



# Estudio de asociación del genoma completo

En genética, un **estudio de asociación del genoma completo** (en inglés, GWAS (*Genome-wide association study*) o WGAS (*Whole genome association study*) es un análisis de una variación genética a lo largo de todo el **genoma humano** con el objetivo de identificar su asociación a un rasgo observable. Los GWAS suelen centrarse en asociaciones entre los **polimorfismos** de un solo **nucleótido** (SNPs) y rasgos como las principales enfermedades.<sup>1</sup>

## Introducción

Para ello se deben estudiar una gran cantidad de individuos, de modo que se puedan comparar datos genéticos entre ellos. De este modo, al comparar los datos genéticos obtenidos de las secuenciaciones de los genomas de diferentes individuos, podemos hallar genes presuntamente ligados a enfermedades o caracteres. Por ejemplo, podemos contrastar cómo quizá se produce la aparición de uno o varios SNPs (variación de un solo par de bases) o una **delección**, repetición, etc. en una secuencia del genoma siempre que aparece el mismo **fenotipo**, pudiendo así concluir que este cambio a nivel genético es probable que se corresponda con un rasgo.

Estas variaciones son asociadas a distintos rasgos, por ejemplo, a enfermedades. En humanos, esta técnica ha permitido descubrir que ciertos genes están asociados a enfermedades como la **degeneración macular asociada a la edad** y la **diabetes**. En el caso de la especie humana, se estudian miles de individuos para hallar **polimorfismos de nucleótido simple** o SNPs (*single-nucleotide polymorphisms*, en inglés). Alrededor de 600 estudios de asociación del genoma completo llevados a cabo en humanos han examinado 150 rasgos y enfermedades, y han encontrado unos 800 SNPs asociados a ellas.<sup>2 3 4 5</sup>

Si ciertas variaciones genéticas son más frecuentes en personas con la enfermedad, se dice que estas variaciones están "asociadas" con esa enfermedad. Estas variaciones son entonces consideradas como señalizadores de la región del genoma humano donde probablemente esté el problema causante de la enfermedad. Se emplean dos estrategias para buscar mutaciones asociadas a la enfermedad: desde la hipótesis o sin hipótesis previa. La estrategia desde la hipótesis comienza con la formulación de la hipótesis de que un determinado gen puede estar asociado con una determinada enfermedad, y trata de buscar esa asociación. La estrategia sin hipótesis previa emplea métodos de "fuerza bruta" para escanear todo el genoma y posteriormente comprobar si algunos genes muestran esa asociación. Los estudios de asociación del genoma completo emplean generalmente la estrategia sin hipótesis previa.<sup>6</sup>

Sorprendentemente, la mayoría de las variaciones de SNPs asociadas con enfermedades no se presentan en las regiones codificantes del ADN, sino que normalmente se sitúan en las amplias regiones no codificantes intergénicas o en los **intrones** que se extraen de la secuencia de ADN cuando las proteínas son procesadas. Estas son presumiblemente secuencias de ADN que controlan otros genes, pero normalmente su función proteica no se conoce.

Aunque los estudios GWAS mayoritariamente se ha realizado para identificar SNPs asociados a enfermedades comunes, también podrían identificar variables genéticas asociadas a caracteres cuantitativos como la estatura o el intervalo QT cardiaco.<sup>7</sup>

## Metodología

Estos estudios suelen basarse en la comparación de dos grandes grupos de individuos, uno formado por personas sanas (que se tomará como control); y otro compuesto por personas afectadas por la enfermedad en cuestión, que será el grupo de casos. Se estudian la mayoría de las SNPs conocidas y más comunes en todos los individuos. El número de SNPs dependerá de la tecnología que se emplee para el estudio, pero normalmente su número oscila alrededor del millón.<sup>8</sup> Para cada SNPs, se analiza si su frecuencia alélica varía de forma significativa entre el grupo control y el grupo de casos; cuando se da esta situación, se determina que la SNP está probablemente relacionada con la enfermedad.<sup>7</sup>

Todo el análisis estadístico se realiza con la ayuda de software bioinformático; además, hay que tener en cuenta varias variables de confusión a la hora de interpretar los resultados. Entre ellas se encuentran el sexo y la edad, así como el origen geográfico y étnico de los individuos, ya que muchas variaciones genéticas están asociadas con la localización geográfica y/o la historia de las poblaciones.

La mayoría de los estudios GWAS, al menos los realizados hasta un pasado reciente (2018), se han basado en el diseño de asociación «casos y controles», en los que se compara las frecuencias alélicas en el grupo de interés con las frecuencias alélicas en el grupo de control (de estar referidos a una enfermedad: grupo de individuos afectados y grupo de no afectados por la enfermedad). Estos estudios tienen menos requerimientos, siendo así menos costosos, que los realizados mediante otros diseños, especialmente si se pueden constituir fácilmente los grupos con un número suficiente de muestras. No obstante, este método ha mostrado debilidades si el número de participantes no es lo suficientemente elevado o el grupo de casos no es representativo, o no se siguen con meticulosidad las metodologías establecidas; algo que en ocasiones ha ocurrido en este tipo de estudios, produciéndose en ellos sesgos importantes comprometiendo los resultados. Actualmente (2018) estos estudios de asociación están principalmente indicados para el estudio de enfermedades raras en los que otros diseños son poco viables.<sup>9</sup>

Existen otros dos diseños (2018): la selección de cohortes y la selección de tríos. Para la selección de cohortes, se estudia la aparición de la enfermedad en un grupo elegido a partir de una población, de esta forma se estima directamente el riesgo y se obtienen menos sesgos, pero se requiere de grupos muy grandes, mucho presupuesto, un tiempo muy largo y no es adecuado para el estudio de enfermedades raras. En cuanto a

la selección de trios, se eligen padres e hijo afecto y se buscan alelos que se hereden en más de la mitad de los hijos afectados, de esta forma, no se requiere el control de las características de la población, es más sencillo el estudio estadístico de los datos y se puede comprobar la herencia mendeliana; sin embargo, es difícil buscar casos para enfermedades de edad adulta.<sup>10</sup>

Muchos estudios GWAS utilizan etapas que involucran diferentes diseños para minimizar el riesgo de falsos positivos.

