollado diversas estrategias para detectar o seleccionar recombinantes homólogos raros. Una estrategia para detectar eventos de alteración de homólogos usa la reacción en cadena de la polimerasa (PCR) para explorar grupos de células transformantes respecto de la inserción de homólogos, seguido de la exploración de clones individuales. Como alternativa, se ha desarrollado una estrategia de selección genética positiva en la que se construye un gen marcador que estará activo únicamente si se produce la inserción de homólogos, permitiendo que estos recombinantes se seleccionen directamente. Una de las estrategias más potentes desarrolladas para seleccionar recombinantes de homólogos es el procedimiento de selección positiva-negativa (PNS) desarrollado para genes para los que no existe una selección directa de la alteración. El procedimiento de PNS es más eficaz para usar como diana genes que no se expresan a altos niveles debido a que el gen marcador tiene su propio promotor. Los recombinantes no homólogos se contraseleccionan usando el gen de timidina cinasa del virus del Herpes Simplex (HSV-TK) y seleccionando contra su inserción no homóloga con fármacos eficaces contra el herpes, tales como ganciclovir (GANC) o (1-(2-desoxi-2-fluoro-B-D-arabinofuranosilo)-5-yodouracilo, (FIAU). Mediante esta contraselección, puede aumentarse el número de recombinantes en los transformantes supervivientes. El uso de ácidos nucleicos que contienen CNV o SNP asociados con el TDAH como casetes insercionales diana proporciona medios para detectar una inserción exitosa, según se visualiza, por ejemplo, mediante la adquisición de reactividad inmunitaria contra un anticuerpo inmunológicamente específico para el polipéptido codificado por el ácido nucleico con CNV o SNP asociados con el TDAH, por lo tanto, facilita la exploración/selección de células ES con el genotipo deseado.

Tal como se usa en el presente documento, un animal *knock-in* (con inserción genética) es uno en el que el gen murino endógeno, por ejemplo, se ha reemplazado por una CNV humana asociada con el TDAH o un fragmento informativo de la misma o un gen que contiene SNP de la invención. Dichos animales *knock-in* proporcionan un sistema modelo ideal para estudiar el desarrollo del TDAH.

Tal como se usa en el presente documento, la expresión de un ácido nucleico que contiene CNV o SNP asociados con el TDAH, un fragmento de CNV informativo parcial o una proteína de fusión asociada con el TDAH en la que la CNV o el SNP está codificado puede dirigirse de "manera específica de tejido" o de "manera específica de tipo celular" usando un vector en el que las secuencias de ácido nucleico que codifican la totalidad o una parte de una CNV o un SNP asociado con el TDAH está unida operativamente a secuencias reguladoras (por ejemplo, promotores y/o potenciadores) que dirigen la expresión de la proteína codificada en un tejido o tipo celular particular. Dichos elementos reguladores pueden usarse ventajosamente para aplicaciones tanto *in vitro* como *in vivo*. Se conocen en la técnica y se describen en el presente documento promotores para dirigir proteínas específicas de tejido.

La secuencia de ácido nucleico que codifica las CNV o los SNP asociados con el TDAH de la invención puede estar unida a una serie de diferentes secuencias promotoras para su expresión en animales transgénicos. Dichos promotores incluyen, pero sin limitación, un promotor de gen de priones, tal como un promotor de priones de hámster y ratón (MoPrP), descrito en la Patente de los Estados Unidos n.º 5.877.399 y en Borchelt y col., Genet. Anal. 13(6) (1996) páginas 159-163; un promotor de enolasa específica neuronal de rata, descrito en las Patentes de los Estados Unidos n.º 5.612.486 y 5.387.742; un promotor de gen de factor de crecimiento B derivado de plaquetas, descrito en la Patente de los Estados Unidos n.º 5.811.633; un promotor de distrofina específico de cerebro, descrito

en la Patente de los Estados Unidos n.º 5.849.999; un promotor Thy-1; un promotor de PGK; un promotor de CMV; un promotor de gen de factor de crecimiento B derivado de plaquetas específico de neuronas; un promotor de NEGR1, un promotor de GRM5, un promotor de cualquier gen listado en las tablas a continuación y un promotor de proteína ácida fibrilar glial (GFAP) para la expresión de transgenes en células gliales.

También se proporcionan en el presente documento procedimientos de uso para los ratones transgénicos de la invención. Los ratones transgénicos en los que se ha introducido un ácido nucleico que contiene CNV o SNP asociados con el TDAH o su proteína codificada son útiles, por ejemplo, para desarrollar procedimientos para explorar agentes terapéuticos para identificar aquellos capaces de modular el desarrollo del TDAH.

V. Terapias farmacéuticas y peptídicas

La dilucidación del papel desempeñado por las CNV y los SNP asociados con el TDAH descritos en el presente documento en la señalización de receptores de ligandos neuroactivos facilita el desarrollo de composiciones farmacéuticas útiles para el tratamiento y el diagnóstico del TDAH. Estas composiciones pueden comprender, además de una de las sustancias anteriores, un excipiente, vehículo, tampón, estabilizante farmacéuticamente aceptable u otros materiales bien conocidos para los expertos en la materia. Dichos materiales no deben ser tóxicos y no deben interferir con la eficacia del principio activo. La naturaleza precisa del vehículo u otro material puede depender de la vía de administración, por ejemplo, las vías oral, intravenosa, cutánea o subcutánea, nasal, intramuscular o intraperitoneal.

Ya sea un polipéptido, anticuerpo, péptido, molécula de ácido nucleico, molécula pequeña u otro compuesto farmacéuticamente útil de acuerdo con la presente invención que se vaya a administrar a un individuo, la administración es preferentemente en una "cantidad profilácticamente eficaz" o una "cantidad terapéuticamente eficaz" (según sea el caso, aunque la profilaxis puede considerarse terapia), siendo esto suficiente para mostrar un beneficio al individuo.

Los materiales y procedimientos expuestos a continuación se proporcionan para facilitar la práctica de los siguientes ejemplos.

Ensayo Illumina Infinium para el descubrimiento de CNV

Los presentes inventores llevaron a cabo un genotipado de SNP por todo el genoma de alto rendimiento usando la tecnología InfiniumII HumanHap550 BeadChip (Illumina, San Diego, CA), en el Centro de Genómica Aplicada en el CHOP (por sus siglas en inglés, Childen's Hospital of Philadelphia). El contenido de datos genotípicos, junto con los datos de intensidad proporcionados por la matriz de genotipado proporcionan llamadas de CNV de alta confianza. De manera importante, el análisis simultáneo de los datos de intensidad y los datos de genotipado en la misma configuración experimental establece una definición altamente precisa para estados diploides normales y cualquier desviación de los mismos. Para llamar las CNV, los presentes inventores usaron el algoritmo PennCNV, que combina múltiples fuentes de información, incluyendo la relación de Log R (LRR) y la frecuencia de alelo B (BAF) en cada marcador de SNP, junto con espaciado de SNP y la frecuencia de población del alelo B para generar las llamadas de CNV. Las cohortes de caso de replicación y de control utilizaron genotipado de SNP por todo el genoma usando las matrices Perlegen 600K, Illumina 1M y Affymetrix 5.0. Los valores de X e Y en bruto se normalizaron con el log(10) y se agruparon para determinar BAF y LRR con el protocolo PennCNV-Affy (**tabla 2**). Las CNV recurrentes raras fueron el foco del presente estudio.

Tabla 2. Muestras de archivo reformateadas de datos de Perlegen para coincidir con el formato de salida de Affymetrix Power Tools.

A) Archivo de llamadas de genotipado (0=AA,1=AB,2=BB,-1=NoCall).				
probeset_id	10009	10010	10021	10022
SNP_rs10000023	1	1	2	1
SNP_rs10000030	1	0	0	1
SNP_rs10000037	0	0	1	1
SNP_rs10000068	2	2	2	2
B) Puntuaciones de confianza de llamadas de genotipado (Todas ajustadas a 1).				
probeset_id	10009	10010	10021	10022
SNP_rs 10000023	1	1	1	1
SNP _rs10000030	1	1	1	1
SNP _rs10000037	1	1	1	1
SNP_rs10000068	1	1	1	1

(continuación)

C) Resumen de intensidad (-A=log10(X), -B=log10(Y) (valor de X e Y de archivos de informe final de muestra individual dbGaP).				
probeset_id	10009	10010	10021	10022
SNP_rs10000023-A	2,85	2,78	2,07	2,89
SNP_rs10000023-B	2,86	2,84	2,98	2,96
SNP_rs10000030-A	2,9	2,99	2,95	3,02
SNP_rs10000030-B	2,91	2,4	2,38	3,05

Llamadas de CNV y revisión de loci significativos

No se observó una "carga de CNV" adicional en los casos frente a controles, en cambio, la distribución de las llamadas efectuadas fue altamente comparable **(figura 1)**. Los presentes inventores determinaron la fiabilidad de las llamadas de CNV en los datos de Illumina y Perlegen observando los patrones mendelianos de herencia. En primer lugar, se verificaron los tríos mediante herencia de genotipo y se analizaron para establecer la calidad de las llamadas de CNV de las plataformas tanto Illumina como Perlegen basándose en la herencia observada. Basándose en todas las llamadas de CNV llamadas por tríos de los datos de Illumina del CHOP, 8.647 CNV observadas en la descendencia fueron heredadas de un progenitor mientras que 437 CNV fueron supuestamente *de novo*, lo que supone una tasa de *de novo* del 4,811%. Basándose en todas las llamadas de CNV llamadas por tríos de los datos de Perlegen IMAGE, 1.862 CNV observadas en la descendencia fueron heredadas de un progenitor mientras que 505 CNV fueron supuestamente *de novo*, lo que supone una tasa de *de novo* del 21,335%. 51 casos de IMAGE, 22 loci de eliminación y 5 loci de duplicación tuvieron múltiples eventos *de novo* debido a la baja calidad de los datos y se excluyeron como valores atípicos; una vez excluidos, 785 CNV fueron heredadas y 63 fueron *de novo*, lo que redujo la tasa de *de novo* observada hasta un nivel aceptable del 7,429%. Basándose en las CNV observadas en progenitores de los datos de Illumina del CHOP, 9.305 CNV fueron transmitidas al hijo mientras que 7.432 CNV no fueron heredadas, dando como resultado una tasa de herencia del 55,595%. Basándose en las CNV observadas en progenitores de los datos de Perlegen IMAGE, 2.114 CNV fueron transmitidas al hijo mientras que 3.789 CNV no fueron heredadas, dando como resultado una tasa de herencia del 35,812%. Los presentes inventores excluyeron 65 muestras de progenitores que eran valores atípicos con 20 o más CNV no heredadas por la descendencia y la eliminación por filtrado de estas muestras dio como resultado 1.204 CNV que se pasaron al niño mientras que 1.221 no fueron heredadas, dando como resultado una tasa de herencia del 49,650%, que aportó confianza para este conjunto de llamadas de CNV.

No es posible revisar todas las llamadas de PennCNV y es antieconómico excluir las CNV menores de un valor umbral. En cambio, los presentes inventores puntuaron estadísticamente los loci basándose en todas las CNV detectadas y la revisión de estas asoció nominalmente los loci de CNVR para solapamiento adecuado, calidad de la señal y herencia mendeliana. Tal como se muestra en la **tabla 3**, todos los loci comunicados muestran al menos un caso con la CNV heredada de un progenitor, en los casos en los que están disponibles ambos progenitores.

Tabla 3. Nuevas CNVR sobre-representadas en pacientes con TDAH

A) Loci significativamente asociados con el TDAH										
CNVR	Casos del CHOP n=1013	Controles del CHOP n=4105	Casos de replicación n=2493	Controles de replicación n=9222	Her	Valor de P combinado	ORCI(95 %)	Tipo	Gen	Distancia exónica
chr11:88269449 -88351661	4‡(3*)	0	6	1	4:1:3 62,5 %	$1,36x10^{-6}$	38,12 (5-298)	Del	*GRM5*	5.858
chr7:126525124 -126536202	3	0	5	0	0:1:0 100%	$3,52x10^{-6}$	infinito	Del	*GRM8*	0
chr3:7183953-7197236	4†(1*)	0	2	0	0:2:0 100%	$8,14x10^{-5}$	infinito	Del	*GRM7*	20.598
chr6:146657076 - 146694047	5	2	3	0	2:0:0 100%	$1,05x10^{-4}$	15,24 (3-72)	Dup	*GRM1*	0
B) Loci de TDAH con significación nominal										
CNVR	Casos del CHOP n=1013	Controles del CHOP n=4105	Casos de replicación n=2493	Controles de replicación n=9222	Her	Valor de P combinado	ORCI(95 %)	Tipo	Gen	Distancia exónica
chr1: 72317292-72328395	4‡	0	1	0	0:3:0 100%	$3,91x10^{-4}$	infinito	Dup	*NEGR1*	10.621
chr7: 153495598-153564827	5(1*)	0	3	2	1:2:0 100%	$4,08x10^{-4}$	15,24 (3-72)	Dup	*DPP6*	68.453
chr5: 65027976-65046520	4	0	2	1	2:0:2 50%	$4,68x10^{-4}$	22,85 (3-190)	Del	*SGTBINLN*	0
chr1: 56053497-56064495	2	0	4	2	1:0:3 25%	$1,54x10^{-3}$	11,42 (2-57)	Del	*USP24**	80.234
chr19: 38427720-38444834	5	2	2	3	2:2:1 80%	$4,95x10^{-3}$	5,33 (2-17)	Del	*SLC7A10**	19.172
chr3: 1844168-1859889	4†	0	3	6	1:3:0 100%	$8,81x10^{-3}$	4,44 (1-13)	Del	*CNTN4**	255.661

(continuación)

B) Loci de TDAH con significación nominal										
CNVR	Casos del CHOP n=1013	Controles del CHOP n=4105	Casos de replicación n=2493	Controles de replicación n=9222	Her	Valor de P combinado	ORCI(95 %)	Tipo	Gen	Distancia exónica
chr2: 81419297-81446082	2	0	2	3	1:0:1 50%	$3,83x10^{-2}$	5,07 (1-23)	Dup	*CTNNA2**	152.417
chr4: 113772340-113788584	2	0	2	3	0:0:0 NA	$3,83x10^{-2}$	5,07 (1-23)	Dup	*LARP7*	0

* Casos presentados en *(13):* ‡ 3 casos están presentes en la misma familia; † 2 casos están presentes en la misma familia; La columna "Her" lista el patrón de herencia de cada CNV de los progenitores a los casos en el formato <heredado de la madre>:<heredado del padre>:<*de novo*>. El porcentaje de herencia se lista a continuación. Obsérvese que no en todos los casos se encontraban disponibles los progenitores. Se comunican variantes raras que fueron recurrentes y se observó que estaban enriquecidas entre los casos de TDAH en relación con las frecuencias de control y se detectaron en múltiples cohortes independientes. Todos los genes GRM se ven afectados directamente por el CNVR. Las regiones listadas representan el solapamiento óptimo de casos y la significación respecto de los controles, tal como se describe en los Procedimientos. Se lista el gen más próximo para cada locus de CNVR, ya que es más probable que se vea afectado. Para los recuentos detallados de cada proyecto contribuyente véase la tabla 16. *No hay ningún gen afectado directamente, por lo que se lista el gen más próximo. Los límites de CNV individuales se proporcionan en la tabla 17. OR: oportunidad relativa; IC: Intervalo de confianza. La replicación representa la combinación de IMAGE, PUWMa, IMAGEII, NIMH y Utah.

ES 2 686 300 T3

En total, hay 3.506 y 13.327 controles, que representan una abundancia de más de tres veces de muestras de control para definir de manera robusta las CNV ausentes o con una frecuencia menor que en las muestras de caso. Aunque el número de CNV detectado por muestra fue de hasta 70, realmente son llamadas diploides (CN=2) normales inferidas que hacen que cada muestra sea equivalente. Estas CNV son muy raras y por lo tanto, el número de llamadas de CNV observadas variará entre muestras.

Validación de CNV mediante PCR cuantitativa (QPCR)

Se seleccionaron sondas de la Biblioteca de Sondas Universal (UPL, por sus siglas en inglés, Roche, Indianápolis, IN) usando el programa informático ProbeFinder v2.41 (Roche, Indianápolis, IN). La PCR cuantitativa se llevó a cabo en un instrumento para PCR en tiempo real ABI 7500 o en un sistema de detección de secuencias ABI Prism™ 7900HT (Applied Biosystems, Foster City, CA). Cada muestra se analizó por cuadruplicado ya sea en una mezcla de reacción de 25 µl (sonda 250 nM, cada cebador a 900 nM, Fast Start TaqMan Probe Master de Roche y 10 ng de ADN genómico) o en una mezcla de reacción de 10 µl (sonda 100 nM, cada cebador a 200 nM, Platinum Quantitative PCR SuperMix-Uracilo-ADN-Glucosilasa (UDG) 1x con ROX de Invitrogen y 25 ng de ADN genómico). Los valores se evaluaron usando Sequence Detection Software v2.2.1 (Applied Biosystems, CA). El análisis de datos se llevó a cabo adicionalmente usando el procedimiento $\Delta\Delta C_T$. Los genes de referencia, seleccionados entre COBL, GUSB y SNCA, se incluyeron basándose en el coeficiente mínimo de variación y los datos se normalizaron posteriormente ajustando un control normal a un valor de 1.

Las llamadas de CNV en la plataforma Perlegen usó un algoritmo altamente similar a los usados en las matrices Illumina, pero las etapas de procesamiento previo de la señal difieren. A diferencia de la plataforma Ilumina, en la que las intensidades de señal normalizadas (relación de Log R y frecuencia del alelo B) pueden exportarse directamente del programa informático BeadStudio, es necesario calcular estas medidas de intensidad en la plataforma Perlegen 600K a partir de la colección de muestras genotipadas basándose en los valores de X e Y en bruto. Para llevar a cabo la normalización de los datos y la extracción de señales de los archivos de informe final en bruto generados en los experimentos de genotipado, los presentes inventores reformatearon en primer lugar los datos a partir de dbGaP en el formato producido por Affymetrix Power Tools: birdseed.calls.txt, birdseed.confidences.txt y quant-norm.pm-only.med-polish.expr.summary.txt (véase la **tabla 2**). Los valores de X e Y proporcionados en la muestra basándose en los archivos de informe de dbGaP se redujeron a un intervalo más finito tomando logaritmos en base 10. Para cada marcador de SNP, los presentes inventores se basaron en la intensidad de señal específica de alelo para los genotipos AA, AB y BB en todas las muestras genotipadas para construir tres grupos genotípicos canónicos en las coordenadas polares theta y R, de manera similar a la estrategia de generación de grupos de Illumina. Se incluyó la opción "-conf 2" al ejecutar generate_affy_geno_cluster.pl ya que 1 se codificó como la mejor puntuación. Una vez se habían construido los grupos de genotipo canónico, los presentes inventores transformaron los valores de intensidad de la señal para cada SNP en valores de relación de Log R (LRR) y frecuencia de alelo B (BAF) usando normalize_affy_geno_cluster.pl. Para detalles más técnicos, véase www.openbioinformatics.org/penncnv/penncnv_tutorial_affy_gw6.html.

Para optimizar el modelo de Markov oculto (MMO), los presentes inventores usaron el archivo de referencia inicial hh550.hmm y ejecutaron "-train" en PennCNV en tres lotes sucesivos de treinta. El primer entrenamiento usó las muestras con la menor desviación estándar de LRR mientras que los otros dos ciclos, usando el archivo creado como nueva referencia, incluyeron muestras representativas más aleatorias. También se crearon archivos de definición proporcionando la distancia entre SNP y la frecuencia de alelo B de la población para informar adicionalmente de las llamadas de CNV adaptadas específicamente para a los datos de Perlegen observados. Esto permitió efectuar llamadas de CNV en 1.887 (642 casos y 1.245 progenitores) de las 2.789 muestras de Perlegen 600K disponibles. Aunque la desviación estándar global de LRR fue menor de 0,2 para la mayoría de muestras (84%), el ruido de los datos de intensidad fue notablemente mayor en las regiones de llamadas de CNV y a menudo mostraron una subpoblación de SNP incapaces de diferenciar una señal de eliminación, quizá debido a la saturación de la PCR durante el procesamiento en el laboratorio. Sin embargo, las características de eliminación y duplicación se siguieron detectando con la conformación de homocigoto y los genotipos AAB/ABB, respectivamente, mostrados para los mismos SNP (véanse las **figuras 2 y 3**).

Por último, se exploraron las llamadas de CNV de Perlegen respecto de su solapamiento con los 11 loci asociados basándose en los datos de Illumina del CHOP. Se revisaron los datos a nivel de SNP subyacentes a cada llamada de CNV para asegurar una calidad limpia de la señal. Para asegurarse de que cada CNV detectada era una característica del ADN verdadera y de modo alguno un falso positivo de la matriz Perlegen 600K usada o a la manipulación bioinformática de los datos, los presentes inventores validaron cada CNV con qPCR en un laboratorio independiente (véase la **figura 4**).

Control de calidad de CNV

Los presentes inventores calcularon las medidas de Control de Calidad (QC) en los datos de HumanHap550 GWAS basándose en distribuciones estadísticas para excluir muestras de ADN de mala calidad y CNV falsas positivas. El primer umbral es el porcentaje de SNP probadas que se genotiparon satisfactoriamente. Solo se incluyeron las muestras con una tasa de llamada > 98%. La señal de intensidad por todo el genoma ha de tener tan poco ruido como sea posible. Solo se incluyeron muestras con la desviación estándar (DE) de la intensidad normalizada (LRR)

< 0,35. Todas las muestras deben tener un origen étnico caucásico, basándose en el análisis de componentes principales (**figura 5**) y se excluyeron todas las demás muestras. Además, se aseguró el emparejamiento de casos y controles calculando un factor de inflación genómica (GIF=1,024) entre los grupos. Se sabe que los falsos positivos de onda que se correlacionan vagamente con el contenido de GC a consecuencia del sesgo de hibridación de una baja cantidad de ADN de longitud completa interfieren con la inferencia precisa de las variaciones del número de copias[43]. Solo se aceptaron muestras en las que el factor de onda de LRR al modelo de onda variaron entre -0,5<x<0,6. En caso de que el recuento de llamadas de CNV efectuadas por PennCNV supere 70 (**figura 1**), la calidad del ADN es normalmente baja. Por lo tanto, solo se incluyeron muestras con un recuento de llamadas de CNV < 70. Cualquier muestra duplicada (como en gemelos monocigóticos) tenía una muestra excluida. La **tabla 4** proporciona el número de muestras excluidas para cualquier medición de control de calidad.

Tabla 4. Exclusión de muestras basándose en las mediciones de control de calidad.

Criterios de exclusión	CHOP	Control
Tasa de llamada < 98%	170	271
DE LRR > 0,35	73	124
Etnia no caucásica	71	48
Factor de onda -0,5>X>0,6	251	1040
Recuento de CNV > 70	197	237
Gemelo monocigótico	31	38
Muestras excluidas basándose en medidas de Control de Calidad (QC) en los datos de HumanHap550 GWAS basándose en distribuciones estadísticas para excluir muestras de ADN de mala calidad y CNV falsas positivas.		

Análisis estadístico de CNV

Se evaluó la frecuencia de CNV entre casos y controles en cada SNP usando la prueba exacta de Fisher. Solo se tomaron en consideración loci que fueron predominantemente significativos entre casos y controles (p<0,05) en los que los casos en la cohorte de descubrimiento del CHOP tenía la misma variación, replicada en IMAGE, PUWMa o IMAGE II o no se observaron en cualquiera de los sujetos de control y se validaron con un procedimiento independiente. Los presentes inventores comunican mínimos locales estadísticos para limitar la asociación en referencia a una región de significación nominal que incluye SNP que residen a menos de 1 Mb entre sí (**figura 4**). Las CNVR nominalmente significativas se excluyeron en caso de que cumpliesen cualquiera de los siguientes criterios: i) encontrarse en citobandas próximas a telómeros o centrómeros; ii) surgir en una "península" de CNV comunes que surgen por una variación en el truncamiento de límites de llamadas de CNV (**figura 7**); iii) regiones genómicas con extremos en el contenido de GC que producen sesgo de hibridación; o iv) muestras que contribuyen a múltiples CNVR. Los presentes inventores efectuaron un ajuste estadístico para el grado de relación de los casos con permutación (1000x). Tres líneas de pruebas determinan la significación estadística: replicación independiente, p<0,05, permutación de observaciones y ausencia de observación de loci con significación de control enriquecida. Los presentes inventores usaron DAVID (por sus siglas en inglés, *Database for Annotation, Visualization, and Integrated Discovery*) [44] para evaluar la significación del agrupamiento de anotación funcional de resultados de CNV asociados de manera independiente en categorías de InterPro.

Permutación para ajustar la significación para el grado de relación

Para la prueba exacta de Fisher inicial, no se controlan los individuos relacionados ya que el objetivo principal es detectar CNV en múltiples muestras independientemente del grado de relación. Las CNV que pasan esta exploración inicial se puntúan respecto de su significación estadística basándose en un valor de P permutado que permuta aleatoriamente los marcadores de caso y control de todas las muestras a condición de que los individuos relacionados tengan el mismo marcador. A cada individuo no relacionado se le asigna un marcador de caso o control y a su hermano relacionado se le asigna el mismo marcador. Basándose en el número de muestras con la CNVR que se está calculando en "casos" y "controles" asignados de manera aleatoria, se asigna un valor de P de la prueba exacta de Fisher. El número de escenarios hipotéticos con una significación igual o mayor (valor de P menor) proporciona el valor de P permutado, que corrige respecto del grado de relación. Por transparencia, se muestran el valor de P de la prueba exacta de Fisher y los recuentos de casos y controles con cada CNVR.

Análisis de asociación por todo el genoma de llamadas de genotipo

Se llevó a cabo una asociación por todo el genoma de genotipo a escala completa y el valor de inflación genómica (GIF) se encontraba a un nivel aceptable (GIF=1,02409). También se comprobó la concordancia de población emparejada para comprobar y filtrar los grados de relación crípticos que podrían dar lugar a CNV raros específicos para subpoblaciones ultra-estratificadas de Europa. Los presentes inventores llevaron a cabo una estadística de prueba de transmisión de desequilibrio (TDT) usando Plink en 397 casos de TDAH con ambos progenitores en los datos de genotipo de Illumina HumanHap550 del CHOP (**tabla 5**). El mejor resultado con más de un SNP significativo en una región fue chr4p12 P(rs1018199)=2,71x10^{-5} y P(rs11724347)=6,19x10^{-5} que tiene impacto sobre *TEC.* También se llevó a cabo una asociación por todo el genoma de genotipo de caso:control en 735 casos y 2.298

controles usando el mismo conjunto de datos de Illumina (**tabla 6**). La señal más fuerte fue chr19p12 P(rs2081051)=4,60x10^{-6} y P(rs399686)=4,72x10^{-6} que reside entre *ZNF66* y *ZNF85.* Por último, se analizaron 623 casos de TDAH con ambos progenitores en los datos de IMAGE Perlegen 600K con la estadística TDT (**tabla 7**). La señal más significativa fue chr5q23.1 P(rs17144308)= 9,70x10^{-6} y P(rs2043053)=3,36x10^{-5} que se encuentra a 237 kb del gen cercano más próximo *DTWD2.* Tomados en conjunto, los SNP que residen alrededor del exón 4 de contactina 3 (*CNTN3*) parecen replicarse de manera más consistente entre las estadísticas de TDT de TDAH de Illumina y Perlegen. SNP rs12488030 es común para ambas plataformas; P=2,51x10^{-3} para Illumina y P=4,97x10^{-3} para Perlegen. Hay dos SNP de soporte estrechamente próximos que también muestran significación en Illumina: P(rs4073942)= 2,78x10^{-3} y P(rs9869828)=8,61x10^{-3} y además en Perlegen: P(rs11915713) =1,86x10^{-5} y P(rs7372975) =7,59x10^{-5}.

Tabla 5. Análisis de TDT de 397 casos de TDAH y progenitores del CHOP genotipados en el chip Illumina HH550.

CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
18	rs8095193	58834095	1	2	167	92	1,815	21,72	3,16E-06
17	rs4357980	13498634	1	2	99	174	0,569	20,6	5,65E-06
18	rs8091710	72897492	1	2	29	73	0,3973	18,98	1,32E-05
14	rs899116	97495185	1	2	101	172	0,5872	18,47	1,73E-05
13	rs9595945	48099556	1	2	245	160	1,531	17,84	2,40E-05
4	rsl018199	47927632	1	2	35	80	0,4375	17,61	2,71E-05
1	rs3795324	157456184	2	1	91	157	0,5796	17,56	2,78E-05
3	rs6444186	188156541	1	2	81	36	2,25	17,31	3,18E-05
9	rs11144627	75654927	2	1	46	14	3,286	17,07	3,61E-05
8	rs1462011	108104653	1	2	199	125	1,592	16,9	3,94E-05
X	rs5991935	100480088	1	2	22	59	0,3729	16,9	3,94E-05
7	rs1013572	78350227	1	2	63	118	0,5339	16,71	4,35E-05
11	rs952619	20316347	1	2	108	177	0,6102	16,71	4,37E-05
4	rs7689018	85116479	1	2	41	87	0,4713	16,53	4,79E-05
18	rs1943825	69128567	2	1	97	162	0,5988	16,31	5,37E-05
4	rs4696821	8473961	1	2	210	135	1,556	16,3	5,39E-05
18	rs1943823	69131624	2	1	157	237	0,6624	16,24	5,57E-05
4	rs11724347	47923023	1	2	26	64	0,4062	16,04	6,19E-05
1	rs7530899	76950752	2	1	89	151	0,5894	16,02	6,28E-05
18	rs4890560	41457783	1	2	93	156	0,5962	15,94	6,54E-05
6	rs2677099	45527900	1	2	220	144	1,528	15,87	6,79E-05
12	rs11067228	113556980	2	1	231	153	1,51	15,84	6,88E-05
6	rs2790102	45540192	1	2	222	146	1,521	15,7	7,44E-05
1	rs4926757	48961624	1	2	192	122	1,574	15,61	7,80E-05
11	rsl7147479	84055504	1	2	137	79	1,734	15,57	7,93E-05
17	rs9913261	12026365	2	1	89	150	0,5933	15,57	7,96E-05
9	rs7041883	135352660	1	2	17	49	0,3469	15,52	8,19E-05
12	rs7309946	103478293	2	1	119	188	0,633	15,51	8,22E-05
7	rs10226468	42907176	2	1	144	219	0,6575	15,5	8,27E-05
5	rs438418	2902436	2	1	78	36	2,167	15,47	8,37E-05
8	rs12682232	108078371	2	1	199	128	1,555	15,42	8,63E-05
X	rs5956634	123092612	2	1	59	110	0,5364	15,39	8,74E-05
7	rs7786719	42850356	1	2	133	205	0,6488	15,34	8,99E-05
6	rs910586	45518290	1	2	221	146	1,514	15,33	9,04E-05
6	rs9395010	44453984	1	2	152	91	1,67	15,31	9,11E-05
14	rs11844273	97489409	1	2	100	163	0,6135	15,09	1,02E-04
2	rs11904235	36288350	1	2	64	27	2,37	15,04	1,05E-04
11	rs487518	131283728	1	2	150	225	0,6667	15	1,08E-04

(continuación)

CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
6	rs6920606	33105652	2	1	164	242	0,6777	14,99	1,08E-04
14	rs2014525	97491178	1	2	109	174	0,6264	14,93	1,12E-04
11	rs7948111	23403649	1	2	65	117	0,5556	14,86	1,16E-04
16	rs12598067	60940038	2	1	65	117	0,5556	14,86	1,16E-04
6	rs9472494	45559814	1	2	223	149	1,497	14,72	1,25E-04
7	rs533486	99085345	2	1	163	240	0,6792	14,71	1,25E-04
8	rs7835921	96345468	1	2	157	96	1,635	14,71	1,26E-04
4	rs827019	8460842	2	1	69	122	0,5656	14,71	1,26E-04
CHR:Número de cromosoma, SNP:Identificador de SNP, A1:Código del alelo menor, A2:Código del alelo mayor, T:Recuento de alelo menor transmitido, U:Recuento de alelo no transmitido, OR:Oportunidad relativa de TDT, CHISQ:Estadística de chi-cuadrado de TDT, P:Valor de p asintótico de TDT									

Tabla 6. Análisis de caso:control de 735 casos de TDAH y 2.298 controles no relacionados del CHOP genotipados en el chip Illumina HH550.

