

Disputas en torno al ADN basura

Laureano Castro Nogueira
20 febrero, 2017



NESSA CAREY

ADN basura

UN VIAJE POR LA MATERIA OSCURA
DEL GENOMA HUMANO



BIBLIOTECA BURIDÁN

ADN basura. Un viaje por la materia oscura del genoma humano

Nessa Carey

Barcelona, Biblioteca Buridán, 2015 356 pp. 30 €

Trad. de Josep Sarret Grau

El cigoto, la célula inicial de la que surge un ser humano, se forma por la fusión de un espermatozoide, aportado por el padre, y de un óvulo, aportado por la madre. Cada uno de los gametos suministra una pequeña cantidad de ADN (unas cinco billonésimas de gramo) en la que están contenidas las instrucciones para fabricar las proteínas y establecer el metabolismo. Las moléculas de ADN se forman mediante la repetición de sólo cuatro unidades o nucleótidos diferentes y se componen de dos cadenas trenzadas, constituyendo una estructura típica, la famosa *dobles hélice*, que fue desvelada por James Francis Crick en 1953.

La información hereditaria está escrita en la secuencia de nucleótidos del ADN de cada gen, que determina, a su vez, la secuencia de aminoácidos de una proteína y, con ella, la estructura que adoptará la proteína en las condiciones ambientales en que se desenvolverán los distintos seres vivos y su capacidad funcional. La relación entre ambos lenguajes es lo que denominamos código genético, desvelado hacia el año 1964 mediante una serie de brillantes experimentos. El ADN se lee en una dirección prefijada en grupos de tres. Cada triplete de nucleótidos codifica para un aminoácido determinado. Como hay sesenta y cuatro tripletes posibles y sólo veinte aminoácidos, varios tripletes pueden determinar el mismo aminoácido. Además, tres de ellos no codifican para ningún aminoácido y se utilizan como señal de terminación. La traducción de un lenguaje a otro se lleva a cabo en unos orgánulos celulares denominados ribosomas. Estos no pueden acceder directamente a los genes, es decir, a las regiones del ADN en que se codifican las distintas proteínas. Es necesario realizar una copia del gen que codifica la proteína que quiere fabricarse a partir de una molécula muy parecida de otro ácido nucleico, denominado ARN mensajero, que sale del núcleo e interacciona con los ribosomas para que traduzcan la información y elaboren la proteína. En dicha traducción se precisa la ayuda de otros ácidos nucleicos denominados ARN transferentes, moléculas que por una zona reconocen un determinado triplete del ARN mensajero y que, en un extremo, llevan el aminoácido correspondiente

a dicho triplete.

Desde el descubrimiento de la estructura del ADN en los años cincuenta, la biología molecular ha tenido una trayectoria brillante. El siguiente gran hito en la investigación molecular ha sido el proyecto de secuenciación del genoma humano, una de las empresas científicas de más envergadura del pasado siglo: tres mil millones de dólares y dos mil quinientos científicos de más de veinte laboratorios de Estados Unidos, Reino Unido y otros países. Se inició en 1990 por un consorcio público internacional, pero, ocho años después, el consorcio privado Celera Genomics, dirigido por Craig Venter, entró en dura competencia con él. La disputa entre ambos consorcios se zanjó en tablas con la presentación conjunta en el año 2000 del primer borrador de la secuencia del genoma humano, en un acto auspiciado por el presidente Bill Clinton y el primer ministro británico Tony Blair que tuvo como escenario la Casa Blanca.