## Limitaciones

---

Los GWAS tienen diversos problemas y limitaciones que pueden subsanarse mediante un control de calidad y un diseño adecuado del estudio. Entre los problemas más comunes se encuentran la selección dificultosa de los pacientes con respecto al fenotipo de la enfermedad que se quiere estudiar, que la muestra tenga un tamaño insuficiente ya que se requiere el análisis de miles de casos y controles; y el control de la estratificación de la población. A tal efecto se ha determinado que el enfoque GWAS puede ser problemático debido a la enorme cantidad de pruebas estadísticas que requiere, ya que presentan un potencial sin precedentes para dar falsos positivos,<sup>11</sup> por lo que se requiere un sistema informático capaz de manejar correctamente esta gran cantidad de información.

Además de estos por problemas que pueden prevenirse, los GWAS han sido objeto de críticas relacionadas con su fundamento, principalmente debido a su suposición de que las variaciones genéticas comunes juegan un papel importante a la hora de explicar la variación hereditaria de las enfermedades más frecuentes. Estos estudios tienen en cuenta los SNPs de forma individual y dejan de lado sus efectos combinatorios. Estos aspectos han llevado a muchos a pensar que los GWAS tradicionales no valen la pena por la gran cantidad de dinero que es necesario invertir en ellos.

Hay que tener en cuenta, al realizar estos estudios, que hay una parte de la población con el mismo fenotipo o carácter estudiado que no puede ser explicado o que no presentan variantes comunes asociadas al mismo, pero que pueden presentar variantes raras, de baja frecuencia, que no pueden ser detectadas en el estudio. Una solución, en este caso, es el que se ha comentado anteriormente: incrementar el tamaño muestral, con el objetivo de detectar un mayor número de individuos con esas variantes poco frecuentes, sin embargo, esto no es sencillo; un ejemplo lo tenemos en los alelos que presentan una herencia recesiva y que en poblaciones de gran tamaño quedan "diluidos" entre los heterocigotos. Una segunda opción es realizar estos estudios en poblaciones aisladas genéticamente, donde es frecuente la endogamia y estos alelos raros verán incrementadas sus frecuencias y podrán ser más fácilmente detectados con una muestra de menor tamaño.

Se ha sugerido como estrategia alternativa los análisis de ligamiento pero, en la actualidad, con la rápida disminución del precio de la secuenciación completa del genoma, ésta se está convirtiendo en una alternativa realista a los de los estudios de GWAS basados en arrays, aunque muchos lo considerarían una técnica distinta. De cualquier forma, la secuenciación de alto rendimiento tiene el potencial para esquivar algunas de las limitaciones de los GWAS tradicionales.<sup>12</sup>

## GWAS y eQTLs

---

Existe millones de variantes genéticas en humanos y la interpretación de sus efectos funcionales en los caracteres humanos es uno de los mayores retos en la genómica.

Los expresión loci de rasgos cuantitativos (*expression quantitative trait loci* (eQTLs)) son loci en el DNA que actúan como elementos regulatorios de la expresión génica, la variación de estos eQTLs y sus consecuencias a nivel de expresión añade una dimensión funcional a nuestros conocimientos, la cual puede ser usada en estudios de enfermedades.

Una proporción de las variantes descritas en los GWAS son eQTLs o trQTLs, por lo que, la integración del genoma secuenciado, eQTLs y el fenotipo celular ayuda a comprender los genes causales de la enfermedad, las variantes genéticas causales que subyacen a los GWAS y los procesos biológicos que intervienen en estos.<sup>13</sup>

## Aplicaciones

---

### Fibrilación Auricular

También se han realizado estudios GWAS en enfermedades cardíacas. Un estudio de meta-análisis realizado en 2018 reveló el descubrimiento de 70 nuevos loci asociados a fibrilación auricular. Se identificaron diferentes variantes asociadas a genes que codifican para factores de transcripción, como TBX3 y TBX5, NKX2-5 o PITX2, implicados en la regulación de la conducción cardíaca, la modulación de canales iónicos y en desarrollo cardíaco. También se han identificado nuevos genes implicados en taquicardia (CASQ2) o asociados a una alteración en la comunicación de cardiomiocitos (PKP2).<sup>14</sup>

### Trastorno Bipolar

El GWAS también ha aportado valiosos conocimientos sobre la etiología genética del trastorno bipolar. En el año 2019, el Psychiatric Genomics Consortium (<https://www.med.unc.edu/pgc/>) (PGC) realizó un estudio que incluyó 29,764 pacientes con trastorno bipolar y 169,118 controles de ascendencia europea.<sup>15</sup> Se identificaron 30 loci que afectan a genes que codifican canales iónicos, transportadores de neurotransmisores y componentes sinápticos. El análisis reveló nueve conjuntos de genes significativamente enriquecidos, incluyendo la regulación de la secreción de insulina y la señalización endocannabinoide.