CHR	SNP	BP	A1	A2	F_A	F_U	OR	CHISQ	P
18	rs16943400	23086102	1	2	0,02778	0,08875	0,2934	57,53	3,33E-14
3	rs7649108	166136126	1	2	0,3156	0,2497	1,386	24,88	6,11E-07
6	rs9390261	145283744	1	2	0,02585	0,009072	2,899	24,54	7,29E-07
X	rs4609327	37790223	2	1	0,1441	0,08032	1,928	24,48	7,50E-07
X	rs5917547	37803525	2	1	0,1578	0,09074	1,878	24,22	8,59E-07
16	rs2278656	54885245	1	2	0,01443	0,04091	0,3432	22,04	2,67E-06
8	rs17834541	2674349	2	1	0,1083	0,1565	0,6545	21,01	4,56E-06
19	rs2081051	20866811	1	2	0,1382	0,1911	0,6786	21	4,60E-06
19	rs399686	20772798	1	2	0,143	0,1962	0,6833	20,95	4,72E-06
X	rs5917937	39750534	2	1	0,1195	0,06572	1,929	20,93	4,76E-06
19	rs10419820	20943636	2	1	0,1789	0,2357	0,7067	20,9	4,84E-06
X	rs10522011	32517409	1	2	0,05924	0,02509	2,447	19,48	1,02E-05
8	rs11203872	17531028	2	1	0,4342	0,37	1,306	19,34	1,09E-05
X	rs9633179	3535471	2	1	0,1089	0,05969	1,925	19,24	1,15E-05
4	rs10519629	143040375	2	1	0,1864	0,1398	1,409	18,81	1,44E-05
19	rs7253306	20951939	2	1	0,219	0,2759	0,736	18,77	1,48E-05
13	rs9569383	55299477	1	2	0,1415	0,1909	0,6984	18,64	1,58E-05
12	rs12229174	62532933	1	2	0,06054	0,03502	1,776	18,56	1,64E-05
19	rs6511169	20893589	1	2	0,1461	0,1961	0,7014	18,51	1,69E-05
11	rs10833476	21190445	1	2	0,1224	0,08502	1,502	18,48	1,72E-05
2	rs1821659	212064488	2	1	0,3109	0,2527	1,334	18,15	2,05E-05
X	rs2480443	53212284	2	1	0,06525	0,02994	2,262	18,1	2,09E-05
7	rs1486173	45965025	2	1	0,1131	0,07764	1,515	17,91	2,32E-05
15	rs4381545	93039961	2	1	0,2296	0,18	1,358	17,8	2,45E-05
7	rs10265665	96175055	1	2	0,0619	0,0365	1,742	17,79	2,46E-05
10	rs11593585	44391199	1	2	0,1286	0,09093	1,475	17,69	2,60E-05
X	rs4134188	17474194	1	2	0,1016	0,05571	1,917	17,62	2,69E-05
4	rs11131363	63013616	2	1	0,2643	0,212	1,335	17,6	2,72E-05
19	rs1469402	20738115	2	1	0,145	0,1934	0,7075	17,52	2,85E-05
11	rs12279152	133861485	1	2	0,02653	0,01139	2,365	17,43	2,98E-05
X	rs5957334	119125665	2	1	0,06667	0,03136	2,206	17,13	3,49E-05
X	rs6632558	36075450	2	1	0,0812	0,04176	2,028	16,94	3,85E-05
1	rs2057594	117348535	1	2	0,2483	0,1983	1,335	16,89	3,96E-05
8	rs17834523	2672777	1	2	0,09592	0,1367	0,6699	16,84	4,06E-05

(continuación)

CHR	SNP	BP	A1	A2	F_A	F_U	OR	CHISQ	P
7	rs10485959	78702412	2	1	0,3007	0,3595	0,7659	16,83	4,09E-05
X	rs5945330	152438289	2	1	0,08807	0,04698	1,959	16,63	4,55E-05
3	rs16854851	145238402	1	2	0,02381	0,009916	2,435	16,62	4,56E-05
8	rs2237826	17519195	2	1	0,4355	0,376	1,28	16,59	4,65E-05
X	rs16987407	35968032	2	1	0,1041	0,05857	1,868	16,5	4,87E-05
X	rs4089885	22878045	2	1	0,1193	0,07027	1,792	16,47	4,94E-05
1	rs2024766	181385290	2	1	0,5027	0,4424	1,274	16,45	4,99E-05
4	rs9312518	173526549	1	2	0,4639	0,4042	1,276	16,45	5,00E-05
4	rs9997484	173517324	2	1	0,4639	0,4042	1,276	16,45	5,00E-05
17	rs4338847	7870502	1	2	0,3102	0,3679	0,7725	16,35	5,28E-05
12	rs17497206	113000660	2	1	0,1537	0,2011	0,7219	16,33	5,32E-05

CHR:Cromosoma, SNP:ID de SNP, BP:Posición física (par de bases), A1:Nombre del alelo menor (basándose en la muestra completa), F_A:Frecuencia de este alelo en casos, F_U:Frecuencia de este alelo en controles, A2:Nombre del alelo mayor, OR:Oportunidad relativa estimada (para A1, es decir, A2 es de referencia), CHISQ:Chi-cuadrado de prueba alélica básica (1df), P:Valor de p asintótico para esta prueba.

Tabla 7. Análisis de TDT de 623 casos de TDAH y progenitores de IMAGE genotipados en la plataforma Perlegen.

CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
12	rs3782309	26750663	1	2	172	99	1,737	19,66	9,23E-06
5	rs17144308	117965870	2	1	244	352	0,6932	19,57	9,70E-06
2	rs7609261	80530821	2	1	199	297	0,67	19,36	1,08E-05
3	rs1344870	21282405	2	1	16	52	0,3077	19,06	1,27E-05
18	rs7244637	17876224	1	2	134	215	0,6233	18,8	1,45E-05
1	rs3850879	48004718	1	2	226	143	1,58	18,67	1,56E-05
14	rs2295426	58446208	2	1	209	307	0,6808	18,61	1,60E-05
16	rs7204253	5576184	2	1	114	189	0,6032	18,56	1,64E-05
4	rs1378945	25382295	2	1	212	310	0,6839	18,4	1,79E-05
3	rs11915713	74568983	1	2	176	266	0,6617	18,33	1,86E-05
12	rs11830382	41718893	2	1	198	122	1,623	18,05	2,15E-05
12	rs4761641	93525817	2	1	137	215	0,6372	17,28	3,22E-05
5	rs2043053	117958083	2	1	126	201	0,6269	17,2	3,36E-05
18	rs12965880	22313077	1	2	235	333	0,7057	16,91	3,92E-05
9	rs17306197	97862011	1	2	162	96	1,688	16,88	3,97E-05
8	rs17668689	96254526	1	2	216	310	0,6968	16,8	4,16E-05
2	rs4852567	80556890	2	1	206	298	0,6913	16,79	4,17E-05
13	rs1002468	93085569	2	1	287	197	1,457	16,74	4,30E-05
1	rs10873925	77234323	2	1	305	212	1,439	16,73	4,31E-05
16	rs12596741	17345435	1	2	228	324	0,7037	16,7	4,39E-05
9	rs2991298	3284851	2	1	81	142	0,5704	16,69	4,41E-05
14	rs1427324	58434446	1	2	206	297	0,6936	16,46	4,96E-05
10	rs11258682	13951273	1	2	204	130	1,569	16,4	5,14E-05
4	rs10520276	175420068	2	1	216	140	1,543	16,22	5,63E-05
1	rs17375519	179499648	1	2	75	133	0,5639	16,17	5,78E-05
1	rs10800069	163296159	1	2	232	327	0,7095	16,14	5,87E-05
7	rs13340504	75277632	1	2	142	82	1,732	16,07	6,10E-05
2	rs6543239	104056246	2	1	251	349	0,7192	16,01	6,31E-05
2	rs4664452	162762970	1	2	30	6	5	16	6,33E-05
4	rs16889099	13341184	2	1	48	96	0,5	16	6,33E-05

(continuación)

CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
5	rs12520147	2000122	1	2	158	237	0,6667	15,8	7,04E-05
11	rs10400283	23523711	1	2	222	314	0,707	15,79	7,07E-05
4	rs1378946	25382548	1	2	197	284	0,6937	15,74	7,28E-05
3	rs7372975	74602140	2	1	169	250	0,676	15,66	7,59E-05
17	rs11654470	74388926	2	1	82	141	0,5816	15,61	7,79E-05
3	rs9878591	121464488	1	2	107	173	0,6185	15,56	8,01E-05
12	rs1553953	28724544	1	2	76	133	0,5714	15,55	8,06E-05
11	rs7121790	45021541	1	2	171	252	0,6786	15,51	8,20E-05
12	rs1452231	83750252	2	1	223	314	0,7102	15,42	8,60E-05
7	rs194847	103560404	1	2	347	251	1,382	15,41	8,65E-05
2	rs11902138	80565100	1	2	173	254	0,6811	15,37	8,86E-05
16	rs12932714	80320240	1	2	150	226	0,6637	15,36	8,88E-05
1	rs1015144	200004976	2	1	204	291	0,701	15,29	9,22E-05
22	rs6009441	47873456	1	2	107	172	0,6221	15,14	9,97E-05
8	rs4734069	104169047	1	2	275	191	1,44	15,14	9,97E-05
20	rs2024946	61678306	2	1	112	61	1,836	15,03	1,06E-04

CHR:Número de cromosoma, SNP:Identificador de SNP, A1:Código del alelo menor, A2:Código del alelo mayor, T:Recuento de alelo menor transmitido, U:Recuento de alelo no transmitido, OR:Oportunidad relativa de TDT, CHISQ:Estadística de chi-cuadrado de TDT, P:Valor de p asintótico de TDT

Criterios del estudio para inclusión en IMAGEN

Diagnóstico del probando: TDAH de subtipo combinado.

Niños con edades comprendidas entre los 6-17 años (inclusive).

Uno o más hermanos en el mismo intervalo de edad.

Disponibilidad para que ambos progenitores proporcionen una muestra de ADN o un progenitor disponible más dos o más hermanos.

CI por encima de 70.

Ausencia de trastornos monogénicos que se sabe que están asociados con el TDAH (por ejemplo, X frágil, fenilcetonuria, hipercalcemia, resistencia a la hormona tiroidea).

Ausencia de enfermedades y daños neurológicos (por ejemplo, hemiplejia y otras parálisis cerebrales, epilepsia, hidrocefalia, síndromes post-encefalíticos, psicosis, discapacidades sensorimotoras).

Convivir con al menos un progenitor biológico y un hermano de padre y madre.

No cumplir con los criterios para autismo o síndrome de Asperger.

Criterios del estudio para inclusión en IMAGE II

Diagnóstico del probando: TDAH según DSM-IV-TR
Entrevista de diagnóstico semi-estructurada: KSADS-PL o Kinder -DIPS
Child Behavior Checklist, Conners parent and teacher Scales o German Teachers Report de síntomas del TDAH según DSM-IV
Niños con edades entre los 6-18 años (probando inicial mayor de 8 años).

CI por encima de 70; peso al nacer > 2000 g; ausencia de eventos médicos importante durante el embarazo; ausencia de abuso de sustancias por parte de la madre durante el embarazo
Ausencia de trastornos monogénicos que se sabe que están asociados con el TDAH (por ejemplo, X frágil, fenilcetonuria, hipercalcemia, resistencia a la hormona tiroidea).

Ausencia de enfermedades y daños neurológicos (por ejemplo, hemiplejia y otras parálisis cerebrales, epilepsia, hidrocefalia, síndromes post-encefalíticos, trastorno neuronal motor, etc.).

No cumplir con los criterios para autismo o síndrome de Asperger, esquizofrenia, trastorno bipolar, episodio depresivo mayor primario y trastorno de ansiedad, síndrome de Tourette.

Controles para IMAGE II

Los sujetos de control usados se extrajeron de sujetos genotipados mediante Affymetrix 6.0 del repositorio de genética NIMH. Se habían recogido mediante un estudio representativo nacional en los Estados Unidos (de aproximadamente 60.000 individuos adultos una sola vez, con renovación constante) determinados mediante marcaje aleatorio. Los participantes fueron cribados respecto de psicosis y trastorno bipolar. Los participantes de control no fueron cribados respecto de TDAH. Se recogió una muestra de sangre mediante un servicio de flebotomía nacional de los Estados Unidos. Los participantes de control otorgaron consentimiento por escrito para que sus materiales biológicos se empleasen para investigación médica a discreción del NIMH. Los controles se genotiparon usando la matriz Affymetrix 6.0, en el Broad Institute National Center for Genotyping and Analysis. Las llamadas de genotipo se efectuaron con el programa BIRDSEED, un módulo del paquete BIRDSUITE.

Análisis de red

Se usó el programa informático Cytoscape [47] para identificar 228 genes con 2 grados de relación con 8 genes de GRM basándose en el interactoma humano agrupado. Los presentes inventores agruparon esta red de genes en 17 grupos modulares distintos basándose únicamente en la topología de la red usando el complemento ClusterViz para el programa informático usando el algoritmo FAG-EC con los parámetros por defecto. Los genes componentes de cada uno de los 17 módulos se remitieron a DAVID[44] para evaluar la significación del enriquecimiento funcional usando anotaciones de GO de *Homo sapiens*.

Los siguientes ejemplos se proporcionan para ilustrar ciertas realizaciones de la invención. Con ellos no se pretende limitar la invención de modo alguno.

Ejemplo I

Alteraciones génicas del receptor de glutamato metabotrópico asociadas con el TDAH

Se han identificado varias CNV recurrentes raras que están sobre-representadas en múltiples cohortes de TDAH independientes que tienen impacto en genes implicados en la neurotransmisión glutaminérgica, un mediador importante para el desarrollo del cerebro y de una función cerebral normal. Estos resultados implican variaciones que abarcan redes génicas glutamatérgicas del cerebro como contribuyentes a la susceptibilidad genética al TDAH.

Participantes del estudio

La cohorte de descubrimiento incluyó un total de 1.013 casos de TDAH de ascendencia europea del norte genotipados en el Childen's Hospital of Philadelphia (CHOP). Esta consistió en 664 casos sin progenitores y 349 casos de tríos completos reclutados en el CHOP (véanse las **tablas 8 y 9**).

Tabla 8. Demográfica clínica de los participantes del estudio.

Cohorte de TDAH	N	Intervalo de edad de los sujetos con TDAH	Ascendencia	Determinación de TDAH
Tríos de TDAH del CHOP	349	6-18	Europea	K-SADS-IVR
Casos de TDAH del CHOP	664	6-18	Europea	Diagnóstico clínico de TDAH y tratamiento con medicamentos para el TDAH; K-SADS-IVR en la mayoría
Tríos de TDAH del NIMH	128	6-12	Europea	DICA; Escalas Conners
Casos UTAH	90	19-60	Europea	WRAADDS, WURS, PRS, criterios DSM-IV estrictos, incluyendo edad de inicio antes de los 7
Tríos de TDAH de IMAGE	642	6-17	Europea	PACS, Conners, SDQ, WISC
Tríos de TDAH de IMAGE II	787	5-14	Europea	Versión alemana de K-SADS, Kinder-DIPS, escalas de padres y profesores de Conners, WISC, K-ABC

(continuación)

Cohorte de TDAH	N	Intervalo de edad de los sujetos con TDAH	Ascendencia	Determinación de TDAH
Tríos de PUWMa	864	6-18	Europea	K-SADS
PACS: Relato parental de los síntomas del niño; Conners: Escalas de puntuación de comportamiento; SDQ: Cuestionario de resistencia y dificultades; WISC: Escala de inteligencia para niños de Wechsler (WISC-IV); KSADS-IVR: Pauta para trastornos afectivos y esquizofrenia para niños en edad escolar-IVR; DICA: Entrevista diagnóstica para niños y adolescentes; Kinder-DIPS: Entrevista diagnóstica para trastornos psiquiátricos en niños, K-ABC: escala de inteligencia de Kaufman-ABC. WRAADDS=Escala de trastorno por déficit de atención en adultos de Wender-Reimherr; WURS=Escala de puntuación Wender Utah; PRS=Escala de puntuación de progenitores.				

Tabla 9. Gravedad de TDAH según K-SADS de los participantes del estudio del CHOP en los dominios de falta de atención, impulsividad e hiperactividad.

Criterios diagnósticos	Puntuación 1	Puntuación 2	Puntuación 3	Puntuación 4
Con frecuencia descuidado	7	40	372	81
Pierde cosas	18	126	277	79
Dificultad para terminar cosas	16	90	311	83
Escucha	10	22	320	148
Concentración*	2	25	337	135
Distraído	1	10	307	182
Organización	19	79	304	98
Esquivo	19	55	278	148
Olvidadizo	19	75	290	116
Interrumpe	28	73	305	94
Actúa antes de pensar	28	112	283	77
Cambia de actividades	72	134	247	47
Habla bruscamente †	135	82	232	48
Dificultad para aguardar turno	80	172	200	48
Hiperactivo	53	127	227	93
Inquieto	15	47	301	137
Dificultad para estar sentado	45	80	287	88
Muy activo	49	89	255	107
Habla en exceso	37	77	255	131
Dificultad para jugar con tranquilidad	98	120	233	49
‘Concentración 1 registro ausente † Habla bruscamente 3 registros ausentes. Las puntuaciones 1 y 2 significan que los síntomas se encuentran dentro del intervalo normal mientras que las puntuaciones 3 y 4 son excesivas.				

Para abordar la replicación, los presentes inventores accedieron a las cohortes de IMAGE, que son parte de la Genetic Association Information Network (GAIN). Hubo 624 muestras de IMAGE que cumplieron con los criterios de control de calidad para el estudio. Se proporcionó acceso a estos genotipos y los datos de intensidad para IMAGE a través de la base de datos de Genotipos y Genotipos (dbGaP). El consorcio PUWMa de la University of California en Los Angeles, el Massachusetts General Hospital y Washington University St. Louis aportaron 864 casos de TDAH y 1.258 progenitores. El consorcio IMAGE II aportó 787 casos de TDAH y 898 controles no relacionados. Además, 128 casos reclutados en el NIMH y 90 casos reclutados en The University of Utah también sirvieron para la replicación. Las muestras de ADN de las cohortes del CHOP, NIMH y Utah se genotiparon usando Illumina Infinium HumanHap550K BeadChip en el CHOP. La cohorte de IMAGE se genotipó usando la plataforma Perlegen 600K. La cohorte de PUWMa se genotipó en el Illumina 1M BeadChip. La cohorte de IMAGE II se genotipó en la matriz Affymetrix 5.0. Para gestionar las diferencias en la detección de CNV entre matrices, los presentes inventores usaron controles genotipados en plataformas coincidentes con las plataformas de caso, incluyendo: 4.105 Illumina 550k del CHOP, 3.297 Perlegen 600k de los proyectos de psoriasis y depresión de GAIN, 3.469 Illumina 1M de progenitores y SAGE de PUWMa y 2.456 controles de Affymetrix 5.0 y 6.0 del repositorio de genética del NIMH y de progenitores

de AGRE.

Tamaño y número de CNV en casos y controles

Para buscar nuevas CNV, se analizaron los 1.013 casos del CHOP como cohorte de descubrimiento en comparación con 4.105 niños de control, todos los cuales fueron de ascendencia europea. Los datos de las cohortes de IMAGE, PUWMa, IMAGE II, NIMH y Utah se usaron para replicación, junto con una cohorte de control independiente de 9.222 genotipados en las mismas plataformas. Por lo tanto, la frecuencia de CNV de control se caracteriza de manera robusta en múltiples cohortes grandes, basándose en las plataformas Illumina, Perlegen y Affymetrix. Los presentes inventores observaron que de las 2.713 muestras de IMAGE (934 casos) disponibles en dbGaP, 1.886 (624 casos) cumplieron de manera estricta los umbrales de calidad de datos para las CNV.

Se usó el programa PennCNV para producir las llamadas de CNV para casos y controles, como se ha descrito anteriormente[10]. La frecuencia de CNV de los sujetos que cumplieron los estándares de calidad, que incluían eliminar valores anómalos sustanciales en el recuento de la métrica de calidad de llamadas de CNV que se desviaban exponencialmente de la distribución de la mayoría de la cohorte, dio como resultado un 93% de sujetos que tenían 8-45 llamadas de CNV (**figura 1**). Los presentes inventores denominaron cuatro estados de número de copias diferentes, incluyendo 3.172 eliminaciones homocigóticas (número de copia o CN=0), 27.810 eliminaciones hemicigóticas (CN =1), 14.806 duplicaciones de una copia (CN =3) y 581 duplicaciones de dos copias (CN =4). La **figura 8** muestra un ejemplo de datos de Illumina en bruto según se ven en el programa informático BeadStudio y las llamadas de CNV resultantes. Las llamadas de CNV abarcaron de 3 a 598 SNP, con una media de 14 SNP por llamada de CNV, con la CNV mayor de 2,2 Mb y un tamaño medio de CNV de 62 kb. La cobertura variable de la sonda permite la detección de CNV de hasta un tamaño físico pequeño, a condición de que haya presencia de al menos 3 SNP y las CNV se validaron experimentalmente usando qPCR.

Los individuos de control examinados también tenían un 93% de sujetos con 8-45 llamadas de CNV (**figura 1**). Entre las llamadas de CNV, los presentes inventores identificaron 4.471 eliminaciones homocigóticas (CN =0), 49.726 eliminaciones hemicigóticas (CN =1), 27.032 duplicaciones de una copia (CN =3) y 1.480 duplicaciones de dos copias (CN =4). Las llamadas de CNV abarcaron de 3 a 708 SNP, con una media de 12,8 SNP por llamada de CNV, con la CNV mayor de 2,9 Mb y un tamaño medio de CNV de 53,6 kb.

Evaluación de la asociación de SNP

Se llevó a cabo un análisis GWA en la cohorte de descubrimiento, aunque no se detectó ni una sola señal de asociación de genotipo de SNP individual que cumpliese los criterios estadísticos para significación en todo el genoma ($p<5x10^{-8}$) (véanse las **tablas 5, 6 y 7**)**.** Sin embargo, se observaron pruebas de replicación de varios SNP de exón terminal en el gen GFOD1 en las familias del CHOP, usando TDT (intervalo de valor de $p = 8x10^{-4}$ - $1x10^{-2}$, para rs1866863, rs9370020, rs2254292 y rs2439565). Los presentes inventores comunican la significación observada para otros SNP comunicados previamente (Lesch y col. 2008; Zhou, y col. 2008) con las pruebas que los respaldan en la **tabla 10.**

Tabla 10. Significación de SNP GWAS de los SNP asociados con TDAH de mejor clasificación comunicados por Lesch y Zhou. A) Datos de TDT de TDAH de Illumina 550k del CHOP; B) Datos de caso:control de TDAH de Illumina 550k del CHOP; C) Datos de TDAH de Perlegen 600k de IMAGE.

A)									
CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
2	rs2241685	1896290	1	2	72	62	1,161	0,7463	0,3877
2	rs13395022	79793915	2	1	136	136	1	0	1
2	rs2587695	120038047	1	2	183	197	0,9289	0,5158	0,4726
2	rs2242073	208819551	2	1	108	106	1,019	0,01869	0,8913
2	rs1110998	217169458	1	2	175	159	1,101	0,7665	0,3813
3	rs10510238	2876647	2	1	84	93	0,9032	0,4576	0,4987
3	rs9879164	54040611	2	1	185	198	0,9343	0,4413	0,5065
3	rs2084358	57457928	2	1	182	198	0,9192	0,6737	0,4118
3	rs10490808	59939739	2	1	175	204	0,8578	2,219	0,1363
3	rs10510850	60542142	1	2	90	83	1,084	0,2832	0,5946
4	rs755403	6507714	2	1	195	180	1,083	0,6	0,4386
4	rs10516182	7143981	2	1	155	169	0,9172	0,6049	0,4367
4	rs7697323	7801488	1	2	180	222	0,8108	4,388	0,0361 9
5	rs173754	65102081	1	2	218	202	1,079	0,6095	0,435
5	rs258082	66166352	1	2	199	205	0,9707	0,08911	0,7653

(continuación)

A)									
CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
6	rs160666	2719051	2	1	179	181	0,989	0,01111	0,9161
6	rs2842643	41758714	2	1	180	149	1,208	2,921	0,0874 4
6	rs3799977	44945334	2	1	209	183	1,142	1,724	0,1891
6	rs8180608	89064414	2	1	178	218	0,8165	4,04	0,0444 2
6	rs1358601	91532294	1	2	180	181	0,9945	0,00277	0,958
6	rs6921403	154156020	2	1	86	90	0,9556	0,09091	0,763
7	rs2237349	28536203	2	1	176	191	0,9215	0,6131	0,4336
7	rs2002865	154132035	2	1	134	157	0,8535	1,818	0,1776
8	rs6991017	5508780	2	1	127	126	1,008	0,003953	0,9499
8	rs2248529	14657354	1	2	188	190	0,9895	0,01058	0,9181
8	rs4961315	142110882	2	1	186	152	1,224	3,42	0,0644 1
9	rs2418326	114759028	1	2	141	142	0,993	0,003534	0,9526
9	rs2502731	128056111	2	1	170	178	0,9551	0,1839	0,668
14	rs10483393	31530235	1	2	146	137	1,066	0,2862	0,5927
15	rs2556560	42609135	2	1	169	171	0,9883	0,01176	0,9136
16	rs8060494	78808972	2	1	190	174	1,092	0,7033	0,4017
17	rs4790372	2701606	2	1	163	169	0,9645	0,1084	0,7419
17	rs12453316	69027654	1	2	177	179	0,9888	0,01124	0,9156
19	rs997669	34996323	2	1	201	183	1,098	0,8438	0,3583
20	rs1555322	33312595	1	2	94	79	1,19	1,301	0,2541
B)									
CHR	SNP	BP	A1	F_A	F_U	A2	OR	CHISQ	P
2	rs2241685	1896290	1	0,0911 6	0,0928 3	2	0,9802	0,03733	0,8468
2	rs13395022	79793915	2	0,2088	0,2095	1	0,9961	0,002865	0,9573
2	rs2587695	120038047	1	0,4973	0,4922	2	1,021	0,1161	0,7333
2	rs2242073	208819551	2	0,1605	0,1568	1	1,029	0,1216	0,7273
2	rs1110998	217169458	1	0,3116	0,2928	2	1,093	1,886	0,1697
3	rs10510238	2876647	2	0,1293	0,1376	1	0,9304	0,6621	0,4158
3	rs9879164	54040611	2	0,4218	0,4359	1	0,9441	0,9084	0,3406
3	rs2084358	57457928	1	0,5184	0,4722	2	1,203	9,597	0,001949
3	rs10490808	59939739	2	0,4068	0,4266	1	0,9218	1,8	0,1797
3	rs10510850	60542142	1	0,1211	0,1116	2	1,097	1,001	0,3172
4	rs755403	6507714	2	0,3985	0,3973	1	1,005	0,007242	0,9322
4	rs10516182	7143981	2	0,2801	0,2954	1	0,9279	1,274	0,259
4	rs7697323	7801488	1	0,3782	0,38	2	0,9927	0,0142	0,9051
5	rs173754	65102081	1	0,4925	0,4915	2	1,004	0,004285	0,9478
5	rs258082	66166352	1	0,4619	0,4521	2	1,04	0,4342	0,5099
6	rs160666	2719051	2	0,2857	0,3025	1	0,9222	1,515	0,2183
6	rs2842643	41758714	2	0,2932	0,2909	1	1,011	0,02797	0,8672
6	rs3799977	44945334	2	0. 4306	0,4076	1	1,099	2,452	0,1174
6	rs8180608	89064414	2	0,4101	0,4441	1	0,8703	5,265	0,02176
6	rs1358601	91532294	1	0,3852	0,3846	2	1,003	0,002076	0,9637
6	rs6921403	154156020	2	0,1373	0,1405	1	0,9736	0,09408	0,7591
7	rs2237349	28536203	2	0,4109	0,4082	1	1,011	0,03276	0,8564
7	rs2002865	154132035	2	0,2075	0,217	1	0,9445	0,6065	0,4361

(continuación)

B)									
CHR	SNP	BP	A1	F_A	F_U	A2	OR	CHISQ	P
8	rs6991017	5508780	2	0,1891	0,1873	1	1,012	0,02315	0,8791
8	rs2248529	14657354	1	0,3604	0,363	2	0,9888	0,03305	0,8557
8	rs4961315	142110882	2	0,2959	0,2995	1	0,983	0,06846	0,7936
9	rs2418326	114759028	1	0,2534	0,252	2	1,007	0,01179	0,9135
9	rs2502731	128056111	2	0,3626	0,3508	1	1,053	0,6767	0,4107
14	rs10483393	31530235	1	0,2272	0,2203	2	1,041	0,3146	0,5749
15	rs2556560	42609135	2	0,419	0,4215	1	0,9899	0,02811	0,8668
16	rs8060494	78808972	2	0,3215	0,3228	1	0,9943	0,008131	0,9282
17	rs4790372	2701606	2	0,3014	0,3112	1	0,9546	0,5122	0,4742
17	rs12453316	69027654	1	0,3612	0,3662	2	0,9788	0,1193	0,7298
19	rs997669	34996323	2	0,4023	0,3876	1	1,064	1,025	0,3114
20	rs1555322	33312595	1	0,1279	0,1277	2	1,002	0,0004034	0,984
C)									
CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P
1	rs2281597	34132445	0	2	0	0	NA	NA	NA
1	rs642969	197590139	0	2	0	0	NA	NA	NA
2	rs2587695	120038287	1	2	320	294	1,088	1,101	0,2941
2	rs2242073	208702290	2	1	185	182	1,016	0,02452	0,8756
3	rs10510850	60542142	1	2	109	115	0,9478	0,1607	0,6885
3	rs17233461	125807474	2	1	305	322	0,9472	0,4609	0,4972
4	rs755403	6440543	2	1	296	278	1,065	0,5645	0,4525
4	rs3857174	7089831	2	1	202	217	0,9309	0,537	0,4637
4	rs7697323	7734317	1	2	269	278	0,9676	0,1481	0,7004
5	rs1457720	110998762	2	1	247	260	0,95	0,3333	0,5637
6	rs160666	2719051	2	1	248	262	0,9466	0,3843	0,5353
6	rs3799977	44945334	2	1	302	282	1,071	0,6849	0,4079
6	rs6921403	154105599	2	1	149	150	0,9933	0,003344	0,9539
8	rs6991017	5508780	2	1	193	191	1,01	0,01042	0,9187
9	rs2418326	116719295	1	2	236	210	1,124	1,516	0,2183
9	rs2416606	119862757	2	1	264	262	1,008	0,007605	0,9305
10	rs16928529	72652991	2	1	277	312	0,8878	2,08	0,1493
10	rs11594082	72969259	1	2	126	138	0,913	0,5455	0,4602
10	rs10786284	98125495	0	1	0	0	NA	NA	NA
10	rs515910	105956394	2	1	300	272	1,103	1,371	0,2417
11	rs3893215	17721406	0	2	0	0	NA	NA	NA
11	rs10830468	87604834	0	2	0	0	NA	NA	NA
12	rs4964805	102716954	0	2	0	0	NA	NA	NA
13	rs7995215	93206507	1	2	279	317	0,8801	2,423	0,1196
14	rs2239627	22705999	0	2	0	0	NA	NA	NA
14	rs10483286	24273582	0	2	0	0	NA	NA	NA
16	rs10514604	83003885	0	2	0	0	NA	NA	NA
17	rs2440129	6847295	0	2	0	0	NA	NA	NA

Análisis comparativo basado en segmentos de CNV

Para identificar nuevos loci genéticos que portan CNV que contribuyen potencialmente al TDAH, los presentes inventores aplicaron una estrategia de puntuación basada en segmentos que escanea el genoma respecto de SNP consecutivos con cambios más frecuentes en el número de copias en los casos en comparación con los controles,

como ya habían descrito previamente (Glessner y col. 2009; Wang, y col. 2007). El espacio genómico para estos SNP consecutivos delinea regiones de variación del número de copias comunes, o CNVR. En las cohortes del CHOP, los presentes inventores identificaron 10 CNVR que se observaron en múltiples casos, pero no en los controles, así como 2 CNVR que tenían una mayor frecuencia en los casos en comparación con los controles. Para asegurar la fiabilidad del presente procedimiento de detección, los presentes inventores validaron todas las CNVR usando PCR cuantitativa (qPCR), un procedimiento de uso común para la validación independiente de CNV (**figura 7**). Por lo tanto, los presentes inventores han aplicado una técnica de validación independiente en todas las CNV comunicadas para asegurar una confirmación positiva. Usando este enfoque, los presentes inventores han identificado y validado experimentalmente un total de 12 loci de CNV que se observaron o bien solo en casos de TDAH o sobre-representados en los casos de TDAH que se tomaron posteriormente para su replicación en cohortes de estudio independientes.