Una de las mayores sorpresas cuando se publicó este primer borrador del genoma humano fue que más del 98% del ADN humano no contiene genes, esto es, no codifica proteínas. Se consideró, en principio, que carecía de utilidad y se lo denominó, de manera quizás un tanto precipitada y casi despectiva, «ADN basura». Los investigadores se quedaron sorprendidos de la pequeña cantidad de ADN codificante frente al total, pero no de que hubiese ADN sin aparente utilidad para la célula, ya que eso encajaba bien con las predicciones de la teoría de la evolución y con algunos datos previos sobre el ADN. La presencia en el genoma de ADN basura puede tener múltiples orígenes, como, por ejemplo, las inserciones de ADN procedente de virus y bacterias que pueden convertirse en elementos móviles, modificando su posición en el genoma celular y, en ocasiones, replicándose por su cuenta (transposones); o los pseudogenes, antiguos genes que han perdido por mutaciones su capacidad de traducirse a proteínas funcionales; o las secuencias de ADN altamente repetido (ADN satélite); o los intrones, secuencias de ADN dentro de los genes que se transcriben a ARN mensajero, pero que se eliminan antes de formar las proteínas; o, sencillamente, las secuencias espaciadoras entre genes. La presencia de ADN sin funcionalidad para la célula no puede extrañar: lo que en realidad sería sorprendente es que no lo hubiese.

También se esperaba encontrar, como así ha sucedido, la existencia de regiones reguladoras encargadas de controlar qué genes se activan en cada momento dentro de una célula o, en los organismos pluricelulares, qué genes se bloquean de forma permanente y cuáles se mantienen potencialmente funcionales para dar lugar a las distintas estirpes celulares que forman los diferentes tejidos y órganos de un organismo. Observe el lector que hablamos de células de un mismo organismo, que contienen los mismos genes, y que, sin embargo, en función de cuáles se mantengan activos, terminan diferenciándose en células tan diferentes en su aspecto y función como una neurona, un hepatocito o una célula muscular. Sí resultó, en cambio, un hallazgo desconcertante constatar que los humanos tenemos más o menos el mismo número de genes (unos veinte mil) que un gusano o un ratón. La complejidad de los organismos pluricelulares parece surgir de la reorganización funcional de los genes disponibles antes que de un incremento en su número.

Nessa Carey, autora del libro *ADN basura*, objeto de esta reseña, es profesora de Biología Molecular

en el Imperial College de Londres. En 2011 publicó un libro, *La revolución epigenética*, también traducido al castellano en la Biblioteca Buridan, que le granjeó cierto prestigio como divulgadora científica. La epigenética consiste en el estudio de las modificaciones en la expresión de genes que no obedecen a una alteración de la secuencia del ADN que codifica las proteínas, sino de otras regiones encargadas de la regulación génica, modificaciones que pueden ser heredables y transmitirse a las células hijas. En el ensayo que ahora nos ocupa, explica la historia del descubrimiento del ADN basura, su origen y el desciframiento, aún en marcha, de las funciones de una parte de este genoma. Carey retoma la ingeniosa metáfora en que se compara el ADN basura con la materia oscura, una entidad invisible y misteriosa que forma la mayor parte de la materia en el universo y de la que dependen muchos de los fenómenos que contemplamos a nuestro alrededor. De manera similar, aventura que el ADN basura podría estar implicado en procesos tales como la respuesta a las infecciones de los virus o la aparición de algunas enfermedades genéticas.

En líneas generales, algunas regiones del ADN basura podrían resultar útiles desempeñando tres clases de funciones. La primera es una función estructural para facilitar que el ADN se empaquete de forma correcta y, en palabras de la autora, no se *deshilache*. Parece que secuencias cortas de ADN altamente repetidas, que forman parte del llamado ADN satélite, podrían estar implicadas en la organización de los centrómeros al formarse los cromosomas. La segunda función vendría dada por algunos restos provenientes del genoma de virus o de otros microorganismos que invadieron a nuestros ancestros y se integraron en la cromatina y que, de forma esporádica, han llegado a ser útiles. Por ejemplo, la proteína sincitina, fundamental para la formación correcta de la placenta en humanos, procede de un resto del retrovirus HERV-W. La tercera función está relacionada con su contribución a la regulación de la expresión de los genes, siendo a veces imprescindibles para su correcta funcionalidad. Se comportan como interruptores que controlan qué genes hay que usar en cada tipo de célula, determinando en parte la diferenciación celular.