El último estudio GWAS realizado por el PGC publicado en *Nature Genetics* (<https://www.nature.com/ng/>) en 2021 se llevó a cabo con 41,917 casos de TB y 371,549 controles de ascendencia europea.<sup>16</sup> Identificaron 64 loci asociados al trastorno bipolar, de los cuales 33 eran nuevos hallazgos, que no se habían asociado previamente a este trastorno. Los alelos de riesgo estaban enriquecidos en los genes relacionados con la señalización sináptica y en genes que se expresan en diferentes tejidos cerebrales, en particular las neuronas piramidales del córtex prefrontal y el hipocampo. Se encontró un enriquecimiento significativo en los genes que codifican las dianas de fármacos como los antipsicóticos, los bloqueadores de los canales de calcio, los antiepilépticos y los anestésicos. La integración de los datos de expresión de locus de rasgos cuantitativos implicó a 15 genes fuertemente vinculados al trastorno bipolar a través de la expresión génica, que codifican dianas farmacológicas como HTR6, MCHR1, DCLK3 y FURIN.

## Trastorno por atracón (TA)

Estudios GWAS en el trastorno por atracón han ayudado a identificar variantes genéticas asociadas a la enfermedad. Dada la fuerte correlación de esta enfermedad con la obesidad, no había genes de riesgo para el TA en concreto, aparte de los genes asociados a la obesidad. Por ello, se busca definir la arquitectura genética del TA independientemente. Esto permite también mejorar los manuales para realizar un correcto diagnóstico diferenciando pacientes con obesidad o sobrepeso de los que padecen el TA. Además, conocer la genética asociada a la enfermedad ayuda a entender los mecanismos de la misma, así como a identificar pacientes que posean factores genéticos de predisposición al TA y actuar en prevención.

Con este objetivo, investigadores del *Center for Disease Neurogenomics*, del Departamento de Psiquiatría de la Facultad de Medicina Icahn (<http://icahn.mssm.edu/research/neurogenomics>) en Mount Sinai, Nueva York, EEUU han desarrollado estudios GWAS excluyendo el IMC. Para el estudio, desarrollan un GWAS bi-ancestral (población con herencia ancestral europea y africana) de los pacientes de la cohorte del *Million Veteran Program (MVP)* (<https://www.research.va.gov/mvp/>). Combinan los resultados de ambos GWAS en un **meta-análisis de efectos fijados (FEMA)** donde logran identificar variantes cerca de los *loci* de los genes *LRP1* (no replicado en los grupos individuales), *HFE* y *MCHR2* (replicados en el grupo ancestral europeo pero no en el africano). Ningún *loci* logró significancia a lo largo del genoma en el grupo ancestral africano. Se sabe que el gen *HFE* está implicado en la patogenia de la hemocromatosis hereditaria o sobrecarga de hierro. Ensayos posteriores para elucidar los mecanismos fisiopatológicos de la enfermedad con **estudios de anotación y mapeo funcional de GWAS (FUMA)** han identificado las vías biológicas alteradas en pacientes con la enfermedad, implicando vías del metabolismo del ácido úrico y del hierro. **Estudios de asociación a lo largo del fenoma (PheWAS)** confirman que en el trastorno por atracón se da un exceso de hierro que no se da en pacientes con un alto IMC.

Posteriormente, realizaron un **meta-análisis en gene-set (MAGMA)** para evaluar la asociación entre conjuntos de genes y los rasgos fenotípicos implicados, con el que identificaron también el *locus* de *APOE*, gen con implicaciones conocidas en enfermedades cardiovasculares, metabólicas y neurodegenerativas como el Alzheimer.

Para validar estos resultados, confirman que el IMC se está excluyendo en su modelo y que no falsea los resultados mediante validación interna para controlar que el modelo no se correlaciona genéticamente al IMC. Además, como validación externa utilizan otras tres cohortes de pacientes diferentes para confirmar que el modelo que utilizan como fenotipo de TA también es capaz de discriminar entre pacientes afectados en otras cohortes. De esta manera, confirman tanto que el modelo es reproducible como que el estudio es capaz de identificar con éxito variantes genéticas de riesgo en *loci* cercanos a genes implicados en la enfermedad del trastorno por atracón, independientemente del IMC de los pacientes.<sup>17</sup>

Tabla con los principales *loci* identificados en los estudios de GWAS y MAGMA para el modelo derivado del trastorno por atracón (umbral de significancia de los estudios GWAS:  $P \text{ valor} = 5.0 \times 10^{-8}$ )<sup>17</sup>