El análisis de replicación se llevó a cabo en cinco cohortes independientes, incluyendo sujetos con TDAH de IMAGE, PUWMa, IMAGE II, NIMH y Utah. Basándose en las 10 CNV específicas de caso de la cohorte de descubrimiento, 3 también fueron exclusivas para los casos de la cohorte de replicación, de manera destacable, GRM7, GRM8 y NEGRI, con valores de P resultantes combinados de $3{,}52x10^{-6}$ y $8{,}14x10^{-5}$, para GRM8 y GRM7, respectivamente (**tabla 3A**)**.** Un tercer gen GRM, GRM5, se observó en 10 casos de TDAH (10/3.506) y uno de control (1/13.327) con un resultado de $P=1{,}36x10^{-6}$ (**tabla 3A**). GRM1 se observó en 8 casos y 2 controles $P=1{,}05x10^{-4}$. Aunque no pudieron estimarse las oportunidades relativas (OR) para GRM7 y GRM8, ya que estas CNV estaban ausentes en el sujeto de control, las OR para GRM5 y GRM1 sumaron 38,12 y 15,24, respectivamente (**tabla 3**) lo que sugiere que la contribución de estas CNV al fenotipo de TDAH es potencialmente elevada. Por lo tanto, estos 4 genes de GRM se vieron afectados por las CNV que estaban asociadas con el TDAH y se replicaron con éxito en las cohortes de TDAH independientes (**tabla 3 y tabla 11**) mientras que también se observó que los otros loci de CNV estaban enriquecidos en los casos de TDAH, aunque con valores de P nominalmente significativos (**tabla 3b y tabla 11**).

Tabla 11. Descubrimiento, replicación y significación combinada de regiones de CNV.

CNVR	Valor de P de descubrimiento	Valor de P de replicación	Valor de P combinado	Valor de P de descubrimiento permutado	Valor de P de replicación permutado	Valor de P combinado permutado	Tipo	Gen
chr11: 88269449-88351661	**$1,53x10^{-3}$**	**$5,29x10^{-4}$**	**$1,36x10^{-6}$**	**0,025**	**0,001**	**0,002**	**Del**	***GRM5***
chr7: 126441593-126621501	**$7,74x10^{-3}$**	**$4,35x10^{-4}$**	**$3,52x10^{-6}$**	**0,013**	**<0,001**	**<0,001**	**Del**	***GRM8***
chr3:7183953-7197236	**$1,53x10^{-3}$**	**$4,53x10^{-2}$**	**$8,14x10^{-5}$**	**0,011**	**0,039**	**<0,001**	**Del**	***GRM7***
chr6: 146657076-146694047	**$4,42x10^{-3}$**	**$9,63x10^{-3}$**	**$1,05x10^{-4}$**	**0,006**	**<0,001**	**<0,001**	**Dup**	***GRM1***
chr1: 72317292-72328395	$1,53x10^{-3}$	$2,13x10^{-1}$	$3,91x10^{-4}$	0,036	0,213	0,011	Dup	*NEGR1*
chr7: 153495598-153564827	$1,53x10^{-3}$	$6,82x10^{-2}$	$4,08x10^{-4}$	<0,001	0,058	<0,001	Dup	*DPP6*
chr5: 65027976-65046520	$1,53x10^{-3}$	$1,17x10^{-1}$	$4,68x10^{-4}$	0,003	0,108	0,001	Del	*SGTBINLN*
chr1: 56053497-56064495	$3,91x10^{-2}$	$2,12x10^{-2}$	$1,54x10^{-3}$	0,035	0,024	<0,001	Del	*USP24*
chr19: 38427720-38444834	$4,42x10^{-3}$	$2,89x10^{-1}$	$4,95x10^{-3}$	0,002	0,262	0,007	Del	*SLC7A10*
chr3:1844168-1859889	$1,53x10^{-3}$	$4,12x10^{-1}$	$8,81x10^{-3}$	0,008	0,416	0,015	Del	*CNTN4*
chr2: 81419297-81446082	$3,91x10^{-2}$	$2,89x10^{-1}$	$3,83x10^{-2}$	0,046	0,294	0,032	Dup	*CTNNA2*
chr4: 113772340-113788584	$3,91x10^{-2}$	$2,89x10^{-1}$	$3,83x10^{-2}$	0,033	0,288	0,042	Dup	*LARP7*

En negrita se muestran los 4 principales loci más significativos.

La **figura 9** muestra las eliminaciones de CNV observadas en el locus de GRM5 (10 casos frente a 1 control), usando el UCSC Genome Browser (12) con la Compilación 36 del genoma humano. La validación experimental de las CNV de IMAGE, PUWMa, IMAGE II, NIMH y Utah, usando qPCR, junto con las gráficas en bruto de BAF y LRR se muestran en las **figuras 2-4**.

Tomados en conjunto, los presentes inventores han descubierto cuatro genes afectados directamente por CNVR en múltiples cohortes independientes que pertenecen a la familia del gen de receptor de glutamato metabotrópico (categoría de InerPro "GPCR, familia 3, receptor de glutamato metabotrópico"; P= $2{,}1 \times 10^{-9}$). También cabe destacar que se observó que tanto GRM2 como GRM6 se ven afectados por eliminaciones en casos de TDAH individuales en la cohorte de IMAGE II y estaban ausentes en los sujetos de control. Adicionalmente, se evaluó la significación de los genes de GRM, usando TDT en la misma cohorte y el mejor soporte se observó para GRM7, P=$8{,}35 \times 10^{-5}$ (**tabla 12**). Además, también se llevó a cabo un análisis para abordar las estadísticas de CNV basadas en familias basándose en el desequilibrio de transmisión y los eventos *de novo* en el subconjunto basado en familias de 311 familias del CHOP y 422 familias de IMAGE (**tablas 12 y 14**).

Tabla 12. GWAS de genotipo de TDAH de genes glutamatérgicos. La asociación de genotipo de SNP más significativa en cada una de las ocho regiones génicas de GRM. A) TDT de TDAH de Illumina 550k del CHOP B) Control:caso de TDAH de Illumina 550k C) IMAGE TDAH de Perlegen 600k.

A)										
CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P	Gen
11	rs4237549	88407924	2	1	31	61	0,5082	9,783	0,001762	GRM5
7	rs17864159	126444172	1	2	22	46	0,4783	8,471	0,003609	GRM8
6	rs3887555	34177040	1	2	208	161	1,292	5,986	0,01442	GRM4
7	rs6943762	86047914	2	1	69	99	0,697	5,357	0,02064	GRM3
3	rs7623055	7485891	1	2	151	193	0,7824	5,128	0,02354	GRM7
6	rs362839	146721428	2	1	125	161	0,7764	4,531	0,03328	GRM1
3	rs4687770	51730105	2	1	114	94	1,213	1,923	0,1655	GRM2
5	rs2078183	178357150	2	1	190	210	0,9048	1	0,3173	GRM6
B)										
CHR	SNP	BP	A1	F_A	F_U	A2	OR	CHISQ	P	Gen
3	rs7623055	7485891	1	0,3582	0,4129	2	0,793 6	15,48	8,35E-05	GRM 7
11	rs1354411	88016449	2	0,0364 3	0,0566	1	0,630 2	10,21	0,00139 6	GRM 5
7	rs2283100	12664329 3	2	0,2281	0,193	1	1,235	9,527	0,00202 4	GRM 8
6	rs1873250	34130718	2	0,2134	0,2455	1	0,833 8	7,062	0,00787 3	GRM 4
7	rs1095289 0	86193151	1	0,0275 3	0,0391 7	2	0,694 5	4,782	0,02877	GRM 3
5	rs2078183	178357150	2	0,4593	0,4897	1	0,8852	4,605	0,03189	GRM 6
6	rs1983635	14670736 5	2	0,316	0,2917	1	1,122	3,515	0,06081	GRM 1
3	rs4687592	51630896	1	0,03442	0,0404 1	2	0,8464	1,191	0,2752	GRM 2
C)										
CHR	SNP	BP	A1	A2	T	U	OR	CHISQ	P	Gen
6	rs12206652	34173960	2	1	265	216	1,227	4,992	0,02547	GRM4
11	rs160195	87932621	2	1	302	253	1,194	4,326	0,03753	GRM5
7	rs11563486	126621501	1	2	130	162	0,8025	3,507	0,06112	GRM8
3	rs11717471	7599469	2	1	238	280	0,85	3,405	0,06498	GRM7
6	rs2300620	146745874	2	1	160	133	1,203	2,488	0,1147	GRM1
7	rs1468413	86271589	1	2	190	162	1,173	2,227	0,1356	GRM3
5	rs7725272	178338994	2	1	289	261	1,107	1,425	0,2325	GRM6
3	rs6445959	51747387	2	1	169	153	1,105	0,795	0,3726	GRM2

En vista del hallazgo anterior, se barajó la hipótesis de que los genes que interactúan con los genes de receptor GRM podrían tener de manera colectiva más casos enriquecidos con CNV en comparación con controles sanos. Los presentes inventores identificaron 228 genes con 2 grados de relación con genes de GRM basándose en el interactoma humano combinado proporcionado por el programa informático Cytoscape (Shannon y col. 2003). Los presentes inventores evaluaron estos genes en 1.231 muestras de caso de TDAH y 4.105 muestras de control, todas las cuales se genotiparon en la misma plataforma en el CHOP, respecto de pruebas de enriquecimiento en las muestras de caso de TDAH ($P<0,05$). Los presentes inventores detectaron 67 genes que interactúan con el receptor GRM que estaban enriquecidos respecto de CNV en los casos, en comparación con 16 genes en los controles, lo que confirma un enriquecimiento de más de 3 veces en las CNV en esta red génica en los casos de TDAH ($P=4,38x10^{-10}$, **figura 10**).

Posteriormente, los presentes inventores agruparon la red de interacción génica de receptor GRM de segundo grado para definir módulos altamente interconectados de genes basándose en la topografía de la red y buscaron respecto de enriquecimiento de anotaciones de ontología génica (OG) en estos módulos. Tal como se muestra en la **figura 11**, Los GRM no forman un gran número de interacciones, pero de manera importante, sirven para coordinar módulos funcionales de otros conjuntos de genes. Por ejemplo, GRM1 porta duplicaciones significativamente enriquecidas en los casos de TDAH y sirve para coordinar módulos funcionales implicados en funciones constitutivas, tales como el metabolismo de carbohidratos, la fosforilación, la apoptosis y la unión a iones. GRM5 y GRM7 portan ambos eliminaciones enriquecidas significativamente en los casos y se agrupan dentro de un módulo funcional implicado en la transmisión sináptica y el corte y empalme alternativo. Específicamente, GRM5 sirve para coordinar el corte y empalme alternativo con la transmisión sináptica y otros procesos neuronales en la densidad post-sináptica, mientras que GRM7 coordina los módulos funcionales que integran procesos neurológicos y la actividad sináptica con las funciones constitutivas, tales como la organización del citoesqueleto y la apoptosis. GRM8 también porta eliminaciones enriquecidas significativamente en los casos y está contenido en sí mismo dentro de un módulo funcional que está implicado en la transmisión sináptica y la neurogénesis. Aunque no está significativamente enriquecido en los casos, GRM3 tiene duplicaciones que se observan con mayor frecuencia en los casos y sirve para coordinar las vías de ubiquitinación, la unión a ARN, el corte y empalme y el procesamiento y la migración neuronal, con los procesos neurológicos que incluyen la transmisión sináptica con los efectos en el comportamiento y la cognición.

Ejemplo II

Panel de SNP multiplexado para el diagnóstico del TDAH

Tal como se ha descrito anteriormente en el ejemplo I, se ha descubierto que varias alteraciones genéticas están asociadas con el fenotipo del TDAH. La información anterior en el presente documento puede aplicarse clínicamente a pacientes para diagnosticar una susceptibilidad aumentada a desarrollar TDAH y una intervención terapéutica. Una realización preferida de la invención comprende la aplicación clínica de la información descrita en el presente documento a un paciente. Pueden diseñarse composiciones diagnósticas, incluyendo micromatrices y procedimientos para identificar las alteraciones genéticas descritas en el presente documento en ácidos nucleicos de un paciente para evaluar la susceptibilidad a desarrollar TDAH. Esto puede producirse después que un paciente llegue a la clínica; se extrae sangre del paciente y usando los procedimientos diagnósticos descritos en el presente documento, un profesional médico puede detectar un SNP en las regiones genéticas listadas en las **tablas 13A y 13B** a continuación. El intervalo de edad típico para explorar a un paciente es de entre 1 y 12 años. La información obtenida de la muestra del paciente (por ejemplo, ácidos nucleicos), que puede amplificarse opcionalmente antes de su evaluación, se usará para diagnosticar a un paciente con una susceptibilidad aumentada o reducida a desarrollar TDAH. En el presente documento también se proporcionan kits para llevar a cabo el procedimiento diagnóstico de la invención. Dichos kits comprenden una micromatriz que comprende al menos uno de los SNP proporcionados en el presente documento y los reactivos necesarios para evaluar las muestras del paciente, como se ha descrito anteriormente. En una realización alternativa, se emplea un panel de SNP multiplexado y la muestra del paciente se evalúa respecto de la presencia o ausencia de todos los SNP listados en las tablas a continuación.

La **tabla 13A** proporciona todos los genes, intervalos genómicos físicos e intervalos de SNP de los marcadores de TDAH desvelados. La **tabla 13B** proporciona una "lista de SNP" que es un conjunto mínimo para una "matriz diagnóstica", tal como veracode, que es una tecnología aprobada por la FDA. De manera óptima, se incluyen los SNP flanqueantes a estas CNV así como los SNP intermedios, ya que pueden usarse todos para capturar la CNV.

Cuando se efectúa la prueba, se llevará a cabo el algoritmo de agrupamiento GenCall en el conjunto de la muestra y se revisarán los SNP poco agrupados o que se desvían significativamente del equilibrio de Hardy Weinberg. Las variaciones en el número de copias (CNV) se detectarán usando el presente algoritmo de variación del número de copias del modelo de Markov oculto de PennCNV (HMM). La normalización de la mediana, que está activada por defecto, se desactiva ya que los datos de las regiones diana en lugar de distribuidas por todo el genoma no proporciona información suficiente para la normalización de la mediana. Para la optimización, puede entrenarse el HMM con datos de buena calidad adecuados para los valores observados de la matriz (tasa de llamada < 98% y desviación estándar de la relación de log r < 0,3). PennCNV detectará regiones de SNP contiguos con una intensidad (relación de log R) tendiente a menos de 0 lo que indica eliminación o tendiente a más de 0, lo que indica duplicación. En caso de que no se observe dicha tendencia, no se realizará una llamada de CNV, lo que indica un

estado diploide normal. En tándem, los datos de genotipo se evalúan en una posición física continua para genotipos homocigotos, lo que indica eliminación o los genotipos AAB y ABB que indican duplicación, ponderada con la correlación positiva para la frecuencia del alelo de control menor.

La identidad de los genes implicados en el TDAH y los resultados del paciente indicarán qué variantes están presentes e identificarán a aquellos que poseen un riesgo alterado para desarrollar TDAH. La información proporcionada en el presente documento permite una intervención terapéutica en instantes más tempranos en la progresión de la enfermedad de lo que antes era posible. Como también se ha descrito anteriormente en el presente documento, la familia de receptor de GRM proporciona nuevas dianas para el desarrollo de nuevos agentes terapéuticos eficaces para el tratamiento del TDAH. En particular, sería deseable modelar la expresión de dichos genes en aquellos pacientes que son más propensos a desarrollar la enfermedad.

Tabla 13A. Panel de SNP multiplexado.

Gen	Intervalo (B36/hgl8)	SNP inicial	SNP final	Recuentos de del (casos: controles)	Recuentos Dup (casos: controles)
GRM5	chr11: 88269449-88351661	rs604179	rs669724	10:1	0:0
GRM8	chr7: 126525124-126536202	rs7794734	rs2237790	8:0	0:0
GRM7	chr3:7183953-7197236	rs1516302	rs6784317	6:0	0:0
GRM1	chr6: 146657076-146694047	rs12200797	rs362949	0:0	8:2
NEGR1	chr1:72317292-72328395	rs12033161	rs2821257	0:0	5:0
DPP6	chr7: 153495598-153564827	rs4389846	rs12703329	0:0	8:2
SGTB/NLN	chr5:65027976-65046520	rs10073281	rs972501	6:1	0:0
USP24	chr1:56053497-56064495	rs7527177	rs4333889	6:2	0:0
SLC7A10	chr19: 38427720-38444834	rs748680	rs4530278	7:5	0:0
CNTN4	chr3:1844168-1859889	rs10510218	rs7625240	7:6	0:0
CTNNA2	chr2:81419297-81446082	rs4430978	rs1595071	4:3	0:0
LARP7	chr4: 113772340-113788584	rs12054518	rs7690429	4:3	0:0
ACAT1	chr11: 107497467-107523485	rs3741049	rs11212525	0:0	1:0
ACCN1	chr17: 28364218-29507938	rs28933	rs11080254	0:0	3:1
ACTR2	chr2:65308405-65351891	rs268859	rs4671124	1:0	0:1
ADCY1	chr7:45580645-45729237	rs4724420	rs3735666	0:0	1:1
ADRBK1	chr11: 66790668-66810933	rs12274774	rs12274774	1:0	0:0
ALDOA	chr16: 29971972-29989236	rs9928448	rs2071390	3:8	2:6
APP	chr21: 26174731-26465003	rs3787620	rs462281	0:0	8:2
ARL15	chr5:53216370-53642160	rs271246	rs35947	1:1	2:0
ATXN7L3	chr17: 39624698-39631055	rs11652516	rs11652516	1:1	0:0

(continuación)

Gen	Intervalo (B36/hgl8)	SNP inicial	SNP final	Recuentos de del (casos: controles)	Recuentos Dup (casos: controles)
BDKRB2	chr14: 95740949-95780538	rs1959053	rs2069591	1:1	0:0
CA8	chr8:61263976-61356508	rs7460476	rs6998745	0:0	1:0
CACNA1B	chr9: 139892061-140136452	rs10867084	rs2606358	0:0	2:2
CACYBP	chr1: 173235193-173247786	rs6425310	rs11590474	1:0	0:0
CALM1	chr14: 89933125-89944363	rs2300497	rs1058903	1:2	0:0
CHRM3	chr1: 237616487-238116519	rs4130463	rs536477	0:0	2:1
CIC	chrl9: 47480656-47491789	rs3826706	rs3826706	1:1	0:0
CNP	chrl7: 37372284-37383280	rs8078650	rs11079028	1:2	0:0
CRHR1	chrl7: 41217448-41268973	rs4792886	rs17763104	1:0	0:0
DISC1	chr1: 229829183-230243641	rs2082552	rs980989	0:0	4:7
DYNLL1	chr12: 119392042-119420681	rs606443	rs580016	0:0	1:0
FPR1	chr19: 56940837-56946962	rs867228	rs4801891	0:0	1:1
GAPDH	chr12:6513917-6517797	rs1060619	rs1060619	0:2	1:1
GNA15	chr19:3087229-3114741	rs1465245	rs1637656	1:1	1:0
GNAI2	chr3:50238727-50271790	rs11716295	rs2236944	2:4	0:0
GNAO1	chr16: 54783648-54948650	rs16956168	rs3790116	0:0	1:1
GNAQ	chr9:79525010-79836012	rs6560613	rs1930543	1:0	0:0
GRIK1	chr21: 29831124-30234153	rs2832390	rs2255821	0:0	8:2
GRIK3	chr1:37039200-37272431	rs528137	rs563293	1:0	0:0
GRM1	chr6: 146390610-146800427	rs12196298	rs2942	0:0	7:2
GRM3	chr7:86111165-86332128	rs701332	rs6967992	0:0	1:0
GRM5	chr11: 87880625-88420888	rs308884	rs7931721	4:0	3:2
GRM7	chr3:6877926-7758217	rs6443074	rs17047886	4:0	0:0

(continuación)

Gen	Intervalo (B36/hgl8)	SNP inicial	SNP final	Recuentos de del (casos: controles)	Recuentos Dup (casos: controles)
GRM8	chr7: 125865892-126670546	rs13240504	rs13246388	3:0	1:1
GSN	chr9: 123003581-123134941	rs1590345	rs306772	1:0	1:0
HOMER1	chr5:78705541-78845456	rs3822568	rs11948804	0:0	1:0
HTR2A	chr13: 46305513-46368995	rs3803189	rs6312	0:0	1:0
MAPK1	chr22: 20443946-20551970	rs2298432	rs2876981	1:0	0:0
MTHFD1	chr14: 63924845-63996474	rs8011839	rs2281603	1:1	0:0
MX1	chr21: 41714311-41753008	rs457920	rs468811	0:0	7:2
NARG1	chr4: 140442125-140531385	rs13147688	rs2060685	1:0	0:0
NMI	chr2: 151835230-151854620	rs446791	rs2113509	0:0	1:0
PCBP3	chr21: 46092504-46186795	rs11701789	rs8133858	3:2	6:3
PDE1C	chr7:31759156-32305466	rs917749	rs215605	1:0	1:1
PPP2R1A	chr19: 57385045-57421483	rs13344984	rs7259175	0:0	1:0
PRPSAP1	chr17: 71818609-71861526	rs407281	rs8075628	1:0	1:1
PSMD11	chr17: 27795614-27832155	rs9889352	rs12162135	2:24	1:0
PSMD13	chr11:226976-242981	rs1045288	rs6598055	0:4	1:2
PXN	chr12: 119132639-119187892	rs10128770	rs1151836	0:0	1:0
QRICH2	chr17: 71781724-71815356	rs347675	rs346789	1:1	0:1
RANBP1	chr22: 18485023-18494706	rs2238798	rs2238798	2:3	0:9
RAP2A	chr13: 96884476-96918245	rs2389908	rs2389908	0:0	1:1
RCC1	chr1:28717331-28738194	rs10915206	rs10915206	0:0	1:0
RGS12	chr4:3285671-3411438	rs12643903	rs16844364	2:0	0:0
RIF1	chr2: 151974645-152040665	rs2444256	rs16830067	0:0	1:0
RUVBL2	chrl9: 54188967-54210994	rs12610125	rs7256033	1:0	0:3
RYR1	chr19: 43616179-43770044	rs919781	rs10408694	1:2	1:1

(continuación)

Gen	Intervalo (B36/hgl8)	SNP inicial	SNP final	Recuentos de del (casos: controles)	Recuentos Dup (casos: controles)
RYR2	chr1: 235272324-236063911	rs1881548	rs6429040	1:0	1:0
SDC3	chr1:31118568-31154067	rs2282440	rs10158813	1:0	0:1
SELE	chr1: 167958405-167969803	rs5368	rs5353	1:0	0:0
SERPINB9	chr6:2832502-2848506	rs318477	rs9503330	0:0	1:0
SETD4	chr21: 36328708-36358576	rs2835239	rs2835263	2:0	8:3
SHANK1	chr19: 55856895-55912007	rs4802724	rs3745530	0:0	1:0
SORD	chr15: 43102643-43154331	rs11636774	rs2854439	0:0	1:0
STRAP	chr12: 15926612-15947677	rs16911383	rs10846246	0:0	1:1
TK1	chr17: 73681775-73694726	rs1065769	rs11653181	2:0	0:2
TNIK	chr3: 172264363-172660546	rs12486818	rs10936688	1:0	0:0
VHL	chr3:10158318-10168746	rs1642742	rs1642742	0:0	1:0

Tabla 13B. Una lista de SNP disponibles para su implementación en una matriz diagnóstica.

Gen	SNP	Gen	SNP	Gen	SNP
GRM5	rs506811	CNP	rs11079028	HOMER1	rs12187625
GRM5	rs604179	GNAQ	rs11145589	SHANK1	rs12460584
GRM5	rs1954979	BDKRB2	rs11160322	CA8	rs12550354
GRM5	rs693008	CACYBP	rs11590474	PDE1C	rs12701140
GRM5	rs594561	GRIK3	rs1160752	CA8	rs12708003
GRM5	rs585423	ATXN7L3	rs11652516	PCBP3	rs13050871
GRM5	rs2047507	TK1	rs11653181	ARL15	rs13164221
GRM5	rs598758	PCBP3	rs11701789	SERPINB9	rs13196459
GRM5	rs547644	GNAI2	rs11716295	PDE1C	rs13238408
GRM5	rs316090	RYR2	rs11810113	PPP2R1A	rs13344984
GRM5	rs656544	SDC3	rs11810325	NMI	rs13383563
GRM5	rs641052	SDC3	rs12085929	RYR2	rs1361115
GRM5	rs11021670	MAPK1	rs12172554	CACNA1B	rs1378954
GRM5	cnvi0116228	ADRBK1	rs12274774	RYR2	rs1409052
GRM5	rs541046	PCBP3	rs12482750	ADCY1	rs1521470
GRM5	rs475872	RUVBL2	rs12610125	PPP2R1A	rs1560092
GRM5	rs694665	TNIK	rs12637875	STRAP	rs1564183
GRM5	rs573912	RGS12	rs12641989	GNA15	rs1637656
GRM5	rs563371	PDE1C	rs12701140	VHL	rs1642742
GRM5	rs533163	RGS12	rs13116176	RIF1	rs16823297
GRM5	rs5027960	RGS12	rs1320763	RIF1	rs16830067

(continuación)

Gen	SNP	Gen	SNP	Gen	SNP
GRM5	rs644170	PDE1C	rs13238408	ARL15	rs16882366
GRM5	rs675010	RYR2	rs1361115	ARL15	rs16882383
GRM5	cnvi0050221	RGS12	rs1406674	STRAP	rs16911383
GRM5	rs518167	RYR2	rs1409052	ARL15	rs169382
GRM5	rs591849	GNAQ	rs1436450	APP	rs17001492
GRM5	rs655683	GNA15	rs1637656	GNAO1	rs17281761
GRM5	rs597462	RGS12	rs16844364	HTR2A	rsl7288723
GRM5	rs539752	GSN	rs16910509	ARL15	rs17413044
GRM5	rs477399	ARL15	rs17267677	DISC1	rs17804007
GRM5	rs597303	CRHR1	rs173365	DISCI	rs7804163
GRM5	rs669724	CRHR1	rs17763104	APP	rs1783016
GRM5	rs677526	GNAQ	rsl7786782	PPP2R1A	rsl7835915
GRM8	rs10954144	PDE1C	rs1860790	APP	rs1787438
GRM8	rs7794734	MAPK1	rs1892846	ACCN1	rs1844737
GRM8	rs12375090	NARG1	rs2060685	PDE1C	rs1860790
GRM8	rs6975798	CNP	rs2070106	ACCN1	rs1985858
GRM8	rs1557644	ALDOA	rs2071390	SETD4	rs2018721
GRM8	rs12706778	RANBP1	rs2238798	FPR1	rs2070746
GRM8	rs2237790	SETD4	rs2255734	RYR1	rs2071085
GRM8	rs11563719	RGS12	rs2269497	DISCI	rs2082552
GRM7	rs9864350	QRICH2	rs2279053	ACCN1	rs2087633
GRM7	rs1516302	QRICH2	rs2279054	NMI	rs2113509
GRM7	rs1400163	TNIK	rs2292005	RIF1	rs2123465
GRM7	rs965170	CALM1	rs2300497	ACCN1	rs2130818
GRM7	rs11131064	CALM1	rs2300502	APP	rs214488
GRM7	rs10866078	PDE1C	rs2302450	PXN	rs2239206
GRM7	rs1400166	RYR1	rs2304150	GRIK1	rs2251388
GRM7	rs17235039	CRHR1	rs242939	SETD4	rs2255734
GRM7	rs10510351	CRHR1	rs242942	GRIK1	rs2268203
GRM7	rs11715681	RYR2	rs2485570	PSMD13	rs2272566
GRM7	rs6784317	RYR2	rs2490365	RYR1	rs2288888
GRM7	rs1963265	RYR2	rs2490371	PDE1C	rs2302450
GRM1	rs6570746	RYR2	rs2490372	RAP2A	rs2389908
GRM1	rs12200797	RYR2	rs2490373	RIF1	rs2432957
GRM1	rs1555084	RYR2	rs2490385	RIF1	rs2444256
GRM1	rs1009085	RYR2	rs2490389	RIF1	rs2444258
GRM1	rs362962	TK1	rs2661679	RIF1	rs2444263
GRM1	rs362949	ACTR2	rs268859	RIF1	rs2444273
GRM1	rs362835	SETD4	rs2835239	RYR2	rs2485570
NEGR1	rs2821267	SETD4	rs2835240	RYR2	rs2490365
NEGR1	rs12033161	SETD4	rs2835244	RYR2	rs2490371
NEGR1	rs988421	PCBP3	rs2839060	RYR2	rs2490372
NEGR1	rs2821255	TK1	rs2854701	RYR2	rs2490373
NEGR1	rs2821257	TK1	rs2854702	RYR2	rs2490385
NEGR1	rs10493493	MAPK1	rs2876981	RYR2	rs2490389
DPP6	rs3115157	GSN	rs306759	ARL15	rs25860
DPP6	rs4389846	GSN	rs306761	GNAO1	rs2587888

(continuación)