Carey, como buena divulgadora, sabe que la mejor forma de captar la atención del lector es describir algunas enfermedades que están causadas por cambios en el ADN basura. La estrategia ha sido estudiar algunas enfermedades conocidas que parecen tener un origen relacionado con la expresión de determinados genes, pero que no siguen las pautas habituales de la transmisión genética. Por ejemplo, la distrofia miotónica, un tipo de atrofia muscular grave con consecuencias a menudo fatales, que se manifiesta normalmente en adultos. Cuando esta enfermedad se transmite de una generación a la siguiente, los síntomas suelen ser más severos y surgen a una edad más temprana. Su forma más grave se da en hijos de madres afectadas. El cambio genético responsable de la enfermedad consiste en la repetición de una pequeña secuencia de tres nucleótidos CGT presente en un intrón de un gen. Esta secuencia puede estar repetida entre cinco y cincuenta copias, de forma que la enfermedad es tanto más grave cuanto mayor es el número de copias repetidas. Los individuos en los que esta secuencia está repetida pocas veces desarrollan la enfermedad de manera tardía, cuando son adultos, pero si tienen hijos pueden transmitir en sus gametos un número mayor de repeticiones de las que ellos mismos poseen, lo que explicaría el agravamiento de la enfermedad en el transcurso de las generaciones, hasta que los hijos mueren antes de reproducirse y la transmisión concluye.

El concepto de ADN basura recibió un varapalo cuando se publicaron en el año 2012 los primeros trabajos del proyecto ENCODE (Encyclopaedia of DNA Elements). Analizando el ADN de casi ciento cincuenta tipos de células, se concluyó que podría asignarse funcionalidad de algún tipo a casi el 80% del genoma, y que era posible que el 20% restante también la tuviese. Lo que muestran estos estudios es que muchas partes del ADN basura tenían actividad bioquímica en al menos algunos tipos de células. Es decir, son regiones que están desplegadas, como lo están los genes activos, y que muestran indicios de que pueden, por ejemplo, transcribirse para formar moléculas de ARN o servir de puntos de anclaje para determinadas proteínas. Sin embargo, hasta el momento no queda claro, salvo en algunos casos, que dicha actividad pueda ser útil para la célula; más bien parecen ser fenómenos irrelevantes y que sólo su alteración puede llegar a veces a ser perjudicial para el organismo, generando enfermedades. La forma en que se han presentado las conclusiones y, sobre todo, el tratamiento que le han dado los medios a estos hallazgos, ha generado una intensa polémica. Es razonable que un proyecto como ENCODE, con un alto coste tanto en fondos públicos (cuatrocientos millones de dólares y cinco años de trabajo) como en la dedicación de laboratorios y personal, trate de vender lo mejor posible sus conclusiones; pero parece que sus expectativas revolucionarias sobre el significado del ADN basura han sido exageradas. El proyecto ENCODE fue llevado a cabo, en su mayor parte, por científicos computacionales que han olvidado que el trabajo de un científico consiste en proponer hipótesis que puedan ponerse a prueba de forma experimental para tratar de refutarlas o, en caso de éxito, para aceptarlas como válidas mientras los resultados se ajusten a lo esperado.

Una parte importante de la controversia está relacionada con el uso del término «funcional» que han hecho los investigadores del proyecto ENCODE. Para los biólogos evolutivos, la palabra función, sea de un órgano, de una célula o de un trozo de ADN, sólo puede aplicarse cuando hablamos de algo que favorece la eficacia biológica (la *fitness*) de sus portadores y que, por tanto, es favorecido por la selección natural al conferir una ventaja adaptativa. Un ejemplo que utiliza Dan Graur, biólogo molecular de la Universidad de Houston y uno de los críticos más acerbos del proyecto ENCODE, para explicar esta distinción, se refiere al corazón: su función biológica es bombear la sangre y, como tal bomba impulsora, la variabilidad genética que afecte a su funcionamiento puede ser objeto de selección natural. Ahora bien, el corazón produce de forma incidental latidos que pueden ser útiles para un médico a la hora de efectuar una exploración diagnóstica y servir para detectar una anomalía, pero carece de sentido considerar dichos latidos como una actividad funcional del corazón en un sentido biológico.