SNP	Cromosoma	Posición	Gen más cercano	Alelo de referencia	Alelo alternativo	Modelo ancestral africano	Modelo ancestral europeo	FEMA bi-ancestral (combinación de ambos GWAS)
<a href="https://www.ncbi.nlm.nih.gov/variation/explorer?db=core;r=6:2609746-2609810;v=rs7922007;vdb=variation;vf=407677810">rs7922007</a> (https://www.ncbi.nlm.nih.gov/variation/explorer?db=core;r=6:2609746-2609810;v=rs7922007;vdb=variation;vf=407677810)	6	26098474	<a href="https://genome.ucsc.edu/cgi-bin/hgGene?hgg_gene=ENS00000357618.10&amp;hgg_chrom=chr6&amp;hgg_start=26087428&amp;hgg_end=26098343&amp;hgg_type=knownGene&amp;db=hg38">HFE</a> (https://genome.ucsc.edu/cgi-bin/hgGene?hgg_gene=ENS00000357618.10&hgg_chrom=chr6&hgg_start=26087428&hgg_end=26098343&hgg_type=knownGene&db=hg38)	T	C	No significativo (P=0.22)	Estadísticamente significativo (P=1.9 x 10 <sup>-9</sup> )	Estadísticamente significativo (P=8.8 x 10 <sup>-10</sup> )
<a href="https://www.ncbi.nlm.nih.gov/variation/explorer?db=core;r=6:100151721-100152721;v=rs17789218;vdb=variation;vf=406988971">rs17789218</a> (https://www.ncbi.nlm.nih.gov/variation/explorer?db=core;r=6:100151721-100152721;v=rs17789218;vdb=variation;vf=406988971)	6	100600097	<a href="https://genome.ucsc.edu/cgi-bin/hgGene?hgg_gene=ENST0000281806.7&amp;hgg_chrom=chr6&amp;hgg_start=99918518&amp;hgg_end=99994223&amp;hgg_type=knownGene&amp;db=hg38">MCHR2</a> (https://genome.ucsc.edu/cgi-bin/hgGene?hgg_gene=ENST0000281806.7&hgg_chrom=chr6&hgg_start=99918518&hgg_end=99994223&hgg_type=knownGene&db=hg38)	T	C	No significativo (P=0.41)	Estadísticamente significativo (P=5.6 x 10 <sup>-10</sup> )	Estadísticamente significativo (P=4.7 x 10 <sup>-10</sup> )
<a href="https://www.ncbi.nlm.nih.gov/variation/explorer?db=core;r=6:149835365-149836365;v=rs2275046;vdb=variation;vf=406278096">rs2275046</a> (https://www.ncbi.nlm.nih.gov/variation/explorer?db=core;r=6:149835365-149836365;v=rs2275046;vdb=variation;vf=406278096)	6	150157001	<a href="https://genome.ucsc.edu/cgi-bin/hgGene?hgg_gene=ENST0000239367.7&amp;hgg_chrom=chr6&amp;hgg_start=149818756&amp;hgg_end=149864359&amp;hgg_type=knownGene&amp;db=hg38">LRP11</a> (https://genome.ucsc.edu/cgi-bin/hgGene?hgg_gene=ENST0000239367.7&hgg_chrom=chr6&hgg_start=149818756&hgg_end=149864359&hgg_type=knownGene&db=hg38)	A	G	No significativo (P=8.3 x 10 <sup>-4</sup> )	No significativo (P=1.3 x 10 <sup>-6</sup> )	Estadísticamente significativo (P=1.1 x 10 <sup>-8</sup> )
<a href="https://www.ncbi.nlm.nih.gov/variation/explorer?db=core;r=19:44908322-44909322;v=rs7412">rs7412</a> (https://www.ncbi.nlm.nih.gov/variation/explorer?db=core;r=19:44908322-44909322;v=rs7412)	19	45412079	<a href="https://genome.ucsc.edu/cgi-bin/hgGene?hgg_gene=ENST00002524">APOE</a> (https://genome.ucsc.edu/cgi-bin/hgGene?hgg_gene=ENST00002524)	C	T	Identificado por MAGMA, no alcanzó significancia en los modelos GWAS (P=0.17)	Identificado por MAGMA, no alcanzó significancia en los modelos GWAS (P=4.4 x 10 <sup>-7</sup> )	Identificado por MAGMA, no alcanzó significancia en los modelos GWAS (P=3.7 x 10 <sup>-7</sup> )

2;vdb=variation;vf=1024563163)		86.9&hg g_chrom=chr19&hg_start=44905795&hg_end=44909393&hg_type=knownGene&db=hg38)				
--------------------------------	--	---	--	--	--	--

## Diabetes

Los estudios de GWAS han permitido asociar diversos alelos con el riesgo de desarrollar diabetes. En concreto, un estudio de 2007 evidenció un mayor riesgo de padecer esta patología en los individuos portadores de determinados SNP en el locus *TCF7L2*.<sup>18</sup> Asimismo, otro estudio reciente asoció la presencia del SNP rs10824307, cercano al locus *LRMDA*, con el desarrollo de diabetes leve asociada a obesidad.<sup>19</sup>

## Dislexia

Numerosos estudios GWAS también se han llevado a cabo en el campo de la dislexia con el objetivo de identificar marcadores genéticos convincentes que puedan ser utilizados tanto para el diagnóstico como para poder comprender los mecanismos biológicos subyacentes a esta condición. En particular, un estudio realizado en 2022<sup>20</sup> analizó a 51,800 pacientes con dislexia (21,513 hombres, 30,287 mujeres) y 1,087,070 pacientes control (446,054 hombres, 641,016 mujeres), logrando identificar **42 loci significativamente asociados con este trastorno**.

De los 42 loci genómicos identificados, 17 ya habían sido descritos previamente en la bibliografía, de los cuales, 15 se relacionan con rasgos cognitivos y educativos. Mientras tanto, los 27 loci restantes aún no han sido mencionados en la literatura, considerándolos nuevos y siendo una fuente importante de investigación futura para determinar su potencial utilidad en el diagnóstico de este trastorno. Con respecto a estos genes identificados, destacan:

- **Gen *AUTS2***: se ha relacionado con autismo y discapacidad intelectual. La variante que se ha descubierto (rs3735260) representa la asociación SNP de *AUTS2* más fuerte identificada con un rasgo del neurodesarrollo hasta la fecha. Este gen tiene un papel fundamental en el neurodesarrollo y en los procesos de migración de las neuronas
- **Gen *TANC2***: se ha relacionado con retraso del lenguaje y discapacidad intelectual. Este gen controla las conexiones entre las células cerebrales.
- **Gen *GGNBP2***: se ha relacionado con retraso del neurodesarrollo y autismo. Este gen es requerido para el desarrollo sináptico de las neuronas motoras.

## Otros ejemplos

Una aplicación clara de este tipo de estudios es la asociación de un síntoma patológico o patología con un determinado locus. Un ejemplo de ellos son los estudios de asociación de la alergia a determinados alérgenos con la susceptibilidad a loci.