Gen	SNP	Gen	SNP	Gen	SNP
DPP6	rs4131646	GSN	rs306784	ARL15	rs26775
DPP6	rs7790046	RGS12	rs3088231	HTR2A	rs2770304
DPP6	rs7794112	CALM1	rs3213718	ARL15	rs277340
DPP6	rs12532924	QRICH2	rs346789	ARL15	rs28033
DPP6	rs4452722	PRPSAP1	rs346794	APP	rs2829984
DPP6	rs11975478	QRICH2	rs347675	APP	rs2829989
DPP6	rs11976255	PCBP3	rs373617	GRIK1	rs2832409
DPP6	rs10280963	RUVBL2	rs3764622	SETD4	rs2835239
DPP6	rs4507681	ACTR2	rs3771099	SETD4	rs2835240
DPP6	rs6955717	CRHR1	rs3785877	SETD4	rs2835244
DPP6	rs7811481	PCBP3	rs3788216	SETD4	rs2835261
DPP6	rs6945869	PCBP3	rs3788217	SETD4	rs2835263
DPP6	rs4074568	ARL15	rs3797251	MX1	rs2838037
DPP6	rs4074817	ARL15	rs3797252	PCBP3	rs2839060
DPP6	rs4726385	ARL15	rs3797255	SORD	rs2854439
DPP6	rs4380850	QRICH2	rs3803737	MX1	rs2898449
DPP6	rs12703323	PDE1C	rs3807618	ACCN1	rs2932925
DPP6	rs9791911	PCBP3	rs381083	SERPINB9	rs318477
DPP6	rs4397308	CIC	rs3826706	SERPINB9	rs318489
DPP6	rs10224365	PCBP3	rs3827268	GRIK1	rs363426
DPP6	rs10267846	PRPSAP1	rs385689	GRIK1	rs363452
DPP6	cnvi0096121	SELE	rs3917410	GRIK1	rs363456
DPP6	cnvi0096122	SELE	rs3917419	GRIK1	rs363463
DPP6	cnvi0096123	ARL15	rs445953	GRIK1	rs363464
DPP6	cnvi0096124	PRPSAP1	rs449114	GRIK1	rs363472
DPP6	cnvi0096125	RGS12	rs4690096	GRIK1	rs363478
DPP6	cnvi0096126	PDE1C	rs4723103	PCBP3	rs373617
DPP6	rs10952466	QRICH2	rs4789267	ACAT1	rs3741049
DPP6	cnvi0096127	CRHR1	rs4792886	SHANK1	rs3745530
DPP6	rs10952467	RUVBL2	rs4802533	SHANK1	rs3745532
DPP6	cnvi0096128	ARL15	rs4865809	RYR1	rs3745844
DPP6	cnvi0096129	BDKRB2	rs4905466	NMI	rs3771886
DPP6	rs10272007	SELE	rs5353	APP	rs3787625
DPP6	cnvi0096130	SELE	rs5361	ARL15	rs3797269
DPP6	cnvi0096131	SELE	rs5367	APP	rs380417
DPP6	cnvi0096132	SELE	rs5368	PDE1C	rs3807618
DPP6	rs4726389	GRIK3	rs550115	PCBP3	rs381083
DPP6	rs12674128	CALM1	rs5871	PCBP3	rs3827268
DPP6	rs10254647	CACYBP	rs6425310	APP	rs383700
DPP6	rs12703329	QRICH2	rs6501880	NMI	rs3856557
DPP6	rs12668613	ACTR2	rs6745303	SHANK1	rs3893128
SGTB/NLN	rs112314	NARG1	rs6848950	APP	rs396969
SGTB/NLN	rs10073281	GSN	rs7046030	PCBP3	rs406179
SGTB/NLN	rs17590975	GNAQ	rs7048503	PRPSAP1	rs407281
SGTB/NLN	rs972501	RUVBL2	rs7256033	PRPSAP1	rs419793
SGTB/NLN	rs252646	RYR1	rs7258075	PCBP3	rs431162
USP24	rs4367814	GRIK3	rs7517274	NMI	rs446791
USP24	rs7527177	SDC3	rs7529390	MX1	rs457920

(continuación)

Gen	SNP	Gen	SNP	Gen	SNP
USP24	rs10888939	PCBP3	rs760436	MX1	rs459498
USP24	rs4512692	RGS12	rs762864	MX1	rs466513
USP24	rs6588574	PDE1C	rs7787057	MX1	rs468105
USP24	rs4333889	MTHFD1	rs8003379	MX1	rs468440
USP24	rs10493190	MTHFD1	rs8011839	MX1	rs468646
SLC7A10	rs752503	BDKRB2	rs8013400	MX1	rs469083
SLC7A10	rs748680	QRICH2	rs8074821	PDE1C	rs4723103
SLC7A10	rs7256230	CNP	rs8078650	PSMD13	rs473151
SLC7A10	rs10500264	MAPK1	rs8136867	PXN	rs4767884
SLC7A10	rs4530278	MAPK1	rs8141815	PXN	rs4767886
SLC7A10	rs736289	TNIK	rs952209	SHANK1	rs4801850
CNTN4	rs9825865	MAPK1	rs9610417	FPR1	rs4801891
CNTN4	rs10510218	TK1	rs9897269	SHANK1	rs4802724
CNTN4	rs12488941	CA8	rs10092625	SHANK1	rs4802731
CNTN4	rs9860556	CA8	rs10108007	GSN	rs4837820
CNTN4	rs17044355	PXN	rs10128770	HTR2A	rs4942587
CNTN4	rs13322503	PDE1C	rs10226190	PSMD13	rs505404
CNTN4	rs6781373	GRM3	rs1024516	PSMD13	rs577259
CNTN4	rs7625240	PDE1C	rs10247918	PSMD13	rs577298
CNTN4	rs1387084	PDE1C	rs1035028	DYNLL1	rs580016
CTNNA2	rs6547363	PPP2R1A	rs10412613	DYNLL1	rs606443
CTNNA2	rs4430978	PPP2R1A	rs10420138	CA8	rs6471849
CTNNA2	rs10208516	PPP2R1A	rs10423794	ACCN1	rs6505377
CTNNA2	rs1595071	PSMD13	rs1045288	PPP2R1A	rs6509626
CTNNA2	rs2862499	CNR1	rs1049353	CA8	rs6984526
LARP7	rs1565010	ACCN1	rs10512455	CA8	rs6986917
LARP7	rs12054518	ACCN1	rs10512456	GRM3	rs701332
LARP7	rs1129065	SERPINB9	rs1052886	PSMD13	rs7128029
LARP7	rs4834296	GAPDH	rs1060619	ACCN1	rs7220709
LARP7	rs4409021	GSN	rs10739593	PPP2R1A	rs7251605
LARP7	cnvi0018439	GSN	rs10760165	FPR1	rs7253284
LARP7	rs4488992	GSN	rs10760167	HTR2A	rs731779
LARP7	rs6533635	CHRM3	rs10802802	CA8	rs7465573
LARP7	rs11722959	STRAP	rs10846246	PDE1C	rs7787057
LARP7	rs4555714	ACAT1	rs10890819	SORD	rs8043226
LARP7	rs11946967	PSMD13	rs10902112	PRPSAP1	rs8078771
LARP7	rs2352050	RCC1	rs10915206	SHANK1	rs8103945
LARP7	rs10031435	CHRM3	rs10925969	PPP2R1A	rs8106271
LARP7	rs7690429	CA8	rs10957123	MX1	rs8132871
LARP7	rs4834302	DISC1	rs11122319	PCBP3	rs8133858
RGS12	rs10027926	CACNA1B	rs11137372	DISC1	rs823161
PCBP3	rs1014446	ACAT1	rs11212525	DISC1	rs823162
SDC3	rs10158813	PXN	rs1151824	DISC1	rs823163
PDE1C	rs10226190	PXN	rs1151832	FPR1	rs867228
PDE1C	rs10247918	PXN	rs1151836	GSN	rs878691
PDE1C	rs1035028	NMI	rs11551174	SETD4	rs880221
ARL15	rs10513040	CHRM3	rs11578320	GNAO1	rs922445
CALM1	rs1058903	SORD	rs11636774	HTR2A	rs927544

(continuación)

Gen	SNP	Gen	SNP	Gen	SNP
RUVBL2	rs1062708	ADCY1	rs11766192	CACNA1B	rs9314645
TK1	rs1065769	CA8	rs11784742	SERPINB9	rs9392442
SDC3	rs10753239	RYR2	rs11810113	HTR2A	rs9534505
GSN	rs10760169	PPP2R1A	rs11881878	HTR2A	rs9534507
GSN	rs1078305	DISC1	rs12030S17	PSMD11	rs9889352
GNAQ	rs10869977	DISC1	rs12084975	ACCN1	rs9903823
CRHR1	rs110402	ADCY1	rs12112953	PCBP3	rs9975850
		CHRM3	rs12124903		

En estudios adicionales, los presentes inventores han extendido su análisis de red de GRM de las CNV a una red más exhaustiva de 335 genes que muestran significación en el TDAH. La red de mGluR original se generó a partir de 271 genes, incluyendo genes que interactúan en primer y segundo grado, según se define por el interactoma humano. El análisis actualizado proporciona una actualización más exhaustiva de la definición de la red de mGluR para capturar mejor las interacciones funcionales de los genes con los GRM.

Para generar estas dianas adicionales, se emplearon las siguientes bases de datos:

1. La base de conocimiento Ingenuity de interacciones biológicas y anotaciones funcionales. Cada interacción de Ingenuity se ha creado manualmente por científicos expertos con publicaciones de apoyo para las moléculas que interactúan específicas.
2. Los resultados de GWAS publicados comunicados en el navegador Human Genome Epidemiology (HuGE), que es una base de conocimientos integrada configurada para búsquedas de asociaciones genéticas y epidemiología genómica humana.
3. La revisión bibliográfica de los presentes inventores de PubMed para interacciones con genes de mGluR.
4. Variaciones raras procedentes de la secuenciación de pacientes de TDAH no halladas en el dominio público

Estas redes actualizadas se usaron para analizar 1292 casos de TDAH en comparación con 7449 controles neurológicamente normales. Se usaron controles negativos para validad los algoritmos de análisis de CNV en el contexto de expectativas biológicamente relevantes, tales como vías de desarrollo (HOX genexes), cáncer (vía del cáncer de pulmón) y la señalización neuronal no implicada en el TDAH (GABA). La red es altamente significativa, con CNV identificadas en un 17% de los casos generales y un 9% de los controles, con una OR de 2,1 y P= $6,5x10^{-17}$). La **tabla 13C** proporciona los nombres de 64 genes adicionales a los 271 presentados (un total de 335) que se identificaron usando el programa informático Ingenuity. El análisis actualizado proporciona interacciones funcionales de genes con GRM basándose en la base de conocimiento Ingenuity de interacciones biológicas y anotaciones funcionales. Cada interacción de Ingenuity se ha creado manualmente por científicos expertos con publicaciones de apoyo para las moléculas que interactúan específicas.

Los resultados del análisis de CNV extendido son los siguientes:

Conjunto de datos	n.º de genes	F_{casos}	$F_{controles}$	P	OR
Original + Ingenuity (GRM de TDAH)	335	0,17	0,09	$6,5x10^{-17}$	2,1
Hox (-ve de control)	38	0,003	0,002	0,43	1,2
Cáncer de pulmón (-ve de control)	421	0,17	0,19	0,99	0,8
Señalización de GABA (-ve de control)	121	0,03	0,07	0,99	0,3

Tabla 13C: a continuación en el presente documento se listan 335 dianas (incluyen las 271 dianas de la tabla 13A).

ACAT1	CACNA1A	ERBB2	GRM2	MIR1236	PCBP3	PSMD11	SET	TRPV1
ACAT2	CACNA1B	ERBB4	GRM3	MIR1245	PCDHA4	PSMD13	SETD4	TUBA1A
ACCN1	CACYBP	ESR1	GRM4	MIR1246	PCMT1	PSMD6	SF3B14	TUBA1B
ACCN2	CALB2	F2R	GRM5	MIR1252	PDCD5	PSME1	SHANK1	TUBA8
ACP1	CALM1	F2RL2	GRM6	MIR1260	PDE1B	PTK2B	SHANK3	TUBB
ACTB	CALM2	F2RL3	GRM7	MIR1262	PDE1C	PXN	SHBG	TUBG1
ACTN1	CAL M3	F3	GRM8	MIR1272	PDE6G	PYGL	SIAH1	TXN
ACTR2	CAMK1	FGF2	GSN	MIR1275	PGM1	PYGM	SIM1	TYMS
ADA	CAMK2B	FKBP3	HBXIP	MIR1276	PHKB	QRICH2	SLC1A2	UBE2I
ADCY1	CAMK4	FLNA	HOMER1	MIR1284	PHKG2	RALA	SLC2A1	UBE2M
ADD1	CASR	FOS	HOMER2	MIR1291	PIAS1	RANBP1	SLC6A3	UBQLN4
ADD2	CAV1	FPR1	HOMER3	MIR1305	PIAS2	RANBP9	SLC9A3R1	UCHL1
ADORA1	CAV3	FSCN1	HSP90AB1	MIR1322	PIAS4	RAP2A	SLC9A3R2	VHL
ADORA2A	CBX7	FURIN	HTR2A	MIR1323	PICK1	RCC1	SNCA	VIPR1
ADRA1B	CCNB1	FYN	HTT	MIR1324	PIK3CA	RCC2	SNRPB2	YWHAQ
ADRA2A	CDC42	GAPDH	IFNG	MIR555	PIK3R1	RGS11	SOCS6	ZAP70
ADRA2C	CHGB	GLP1R	IMPDH2	MIR559	PLA2G7	RGS12	SOCS7	
ADRB2	CHP	GLP2R	IQGAP2	MIR591	PLCB1	RGS2	SORD	
ADRBK1	CHRM2	GNA15	ITGB1	MIR610	PLCB3	RGS3	SRC	
ADRBK2	CHRM3	GNAI1	ITGB7	MIR637	PLCG2	RGS4	STAU1	
ALDOA	CIC	GNAI2	ITPR1	MIR641	PLD1	RGS9BP	STRAP	
ANXA2	CNP	GNAI3	KIAA0090	MIR769	POMC	RHOA	STX12	
APTX	CNR1	GNAO1	KIAA1683	MRPL14	PPIH	RIF1	SUMO1	
AQP1	COPB2	GNAQ	KLHL17	MRPS16	PPM1A	ROCK2	SYK	
ARHGAP24	CRHR1	GNB2L1	KPNA1	MTHFD1	PPM1B	RPA2	TBCA	
ARL15	CYCS	GNB5	KPNA3	MTNR1A	PPM1D	RPLP2	TBXA2R	
ARNT2	CYTH2	GOPC	LAMA4	MTNR1B	PPM1G	RPN2	TCP1	
ARRB1	CYTIP	GOT1	LRP2BP	MX1	PPP1CC	RPS14	TEAD3	
ARRB2	DCN	GP1BA	LRRC59	MYO6	PPP2R1A	RRM1	TFAM	
ATXN7L3	DHCR7	GPR26	LTA	NANS	PRDX1	RUVBL2	TGFB1	
BDKRB1	DLST	GRASP	LYAR	NCK1	PRKCA	RYR1	TGM2	
BDKRB2	DNM3	GRB7	LYN	NFKBIA	PRKCG	RYR2	TJP1	

(continuación)

BDNF	DRD2	GRIA1	MAGI2	NMI	PRLHR	S100A6	TK1	
BTBD2	DRD3	GRIK3	MAP4	NPY2R	PRMT1	SACS	TLR10	
BTG2	DSTN	GRIN1	MAPK1	NR3C1	PRPSAP1	SARS	TNIK	
C17orf44	DYNLL1	GRIP1	MAPK3	NUDC	PSAT1	SCTR	TPI1	
C1orf116	ECHS1	GRK4	MARK4	OPRD1	PSEN1	SDC3	TRAF2	
C7orf25	EFNB2	GRK5	MC4R	OPTN	PSMA1	SDCBP	TRMT112	
CA8	EGFR	GRK6	MIR1200	PAFAH 1B3	PSMC1	SELE	TRPC1	
ACAT1	EPHB1	GRM1	MIR1207	PCBP1	PSMD1	SERPINB9	TRPC3	

De acuerdo con la presente invención, se ha observado que un 10% de los pacientes con TDAH portan tipos específicos de mutaciones de genes que codifican receptores de glutamato metabotrópicos (mGluR). Estas mutaciones son biomarcadores sensibles y específicos para seleccionar y tratar el TDAH causado por vías de mGluR defectuosas. Además, los presentes inventores han identificado candidatos a fármacos que activan específicamente los mGluR, restaurando potencialmente la neurofisiología normal en pacientes de TDAH con mutaciones en la familia *GRM* de genes de mGluR. Véase la tabla 1.

Por ejemplo, los compuestos que pueden administrarse al implementar la prueba y tratar el paradigma descritos en el presente documento incluyen la familia de piracetam de agentes nootrópicos, como se describe en F. Gualtieri y col., Curr. Pharm. Des., 8: 125-38 (2002). Más preferentemente, el agente de tratamiento es una piroglutamida. Los detalles referentes a la preparación y formulación de las piroglutamidas que pueden usarse en la práctica de la presente invención se proporcionan en la Patente de los Estados Unidos 5.102.882 de Kimura y col. Un agente particularmente preferido para el tratamiento del TDAH en pacientes que se determina que tienen uno o más SNP que indican la presencia de una variación del número de copias asociada con el TDAH, tal como se expone en la tabla 13, es monohidrato de (+)-5-oxo-D-prolinepiperidinamida (NS-105).

Ejemplo III

Los genes de mGluR que contienen CNV anteriormente identificados implicados en la patogénesis del TDAH también proporcionan nuevas dianas para el desarrollo de agentes terapéuticos eficaces para el tratamiento del TDAH. Con ese fin, pueden llevarse a cabo procedimientos para la selección de fármacos candidatos (agente o compuesto) que modulen las interacciones de proteína de mGluR y la patología asociada basándose en la información proporcionada en el presente documento. Los fármacos candidatos representativos incluyen ácidos nucleicos, polipéptidos, compuestos de molécula pequeña y peptidomiméticos.

En algunos casos, pueden seleccionarse agentes genéticos poniendo en contacto la célula de levadura con una construcción de ácido nucleico que codifica un gen. Por ejemplo, se pueden explorar bibliotecas de ADNc que expresan una serie de genes, para identificar otros genes que modulan dichas interacciones. Por ejemplo, los fármacos identificados pueden modular la señalización neuronal asociada con el glutamato, la localización de proteínas subcelulares y/o la morfología o viabilidad de células neuronales. Por consiguiente, independientemente del mecanismo de acción exacto, se espera que los fármacos identificados mediante los procedimientos de exploración descritos en el presente documento proporcionen un beneficio terapéutico a los pacientes que padecen TDAH.

Los procedimientos de exploración adecuados pueden emplear una variedad de tipos celulares neuronales que pueden obtenerse de la ATCC. Los fármacos candidatos pueden seleccionarse de grandes bibliotecas de compuestos sintéticos o naturales. Un ejemplo es una biblioteca aprobada por la FDA de compuestos que pueden usarse por seres humanos. Además, se encuentran disponibles en el mercado bibliotecas de compuestos de una serie de compañías, incluyendo, pero sin limitación, Maybridge Chemical Co. (Trevillet, Cornualles, R. U.), Comgenex (Princeton, NJ), Microsource (New Milford, CT), Aldrich (Milwaukee, WI), AKos Consulting and Solutions GmbH (Basilea, Suiza), Ambinter (París, Francia), Asinex (Moscú, Rusia), Aurora (Graz, Austria), BioFocus DPI, Suiza, Bionet (Camelford, R. U.), ChemBridge, (San Diego, CA), ChemDiv, (San Diego, CA), Chemical Block Lt, (Moscú, Rusia), ChemStar (Moscú, Rusia), Exclusive Chemistry, Ltd (Obninsk, Rusia), Enamine (Kiev, Ucrania), Evotec (Hamburgo, Alemania), Indofme (Hillsborough, NJ), Interbioscreen (Moscú, Rusia), Interchim (Montlucon, Francia), Life Chemicals, Inc. (Orange, CT), Microchemistry Ltd. (Moscú, Rusia), Otava, (Toronto, ON), PharmEx Ltd. (Moscú, Rusia), Princeton Biomolecular (Monmouth Junction, NJ), Scientific Exchange (Center Ossipee, NH), Specs (Delft, Países Bajos), TimTec (Newark, DE), Toronto Research Corp. (North York, ON), UkrOrgSynthesis (Kiev, Ucrania), Vitas-M, (Moscú, Rusia), Zelinsky Institute, (Moscú, Rusia) y Bicoll (Shanghái, China). Las bibliotecas combinatorias se encuentran disponibles y pueden prepararse. Se encuentran disponibles comercialmente bibliotecas de compuestos naturales en forma de extractos bacterianos, fúngicos, vegetales y animales o pueden prepararse fácilmente mediante procedimientos bien conocidos en la técnica. Se propone ensayar compuestos aislados de fuentes naturales, tales como animales, bacterias, hongos, fuentes vegetales, incluyendo hoja y corteza y muestras marinas como candidatos para la presencia de agentes farmacéuticos potencialmente útiles. Se entenderá que los agentes farmacéuticos a explorar también podrían proceder o sintetizarse a partir de composiciones químicas o compuestos elaborados por el hombre.

Por ejemplo, las células neuronales pueden incubarse en presencia y ausencia de un compuesto de ensayo, tal como piroglutamidas (véase, por ejemplo, la Patente de los Estados Unidos 5.102.882) y otros miembros de la familia de piracetam de agentes nootrópicos, tras lo cual se evalúa el efecto del compuesto en la señalización del glutamato. Los agentes identificados de este modo podrían ensayarse en modelos animales completos de TDAH para evaluar la eficacia *in vivo*.

Los agentes identificados usando los ensayos de exploración descritos en el presente documento también se encuentran abarcados por la presente invención.

Discusión

En la actualidad, hay una escasez notable de estudios de asociación por todo el genoma en el TDAH y ningún estudio ha comunicado CNV que estén asociados significativamente con el TDAH. Como tal, el presente estudio representa en primer escaneo a gran escala, sin sesgo en dos etapas de CNV en el TDAH. Aunque los presentes inventores habían comunicado previamente eliminación de *GRM5* en una sola familia con TDAH con tres niños afectados y una familia con eliminación de *GRM7*, junto con otros 57 genes no receptores de GRM, la mayoría de los cuales eran acontecimientos individuales [13], se ha demostrado por primera vez que los genes de la familia de receptor de glutamato metabotrópico (*GRM5, GRM7, GRM8* y *GRM1)* se ven afectados por las CNV que se asocian significativamente con el TDAH y se observa que están replicados en múltiples conjuntos de datos de control de caso independientes.

Los receptores de glutamato metabotrópicos (GRM o mGluR) son una clase de receptores acoplados a proteína G que poseen siete regiones transmembrana implicadas en la modulación de la transmisión sináptica excitatoria en el sistema nervioso [14]. Hay siete grupos de receptores, basándose en la homología de secuencia, los supuestos mecanismos de transducción de señales y las propiedades farmacológicas [15]. *GRM5* y *GRM1* son miembros del Grupo I expresado particularmente en los ganglios basales y el cerebelo [16], áreas relevantes del cerebro para el TDAH. Se ha demostrado que estos receptores activan la fosfolipasa C y se ha postulado que pueden tener un papel en la adicción, la ansiedad y los trastornos del comportamiento [17]. *GRM7* y *GRM8* son miembros del Grupo III que está relacionado con la inhibición de la cascada de AMP cíclico. *GRM7* se ha relacionado con la ansiedad [18] y es el más conservado de todos los subtipos de mGluR entre diferentes especies de mamífero [19].

Las evidencias de la implicación glutaminérgica en el TDAH surge de diversos campos. Aunque los estudios de asociación que investigan variantes en los receptores y transportadores glutamatérgicos han arrojado resultados mixtos [20-23], un estudio de asociación por todo el genoma que investiga la respuesta al metilfenidato en niños con TDAH detectó pruebas nominales de asociación de varios SNP, incluyendo SNP en *GRM7* (rs3792452) [24]. Según se ha publicado en un estudio de relación genética, *GRIN2A* está asociado con el TDAH [20] y *GRIN2B* se asoció mediante TDT [25]. Los estudios de espectroscopía de resonancia magnética han demostrado un mayor tono glutamatérgico en las regiones cerebrales frontales y estriadas de sujetos con TDAH[26-28] que se normalizan con estimulantes y atomoxetina [29]. El ratón *SLC6A3-KO* (DAT-KO), un modelo animal para TDAH, sigue siendo sensible al metilfenidato a pesar de la ausencia de transportador de dopamina[30] y puede aumentarse la hiperactividad en estos ratones mediante bloqueadores del receptor de NMDA y suprimirse mediante fármacos que aumentan la transmisión glutamatérgica[31]. Se ha comunicado un aumento de la expresión de *SLC6A3* y *DRD4* en el cerebro medio de ratas, en las que se hallaron aumentos en el transportador de glutamato en el estriado [32] lo que sugiere que las reducciones en la dopamina pueden alterar la señalización del glutamato. Además, la alteración del gen de subunidad de glutamato *(GRIN2A)* aumentó el metabolismo de DA y serotonina en el córtex frontal y el estriado de ratones y aumentó la actividad locomotora que se redujo mediante antagonistas del receptor de dopamina o serotonina [33]. Además, se ha observado expresión desregulada de genes en las vías glutamatérgicas en el modelo de rata expuesta a SHR [34-37] y a PCB para TDAH [36]. Se han comunicado niveles aumentados de glutamato en el neurometabolismo de cerebros con TDAH, lo que sugiere que la transmisión alterada de glutamato puede ser importante en el TDAH[28]. Aunque los receptores de glutamato que se asociaron con el TDAH en el presente estudio se eliminaron en tres casos y se duplicaron en un caso, las perturbaciones resultantes en la señalización de glutamato en los casos eliminados podría promover el TDAH mediante un bucle de retroalimentación que libera glutamato adicional en un intento de compensar la disparidad de las señales de neurotransmisión enviadas y recibidas.

Aparte de la familia de genes de GRM, los presentes inventores han detectado la asociación de otros ocho loci con el TDAH, cuatro de los cuales tienen impacto directo sobre genes (tabla 3B). Entre estos, hay genes con una biología intrigante respecto al TDAH. *DPP6* se ha asociado previamente con la esclerosis lateral amiotrófica (ELA) en estudios de asociación por todo el genoma [38,39], y se han comunicado CNV que afectan a *DPP6* en relación con el autismo [40]. *DPP6* y *CTNNA2* (aunque la presente asociación no tiene impacto directo sobre *CTNNA2)* se han visto implicados por GWAS de genotipado de SNP de TDAH [9]. *NLN* es un candidato interesante responsable de la inactivación metabólica de péptidos neuronales, tales como el neuropéptido Y (NPY) que se ha relacionado anteriormente con el TDAH [45,46]. Se ha demostrado que *SLC7A10* desempeña un papel en la modulación de la transmisión glutamatérgica mediante la movilización de D-serina en la sinapsis glutamatérgica. *LARP7* es importante para la integridad de snRNP, un complejo proteico responsable del corte y empalme postranscripcional. *NEGR1* codifica una molécula de adhesión de células neurales y un factor promotor del crecimiento trans-neural en el brote regenerativo de axones y el crecimiento neuronal en el cerebro de mamíferos. De forma interesante, este gen neuronal se asoció recientemente con la obesidad [41].

En las cohortes de descubrimiento del CHOP, la familia 230 se ve afectada tanto por la eliminación de *GRM5* heredada de la madre como por la duplicación de *NEGR1* heredada del padre en los tres casos de TDAH en la familia. A pesar de los niveles de CI superiores, estos 3 niños tienen una discapacidad grave. Estas fueron las únicas regiones de CNV observadas en los tres casos familiares y no observadas en los controles. La evaluación de la madre usando una Escala de autoinforme de TDAH en adultos[42] indicó probabilidad de TDAH.

Hay ocho CNVR presentadas que alteran directamente el gen respectivo en estas regiones (incluyendo *GRM5,*

GRM7, GRM8, *GRM1, NEGR1, DPP6, SGTB/NLN* y *LARP7)* mientras que los demás se anotan con los más próximos (tabla 3A y 3B). Además, las CNV de *GRM8, GRM1, SGTB/NLN,* y *LARP7* son exónicas. Se llevarán a cabo estudios funcionales adicionales para caracterizar por completo la función de los genes asociados en relación con los fenotipos de TDAH. Por lo tanto, el enfoque sin sesgos de los presentes inventores para evaluar el genoma completo en múltiples cohortes independientes ha revelado CNV en nuevos genes que no se habían estudiado con anterioridad respecto de cualquier impacto biológico o fisiológico en el cerebro en el TDAH y quedan a la espera de una caracterización adicional.

Dada la importancia de los cuatro genes de receptor GRM comunicados en la patogénesis del TDAH y de la rareza de las CNV en cada uno de los loci, se eligió evaluar los genes que interactúan con el receptor de GRM respecto de su frecuencia de observación de CNV en los casos y los controles. Esto permite incluir loci marginalmente significativos, dado el conocimiento previo de una fuerte asociación de la familia génica de receptor de GRM. Normalmente las CNV son muy escasas (<1%) en un locus dado pero sus asociaciones proporcionan una correlación directa con la patología más fuerte que las variantes comunes, según se evidencia por sus impresionantes tamaños del efecto (véanse las OR en la tabla 3). Basándose únicamente en loci individualmente significativos, un 3,66% de los casos de TDAH están fuertemente relacionados con las CNV descubiertas. Si se extienden las observaciones de la familia de GRM de confianza a redes génicas de genes que interactúan/señalizan con receptores de GRM se obtiene un 9,94% de casos de TDAH con caracterización genética de su enfermedad tras ajustar respecto de la frecuencia del control (es decir, impacto neto en los casos). Los nodos de soporte principales de esta red incluyen *TNIK* [48], *GNAQ* [49], y *CALM1* [50] (**figura 11**), asociados previamente con la esquizofrenia y la epilepsia. De forma interesante, el gen de la red de receptor GRM, *GRIK1,* también se ha asociado con los síntomas de hiperactividad/impulsividad del TDAH [8].

Tomados en conjunto, el presente análisis de CNV y de enriquecimiento funcional de la red de interacción génica del receptor GRM sugiere que los GRM no forman un gran número de interacciones, pero sirven para coordinar módulos funcionales de otros conjuntos de genes. De manera alentadora, algunos de estos módulos son importantes en el proceso de la transmisión sináptica, la neurogénesis y otros procesos neuronales que se cree que son defectuosos en el TDAH. Por lo tanto, mediante un análisis de redes de CNV que tienen impacto en el TDAH, los presentes inventores han identificado módulos que son importantes en procesos tales como la unión a ARN, su procesamiento y el corte y empalme alternativo, que se ha demostrado que tienen influencia en la actividad sináptica específica del cerebro[51,52]. Además, los presentes inventores han identificado módulos funcionales implicados en la ubiquitinación, un proceso que se ha vinculado previamente con el autismo[1], que comparte ciertos rasgos fenotípicos con el TDAH. Además, la conectividad cerebral funcional anormal es un factor potencial para los trastornos del desarrollo cerebral asociados con la disfunción cognitiva, incluyendo el TDAH[53,54]. Por lo tanto, el impacto de las CNVR entre la familia de receptores de GRM y en particular en *GRM5* y *GRM7,* puede ser importante para la etiología molecular subyacente del TDAH.

En conclusión, usando una estrategia de asociación por todo el genoma en dos etapas para la detección de alta resolución de CNV, los presentes inventores han identificado 12 loci que demuestran enriquecimiento de CNV en los casos de TDAH en comparación con los controles y han replicado con éxito 4 de estos usando conjuntos de datos independientes de casos de TDAH y controles sanos genotipados en tres plataformas diferentes emparejadas para casos y controles. Cuatro de los genes afectados pertenecen a la familia del receptor de glutamato metabotrópico. La red de más de 200 genes que interactúan con los receptores de glutamato se ve afectada de manera colectiva por las CNV y captura la diversidad genética de aproximadamente un 10% de todos los casos de TDAH. Además, esta red de genes que interactúan con los receptores de glutamato metabotrópicos define un conjunto de módulos funcionales con funciones neuronales significativas, cuyos defectos se cree que subyacen al TDAH y otros trastornos del neurodesarrollo. Por lo tanto, el enriquecimiento de genes en este sistema molecular respecto de CNV asociadas con el TDAH sugiere nuevos mecanismos de susceptibilidad para la enfermedad y fomentará la evaluación de variaciones adicionales, incluyendo variaciones estructurales y cambios de una sola base en genes candidatos dentro de estas redes moleculares. Los presentes resultados sugieren ensayos de expresión y funcionales para evaluar los efectos biológicos de las CNV en estos genes candidatos.

Tabla 14. Desequilibrio de transmisión basado en familias de CNV de TDAH y pruebas estadísticas *de novo*.