Cuando se compara nuestro genoma con el de otras especies, incluidas aquellas más cercanas evolutivamente, vemos que la cantidad de ADN codificante es similar, pero que la cantidad de ADN no codificante fluctúa mucho entre especies, incluso aunque sean próximas filogenéticamente. En general, existe una correlación positiva entre la complejidad de un organismo y su contenido en ADN no codificante, ya que parece razonable que la selección natural pueda poner freno más fácilmente a la proliferación de ADN basura en los organismos más simples. Sin embargo, aun admitiendo que esto sea así, existen notables excepciones, como el genoma del protozoo unicelular *Amoeba dubia*, que contiene más de doscientas veces la cantidad de ADN que una célula humana, o el ADN presente en

muchas especies vegetales. Muchos evolucionistas asumen que gran parte del genoma no está bajo la influencia de la selección natural y, en este sentido, no es funcional. Sólo el 5% del genoma está conservado entre las distintas especies de mamíferos surgidas a partir de un ancestro común, hace unos cien millones de años. Esto explicaría que la cantidad de ADN no codificante varíe mucho de unas especies a otras, incluso dentro de un mismo grupo evolutivo, mientras que el ADN codificante apenas lo hace. En la última reunión del proyecto ENCODE, que se celebró en el año 2015, y no recogida por ello en el libro de Carey, el porcentaje del genoma que se considera funcional se ha rebajado del 80% al 50% y, además, al término «funcional» se le ha dado una interpretación más laxa: se trataría de regiones con una cierta actividad bioquímica que, aunque no presenten funcionalidad para la célula, pudieran tener relevancia médica por su papel en determinadas enfermedades. No obstante, el debate continúa a la espera de nuevas investigaciones que permitan precisar el significado funcional del ADN no codificante.

Esta controversia sobre el ADN basura ha incidido también en las disputas que los biólogos evolucionistas mantienen con los creacionistas, sobre todo en Estados Unidos. Un argumento utilizado contra los creacionistas de última generación, los defensores del llamado Diseño Inteligente, ha sido el hecho de que gran parte del genoma no sirviera para nada. No tiene sentido en apariencia que un diseñador inteligente construyese la vida con tal cantidad de ADN carente de función. Por ello, los primeros resultados del proyecto ENCODE dieron alas a los creacionistas, argumentando precisamente que no cabe esperar que la selección natural pueda mantener un 80% del genoma funcional a pesar de ser no codificante. Sin embargo, a día de hoy, y por lo que ya se ha comentado, este argumento es endeble y puede afirmarse que la genómica sigue planteando dificultades insuperables a los creacionistas. En los años treinta del siglo pasado se hizo popular la interpelación que el famoso biólogo evolutivo John Burdon Sanderson Haldane hacía a los creacionistas: «Dado que en la naturaleza, una de cada cuatro especies conocidas es un coleóptero (hay en torno a cuatrocientas mil especies), si Dios es el autor de todas las criaturas, hay que reconocer que siente un extraordinario cariño por los escarabajos». Hoy, algunos biólogos moleculares malévolos, retomando la ironía de Haldane, argumentan que, puesto que el genoma de la cebolla es cinco veces mayor que el humano, Dios debe sentir una especial predilección por las cebollas.