No podemos olvidar que también pueden encontrarse asociaciones con loci que confieren ventajas o protección frente a un posible rasgo patológico. Un ejemplo lo encontramos en un estudio reciente publicado en la revista Nature, en el que tras secuenciar el genoma de 2120 individuos de la isla de Cerdeña (considerada dicha población como aislado genético), se realizaron estudios de asociación de las variantes genéticas raras, con el perfil lipídico y marcadores de inflamación dentro de la población. Así pudieron encontrar, entre otros, una variante en el gen ***APOA5*** (Arg282Ser) que se asociaba a menores niveles de triglicéridos en el suero de sus portadores, lo que sugiere un posible papel protector frente a enfermedades cardiovasculares.<sup>21</sup> Otro ejemplo sería el **Proyecto UK10K** (10.000 genomas de Reino Unido), en el que se combinan estudios WGS (Secuenciación del Genoma Completo) y GWAS para intentar identificar variantes genéticas raras o poco frecuentes relacionadas con problemas como la obesidad, el autismo, la esquizofrenia o algunas enfermedades raras.<sup>22</sup>

La asociación de un locus a una determinada patología nos permite realizar un seguimiento más completo a los pacientes que tengan presente este alelo de riesgo y por tanto nos permitirá un diagnóstico precoz de muchas enfermedades que puede salvar vidas como es el caso del diagnóstico precoz en cáncer de páncreas.<sup>23</sup>

En la revista Nature, se ha publicado un artículo que se basa en el GWAS con el objetivo de determinar y encontrar cuantos loci genéticos están asociados con el peso al nacer. Gracias a este tipo de análisis genómico se han identificado 60 loci que están relacionados, y ha permitido concluir que existe una fuerte correlación entre bajo peso al nacer y la probabilidad de padecer enfermedades metabólicas en el adulto.<sup>24</sup>

También a través de GWAS se llevó a cabo el estudio de genes que podían estar relacionados con el volumen intracraneal de las personas, descubriéndose 5 loci anteriormente desconocidos (6q21, 10q24, 3q28, 12q14, 12q23). Este trabajo, publicado en la revista Nature, permitió además establecer relaciones entre dicho volumen intracraneal y otros rasgos antropométricos, función cognitiva, enfermedades neurodegenerativas (Parkinson), etc.<sup>25</sup>

De igual modo, también se ha logrado determinar mediante GWAS en un estudio publicado en la revista *Nature Genetics* en 2018 la existencia de 190 nuevos loci relacionados con la inteligencia que suman un total de 205 loci con los previamente descubiertos y, así mismo, 939 genes nuevos para un total de 1016 genes relacionados con la inteligencia. De esta manera, también se ha podido establecer que dichos genes asociados se encuentran mayormente expresados en el cerebro, específicamente en tejido neuronal estriatal y piramidal del hipocampo.

Mediante GWAS también se pueden estudiar genes que pueden estar relacionados con el comportamiento humano reproductivo, mediante la identificación de determinados loci que influyen en la actitud de las personas en referencia a la reproducción. En la revista *Nature*<sup>26</sup> se publicó un artículo en el que se identificaron 12 loci mediante GWAS, localizados en los cromosomas 1,2,3,5,6,7,8 y 20, que presentaban una función potencial en la reproducción y la fertilidad.

En un estudio publicado en la revista *Nature Communications* en 2019, se presentan los resultados de un extenso GWAS realizado sobre población de Reino Unido con ascendencia europea, que reveló la asociación de 42 loci a lo largo del genoma humano relacionados con la somnolencia diurna de los individuos y que además, agrupaban a los individuos en 2 subtipos diferentes: propensión al sueño y fragmentación del sueño.<sup>27</sup>

En otro estudio GWAS publicado en 2019 en la revista *Nature Genetics*, se identifican nuevos loci y vías de señalización implicadas en el riesgo de *Alzheimer*. El meta-análisis es llevado a cabo en un total de 455.000 personas europeas de diferentes cohortes, entre las cuales se incluyen pacientes con *Alzheimer* o *Alzheimer-by-proxy* y controles o *controles-by-proxy*, con el fin de ampliar el tamaño muestral. Esto permite a los autores identificar 29 loci de riesgo, de los cuales 9 aún no se habían asociado al *Alzheimer*. Asimismo, se observa que estos loci de riesgo afectan a 215 genes, de los cuales 16 fueron identificados mediante las cuatro estrategias de mapeo génico empleadas en este artículo. Por último, los autores encuentran que estos genes están altamente expresados en tejidos y tipos celulares relacionados con el sistema inmune (bazo, hígado y microglía) y están implicados en procesos lipídicos y en la degradación de proteínas precursoras amiloides, fortaleciendo la hipótesis de que cambios lipídicos podrían estar implicados en la patogénesis del *Alzheimer*. Por todo ello, estos resultados obtenidos en este estudio GWAS sugieren nuevos genes prometedores sobre los que realizar futuros experimentos, desarrollar fármacos o estratificar pacientes.<sup>28</sup>

El último experimento mediante el estudio de la asociación del genoma completo ha sido publicado en enero de 2020, en la revista *The American Journal of Psychiatry* (<https://ajp.psychiatryonline.org/>), en el que se analizaron casi 200.000 genomas de participantes en el "*Million Veteran Program* (<https://www.research.va.gov/mvp/>)", una de las bases de datos genéticas más grandes del mundo. Este estudio tenía como objetivo examinar la arquitectura genética de la ansiedad y sus síntomas. Los autores han identificado 5 señales significantes en todo el genoma para americanos europeos, y 1 para los afroamericanos. Los resultados también evidenciaron la correlación genética entre la ansiedad, depresión y neuroticismo.<sup>29</sup>