A) Eliminaciones Illumina del CHOP enriquecidas por herencia							
CNVR	Recuento de SNP	TDTDel	HerDel	Del *de novo*	ParDel NoHer	Gen	Distancia
chr18:74258734-74260996	3	0,001953	9	0	0	*SALL3*	580267
chr7:120092385-120099982	3	0,001953	9	0	0	*KCND2*	0
chr4:92499956-92502794	8	0,001953	9	0	0	*KIAA1680*	0
chr11:69755529-69759313	12	0,007813	7	0	0	*FADD*	24395
chr4:42400885-42403451	15	0,007813	7	0	0	*ATP8A1*	47238
chr5:104463047-104518786	17	0,007813	7	0	0	*NR_000039*	0
chr13:69637654-69666685	18	0,015625	6	0	0	*NR_002717*	25969
chr3:195971510-195982215	5	0,03125	5	1	0	*FAM43A*	80455
chr19:44369918-44376749	3	0,03125	5	1	0	*LOC342897*	2695
chr1:2349841-2356176	4	0,03125	5	1	0	*PEX10*	15971
chr21:45777720-45782727	3	0,03125	5	0	0	*SLC19A1*	0
chr10:67748487-67785209	30	0,03125	5	0	0	*CTNNA3*	0
B) Duplicaciones Illumina del CHOP enriquecidas por herencia							
CNVR	Recuento de SNP	TDTDup	DupHer	Dup *de novo*	Par No Dup Her	Gen	Distancia
chr20:59015708-59022667	4	0,007813	7	0	0	*CDH4*	238287
chr12:72808323-72832667	5	0,015625	6	0	0	*BC061638*	0
chr6:73021641-73023171	3	0,03125	5	0	0	*RIMS1*	0
chr17:74089903-74106726	9	0,03125	5	0	0	*DNAHL1*	10904
chr1:9243828-9310031	22	0,03125	5	0	0	*H6PD, SPSB1*	0
C) Eliminaciones Illumina del CHOP enriquecidas para *de novo*							
CNVR	Recuento de SNP	TDTDel *de novo*	HerDel	Del *de novo*	ParDel NoHer	Gen	Distancia
chr16:87694595-87778383	16	3,02E-05	32	2	21	*AX748415, CDH15, LO C197322*	0
chr18:65358832-65367619	18	3,02E-05	33	2	21	*DOK6*	0
chr12:55902280-55923860	3	0,000367	9	3	19	*NDUFA4L2, NXPH4, SH MT2, STAC3*	0
chrl7:71112486-71120734	4	0,001848	12	3	16	*KIAA1783*	0
chr22:38384374-38403731	8	0,018158	4	4	13	*CACNA1I*	0
chr19:15992679-15997923	2	0,025875	15	6	15	*LOC126536*	0

(continuación)

D) Duplicaciones Illumina del CHOP enriquecidas para *de novo*							
CNVR	Recuento de SNP	TDTDup *de novo*	HerDup	Dup *de novo*	ParDup NoHer	Gen	Distancia
chr19:59423491-59428132	12	4,85E-09	74	3	38	*ULRB3,UR-3*	0
chr8:145217675-145247517	4	3,05E-05	19	0	15	*CYC1, MAF1, SHARPIN, hSIPL1A*	0
chr18:64897188-64906488	48	0,000122	9	0	13	*CCDC102B*	23782
chr14:104225150 -104339273	35	0,00293	7	1	11	*ADSS, ADSSL1, AKT1, SIVA1*	0
chr9:138606913-138647195	17	0,005371	10	1	10	*AF161442*	15688
chr16:650256-2028586	41	0,015625	8	0	6	*Muchos*	0
chr20:61642713-61668792	11	0,03125	4	1	7	*C20orf195,PRIC285, SRMS*	0
chr16:87399730-87430019	22	0,03125	7	1	7	*APRT, CDT1, FLJ00319, GALNS*	0
chr16:3553005-3590430	20	0,03125	8	0	5	*BTBD12,NLRC3*	0
chr22:17257787-17355587	60	0,03125	3	0	5	*DGCR6,KIAA1647, PRODH*	0
E) Eliminaciones Perlegen IMAGE enriquecidas por herencia							
CNVR	Recuento de SNP	TDTDel	HerDel	Del *de novo*	ParDel NoHer	Gen	Distancia
chr2:180271795-180274556	5	0,003204	2	1	13	*ZNF533*	0
chr14:79919894-79924934	5	0,03125	1	0	7	*BC039670*	0
chr7:19828746-19840916	7	0,041656	4	0	11	*MGC42090*	49005
F) Duplicaciones Perlegen IMAGE enriquecidas por herencia							
CNVR	Recuento de SNP	TDTDup	HerDup	Dup *de novo*	ParDup NoHer	Gen	Distancia
chr22:17361563-17369020	3	0,015625	6	0	0	*CR623368, KIAA1647*	0
chr15:30088094-30090949	3	0,03125	5	1	0	*CHRNA7*	19069
chr7:71664963-71712086	5	0,03125	5	0	0	*MGC87315*	0
G) Eliminaciones Perlegen IMAGE enriquecidas para *de novo*							
CNVR	Recuento de SNP	TDTDel *de novo*	HerDel	Del *de novo*	ParDel NoHer	Gen	Distancia
chr2:180271795-180274923	6	0,000854	2	1	13	*ZNF533*	0
chr10:85445139-85446804	7	0,03125	5	1	7	*GHITM*	442361

(continuación)

H) Duplicaciones Perlegen IMAGE enriquecidas para *de novo*							
CNVR	Recuento de SNP	TDTDup *de novo*	HerDup	Dup *de novo*	ParDup NoHer	Gen	Distancia
chr12:31276361-31285014	9	6,87E-05	15	1	17	*OVOS2*	26006
chr10:47089854-47154881	31	6,87E-05	11	1	17	*AK057316*	0
chr7:140018-162903	13	0,005371	10	1	10	*AL137655*	23529
chr8:2437197-2492653	23	0,03125	4	1	7	*BC045738*	0
chr6:168234697-168295618	13	0,043945	5	2	8	*FLJ00181*	9639

ES 2 686 300 T3

Tabla 15. Desequilibrio de transmisión basado en familias de CNV de TDAH y pruebas estadísticas *de novo*.

CNVR (hg 18/B36/ Mar2006)	Tipo	TDTDel	TDTDup	TDTDel *de novo*	TDTDup *de novo*	HerDel	Del *de novo*	ParDel NoHer	HerDup	Dup *de novo*	ParDup NoHer
chr7: 126441593-126621501	Del	1	1	1	1	0	0	0	0	0	0
chr11: 88269449-88351661	Del	0,125	1	0	1	3	0	0	0	0	0
chr3:7183953-7197236	Del	0,25	1	1	1	2	0	0	0	0	0
chr6: 146657076-146694047	Dup	1	1	1	1	0	0	0	0	0	0
chr7: 153495598-153564827	Dup	0,205	1	0,016	1	4	0	6	0	0	0
chr5: 65027976-65046520	Del	1	0,5	1	1	0	0	0	1	0	0
chr1:56053497-56064495	Del	1	1	1	1	0	0	0	0	0	0
chr1: 72317292-72328395	Dup	1	1	1	1	0	0	0	0	0	0
chr19: 38427720-38444834	Del	0,183	1	0,004	1	6	0	8	0	0	0
chr3:1844168-1859889	Del	0,063	1	0	1	4	0	0	0	0	0
chr2: 81419297-81446082	Dup	1	0,5	1	1	0	0	0	1	0	0
chr4: 113772340-113788584	Dup	0,375	1	0,5	1	2	0	1	0	0	0

Tabla 16. Contribuciones de fuente de muestra a loci de CNV con impacto.

CNVR	Casos del CHOP	Controles del CHOP	Casos NIMH	Casos Utah	Casos IMAGE	Control por psoriasis	Control por depresión	Casos PUWMa	Progenitores PUWMa	Casos IMAGE II	Controles IMAGE II	Controles de Illumina 1M SAGE	Controles de progenitores Affy 5.0 AGRE	Tipo	Gen
chr11:88269449-88351661	4	0	0	0	5	0	0	1	1	0	0	0	0	Del	*GRM5*
chr7:126441593-126621501	3	0	0	0	3	0	0	2	0	0	0	0	0	Del	*GRM8*
chr3:7183953-7197236	4	0	0	0	2	0	0	0	0	0	0	0	0	Del	*GRM7*
chr6:146657076-146694047	5	2	1	1	0	0	0	0	0	1	0	0	0	Dup	*GRM1*
chr1:72317292-72328395	4	0	0	0	0	0	0	1	0	0	0	0	0	Dup	*NEGR1*
chr7:153495598-153564827	5	0	1	0	0	0	0	2	0	0	1	0	1	Dup	*DPP6*
chr5:65027976-65046520	4	0	0	0	1	0	0	0	0	1	1	0	0	Del	*SGTBINLN*
chr1:56053497-56064495	2	0	0	0	3	0	0	0	0	1	0	0	2	Del	*USP24*
chr19:38427720-38444834	5	2	0	0	1	0	0	1	3	0	0	0	0	Del	*SLC7A10*
chr3:1844168-1859889	4	0	0	0	0	0	0	2	2 (her)	1	1	4	1	Del	*CNTN4*
chr2:81419297-81446082	2	0	0	0	1	0	3	0	0	1	0	0	0	Dup	*CTNNA2*
chr4:113772340-113788584	2	0	0	0	1	0	0	1	1	0	0	1	1	Dup	*LARP7*

Tabla 17. Límites de las CNV individuales en la tabla 1A y 1B.

CNVR	Gen	Tipo	ID de muestra	Región llamada en la muestra	Distancia exónica*	Ciclo de validación de muestra
chr11:88269449-88351661	*GRM5*	Del	230-3	chr11:88269449-88351661	5.858	Y
chr11:88269449-88351661	*GRM5*	Del	230-4	chr11:88269449-88351661	5.858	Y
chr11:88269449-88351661	*GRM5*	Del	230-5	chr11:88269449-88351661	5.858	Y
chr11:88269449-88351661	*GRM5*	Del	497	chr11:83876556-91038751	0	Y
chr11:88269449-88351661	*GRM5*	Del	16794	chr11:87996654-88837360	0	Y
chr11:88269449-88351661	*GRM5*	Del	13304	chr11:88109331-88827923	0	Y
chr11:88269449-88351661	*GRM5*	Del	13270	chr11:88115425-88481107	0	Y
chr11:88269449-88351661	*GRM5*	Del	13761	chr11:88305340-88385387	0	Y
chr11:88269449-88351661	*GRM5*	Del	17580	chr11:88305340-88385387	0	N^
chr11:88269449-88351661	*GRM5*	Del	M.Of.M.Cs.60 4401	chr11:88324615-88342595	14.924	Y
chr7:126441593-126621501	*GRM8*	Del	1953313026_ A	chr7:126532786-126536202	0	Y
chr7:126441593-126621501	*GRM8*	Del	1965040688_ A	chr7:126463602-126478050	54.536	Y
chr7:126441593-126621501	*GRM8*	Del	4011452014_ A	chr7:126532786-126536202	0	Y
chr7:126441593-126621501	*GRM8*	Del	14125	chr7:125660695-126036276	0	N^
chr7:126441593-126621501	*GRM8*	Del	16794	chr7:125660695-126036276	0	N^
chr7:126441593-126621501	*GRM8*	Del	11804	chr7:125679479-125937528	0	N^
chr7:126441593-126621501	*GRM8*	Del	987314	chr7:126503602-126563602	0	Y
chr7:126441593-126621501	*GRM8*	Del	987124	chr7:126463602-126603602	0	Y
chr3:7183953-7197236	*GRM7*	Del	2023340146	chr3:7053179-7144453	18.686	Y
chr3:7183953-7197236	*GRM7*	Del	068-3	chr3:7183954-7197236	20.599	Y
chr3:7183953-7197236	*GRM7*	Del	068-4	chr3:7183954-7197236	20.599	Y
chr3:7183953-7197236	*GRM7*	Del	4079019863_ A	chr3:7183954-7197236	20.599	Y
chr3:7183953-7197236	*GRM7*	Del	11891	chr3:6979874-7003319	101.280	Y
chr3:7183953-7197236	*GRM7*	Del	11923	chr3:6980446-7001696	101.852	Y
chr6:146657076-146694047	*GRM1*	Dup	388-3	chr6:146657077-146675511	0	Y
chr6:146657076-146694047	*GRM1*	Dup	387-3	chr6:146657077-146675511	0	Y
chr6:146657076-146694047	*GRM1*	Dup	386-3	chr6:146657077-146675511	0	Y
chr6:146657076-146694047	*GRM1*	Dup	4301337678_ R02C01	chr6:146657077-146675511	0	Y
chr6:146657076-146694047	*GRM1*	Dup	4305910011_ R01C02	chr6:146657077-146675511	0	Y
chr6:146657076-146694047	*GRM1*	Dup	1181	chr6:146657077-146694047	0	Y
chr6:146657076-146694047	*GRM1*	Dup	83158	chr6:146657077 -146694047	0	Y

(continuación)

CNVR	*Gen*	Tipo	ID de muestra	Región llamada en la muestra	Distancia exónica*	Ciclo de validación de muestra
chr6:146657076-146694047	*GRM1*	Dup	b3_SF_0181	chr6:146685878-146701196	13.883	Y
chr1:72317292-72328395	*NEGR1*	Dup	230-3	chr1:72317292-72328395	10.621	Y
chr1:72317292-72328395	*NEGR1*	Dup	230-4	chr1:72317292-72328395	10.621	Y
chr1:72317292-72328395	*NEGR1*	Dup	230-5	chr1:72317292-72328395	10.621	Y
chr1:72317292-72328395	*NEGR1*	Dup	TD207.1	chr1:71648994-73025013	0	Y
chr1:72317292-72328395	*NEGR1*	Dup	M.Of.M.Cs.63 08601	chr1:72322424-72328395	10.621	Y
chr7:153495598-153564827	*DPP6*	Dup	332-3	chr7:153495598-153578582	54.698	Y
chr7:153495598-153564827	*DPP6*	Dup	4079019863_ A	chr7:153495598-153564827	68.453	Y
chr7:153495598-153564827	*DPP6*	Dup	4193372403_ B	chr7:153495598-153554210	79.070	Y
chr7:153495598-153564827	*DPP6*	Dup	4243114113_ R01C02	chr7:153495598-153577484	55.796	Y
chr7:153495598-153564827	*DPP6*	Dup	1135	chr7:153495598-153576455	56.825	N^
chr7:153495598-153564827	*DPP6*	Dup	8201671744	chr7:153118878-153338318	0	Y
chr7:153495598-153564827	*DPP6*	Dup	W.Of.F.Cs.140 002	chr7:153502896-153517548	115.317	Y
chr7:153495598-153564827	*DPP6*	Dup	W.Of.M.Cs.23 4002	chr7:153545279-153559377	73.903	Y
chr5:65027976-65046520	*SGTBINL N*	Del	067-3	chr5:65027976-65046520	0	Y
chr5:65027976-65046520	*SGTBINL N*	Del	117-3	chr5:65027976-65046520	0	Y
chr5:65027976-65046520	*SGTBINLN*	Del	152-3	chr5:65027976-65046520	0	Y
chr5:65027976-65046520	*SGTBINL N*	Del	1670639198_A	chr5:65027976-65046520	0	Y
chr5:65027976-65046520	*SGTBINL N*	Del	15962	chr5:64483534-65101307	0	Y
chr5:65027976-65046520	*SGTBINL N*	Del	b11 SF_1055	chr5:65020291-65030503	3.236	Y
chr1:56053497 -56064495	*USP24*	Del	4147907208_ B	chr1:56053497 -56064495	80.234	Y
chr1:56053497 -56064495	*USP24*	Del	393-3	chr1:56053497 -56064495	80.234	Y
chr1:56053497 -56064495	*USP24*	Del	11411	chr1:56040939-56132401	67.676	Y
chr1:56053497 -56064495	*USP24*	Del	11804	chr1:56040939-56263366	67.676	Y
chr1:56053497 -56064495	*USP24*	Del	11727	chr1:56053497 -56064840	80.234	Y
chr1:56053497 -56064495	*USP24*	Del	b2_SF_0094	chr1:56051215-56057576	77.952	Y
chr19:38427720-38444834	*SLC7A10*	Del	120-3	chr19:38415546-38444834	6.998	Y
chr19:38427720-38444834	*SLC7A10*	Del	224-3	chr19:38415546-38444834	6.998	Y
chr19:38427720-38444834	*SLC7A10*	Del	305-3	chr19:38415545-38434210	6.997	Y
chr19:38427720-38444834	*SLC7A10*	Del	134-4	chr19:38418216-38444834	9.668	Y
chr19:38427720-38444834	*SLC7A10*	Del	168-3	chr19:38423641-38444834	15.093	Y

(continuación)

CNVR	*Gen*	Tipo	ID de muestra	Región llamada en la muestra	Distancia exónica*	Ciclo de validación de muestra
chr19:38427720-38444834	*SLC7A10*	Del	11931	chr19:38427721-38455315	19.173	Y
chr19:38427720-38444834	*SLC7A10*	Del	W.Of.F.Cs.121 001	chr19:38423391 -38442154	14.843	Y
chr3:1844168-1859889	*CNTN4*	Del	078-3	chr3:1273990-1859889	0	Y
chr3:1844168-1859889	*CNTN4*	Del	078-4	chr3:1273990-1859889	0	Y
chr3:1844168-1859889	*CNTN4*	Del	141-3	chr3:1756625-1928413	187.137	Y
chr3:1844168-1859889	*CNTN4*	Del	177-3	chr3:1844168-1936623	178.927	Y
chr3:1844168-1859889	*CNTN4*	Del	M.Of.F.Cs.537 01	chr3:1793056-1956567	158.983	Y
chr3:1844168-1859889	*CNTN4*	Del	U.Of.F.Cs.852 301	chr3:1835561-1852134	263.416	Y
chr3:1844168-1859889	*CNTN4*	Del	b3_SF_0253	chr3:1797102-1930071	185.479	Y
chr2:81419297-81446082	*CTNNA2*	Dup	134-4	chr2:81035643-81654296	0	Y
chr2:81419297-81446082	*CTNNA2*	Dup	144-3	chr2:81035643-81654296	0	Y
chr2:81419297-81446082	*CTNNA2*	Dup	11484	chr2:81419297-81446082	152.417	Y
chr2:81419297-81446082	*CTNNA2*	Dup	b10_SF_0900	chr2:81352586-81386102	85.706	Y
chr4:113772340-113788584	*LARP7*	Dup	303-3	chr4:113744172-113798058	0	Y
chr4:113772340-113788584	*LARP7*	Dup	314-3	chr4:113744172-113798058	0	Y
chr4:113772340-113788584	*LARP7*	Dup	17190	chr4:113772340-113788584	0	Y
chr4:113772340-113788584	*LARP7*	Dup	M.Fa.M.Cs.63 00503	chr4:113769438-113801755	0	Y

*una distancia exónica de "0" indica que el exón se ve afectado por la CNV
^muestra no disponible para validación por qPCR (muestra validad visualmente en BeadStudio).

Tabla 18. Frecuencia de CNV en genes que interactúan con receptores de GRM en casos de TDAH y controles.

Gen	Recuentos de del (casos:controles)	Recuentos de dup (casos:controles)	Enriquecimiento de TDAH
ACAT1	0:0	1:0	Sí
ACCN1	0:0	3:1	Sí
ACTR2	1:0	0:1	Sí
ADCY1	0:0	1:1	Sí
ADRBK1	1:0	0:0	Sí
ALDOA	3:8	2:6	Sí
APP	0:0	8:2	Sí
ARL15	1:1	2:0	Sí
ATXN7L3	1:1	0:0	Sí
BDKRB2	1:1	0:0	Sí
CA8	0:0	1:0	Sí
CACNA1B	0:0	2:2	Sí
CACYBP	1:0	0:0	Sí
CALM1	1:2	0:0	Sí
CHRM3	0:0	2:1	Sí
C/C	1:1	0:0	Sí
CNP	1:2	0:0	Sí
CRHR1	1:0	0:0	Sí
DISC1	0:0	4:7	Sí
DYNLL1	0:0	1:0	Sí
FPR1	0:0	1:1	Sí
GAPDH	0:2	1:1	Sí
GNA15	1:1	1:0	Sí
GNAI2	2:4	0:0	Sí
GNAO1	0:0	1:1	Sí
GNAQ	1:0	0:0	Sí
GRIK1	0:0	8:2	Sí
GRIK3	1:0	0:0	Sí
GRM1	0:0	7:2	Sí
GRM2	1:0	1:0	Sí
GRM3	0:0	1:0	Sí
GRM5	4:0	3:2	Sí
GRM6	1:0	0:4	Sí
GRM7	4:0	0:0	Sí
GRM8	3:0	1:1	Sí
GSN	1:0	1:0	Sí
HOMER1	0:0	1:0	Sí
HTR2A	0:0	1:0	Sí
MAPK1	1:0	0:0	Sí
MTHFD1	1:1	0:0	Sí
MX1	0:0	7:2	Sí
NARG1	1:0	0:0	Sí
NMI	0:0	1:0	Sí
PCBP3	3:2	6:3	Sí
PDE1C	1:0	1:1	Sí
PPP2R1A	0:0	1:0	Sí

(continuación)

Gen	Recuentos de del (casos:controles)	Recuentos de dup (casos:controles)	Enriquecimiento de TDAH
PRPSAP1	1:0	1:1	Sí
PSMD11	2:24	1:0	Sí
PSMD13	0:4	1:2	Sí
PXN	0:0	1:0	Sí
QRICH2	1:1	0:1	Sí
RANBP1	2:3	0:9	Sí
RAP2A	0:0	1:1	Sí
RCC1	0:0	1:0	Sí
RGS12	2:0	0:0	Sí
RIF1	0:0	1:0	Sí
RUVBL2	1:0	0:3	Sí
RYR1	1:2	1:1	Sí
RYR2	1:0	1:0	Sí
SDC3	1:0	0:1	Sí
SELE	1:0	0:0	Sí
SERPINB9	0:0	1:0	Sí
SETD4	2:0	8:3	Sí
SHANK1	0:0	1:0	Sí
SORD	0:0	1:0	Sí
STRAP	0:0	1:1	Sí
TK1	2:0	0:2	Sí
TNIK	1:0	0:0	Sí
VHL	0:0	1:0	Sí
BTBD2	0:7	1:15	No
ECHS1	0:1	1:22	No
F2RL3	1:16	0:0	No
GNB2L1	0:0	0:4	No
HOMER3	1:12	1:9	No
ITGB7	0:5	1:12	No
KIAA1683	3:18	0:3	No
PDE6G	3:26	1:12	No
PLCB3	0:5	0:2	No
PYGM	3:29	0:4	No
RPLP2	8:92	1:6	No
SLC6A3	0:0	0:11	No
SRC	1:19	0:2	No
TBCA	1:10	0:0	No
TRAF2	5:24	1:11	No
40425	0:0	0:0	Sin SNP en el gen
ADRA2A	0:0	0:0	Sin SNP en el gen
ADRA2C	1:1	0:1	Sin SNP en el gen
C17orf44	0:0	0:0	Sin SNP en el gen
C7orf25	0:0	0:0	Sin SNP en el gen
F2RL2	0:3	0:0	Sin SNP en el gen
FKBP3	0:0	0:0	Sin SNP en el gen
FSCN1	0:0	0:1	Sin SNP en el gen

(continuación)

Gen	Recuentos de del (casos:controles)	Recuentos de dup (casos:controles)	Enriquecimiento de TDAH
GRB7	0:0	0:0	Sin SNP en el gen
H5P90AB1	1:0	0:0	Sin SNP en el gen
IMPDH2	0:0	0:0	Sin SNP en el gen
LOC642393	0:0	1:4	Sin SNP en el gen
LOC653098	0:0	0:0	Sin SNP en el gen
MC4R	0:0	0:0	Sin SNP en el gen
MGC11082	0:0	0:0	Sin SNP en el gen
MRPS16	0:0	0:0	Sin SNP en el gen
NPY2R	1:0	0:0	Sin SNP en el gen
PAFAH1B3	1:1	0:0	Sin SNP en el gen
PCBP1	0:0	0:0	Sin SNP en el gen
PCMT1	0:0	0:0	Sin SNP en el gen
PHKG2	0:0	0:0	Sin SNP en el gen
PRLHR	0:0	0:0	Sin SNP en el gen
PSME1	0:0	0:0	Sin SNP en el gen
RAB2	2:2	0:1	Sin SNP en el gen
RGS2	0:0	0:0	Sin SNP en el gen
S100A6	0:0	0:0	Sin SNP en el gen
SET	0:0	0:0	Sin SNP en el gen
sF3B14	0:0	0:0	Sin SNP en el gen
TBXA2R	10:44	0:10	Sin SNP en el gen
TMEM4	0:0	0:0	Sin SNP en el gen
TPI1	0:0	1:1	Sin SNP en el gen
TRMT112	0:1	0:2	Sin SNP en el gen
TUBA1	0:0	0:0	Sin SNP en el gen
TUBA1A	0:0	0:0	Sin SNP en el gen
TUBA2	0:1	0:0	Sin SNP en el gen
TUBB	0:0	0:0	Sin SNP en el gen
TUBG1	0:1	0:0	Sin SNP en el gen
ACAT2	0:0	0:0	
ACCN2	0:2	0:0	
ACP1	0:0	0:3	
ACTB	0:0	0:0	
ADA	0:0	0:0	
ADD1	0:0	0:0	
ADD2	0:0	0:0	
ADORA1	0:0	0:1	
ADRA1B	0:0	0:0	
ADRB2	0:0	0:0	
ANXA2	0:0	0:0	
APTX	0:0	0:0	
AQP1	0:0	0:1	
ARHGAP24	0:0	0:0	
ARRB1	0:0	0:0	
ARRB2	0:0	0:1	
BDKRB1	0:0	0:0	
BTG2	0:0	0:1	

(continuación)

Gen	Recuentos de del (casos:controles)	Recuentos de dup (casos:controles)	Enriquecimiento de TDAH
C1orf116	0:0	0:1	
CALB2	0:0	0:0	
CALM2	0:0	0:0	
CALM3	0:0	0:0	
CAMK1	0:0	0:0	
CAMK2B	0:0	0:0	
CAMK4	0:0	0:0	
CCNB1	0:0	0:0	
CDC42	0:0	0:0	
CENTG1	0:1	0:0	
CHGB	0:0	0:0	
CHP	0:0	0:0	
CHRM2	0:0	0:0	
CMPK	0:0	0:0	
CNR1	0:0	3:8	
COPB2	0:0	0:0	
CYCS	0:0	0:0	
DCN	0:0	0:0	
DHCR7	0:0	0:1	
DLST	0:0	0:0	
DRD2	0:0	0:0	
DRD3	0:0	0:0	
DSTN	0:0	0:0	
EGFR	0:0	0:0	
EIF3S3	0:0	0:1	
ERBB2	0:0	0:0	
F2R	0:0	0:0	
F3	0:0	0:0	
FURIN	0:0	0:0	
FYN	0:0	0:0	
GLP1R	0:0	0:0	
GLP2R	0:0	0:0	
GNAI1	0:0	0:0	
GNAI3	0:0	0:0	
GOT1	0:0	0:0	
GP1BA	0:0	0:0	
GPR26	0:0	0:0	
GRB2	0:0	0:0	
GRIA1	0:0	0:0	
GRM4	0:0	0:0	
HBXIP	0:0	0:0	
HD	0:0	0:0	
HNRPA3	0:0	0:0	
IL8RB	0:0	0:0	
IQGAP2	0:0	0:0	
ITGB1	0:0	0:0	
ITPR1	0:0	0:0	

(continuación)

Gen	Recuentos de del (casos:controles)	Recuentos de dup (casos:controles)	Enriquecimiento de TDAH
KIAA0090	0:1	0:0	
LAMA4	0:0	0:0	
LRP2BP	0:3	0:0	
LRRC59	0:0	0:0	
LTA	0:0	0:0	
LYAR	0:0	0:0	
LYN	1:3	0:0	
MAP4	0:0	0:0	
MAPT	0:0	0:0	
MARK4	0:0	0:0	
MRPL14	0:0	0:0	
MTNR1A	0:3	0:0	
MTNR1B	0:0	0:0	
MYC	0:1	0:0	
MY06	0:0	0:0	
NANS	0:0	0:0	
NCK1	0:0	0:0	
NFKBIA	0:0	0:0	
NUDC	0:0	0:1	
OPRD1	3:13	0:0	
PCDHA4	0:0	0:0	
PCID1	0:0	0:0	
PDCD5	0:0	0:0	
PDE1B	0:0	0:0	
PGM1	0:0	0:0	
PHKB	0:0	0:0	
PICK1	0:3	0:1	
PIK3CA	0:0	0:0	
PIK3R1	0:0	0:0	
PLA2G7	0:0	0:0	
PLCB1	0:0	0:0	
PLCG2	0:0	0:0	
PPIH	0:0	0:0	
PRDX1	0:0	0:0	
PRKCA	0:0	0:0	
PRMT1	0:0	0:1	
PSAT1	0:0	0:0	
PSEN1	0:0	0:0	
PSMA1	0:0	0:1	
PSMC1	0:0	0:0	
PSMD1	0:0	0:0	
PSMD6	0:0	0:0	
PTHR2	0:0	0:0	
PYGL	0:0	0:0	
RALA	0:0	0:0	
RCC2	0:0	0:0	
RHOA	0:0	0:0	

(continuación)

Gen	Recuentos de del (casos:controles)	Recuentos de dup (casos:controles)	Enriquecimiento de TDAH
RPA2	0:0	0:0	
RPN2	0:0	0:0	
RPS14	0:0	0:0	
RRM1	0:0	0:0	
SACS	0:0	0:1	
SARS	0:0	0:0	
SCTR	0:0	0:0	
SHBG	0:0	0:0	
SIAH1	0:0	0:0	
SLC2A1	0:0	0:0	
SNCA	0:0	0:0	
SNRPB2	0:0	0:0	
SOCS6	0:0	0:0	
SOCS7	0:0	0:0	
STAU1	0:0	0:0	
STX12	0:0	0:0	
SYK	0:0	0:0	
TCP1	0:0	0:0	
TEAD3	0:0	0:0	
TFAM	0:0	0:0	
TGM2	0:0	0:3	
TJP1	0:0	0:2	
TLR10	0:0	0:0	
TUBA1B	0:0	0:0	
TXN	0:0	0:0	
TXNDC4	0:2	0:1	
TXNL2	0:0	0:1	
TYMS	0:0	0:2	
UBQLN4	0:0	0:0	
UCHL1	0:0	0:0	
VIPR1	0:0	0:0	
YWHAQ	0:0	0:0	
ZAP70	0:0	0:0	

Tabla 19. Grupos de genes basándose en la red de genes que interactúan

N.º de grupo	Genes
1	*SET, HNRPA3, RRM1, SORD, PSMC1, MTHFD1, CACYBP, PCBP1, TXNL2, 40425, SARS, PCID1, GSN, PSMD6, TBCA, MRPS16, RCC2, COPB2, RANBP1, PRMT1, ANXA2, FSCN1, RCC1, ACAT1, NUDC, EIF3S3, UCHL1, FKBP3, PDCD5, ACTR2, PSAT1, LYAR, PCBP3, SF3B14, LRRC59, ACP1, ACAT2, RUVBL2, GPR26, MAPK1, CYCS, MGC11082, STRAP, RAP2A, IMPDH2, ACTR2, PSMD1, SETD4, TRMT112, CMPK, MRPL14, SNRPB2, TEAD3, TMEM4, TFAM, DSTN, PRPSAP1, KIAA0090, PPIH, PSMA1, RPS14, DHCR7, PSMD13, TRAF2, TNIK, RPN2, TYMS, NCK1, NANS, NARG1, PPP2R1A, ECHS1, GOT1, PCMT1*
2	*GRB7, PYGL, CRHR1, PDE1C, CALM1, GLP1R, PYGM, PHKG2, PTHR2, PDE1B, GLP2R, ADD2, ADCY1, SCTR, PHKB, VIPR1, ADD1, PGM1, PGM1, IQGAP2*
3	*HBXIP, S10OA6, TXN, SLC2A1, CAMK1, RAB2, PCDHA4, QRICH2, GAPDH, BTBD2, PAFAH1B3, SERPINB9, PSMD11, PRDX1, RPA2, CAMK2B, LAMA4, ARL15, TPI1, CAMK4, TK1, FYN, PGM1, ACTB, CHP*

(continuación)

N.º de grupo	*Genes*
4	*SLC6A3, UBQLN4, PRLHR, PICK1, CIC, APTX, ERBB2, ATXN7L3, ACCN2, AQP1, GRIA1, ACCN1, ECHS1, SACS, BTG2, LRP2BP, PRKCA*
5	*RALA, CDC42, DRD3, ITGB1, ITGB7, TLR10, HSP90AB1, TJP1, FURIN, VHL, MTNR1B, PSEN1, SHBG, DCN, F3, GRIK3, GP1BA, RHOA, SELE, DRD2, ARHGAP24, MTNR1A, FKBP3, ARRB2, GRM8*
6	*NPY2R, RGS12, GNAI3, ADRA2C, GNAI2, GNAO1, CACNA1B, GNAI1, GRM6, IL8RB, PLCB3*
7	*ADRA2A, PDE6G, SRC, MC4R, ARRB1, SNCA, RPLP2, FPR1, BDKRB2, ADRBK1, OPRD1*
8	*PLCB1, TXNDC4, ITPR1, CCNB1, LYN, CA8, PLCG2*
9	*F2RL3, HTR2A, ADRA1B, F2R, RGS2, HTR2A, GNAQ, F2RL2, CHRM3, PIK3CA, BDKRB2, TBXA2R, BDKRB1*
10	*GNB2L1, CNP, STAU1, CHGB, PSME1, SOCS7, DLST, ALDOA, SYK, SDC3, TUBB, TGM2, HD, MARK4, MAP4, MX1, TUBA1A, SOCS6, C7orf25, PLA2G7*
11	*HOMER1, STX12, CENTG1, RYR2, LOC653098, HOMER3, C1orf116, SHANK1, RYR1*
12	*CNR1, GNA15, CHRM2, ADRB2*
13	*DYNLL1, PIK3R1, NMI, TUBA2, PXN, TUBG1, NFKBIA, TUBA1B, YWHAQ*
14	*HRPT2, RIF1, GRM3*
15	*CALM3, GRM5, MYO6, KIAA1683, GRM7, LOC642393, C17orf44, CALM2*
16	*CALB2, TCP1, LTA, TUBA1, ZAP70*
17	*ADA, ADORA1*

Ejemplo IV

Se examinaron los efectos biológicos de una CNV de GRM5 (eliminación) identificada en una familia de pacientes con TDAH que portan una gran eliminación en una copia de sus genes de GRM5, en líneas celulares derivadas de PBMC transformadas mediante el virus de Epstein-Barr (VEB). Se ha comunicado en las referencias que las PBMC expresan tanto mGluR1 como mGluR5 y los mGluR pueden desempeñar un papel en la activación de los linfocitos T. Los presentes inventores examinaron en primer lugar la expresión de mGluR5 en líneas celulares transformadas con VEB procedentes de sujetos sanos mediante qRT-PCR y citometría de flujo. Los resultados confirmaron la expresión de mGluR5. Posteriormente, los presentes inventores compararon cuatro líneas celulares procedentes de la familia con TDAH con cuatro líneas celulares de control mediante análisis cuantitativo de la señal de fluorescencia mediante citometría de flujo. Se calculó la proporción del valor de fluorescencia entre la tinción de Ac para mGluR5 y la tinción con Ac de control. La relación media del grupo de TDAH fue de 4,6 (± 0,4), mientras que la media del control fue de 7,3 (± 1,7). La diferencia es estadísticamente significativa (prueba de la t, p=0,024), que representa una reducción de aproximadamente el 36% de la expresión de mGluR5 en los casos con eliminación. Este resultado proporciona la primera prueba de que esta CNV en el gen GRM5 da como resultado una reducción de la expresión de mGluR5.