Carey pasa revista a los distintos tipos de ADN basura. El premio Nobel de Medicina de 1993 se concedió a dos investigadores, Richard J. Roberts y Philip A. Sharp, que mostraron que la mayor parte de los genes de las células eucariotas, provistas de núcleo, están formados por dos tipos de fragmentos de ADN: unos que codifican los aminoácidos que debe llevar la proteína (exones) y otros, no codificantes, que carecen de sentido (intrones). Cuando la célula transcribe el ARN mensajero, copia todas las bases del gen, pero luego elimina los fragmentos no codificantes, de forma que la versión final del ARN mensajero sólo contiene las instrucciones para producir la proteína correcta, fenómeno conocido como *splicing*. Los genes humanos contienen en promedio ocho exones codificantes separados entre sí por los intrones de ADN basura. La maquinaria del *splicing* está formada por un conjunto de moléculas denominado espliceosoma. Algunas enfermedades están relacionadas con la gestión de este proceso de cortar y pegar los fragmentos de ARN útiles. Por ejemplo, una mutación que afecta al funcionamiento del espliceosoma es responsable de algunos

tipos de retinosis pigmentosa, un tipo de ceguera progresiva que a menudo empieza en la adolescencia y que con la edad se agrava, porque las células de la retina que captan la luz van muriéndose gradualmente. Otro ejemplo lo constituye el ya mencionado de la distrofia miotónica que se produce por la repetición de una pequeña secuencia de nucleótidos en un intrón. La presencia de intrones en los genes eucarióticos ha tenido repercusiones evolutivas, ya que el «corta y pega» de fragmentos permite que el organismo pueda fabricar a partir de la secuencia de un gen más de una proteína según elimine realmente todos los intrones o deje alguno sin cortar. Se piensa que, como mínimo, el 70% de los genes humanos crean al menos dos proteínas diferentes por este método.

El último capítulo está dedicado a los microARN o ARN pequeños, formados por poco más de veinte nucleótidos. Se trata de fragmentos de ARN de una sola cadena que se fabrican copiando determinadas regiones de ADN no codificante. Aunque su funcionamiento sigue siendo muy debatido, parece que influyen en la regulación de más de diez mil genes, incluyendo algunos que se expresan en las células cancerosas, las cardiovasculares y las de la piel. Su impacto proviene de que estos microARN pueden unirse a regiones de los ARN mensajeros de muchas proteínas, de tal manera que los destruyen o, al menos, dificultan su traducción. Tienen, por tanto, la capacidad de regular la expresión de otros genes, normalmente inhibiendo la síntesis proteica. También pueden contribuir a causar enfermedades por un mecanismo realmente notable, conocido como secuestro de los ARN pequeños por parte de los virus. Es el caso, por ejemplo, de la encefalitis equina oriental de Norteamérica, transmitida por la picadura de un mosquito que puede afectar a caballos y a humanos. Nada más entrar en el torrente sanguíneo, el virus es absorbido por los glóbulos blancos, que inician una serie de reacciones de defensa. Sin embargo, a veces un ARN pequeño de los propios glóbulos blancos se une a un extremo del ARN del virus que se queda inactivo. Aunque pudiera parecer una buena noticia, en realidad no lo es, porque el virus escapa al control del sistema inmunológico y puede moverse libremente hasta llegar al sistema nervioso, donde desencadena una reacción mortal. En definitiva, este libro ayudará sin duda a personas con conocimientos en la materia a mantenerse al día sobre las últimas investigaciones en el fascinante mundo del ADN basura. Sin embargo, estamos ante un libro que podríamos llamar de alta divulgación científica y, para disfrutar de verdad de lo que allí se dice, pensamos que se requieren unos mínimos conocimientos previos de genética molecular, ya que se utilizan un gran número de conceptos técnicos sin los cuales resulta difícil una correcta comprensión del texto.

Laureano Castro Nogueira es catedrático de Bachillerato y profesor-tutor de la UNED. Es coautor, junto con Luis y Miguel Ángel Castro Nogueira, del libro *¿Quién teme a la naturaleza humana?* (Madrid, Tecnos, 2008).

Miguel Ángel Toro es catedrático de Producción Animal en la Universidad Politécnica de Madrid. Es coautor, con Carlos López Fanjul y Laureano Castro, de *A la sombra de Darwin. Las aproximaciones evolucionistas al comportamiento humano* (Madrid, Siglo XXI, 2003).