Además de sus aplicaciones en humanos, vinculadas sobre todo al campo de la medicina, los GWAS se han utilizado en animales domésticos<sup>30</sup> y en plantas de interés agrícola, con objetivo de identificar loci relacionados con características valiosas desde el punto de vista productivo. Entre otros ejemplos, en ganado se han hecho GWAS estudiando la heredabilidad de los rasgos relacionados con la calidad del producto,<sup>31</sup> y también se ha estudiado la resistencia a enfermedades;<sup>32</sup> mientras, en plantas se pueden encontrar estudios de asociación centrados en los loci que controlan la producción de biomasa.<sup>33 34</sup>

## Referencias

1. Manolio, Teri A. (2 de julio de 2010). «Genomewide Association Studies and Assessment of the Risk of Disease» (<https://www.nejm.org/doi/10.1056/NEJMra0905980>). <https://doi.org/10.1056/NEJMra0905980> (en inglés). doi:10.1056/nejmra0905980 (<https://dx.doi.org/10.1056%2Fnejmra0905980>). Consultado el 16 de julio de 2021.
2. Roberts, Paul (15 de marzo de 2005). «WebTV Virus Writer Sentenced to Prison» ([https://web.archive.org/web/20120707060121/http://www.pcworld.com/article/120050/webtv\\_virus\\_writer\\_sentenced\\_to\\_prison.html](https://web.archive.org/web/20120707060121/http://www.pcworld.com/article/120050/webtv_virus_writer_sentenced_to_prison.html)) (en inglés). IDG News. Archivado desde el original ([http://www.pcworld.com/article/120050/webtv\\_virus\\_writer\\_sentenced\\_to\\_prison.html](http://www.pcworld.com/article/120050/webtv_virus_writer_sentenced_to_prison.html)) el 7 de julio de 2012. Consultado el 15 de julio de 2010.
3. Manolio TA (2010). «Genomewide association studies and assessment of the risk of disease» (<http://content.nejm.org/cgi/content/full/363/2/166>). *N Engl J Med* **363**: 166.
4. Manolio TA; Pearson TA (2008). «How to interpret a genome-wide association study» (<http://jama.ama-assn.org/cgi/content/full/299/11/1335>). *JAMA* **299** (11): 1335–1344. PMID 18349094 (<https://www.ncbi.nlm.nih.gov/pubmed/18349094>). doi:10.1001/jama.299.11.1335 (<https://dx.doi.org/10.1001%2Fjama.299.11.1335>).
5. Genome.gov (ed.). «Genome-Wide Association Studies» (<http://www.genome.gov/20019523>). *National Human Genome Research Institute, National Institutes of Health*. (Explicación más simple)
6. Hunter DJ, Altshuler D, Rader DJ (junio de 2008). «From Darwin's Finches to Canaries in the Coal Mine — Mining the Genome for New Biology» (<http://content.nejm.org/cgi/content/full/358/26/2760>). *N Engl J Med* **358**: 2760-2763.
7. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (September 2007). "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950838>). *Am. J. Hum. Genet.* **81** (3): 559–75. doi: 10.1086/519795 (<https://dx.doi.org/10.1086%2F519795>). PMC 1950838 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1950838>). PMID 17701901 (<https://www.ncbi.nlm.nih.gov/pubmed/17701901>).
8. Bush WS, Moore JH (2012). "Chapter 11: genome-wide association studies" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3531285>). In Lewitter, Fran; Kann, Maricel. *PLoS Comput Biol* **8** (12): e1002822. doi: 10.1371/journal.pcbi.1002822 (<https://dx.doi.org/10.1371%2Fjournal.pcbi.1002822>). PMC 3531285 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3531285>). PMID 23300413 (<https://www.ncbi.nlm.nih.gov/pubmed/23300413>)