Los siguientes materiales y procedimientos se proporcionan para facilitar la puesta en práctica del ejemplo IV.

qRT-PCR. Se adquirieron sondas y cebadores TaqMan de Applied Biosystems (n.º de cat. Hs00168275-ml). La PCR en tiempo real se llevó a cabo en placas de 384 pocillos usando el sistema 7900HTreal-time PCR (Applied Biosystems). Se sintetizó el ADNc de los moldes a partir del ARN total y un cebador aleatorio usando un kit para síntesis RT de ADNc (Applied Biosystems, n.º de cat. 4374867). Se aisló el ARN de las células cultivadas en medio RPMI 1640 que contenía FBS al 10%.

Citometría de flujo. Se fijaron las células con paraformaldehído y se tiñeron con anticuerpo monoclonal específico para mGluR5 (R&D, n.º de cat. MAB4514) o anticuerpo de control de isotipo emparejado (R&D, n.º de cat. MAB002) seguido de un anticuerpo anti-ratón conjugado con ficoeritrina. Las células teñidas se analizaron usando un citómetro de flujo BD FACS Calibur.

RESULTADOS

Las células de cada sujeto se dividieron en tres muestras. Una de ellas se usó como control de fondo con tinción y dos de ellas se tiñeron con Ac para mGluR5 y Ac de control, respectivamente. En la tabla A se muestra la fluorescencia media de las células teñidas con Ac para mGluR5 y Ac de control para cada sujeto. Se calculó la relación de la tinción con Ac para mGluR y la tinción con Ac de control y también se presenta en la tabla 1. La relación media del grupo de TDAH fue de 4,6 (± 0,4), mientras que la media del control fue de 7,3 (± 1,7). La diferencia es estadísticamente significativa (prueba de la t, p=0,024), que representa una reducción de aproximadamente el 36% de la expresión de mGluR5.

Estos datos demuestran que las consecuencias biológicas de las CNV en las diversas dianas génicas proporcionadas en el presente documento pueden determinarse experimentalmente y verificarse en sistemas biológicos, facilitando de este modo la identificación y caracterización de agentes terapéuticos beneficiosos que pueden, por ejemplo, restaurar la actividad biológica que se pierde a consecuencia de estas eliminaciones.

Tabla A. Resultados del análisis por citometría de flujo de líneas celulares procedentes de PBMC

Sujeto	Fluorescencia		Relación
	Ac para mGluR5	Ac de control	
ADHD-1	52,8	11,3	4,7
ADHD-2	50,8	11,2	4,5
ADHD-3	79,8	15,4	5,2
ADHD-4	56,5	13,3	4,2
Control-1	120	19,3	6,2
Control-2	92	13,3	6,9
Control-3	56,7	9,01	6,3
Control-4	116	11,7	9,9

Referencias

1. Glessner JT, Wang K, Cai G, y col. Autism genome-wide copy number variation reveals ubiquitin and neuronal genes. Nature 2009;459:569-573.
2. Derks EM, Hudziak JJ, Dolan CV, van Beijsterveldt TC, Verhulst FC, Boomsma DI. Genetic and environmental influences on the relation between attention problems and attention deficit hyperactivity disorder. Behav Genet 38(1), 11-23 (2008).
3. Wood AC, Rijsdijk F, Saudino KJ, Asherson P, Kuntsi J. High heritability for a composite index of children's activity level measures. Behav Genet 38(3), 266-276 (2008).
4. Haberstick BC, Timberlake D, Hopfer CJ, Lessem JM, Ehringer MA, Hewitt JK. Genetic and environmental contributions to retrospectively reported DSM-IV childhood attention deficit hyperactivity disorder. Psychol Med 38,1057-1066 (2008).
5. Wang K, Zhang H, Ma D, y col., Common genetic variants on 5p14.1 associate with autism spectrum disorders. Nature 459, 528-533 (2009).
6. Franke B, Neale BM, Faraone SV, Genome-wide association studies in ADHD. Hum Genet, doi:10.1007/s00439-009-0663-4 (2009).
7. Neale BM, Lasky-Su J, Anney R, y col. Genome-wide association scan of attention deficit hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet 147B, 1337-1344. (2008).
8. Lasky-Su J, Neale BM, Franke B y col., Genome-wide association scan of quantitative traits for attention deficit hyperactivity disorder identifies novel associations and confirms candidate gene associations. American Journal of Medical Genetics Parte B: Neuropsychiatric Genetics 147B(8), 1345-1354 (2008).
9. Lesch KP, Timmesfeld N, Renner TJ, y col. Molecular genetics of adult ADHD: converging evidence from genomewide association and extended pedigree linkage studies. J Neural Transm 115, 1573-1585 (2008).
10. Wang K, Li M, Hadley D, y col. PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. Genome Res. 17, 1665-1674 (2007).
11. Zhou K, Dempfle A, Arcos-Burgos M y col. Meta-analysis of genome-wide linkage scans of attention deficit hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet. 147B(8), 1392-8 (2008).
12. Kent WJ, Sugnet CW, Furey TS, y col. The human genome browser at UCSC. Genome Res. 12(6), 996-1006 (2002).
13. Elia, J., Gai, X., y col. Rare Structural Variants found in Attention-Deficit Hyperactivity Disorder are Preferentially Associated with Neurodevelopmental Genes. Molecular Psychiatry June 23 2009 Epub.
14. Taniura, H., Sanada, N., Kuramoto, N., Yoneda, Y. Metabotropic Glutamate Receptor Family Gene in Dictyostelium discoideum. J. Biol. Chem., 281(18), 12336-12343, (2006).
15. Conn, P. J. y Pin, J., Phamacology and Functions of Metabotropic Glutamate Receptors. Annu. Rev. Pharmacol. Toxicol. 37, 205-37 (1997).
16. Berthele A, Platzer S, Laurie DJ, y col. Expression of metabotropic glutamate receptor subtype mRNA (mGluRl-8) in human cerebellum. NeuroReport 10(18), 3861-3867 (1999).
17. Koob, G. F., Sanna, P.P., Bloom, F. E., Neuroscience of Addiction. Neuron, 21.467-476, (1998).
18. Cryan JF, Kelly PH, Neijt HC, Sansig G, Flor PJ, van Der Putten H. Antidepressant and anxiolytic-like effects in mice lacking the group III metabotropic glutamate receptor mGluR7. European Journal of Neuroscience. 17(11):2409-2417, (2003).
19. Makoff, A., Pillinga, C., Harrington, K., Emson, P. Human metabotropic glutamate receptor type 7: Molecular cloning and mRNA distribution in the CNS. Molecular Brain Research. 40(1), 165-170 (1996).
20. Turic D, Langley K, Mills S, y col. Follow-up of genetic linkage findings on chromosome 16p13: evidence of association of N-methyl-D aspartate glutamate receptor 2A gene polymorphism with ADHD. Mol Psychiatry 9,

169-173 (2004).
21. Mick, E. y Faraone, S.V., Genetics of attention deficit hyperactivity disorder. Child Adolesc Psychiatr Clin N Am 17, 261-284, vii-viii. (2008).
22. Turic D, Langley K, Williams H, y col. A family based study implicates solute carrier family 1-member 3 (SLC1A3) gene in attention-deficit/hyperactivity disorder. Biol Psychiatry. 1 de junio de 2005;57(11):1461-6.
23. Elia J, Capasso M, Zaheer Z, y col. Candidate gene analysis in an on-going genome-wide association study of attention-deficit hyperactivity disorder: suggestive association signals in ADRA1A. Psychiatr Genet. PMID: 19352218 (2009).
24. Mick E, Neale B, Middleton FA, McGough JJ, Faraone SV. Genome-wide association study of response to methylphenidate in 187 children with attention-deficit/hyperactivity disorder. Am J Med Genet B Neuropsychiatr Genet 147B,1412-1418 (2008).
25. Dorval KM, Wigg KG, Crosbie J, y col., Association of the glutamate receptor subunit gene GRIN2B with attention-deficit/hyperactivity disorder. Genes Brain Behav. 6(5) (2007).
26. Jin Z, Zang YF, Zeng YW, Zhang L, Wang YF. Striatal neuronal loss or dysfunction and choline rise in children with attention-deficit hyperactivity disorder: a 1H-magnetic resonance spectroscopy study. Neurosci Lett 315, 45-48 (2001).
27. MacMaster FP, Carrey N,. Sparkes S, Kusumakar V. Proton spectroscopy in medication-free pediatric attention-deficit/hyperactivity disorder. Biol Psychiatry 53:184-187. (2003).
28. Courvoisie H, Hooper SR, Fine C, Kwock L, Castillo M. Neurometabolic functioning and neuropsychological correlates in children with ADHD-H: preliminary findings. J Neuropsychiatry Clin Neurosci 16,63-69 (2004).
29. Carrey N, MacMaster FP, Fogel J, y col. Metabolite changes resulting from treatment in children with ADHD: a 1H-MRS study. Clin Neuropharmacol 26, 218-221 (2003).
30. Gainetdinov RR, Wetsel WC, Jones SR, Levin ED, Jaber M, Caron MG. Role of serotonin in the paradoxical calming effect of psychostimulants on hyperactivity. Science 283, 397-401 (1999).
31. Gainetdinov, R.R., Mohn, A.R., Bohn, L.M., Caron, M.G., Glutamatergic modulation of hyperactivity in mice lacking the dopamine transporter. Proc Natl Acad Sci U S A 98,11047-11054 (2001).
32. Masuo, Y. Ishido, M., Morita, M., Oka, S., Effects of neonatal 6-hydroxydopamine lesion on the gene expression profile in young adult rats. Neurosci Lett 335, 124-128 (2002).
33. Miyamoto K, Nakanishi H, Moriguchi S, y col. Involvement of enhanced sensitivity of N-methyl-D-aspartate receptors in vulnerability of developing cortical neurons to methylmercury neurotoxicity. Brain Res 901, 252-258 (2001).
34. Russell V, Allie S, Wiggins T. Increased noradrenergic activity in prefrontal cortex slices of an animal model for attention-deficit hyperactivity disorder--the spontaneously hypertensive rat. Behav Brain Res 117,69-74 (2000).
35. Russell VA. Dopamine hypo function possibly results from a defect in glutamate-stimulated release of dopamine in the nucleus accumbens shell of a rat model for attention deficit hyperactivity disorder--the spontaneously hypertensive rat. Neurosci Biobehav Rev 27, 671-682 (2003).
36. DasBanerjee T, Middleton FA, Berger DF, Lombardo JP, Sagvolden T, Faraone SV. A comparison of molecular alterations in environmental and genetic rat models of ADHD: a pilot study. Am J Med Genet B Neuropsychiatr Genet 147B, 1554-63 (2008).
37. Sagvolden T, Johansen EB, Wøien G, y col. The spontaneously hypertensive rat model of ADHD--the importance of selecting the appropriate reference strain. Neuropharmacology 57, 619-26 (2009).
38. Del Bo R, Ghezzi S, Corti S, y col., DPP6 gene variability confers increased risk of developing sporadic amyotrophic lateral sclerosis in Italian patients. Journal of Neurology, Neurosurgery & Psychiatry. 79(9), 1085 (2008).
39. Cronin S, Tomik B, Bradley DG, Slowik A, Hardiman O. Screening for replication of genome-wide SNP associations in sporadic ALS. European Journal of Human Genetics 17, 213-218 (2009).
40. Marshall CR, Noor A, Vincent JB, y col. Structural variation of chromosomes in autism spectrum disorder. Am. J. Hum. Genet. 82, 477-88 (2008).
41. Renström F, Payne F, Nordström A, y col. Replication and extension of genome-wide association study results for obesity in 4923 adults from northern Sweden Human Molecular Genetics 18(8):1489-1496 (2009).
42. Kessler RC, Adler LA, Barkley R, y col., Patterns and predictors of attention-deficit/hyperactivity disorder persistence into adulthood: results from the national comorbidity survey replication. Biological Psychiatry 57(11), 1442-1451 (2005).
43. Diskin S, Li M, Hou C, y col., Adjustment of genomic waves in signal intensities from whole-genome SNP genotyping platforms. Nucleic Acids Research. 36(19) (2008).
44. Dennis G Jr, Sherman BT, Hosack DA, y col., DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biology. 4(9), (2003).
45. Lesch KP, Selch S, Renner TJ, y col. Genome-wide copy number variation analysis in ADHD: association with neuropeptide Y gene dosage in an extended pedigree. Mol Psychiatry, Epub antes de impresión (2010).
46. Oades RD, Daniels R, Rascher W. Plasma neuropeptide Y levels, monoamine metabolism, electrolyte excretion, and drinking behavior in children with attention-deficit hyperactivity-disorder (ADHD). Psychiat. Res., 80, 77-186. (1998).
47. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. Noviembre de 2003;13(11):2498-504.

48. Potkin SG, Turner JA, Guffanti G, Lakatos A, Fallon JH, Nguyen DD, Mathalon D, Ford J, Lauriello J, Macciardi F; FBIRN. A genome-wide association study of schizophrenia using brain activation as a quantitative phenotype. Schizophr Bull. Enero de 2009;35(1):96-108. Epub 2008 Nov 20.

49. Wang X, Bao X, Pal R, Agbas A, Michaelis EK. Transcriptomic responses in mouse brain exposed to chronic excess of the neurotransmitter glutamate. BMC Genomics. 7 de junio de 2010;11:360.

50. C. de Lanerolle, Nihal; Eid, Tore; Lee, Tih-Shih. Genomic Expression in the Epileptogenic Hippocampus and Psychiatric Co-Morbidities. Current Psychiatry Reviews, Volumen 6, Número 2, mayo de 2010, págs. 135-144(10).

51. Ule, J., Stefani, G., Mele, A., Ruggiu, M., Wang, X., Taneri, B., y col. (2006). An RNA map predicting Nova-dependent splicing regulation. Nature, 444(7119), 580-6. doi: 10.1038/nature05304.

52. Ule, J., Ule, A., Spencer, J., Williams, A., Hu, J., Cline, M., y col. (2005). Nova regulates brain-specific splicing to shape the synapse. Nature genetics, 37(8), 844-52. doi: 10.1038/ng1610.

53. Murias, M., Swanson, J. M., y Srinivasan, R. (2007). Functional connectivity of frontal cortex in healthy and ADHD children reflected in EEG coherence. Cerebral cortex (Nueva York, N.Y. : 1991), 17(8), 1788-99. doi: 10.1093/cercor/bh1089.

54. Wang, L., Zhu, C., He, Y., Zang, Y., Cao, Q., Zhang, H., y col. (2009). Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder. Human brain mapping, 30(2), 638-49. doi: 10.1002/hbm.20530.

Aunque se han descrito e ilustrado de manera específica algunas de las realizaciones preferidas de la presente invención, no se pretende que la invención quede limitada a dichas realizaciones. Será evidente para un experto en la materia que pueden efectuarse diversos cambios y modificaciones en la misma sin desviarse del ámbito de la presente invención, tal como se expone en las siguientes reivindicaciones.

REIVINDICACIONES

1. Un procedimiento para detectar un riesgo aumentado de desarrollar trastorno por déficit de atención e hiperactividad (TDAH) en un sujeto de ensayo, que comprende,

a) proporcionar una muestra de ácido nucleico de dicho sujeto; y
b) determinar si dicha muestra contiene al menos una variación en el número de copias (CNV) en cualquiera de los genes seleccionados entre GRM5, GRM7, GRM8, GRM1, NEGR1, DPP6, USP24, SLC7A10, CNTN4, CTNNA2, SGTB/NLN y LARP7,

en el que si está presente dicha CNV, dicho sujeto tiene un riesgo aumentado de desarrollar TDAH.

2. El procedimiento según la reivindicación 1, en el que el ácido nucleico se amplifica antes de la detección.

3. El procedimiento de la reivindicación 1, en el que la etapa de detectar la presencia de dicha CNV se lleva a cabo usando un procedimiento seleccionado entre el grupo que consiste en: detección de hibridación específica; medición del tamaño del alelo; análisis de polimorfismo de longitud de fragmento de restricción; análisis de hibridación específica de alelo; reacción de extensión de cebador de una sola base; y secuenciación de un polinucleótido amplificado.

4. El procedimiento según la reivindicación 1, en el que el ácido nucleico es ADN.

5. El procedimiento de cualquiera de las reivindicaciones anteriores, en el que dicha CNV es una eliminación en al menos uno de los siguientes: receptor de glutamato metabotrópico 5 (GRM 5), receptor de glutamato metabotrópico 7 (GRM 7) o receptor de glutamato metabotrópico 8 (GRM 8) o una duplicación en el receptor de glutamato metabotrópico 1 (GRM 1).

6. El procedimiento de cualquiera de las reivindicaciones anteriores, en el que se ha identificado previamente que el sujeto tiene o tiene probabilidad de tener TDAH.

7. El procedimiento de cualquiera de las reivindicaciones anteriores, en el que los ácidos nucleicos que comprenden la CNV se obtienen de una célula aislada de un sujeto de ensayo humano.

8. El procedimiento de cualquiera de las reivindicaciones anteriores, en el que el sujeto tiene entre 1 y 12 años de edad.

9. (+)-5-oxo-D-prolinepiperidinamida para su uso en un procedimiento para tratar el trastorno por déficit de atención e hiperactividad (TDAH) en un sujeto humano en el que se determina que tiene al menos un polimorfismo de un solo nucleótido (SNP) que indica la presencia de una variación del número de copias (CNV) asociada con el TDAH, en el que dicha al menos una CNV es en cualquiera de los genes seleccionados entre GRM5, GRM7, GRM8, GRM1, NEGR1, DPP6, USP24, SLC7A10, CNTN4, CTNNA2, SGTB/NLN y LARP7.

10. (+)-5-oxo-D-prolinepiperidinamida para su uso en un procedimiento para tratar el trastorno por déficit de atención e hiperactividad (TDAH) en un sujeto humano en el que se determina que tiene al menos una variación del número de copias (CNV) asociada con el TDAH en uno o más de los genes seleccionados entre GRM5, GRM7, GRM8, GRM1, NEGR1, DPP6, USP24, SLC7A10, CNTN4, CTNNA2, SGTB/NLN y LARP7.

11. La (+)-5-oxo-D-prolinepiperidinamida para su uso de acuerdo con la reivindicación 9 o la reivindicación 10, en el que dicha CNV es una eliminación en al menos uno de los siguientes: receptor de glutamato metabotrópico 5 (GRM 5), receptor de glutamato metabotrópico 7 (GRM 7), receptor de glutamato metabotrópico 8 (GRM 8).

12. La (+)-5-oxo-D-prolinepiperidinamida para su uso de acuerdo con la reivindicación 9 o la reivindicación 10, en el que dicha CNV es una duplicación del receptor de glutamato metabotrópico 1 (GRM1).

13. La (+)-5-oxo-D-prolinepiperidinamida para su uso de acuerdo con cualquiera de las reivindicaciones 9 a 12, en el que dicha (+)-5-oxo-D-prolinepiperidinamida es monohidrato de (+)-5-oxo-D-prolinepiperidinamida (NS-105).

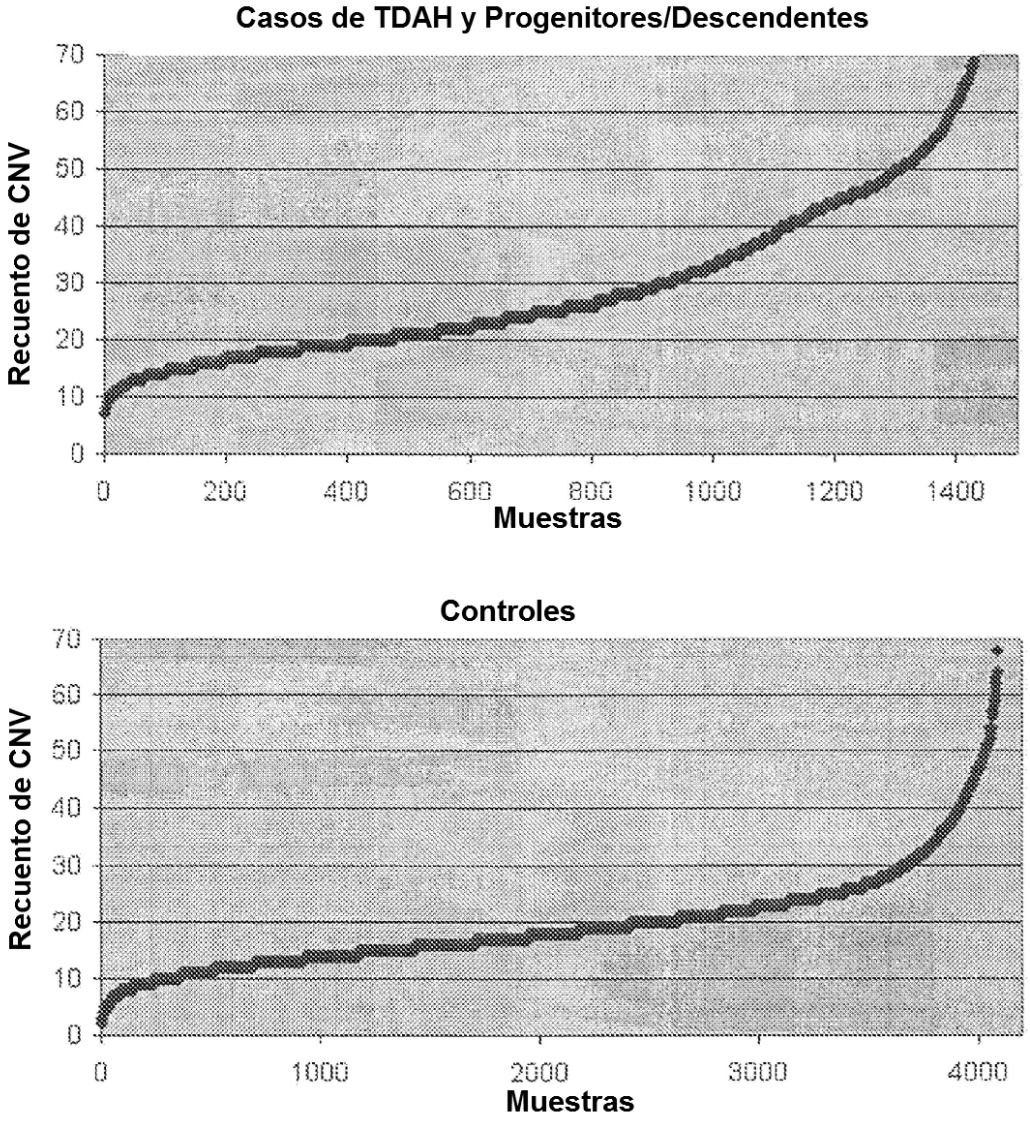

Casos de TDAH y Progenitores/Descendentes
Recuento de CNV
70
60
50
40
30
20
10
0
0
200
400
600
800
1000
1200
1400
Muestras
Controles
Recuento de CNV
70
60
50
40
30
20
10
0
0
1000
2000
3000
4000
Muestras

Figura 1

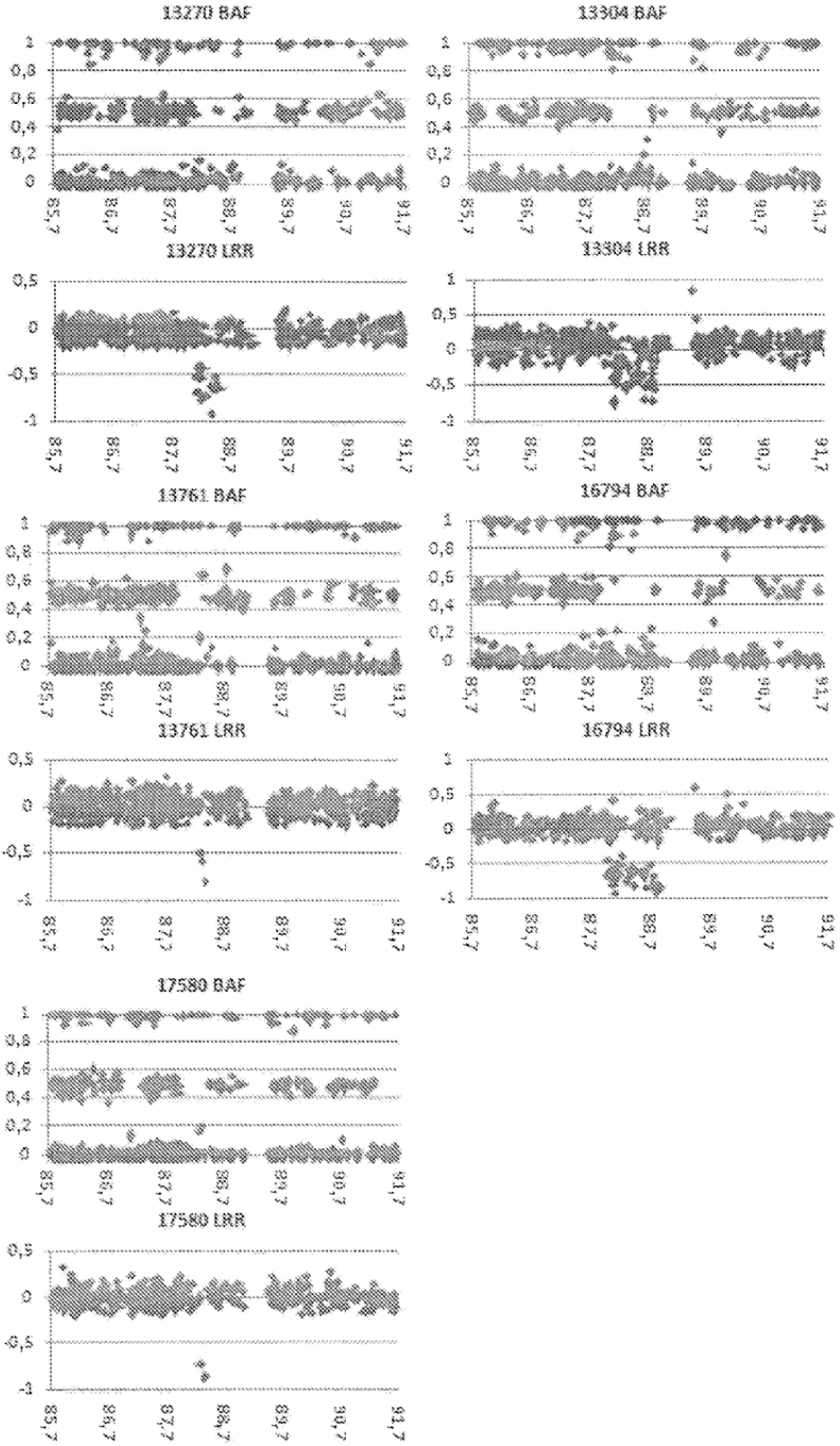

13270 BAF
13304 BAF
13270 LRR
13304 LRR
13761 BAF
16794 BAF
13761 LRR
16794 LRR
17580 BAF
17580 LRR

Figura 2

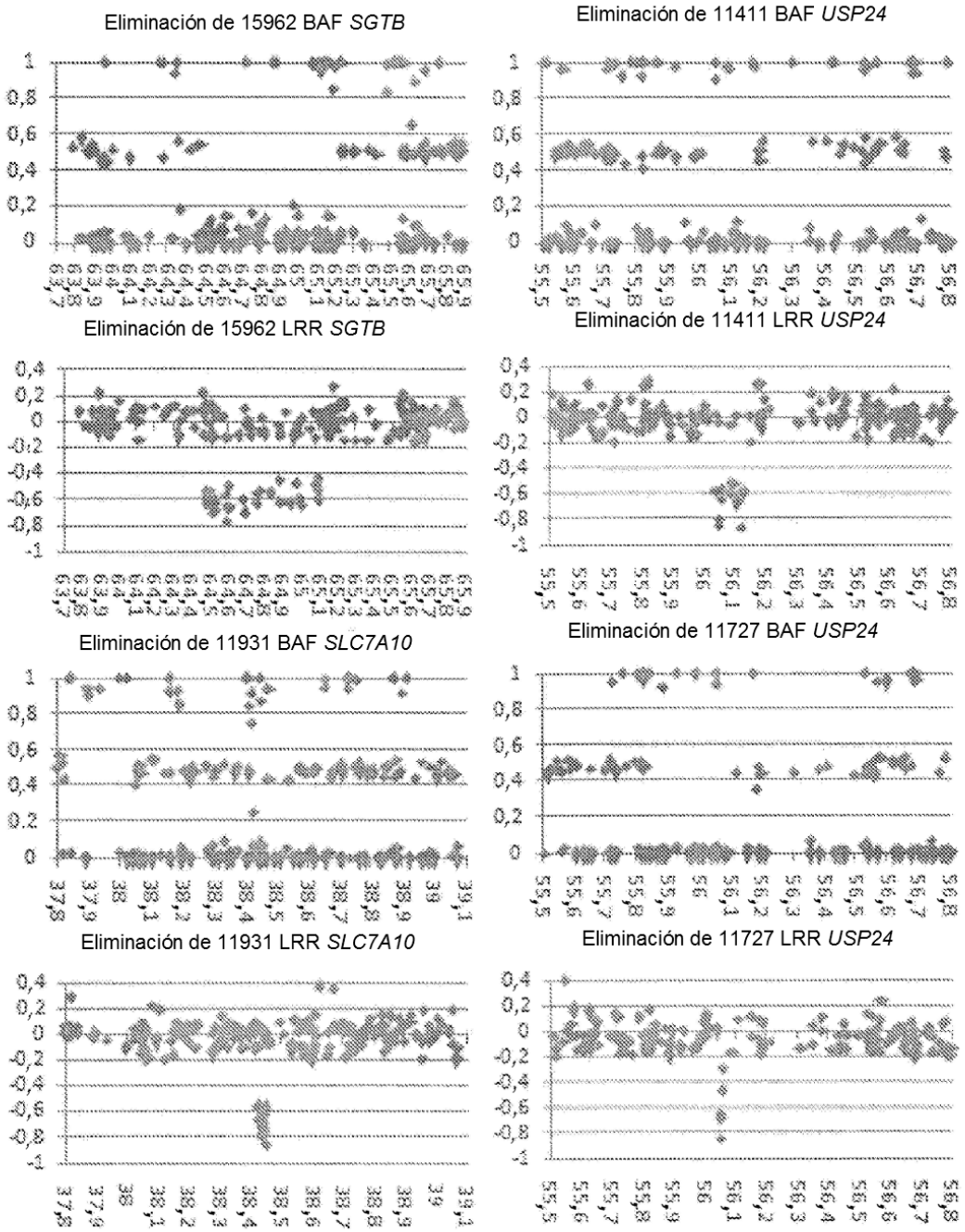

Eliminación de 15962 BAF *SGTB*
Eliminación de 11411 BAF *USP24*
Eliminación de 15962 LRR *SGTB*
Eliminación de 11411 LRR *USP24*
Eliminación de 11931 BAF *SLC7A10*
Eliminación de 11727 BAF *USP24*
Eliminación de 11931 LRR *SLC7A10*
Eliminación de 11727 LRR *USP24*

Figura 3A

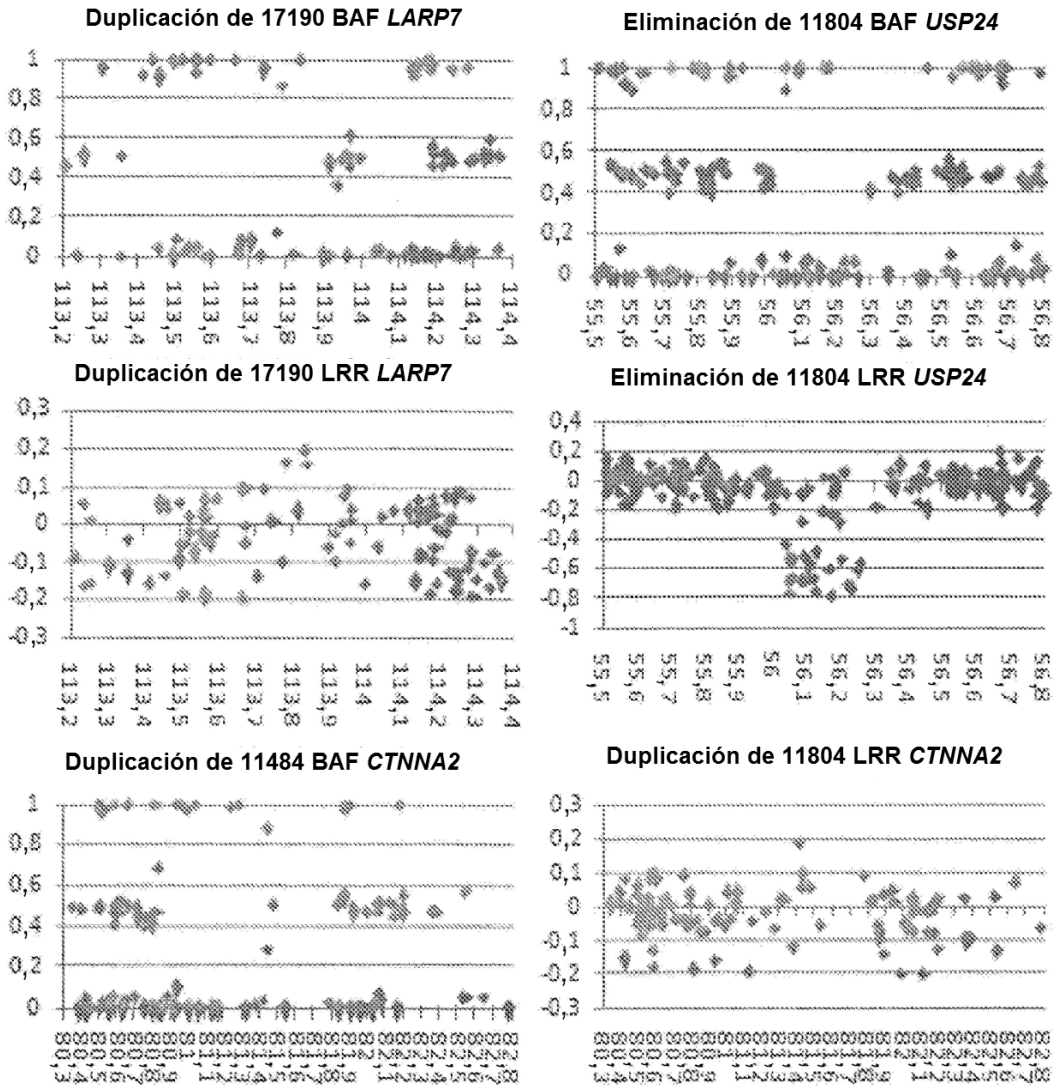


Figura 3A (continuación)

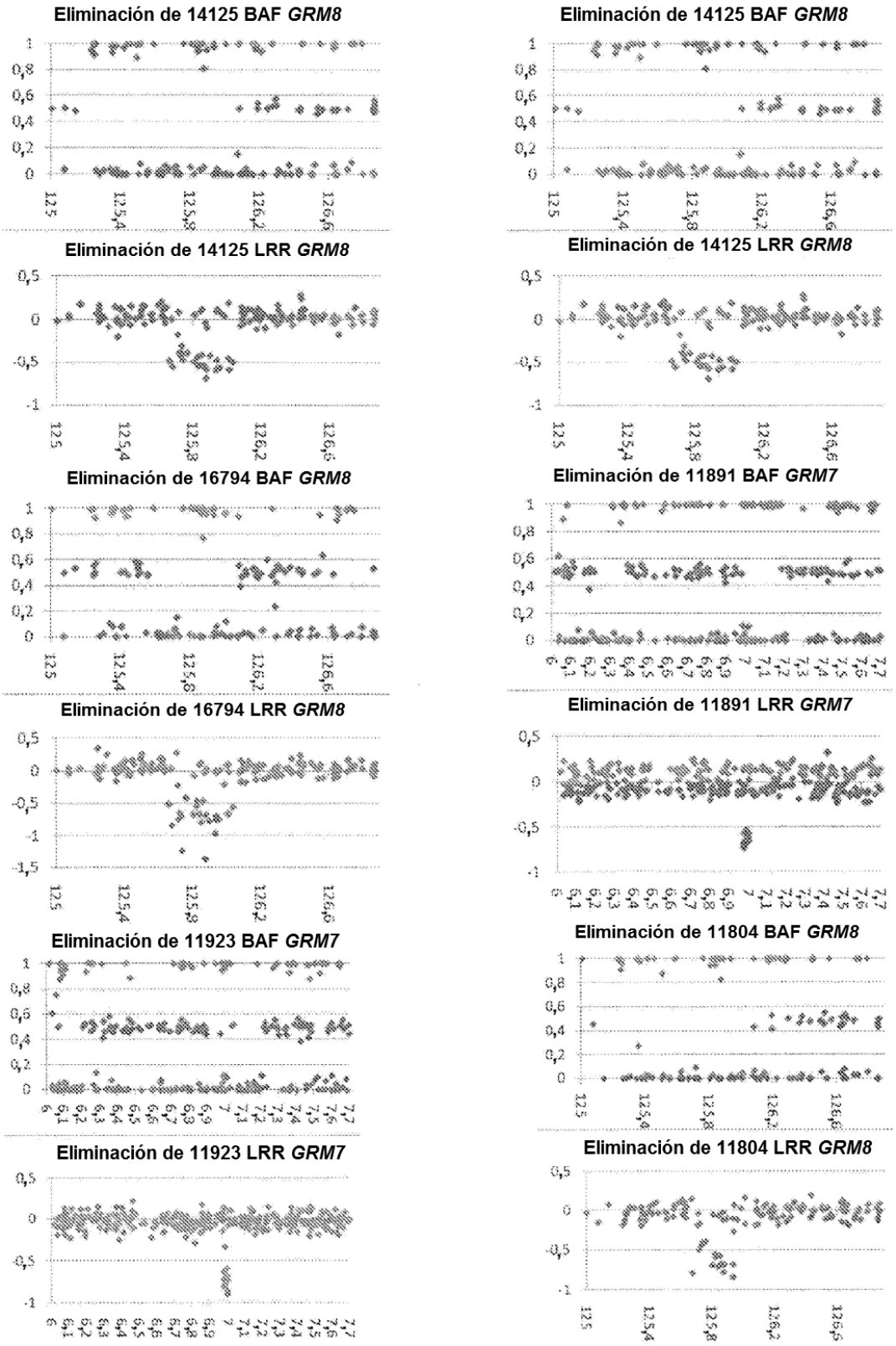


Figura 3A (continuación)

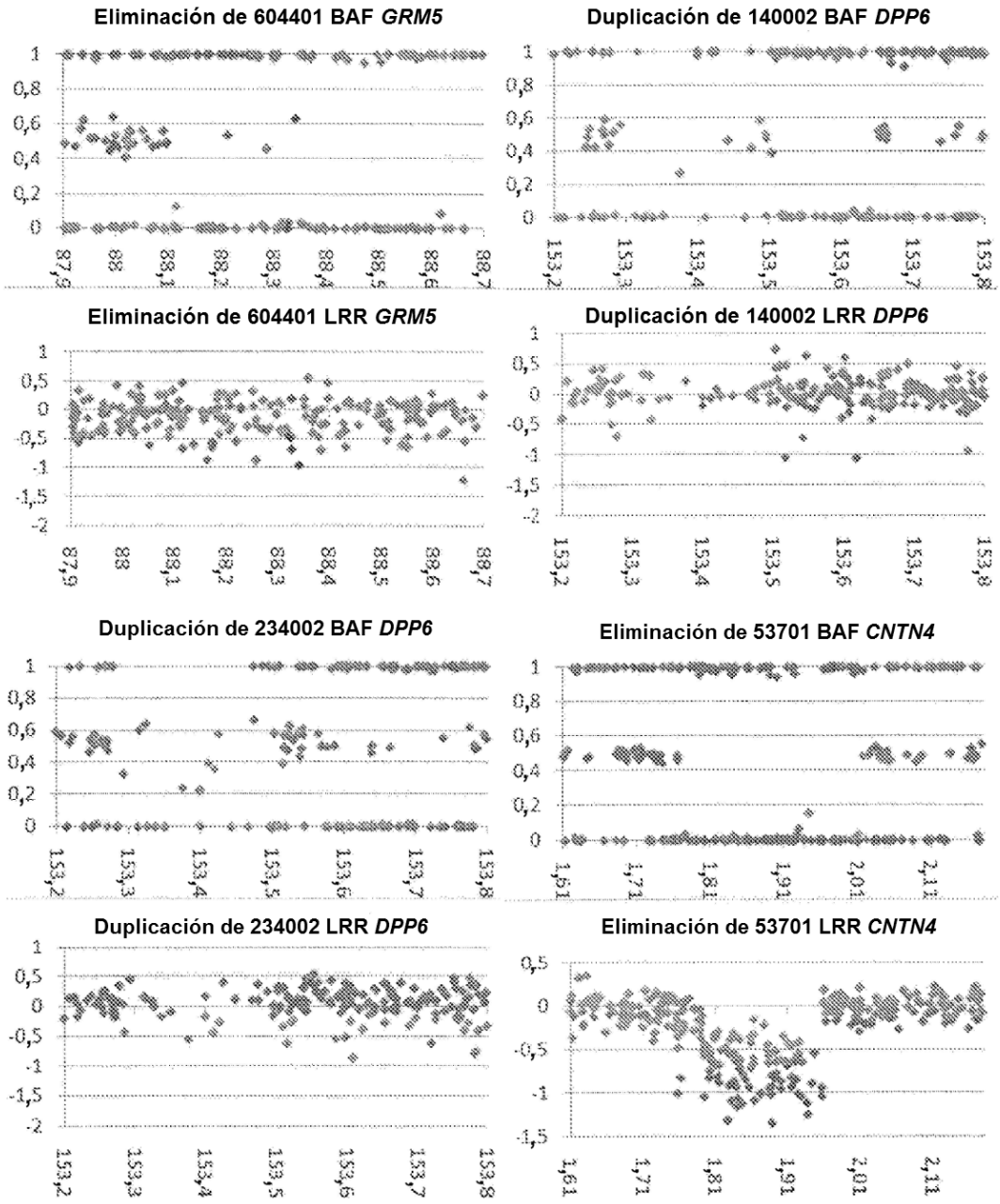

Eliminación de 604401 BAF *GRM5*
Duplicación de 140002 BAF *DPP6*
Eliminación de 604401 LRR *GRM5*
Duplicación de 140002 LRR *DPP6*
Duplicación de 234002 BAF *DPP6*
Eliminación de 53701 BAF *CNTN4*
Duplicación de 234002 LRR *DPP6*
Eliminación de 53701 LRR *CNTN4*

Figura 3B

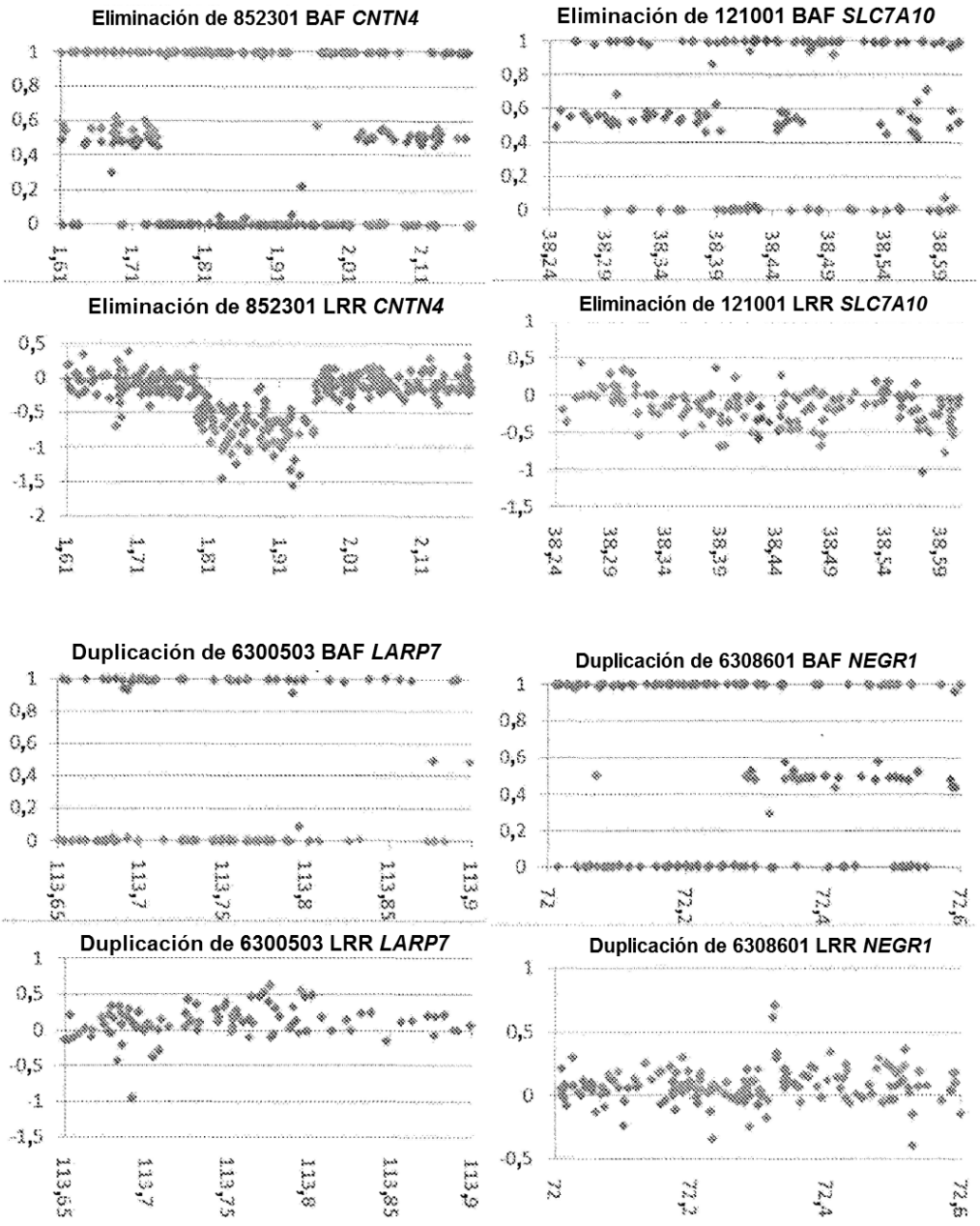


Figura 3B (continuación)

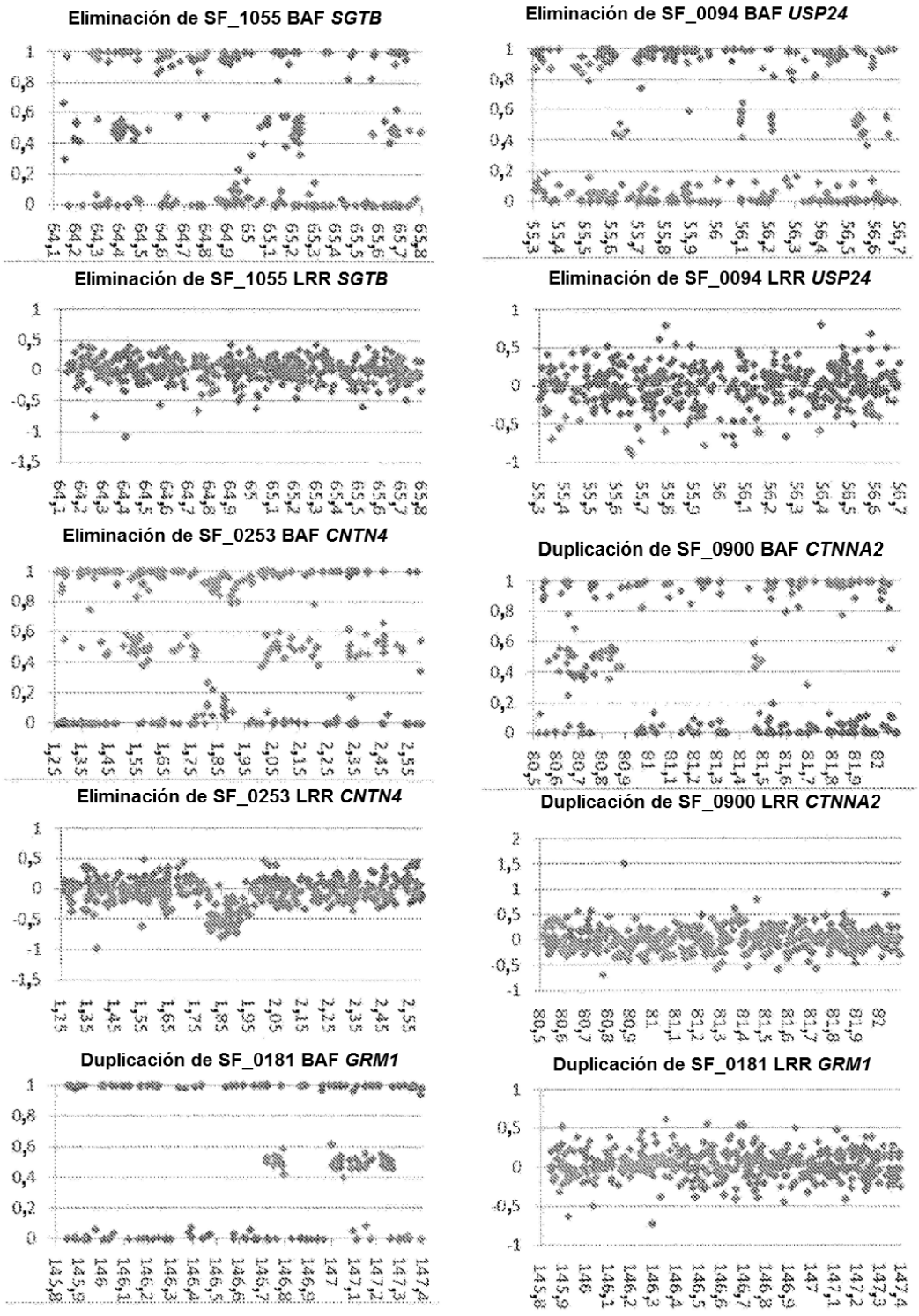

Eliminación de SF_1055 BAF SGTB
Eliminación de SF_0094 BAF USP24
Eliminación de SF_1055 LRR SGTB
Eliminación de SF_0094 LRR USP24
Eliminación de SF_0253 BAF CNTN4
Duplicación de SF_0900 BAF CTNNA2
Eliminación de SF_0253 LRR CNTN4
Duplicación de SF_0900 LRR CTNNA2
Duplicación de SF_0181 BAF GRM1
Duplicación de SF_0181 LRR GRM1

Figura 3C

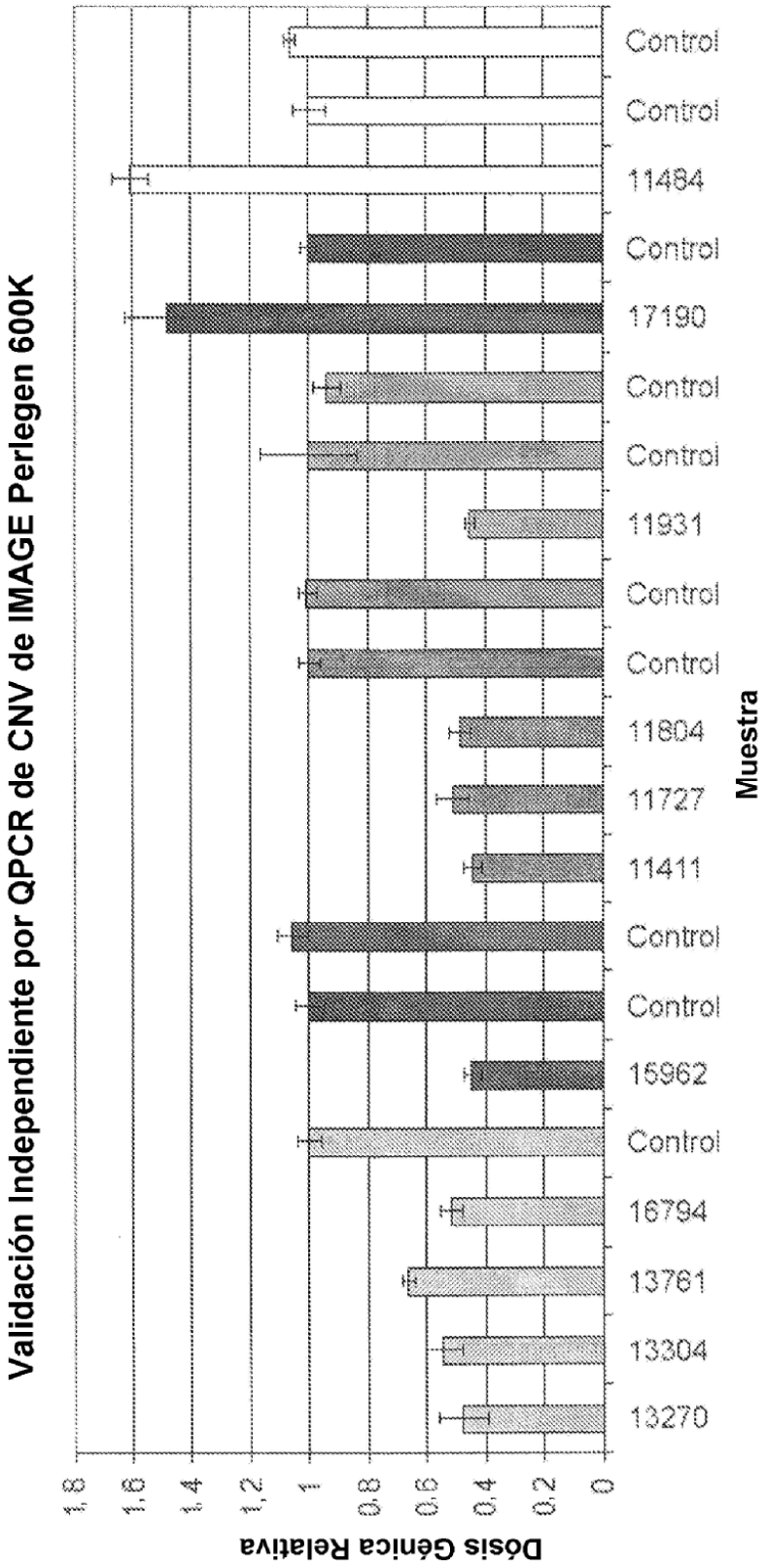
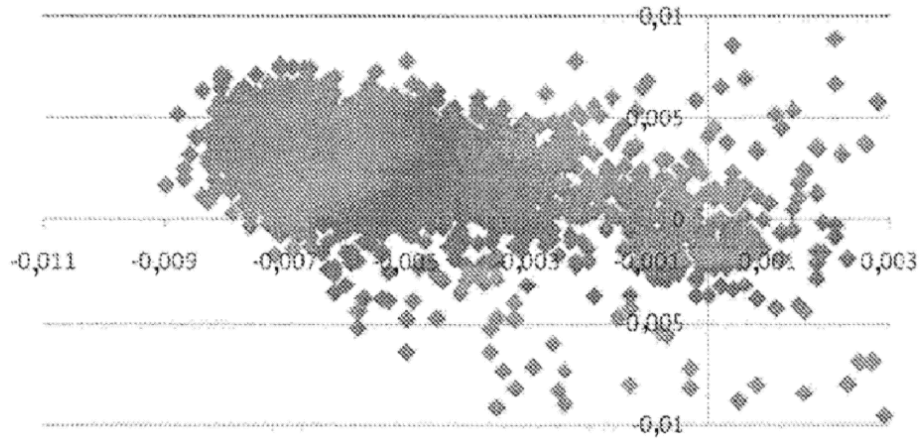

Validación Independiente por QPCR de CNV de IMAGE Perlegen 600K
Dósis Génica Relativa
1,8
1,6
1,4
1,2
1
0,8
0,6
0,4
0,2
0
Muestra
Control
Control
11484
Control
17190
Control
Control
11931
Control
Control
11804
11727
11411
Control
Control
16962
Control
16794
13761
13304
13270

Figura 4

Análisis de componentes principales de la población Illumina del CHOP



Análisis de componentes principales de la población IMAGE Perlegen

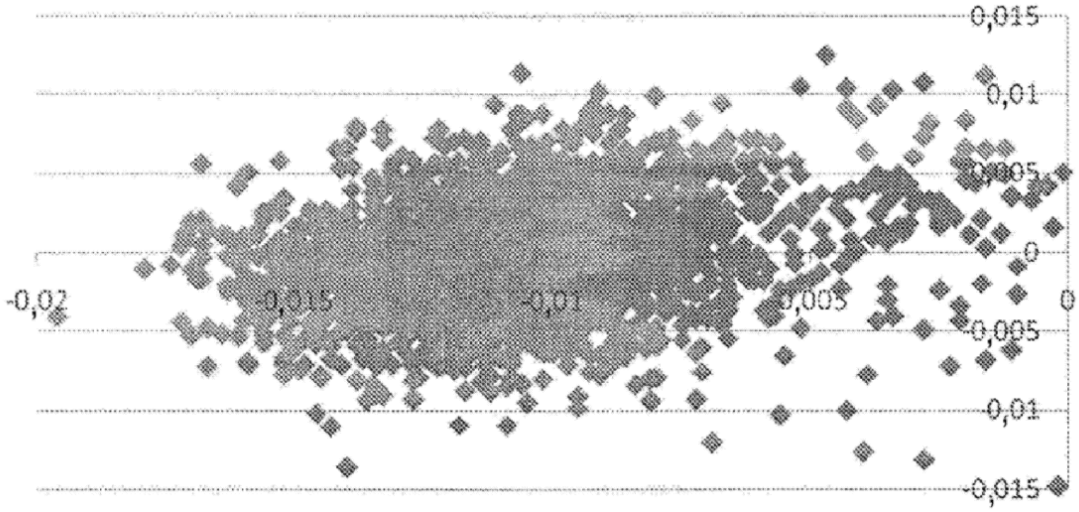


Figura 5

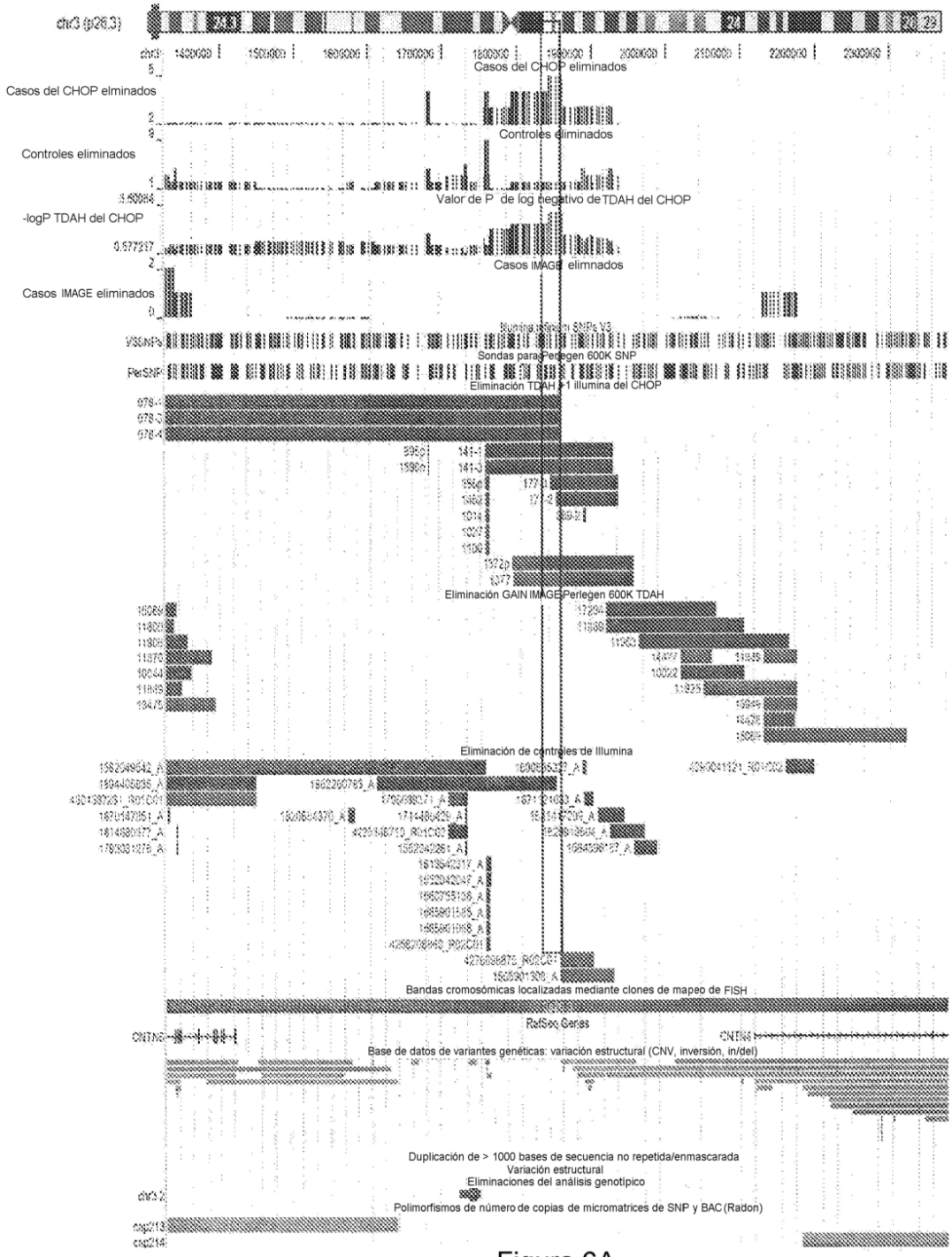

Casos del CHOP eliminados
Casos del CHOP elminados
Controles eliminados
Controles eliminados
Valor de P de log negativo de TDAH del CHOP
-logP TDAH del CHOP
Casos IMAGE elimnados
Casos IMAGE eliminados
Sondas para Perlegen 600K SNP
Eliminación TDAH +1 Illumina del CHOP
Eliminación GAIN IMAGE Perlegen 600K TDAH
Eliminación de controles de Illumina
Bandas cromosómicas localizadas mediante clones de mapeo de FISH
RefSeq Genes
Base de datos de variantes genéticas: variación estructural (CNV, inversión, in/del)
Duplicación de > 1000 bases de secuencia no repetida/enmascarada
Variación estructural
Eliminaciones del análisis genotípico
Polimorfismos de número de copias de micromatrices de SNP y BAC (Radon)