9. Pearson TA, Manolio TA. How to Interpret a Genome-wide Association Study. *JAMA*. 2008; 299 (11): 1335-1344. doi: 10.1001/jama.299.11.1335, pp 2,3. <https://jamanetwork.com/journals/jama/article-abstract/181647> (consultado: 05/03/2018).
10. Apuntes de la asignatura de Diagnóstico Molecular y Genético de 3º del Grado en Biotecnología de la Universidad Pablo de Olavide.
11. Pearson TA, Manolio TA (March 2008). "How to interpret a genome-wide association study". *JAMA* **299** (11): 1335–44. doi: 10.1001/jama.299.11.1335 (<https://dx.doi.org/10.1001%2Fjama.299.11.1335>). PMID 18349094 (<https://www.ncbi.nlm.nih.gov/pubmed/18349094>)
12. Visscher PM, Goddard ME, Derks EM, Wray NR (June 2011). "Evidence-based psychiatric genetics, AKA the false dichotomy between common and rare variant hypotheses". *Mol Psychiatry* **17** (5): 474–85. doi: 10.1038/mp.2011.65 (<https://dx.doi.org/10.1038%2Fmp.2011.65>). PMID 21670730 (<https://www.ncbi.nlm.nih.gov/pubmed/21670730>)
13. Lappalainen T. et al. (2013). "Transcriptome and genome sequencing uncovers functional variation in humans". *Nature* (501): 506-511.
14. Roselli, C., Chafin, M., Weng, L., et al. (2018). «Multi-ethnic genome-wide association study for atrial fibrillation». *Nature Genetics* **50** (9): 1225-1233. PMID 29892015 (<https://www.ncbi.nlm.nih.gov/pubmed/29892015>). doi:10.1038/s41588-018-0133-9 (<https://dx.doi.org/10.1038%2Fs41588-018-0133-9>).
15. Stahl, E. A. et al. (2019). «Genome-wide association study identifies 30 loci associated with bipolar disorder». *Nature Genetics* **51**, 793–803. doi:10.1038/s41588-019-0397-8 (<https://dx.doi.org/10.1038%2Fs41588-019-0397-8>).
16. Mullins, N., Forstner, A.J., O'Connell, K.S. et al. (2021). «Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology». *Nature Genetics* **53**, 817–829. doi:10.1038/s41588-021-00857-4 (<https://dx.doi.org/10.1038%2Fs41588-021-00857-4>).
17. Burstein, David; Griffen, Trevor C.; Therrien, Karen; Bendl, Jaroslav; Venkatesh, Sanan; Dong, Pengfei; Modabbernia, Amirhossein; Zeng, Biao et al. (2023-09). «Genome-wide analysis of a model-derived binge eating disorder phenotype identifies risk loci and implicates iron metabolism» (<https://www.nature.com/articles/s41588-023-01464-1>). *Nature Genetics* (en inglés) **55** (9): 1462-1470. ISSN 1546-1718 (<https://portal.issn.org/resource/issn/1546-1718>). doi:10.1038/s41588-023-01464-1 (<https://dx.doi.org/10.1038%2Fs41588-023-01464-1>). Consultado el 1 de febrero de 2024.
18. Billings, Liana K.; Florez, Jose C. (de noviembre de 2010). «The genetics of type 2 diabetes: what have we learned from GWAS?: Genetics of type 2 diabetes». *Annals of the New York Academy of Sciences* **1212** (1): 59-77. doi:10.1111/j.1749-6632.2010.05838.x (<https://dx.doi.org/10.1111%2Fj.1749-6632.2010.05838.x>).
19. Mansour Aly, Dina; Dwivedi, Om Prakash; Prasad, Rashmi B.; Käräjämäki, Annemari; Hjort, Rebecka; Thangam, Manonanthini; Åkerlund, Mikael; Mahajan, Anubha; Udler, Miriam S.; Florez, Jose C.; McCarthy, Mark I.; Abecasis, Goncalo; Baras, Aris; Cantor, Michael; Coppola, Giovanni; Economides, Aris; Lotta, Luca A.; Overton, John D.; Reid, Jeffrey G.; Shuldiner, Alan; Beechert, Christina; Forsythe, Caitlin; Fuller, Erin D.; Gu, Zhenhua; Lattari, Michael; Lopez, Alexander; Schleicher, Thomas D.; Padilla, Maria Sotiropoulos; Widom, Louis; Wolf, Sarah E.; Pradhan, Manasi; Manoochchhari, Kia; Ulloa, Ricardo H.; Bai, Xiaodong; Balasubramanian, Suganthi; Blumenfeld, Andrew; Eom, Gisu; Habegger, Lukas; Hawes, Alicia; Khalid, Shareef; Maxwell, Evan K.; Salerno, William; Staples, Jeffrey C.; Jones, Marcus B.; Mitnaul, Lyndon J.; Brosnan, Julia; Melander, Olle; Carlsson, Sofia; Hansson, Ola; Tuomi, Tiinamaija; Groop, Leif; Ahlqvist, Emma (de noviembre de 2021). «Genome-wide association analyses highlight etiological differences underlying newly defined subtypes of diabetes». *Nature Genetics* **53** (11): 1534-1542. doi:10.1038/s41588-021-00948-2 (<https://dx.doi.org/10.1038%2Fs41588-021-00948-2>).
20. Doust, Catherine; Fontanillas, Pierre; Eising, Else; Gordon, Scott D.; Wang, Zhengjun; Alagöz, Gökberk; Molz, Barbara; Pourcain, Beate St; Francks, Clyde; Marioni, Riccardo E.; Zhao, Jingjing; Paracchini, Silvia; Talcott, Joel B.; Monaco, Anthony P.; Stein, John F.; Gruen, Jeffrey R.; Olson, Richard K.; Willcutt, Erik G.; DeFries, John C.; Pennington, Bruce F.; Smith, Shelley D.; Wright, Margaret J.; Martin, Nicholas G.; Auton, Adam; Bates, Timothy C.; Fisher, Simon E.; Luciano, Michelle (de noviembre de 2022). «Discovery of 42 genome-wide significant loci associated with dyslexia». *Nature Genetics* **54** (11): 1621-1629. doi:10.1038/s41588-022-01192-y (<https://dx.doi.org/10.1038%2Fs41588-022-01192-y>).
21. Genome sequencing elucidates Sardinian genetic architecture and augments association analyses for lipid and blood inflammatory markers (<http://www.nature.com/ng/journal/v47/n11/full/ng.3368.html>) (enlace roto disponible en Internet Archive; véase el historial ([https://web.archive.org/web/\\*/http://www.nature.com/ng/journal/v47/n11/full/ng.3368.html](https://web.archive.org/web/*/http://www.nature.com/ng/journal/v47/n11/full/ng.3368.html)), la primera versión (<https://web.archive.org/web/1/http://www.nature.com/ng/journal/v47/n11/full/ng.3368.html>) y la última (<https://web.archive.org/web/2/http://www.nature.com/ng/journal/v47/n11/full/ng.3368.html>)).
22. «UK10K» (<http://www.uk10k.org/>). Consultado el 29 de marzo de 2016.
23. M Wolpin, Brian; Rizzato, Cosmeri (2014). «Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer» (<http://www.ncbi.nlm.nih.gov/pubmed/25086665>). *Nature genetics* **46**: 994–1000. PMID 25086665 (<https://www.ncbi.nlm.nih.gov/pubmed/25086665>).
24. Horikoshi et al. (13 de octubre de 2016). «Genome-wide associations for birth weight and correlations with adult disease» (<http://www.ncbi.nlm.nih.gov/pubmed/27680694>). *Nature* **538**: 248-267. doi:10.1038/nature19806 (<https://dx.doi.org/10.1038%2Fnature19806>). Consultado el 12 de febrero de 2017.
25. Adams, Hieab H H; Hibar, Derrek P; Chouraki, Vincent; Stein, Jason L; Nyquist, Paul A; Rentería, Miguel E; Trompet, Stella; Arias-Vasquez, Alejandro et al.. «Novel genetic loci underlying human intracranial volume identified through genome-wide association» (<http://www.nature.com/doi/10.1038/nn.4398>). *Nature Neuroscience* **19** (12): 1569-1582. PMID 27694991 (<http://www.ncbi.nlm.nih.gov/pubmed/27694991>). doi:10.1038/nn.4398 (<https://dx.doi.org/10.1038%2Fnn.4398>).
26. <http://www.nature.com/ng/journal/v48/n12/full/ng.3698.html>
27. Wang, Heming; Lane, Jacqueline M.; Jones, Samuel E.; Dashti, Hassan S.; Ollila, Hanna M.; Wood, Andrew R.; van Hees, Vincent T.; Brumpton, Ben et al. (2019-12). «Genome-wide association analysis of self-reported daytime sleepiness identifies 42 loci that suggest biological subtypes» (<http://www.nature.com/articles/s41467-019-11456-7>). *Nature Communications* (en inglés) **10** (1): 3503. ISSN 2041-1723 (<https://portal.issn.org/resource/issn/2041-1723>). doi:10.1038/s41467-019-11456-7 (<https://dx.doi.org/10.1038%2Fs41467-019-11456-7>). Consultado el 19 de diciembre de 2019.
28. Jansen, I.E.; Savage, J.E.; Watanabe, K.; Bryois, J.; Williams, D.M.; Steinberg, S.; Sealock, J.; Karlsson, I.K. et al. (2019). «Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk» (<https://www.nature.com/articles/s41588-018-0311-9>). *Nature Genetics* **51** (3): 404-413. doi:10.1038/s41588-018-0311-9 (<https://dx.doi.org/10.1038%2Fs41588-018-0311-9>). Consultado el 16 de diciembre de 2019.