Figura 6A

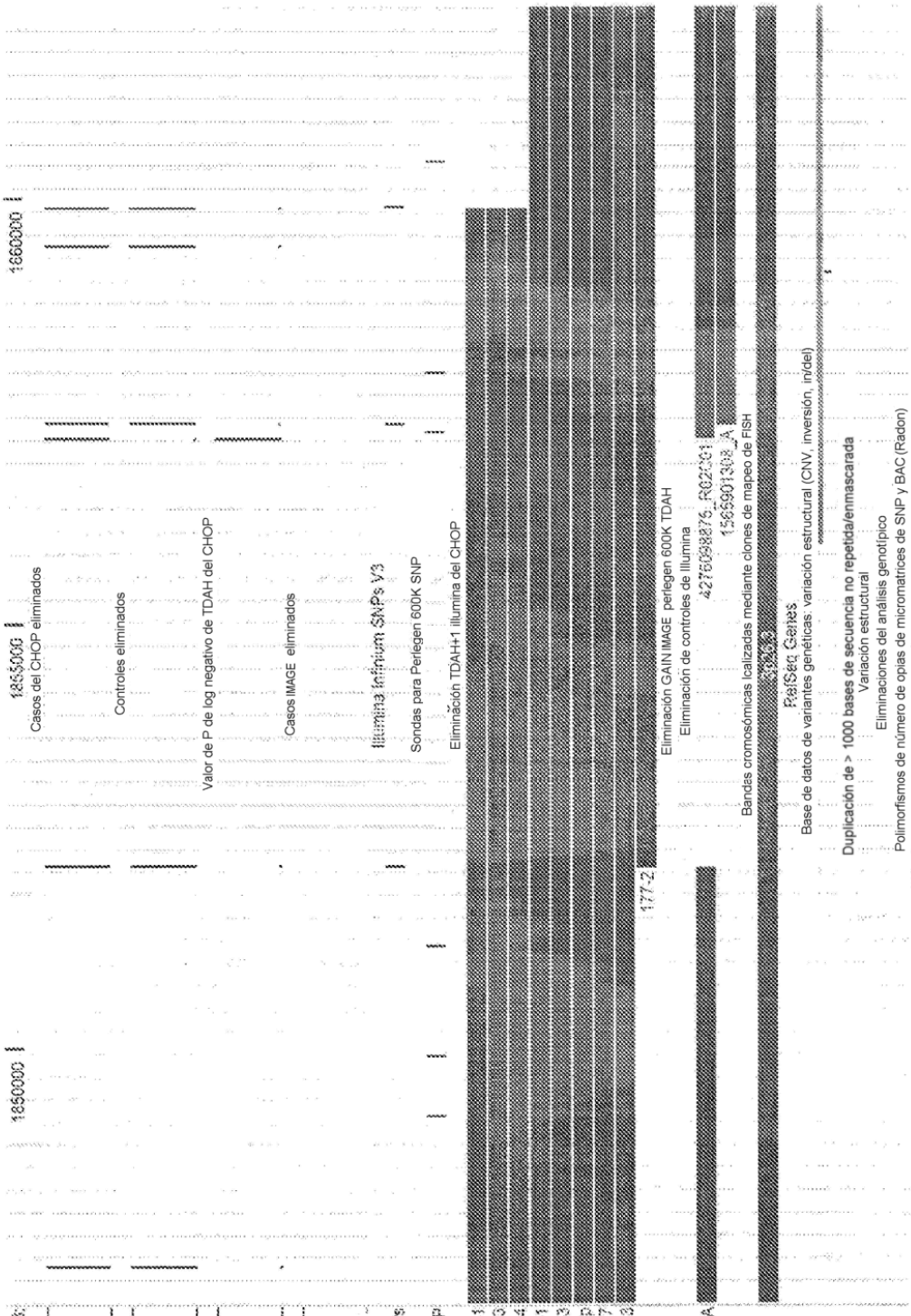

chr3:
1850000
1855000
1860000
Casos del CHOP eliminados
Controles eliminados
-logP TDAH del CHOP
Valor de P de log negativo de TDAH del CHOP
Casos IMAGEeliminados
Casos IMAGE eliminados
Sin datos
Illumina Infinium SNPs V3
V3SNPs
Sondas para Perlegen 600K SNP
PerSNP
Eliminación TDAH+1 illumina del CHOP
078-1
078-3
078-4
141-1
141-3
1372b
1377
177-3
177-2
Eliminación GAIN IMAGE perlegen 600K TDAH
Eliminación de controles de Illumina
Bandas cromosómicas localizadas mediante clones de mapeo de FISH
PerSeq Genes
Base de datos de variantes genéticas: variación estructural (CNV, inversión, in/del)
Duplicación de > 1000 bases de secuencia no repetida/enmascarada
Variación estructural
Eliminaciones del análisis genotípico
Polimorfismos de número de copias de micromatrices de SNP y BAC (Radon)

Figura 6B

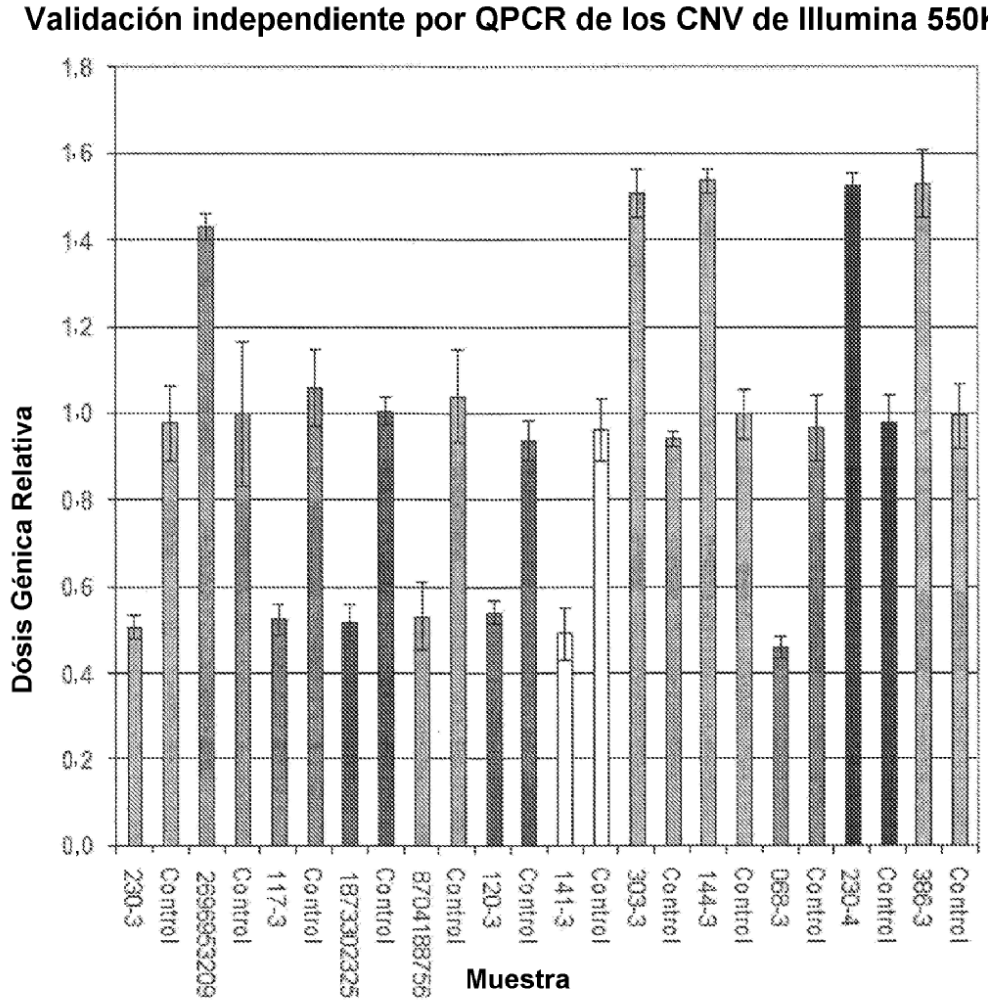

Validación independiente por QPCR de los CNV de Illumina 550K
Dósis Génica Relativa
1,8
1,6
1,4
1,2
1,0
0,8
0,6
0,4
0,2
0,0
230-3
Control
2696953209
Control
117-3
Control
1873302325
Control
8704188756
Control
120-3
Control
141-3
Control
303-3
Control
144-3
Control
068-3
Control
230-4
Control
386-3
Control
Muestra

Figura 7

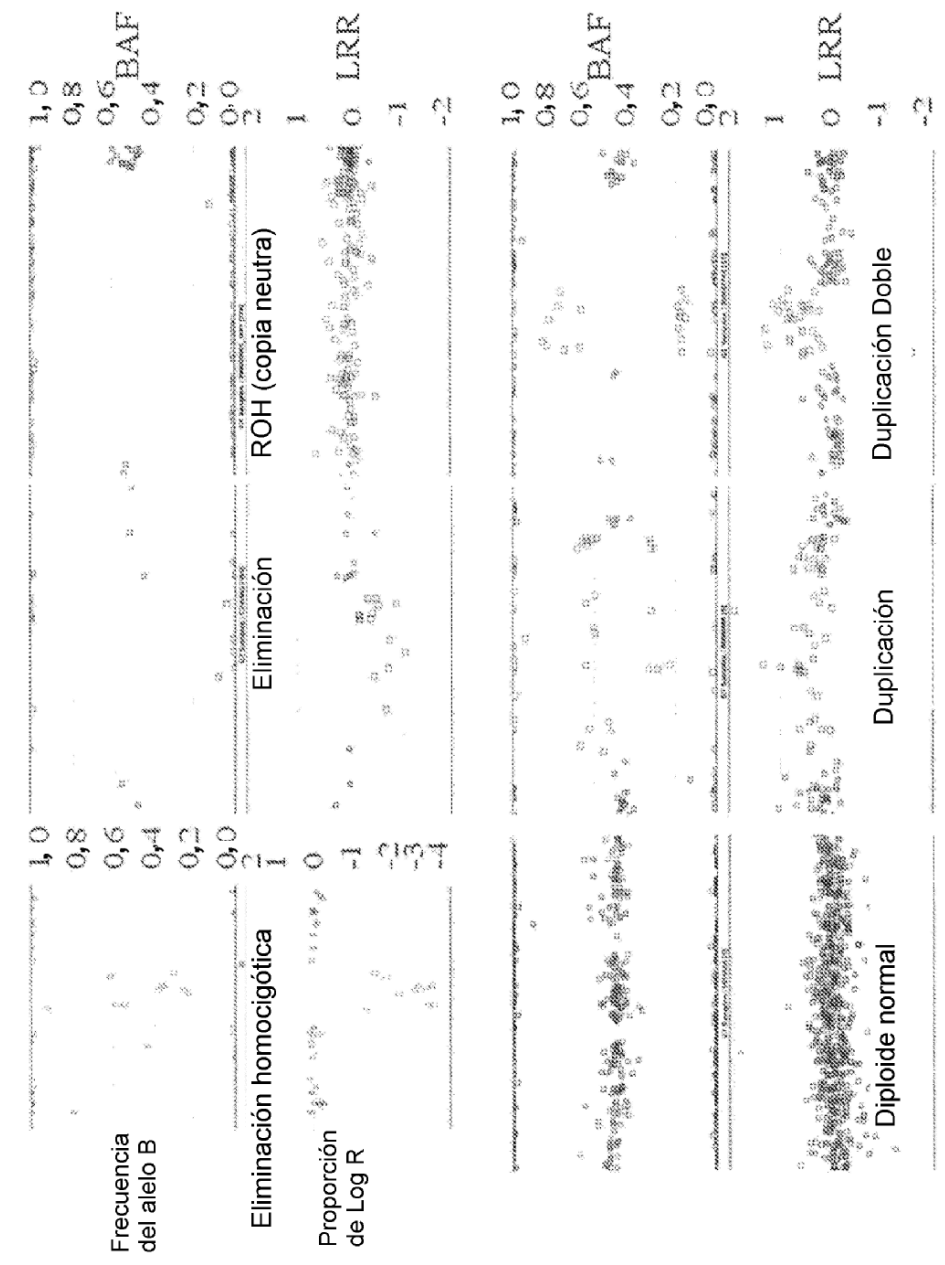

Frecuencia del alelo B
1,0
0,8
0,6
0,4
0,2
0,0
BAF
Eliminación homocigótica
Eliminación
ROH (copia neutra)
Proporción de Log R
0
-1
-2
-3
-4
LRR
Diploide normal
Duplicación
Duplicación Doble

Figura 8

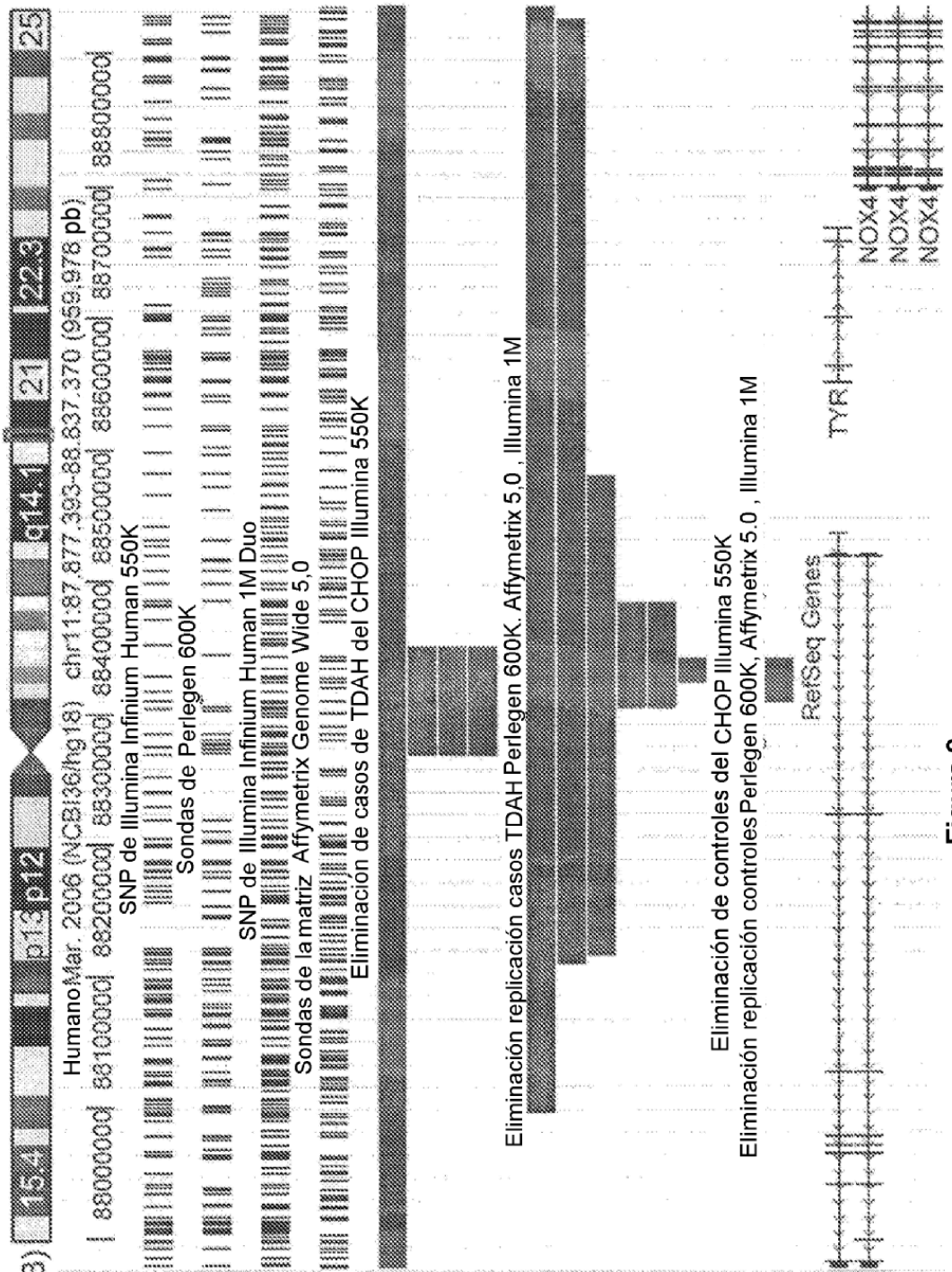

chr11 (q14.2-q14.3)
15.4
p13
p12
q14.1
21
22.3
25
Posición de la ventana
chr11:
HumanoMar. 2006 (NCBI36/hg18) chr11:87.877.393-88.837.370 (959.978 pb)
88000000
88100000
88200000
88300000
88400000
88500000
88600000
88700000
88800000
550SNP
SNP de Illumina Infinium Human 550K
Per600kSNP
Sondas de Perlegen 600K
1MSNP
SNP de Illumina Infinium Human 1M Duo
gwSNP
Sondas de la matriz Affymetrix Genome Wide 5,0
Eliminación de casos de TDAH del CHOP Illumina 550K
497
230-3
230-4
230-5
Eliminación replicación casos TDAH Perlegen 600K. Affymetrix 5,0 , Illumina 1M
16794
13304
13270
13761
17580
M.Of.M.Cs.604401
Eliminación de controles del CHOP Illumina 550K
Eliminación replicación controles Perlegen 600K, Affymetrix 5.0 , Illumina 1M
W.Fa.M.Cn.165004
RefSeq Genes
GRM5
GRM5
TYR
NOX4
NOX4
NOX4

Figura 9

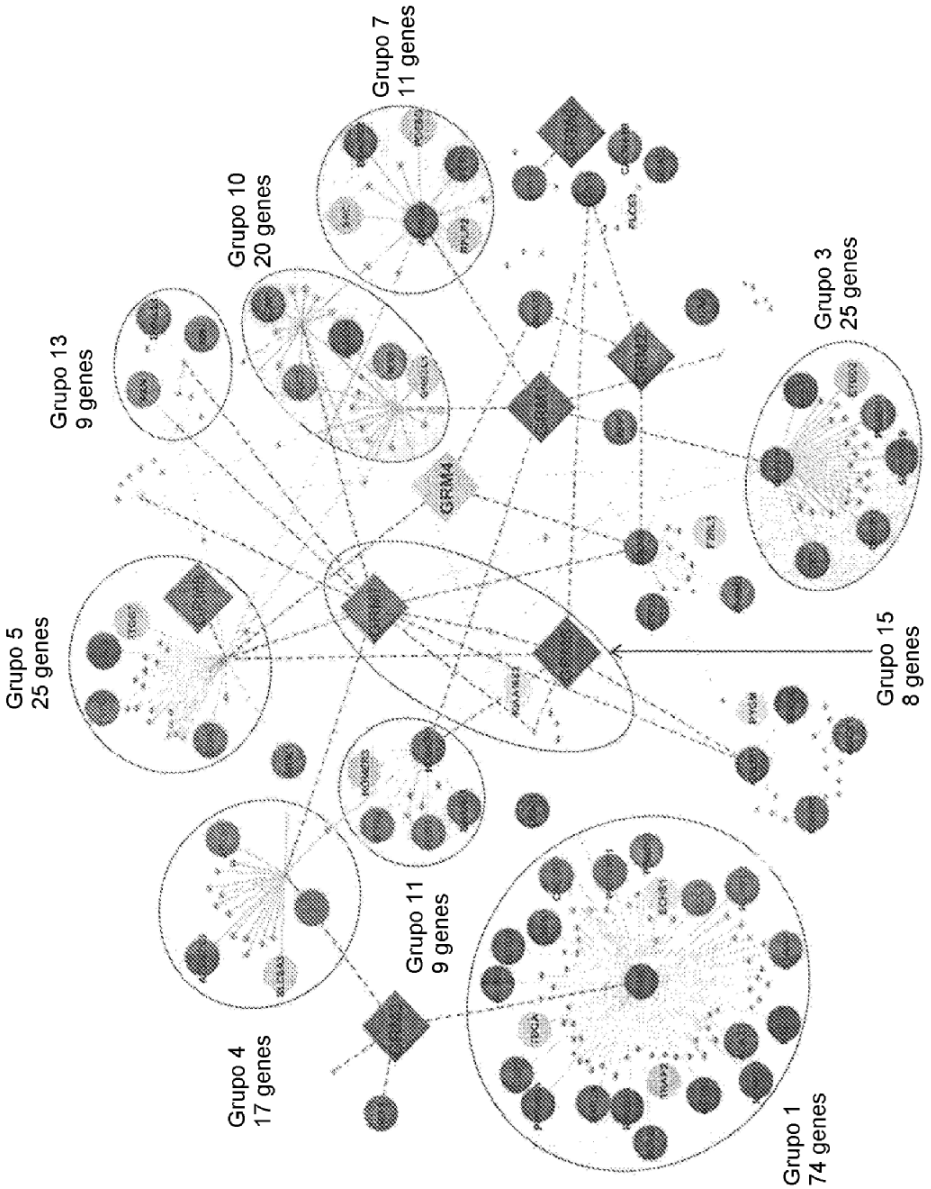

Grupo 7
11 genes
Grupo 10
20 genes
Grupo 13
9 genes
Grupo 3
25 genes
Grupo 5
25 genes
Grupo 15
8 genes
Grupo 11
9 genes
Grupo 4
17 genes
Grupo 1
74 genes

Figura 10

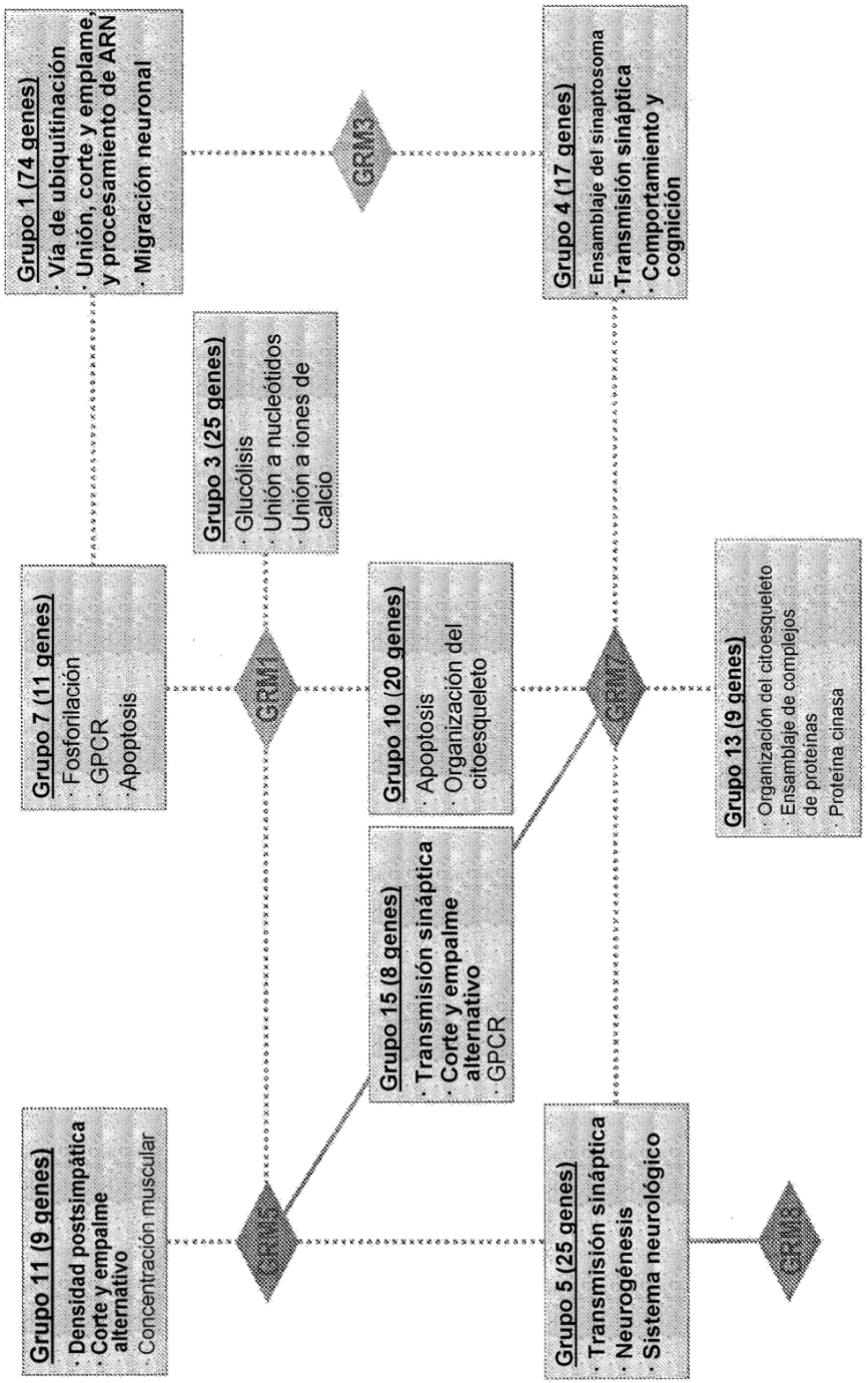

Grupo 1 (74 genes)
· Vía de ubiquitinación
· Unión, corte y emplame, y procesamiento de ARN
· Migración neuronal
GRM3
Grupo 4 (17 genes)
· Ensamblaje del sinaptosoma
· Transmisión sináptica
· Comportamiento y cognición
Grupo 3 (25 genes)
· Glucólisis
· Unión a nucleótidos
· Unión a iones de calcio
Grupo 7 (11 genes)
· Fosforilación
· GPCR
· Apoptosis
GRM1
Grupo 10 (20 genes)
· Apoptosis
· Organización del citoesqueleto
GRM7
Grupo 13 (9 genes)
· Organización del citoesqueleto
· Ensamblaje de complejos de proteínas
· Proteína cinasa
Grupo 15 (8 genes)
· Transmisión sináptica
· Corte y empalme alternativo
· GPCR
Grupo 11 (9 genes)
· Densidad postsimpática
· Corte y empalme alternativo
· Concentración muscular
GRM5
Grupo 5 (25 genes)
· Transmisión sináptica
· Neurogénesis
· Sistema neurológico
GRM8

Figura 11

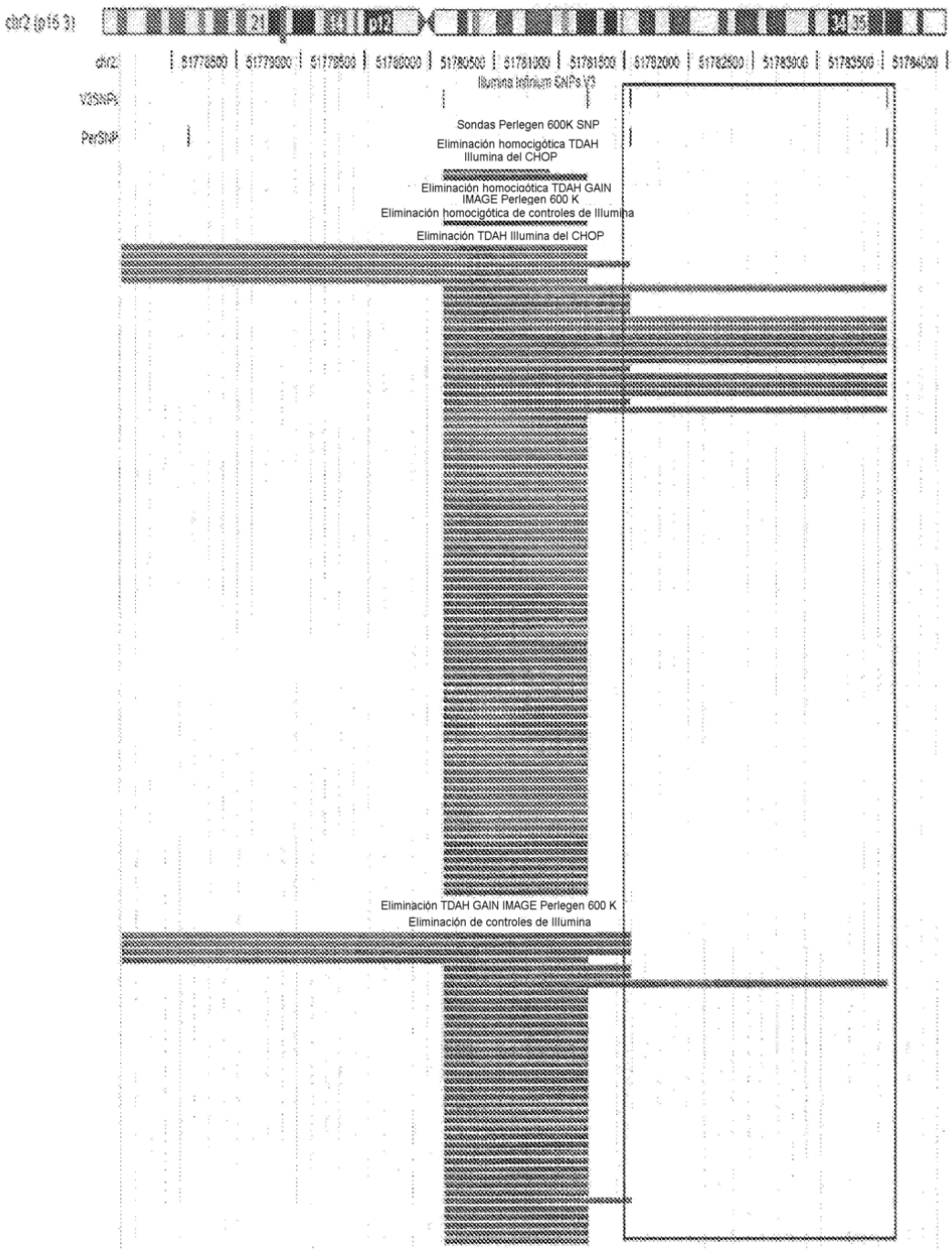

chr2 (p16.3)
chr2:
51778500
51779000
51779500
51780000
51780500
51781000
51781500
51782000
51782500
51783000
51783500
51784000
Illumina Infinium SNPs V3
V3SNPs
PerSNP
Sondas Perlegen 600K SNP
Eliminación homocigótica TDAH Illumina del CHOP
Eliminación homocigótica TDAH GAIN IMAGE Perlegen 600 K
Eliminación homocigótica de controles de Illumina
Eliminación TDAH Illumina del CHOP
Eliminación TDAH GAIN IMAGE Perlegen 600 K
Eliminación de controles de Illumina

Figura 12