29. «Reproducible Genetic Risk Loci for Anxiety: Results From 200.000 Participants in the Million Veteran Program» (<https://ajp.psychiatryonline.org/doi/10.1176/appi.ajp.2019.19030256>). *The American Journal of Psychiatry*. 7 de enero de 2020. doi:10.1176/appi.ajp.2019.19030256 (<https://dx.doi.org/10.1176%2Fappi.ajp.2019.19030256>).
30. Zhang, H; Wang, Z (2012). «Progress of genome wide association study in domestic animals» (<https://www.ncbi.nlm.nih.gov/pubmed/22958308>). *Journal of Animal Science and Biotechnology*. PMID 22958308 (<https://www.ncbi.nlm.nih.gov/pubmed/22958308>). doi:10.1186/2049-1891-3-26 (<https://dx.doi.org/10.1186%2F2049-1891-3-26>).
31. Lee, Young-Sup; Jeong, Hyeonsoo (2015). «Genome-wide Association Study (GWAS) and Its Application for Improving the Genomic Estimated Breeding Values (GEBV) of the Berkshire Pork Quality Traits» (<http://ajas.info/journal/view.php?number=23399>). *Asian-Australasian Journal of Animal Sciences*. doi:10.5713/ajas.15.0287 (<https://dx.doi.org/10.5713%2Fajas.15.0287>).
32. Lee, Bo-Young; Lee, Kwang-Nyeong (2015). «Bovine Genome-wide Association Study for Genetic Elements to Resist the Infection of Foot-and-mouth Disease in the Field» (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4283160/>). *Asian-Australasian Journal of Animal Sciences*. doi:10.5713/ajas.14.0383 (<https://dx.doi.org/10.5713%2Fajas.14.0383>).
33. Salas Fernández, María G.; Becraft, Philip W.; Yin, Yanhai; Lübberstedt, Thomas (2009). «From dwarves to giants? Plant height manipulation for biomass yield» (<http://www.sciencedirect.com/science/article/pii/S1360138509001666>). *Trends in Plant Science*. PMID 22958308 (<https://www.ncbi.nlm.nih.gov/pubmed/22958308>). doi:10.1016/j.tplants.2009.06.005 (<https://dx.doi.org/10.1016%2Fj.tplants.2009.06.005>).
34. Salas Fernández, María G.; Mantilla Pérez, María B.; Zhao, Jing; Hu, Jieyun (2016). «Genome-Wide Association Study for Nine Plant Architecture Traits in Sorghum» (<https://web.archive.org/web/20170222202024/https://dl.sciencesocieties.org/publications/tpg/pdfs/0/0/plantgenome2015.06.0044>). *Plant Genome*. doi:10.3835/plantgenome2015.06.0044 (<https://dx.doi.org/10.3835%2Fplantgenome2015.06.0044>). Archivado desde el original (<https://dl.sciencesocieties.org/publications/tpg/pdfs/0/0/plantgenome2015.06.0044>) el 22 de febrero de 2017.

## Enlaces externos

---

- [Un catálogo de los estudios de asociación del genoma completo publicados \(http://www.genome.gov/gwastudies/\)](http://www.genome.gov/gwastudies/)
- 

Obtenido de «[https://es.wikipedia.org/w/index.php?title=Estudio\\_de\\_asociaci3n\\_del\\_genoma\\_completo&oldid=160127603](https://es.wikipedia.org/w/index.php?title=Estudio_de_asociaci3n_del_genoma_completo&oldid=160127603)»

■