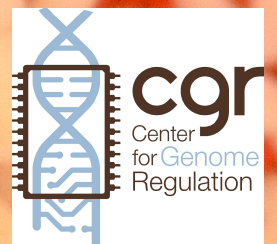


MIGUEL L ALLENDE

**LA LÓGICA DE LOS
GENOMAS**

CENTRO DE REGULACIÓN DEL GENOMA



La Lógica De Los Genomas

Miguel L Allende

Índice

Prólogo 4

Introducción 8

- **Conocer los genomas** nos permite ubicar a las especies en su contexto evolutivo y entender la diversidad.
- **Genes y Genomas:** el balance entre conservación y cambio.
- **Conciliando el determinismo y la plasticidad genética**

I. Genomas 20

- **Transcripción:** los genes se manifiestan...a veces.
- **Cromatina:** el genoma eucariótico se encuentra empaquetado en una compleja estructura
- **ADN intergénico:** ¿basura evolutiva o mecanismo regulatorio y estructural?

II. Tecnologías 37

- **¿Cómo se secuencia un genoma?**
- **Genomas, transcriptomas, metagenomas, microbiomas, metabolomas...** las ciencias “ómicas”.
- **Accediendo a los datos:** la Bioinformática
- **Tecnologías de manipulación de genes**
- **CRISPR/Cas y la edición de genomas:** la revolución que se viene

III. Consecuencias 63

- **Alcances de la manipulación genética de las especies.**
- **Transgénicos 2.0**
- **Gracias a la vida,** que me ha dado TAAATCGCGTG
- **Alquimia genómica:** ¿un quiebre en nuestra evolución?

Agradecimientos 88

Sobre el Autor 90

We're all clones, one is one and one are all...

(*Clones*. Alice Cooper, del Album *Flush The Fashion*. 28 de abril, 1980)

Prólogo

Terrorífico. Blasfemia. Monstruoso. Irresponsabilidad. Y las aparentemente contradictorias “eso es obra del demonio” o “estas jugando a ser Dios”. Son conceptos comunmente expresados en una conversación sobre manipulación genética, sobre todo si se trata de humanos como objeto de ella. Es probable que una buena cuota de la desconfianza que el ciudadano común nos tiene a los científicos que trabajamos en este tema se justifique. Con la liviandad con que el gáster anuncia que “hay que entrar a picar” decimos cosas como, “para corregir esto habría que meterse al ADN y modificarlo”. Si la Naturaleza es tan sabia y perfecta, ¿porque jugar con ella? ¿No estan bien las cosas tal cuál nos llegan?

Al mismo tiempo, he recibido cartas y preguntas de personas desconocidas que quieren saber si hay alguna posibilidad de que un ser querido sane, si esa enfermedad incurable y devastadora llegará algún día a ser revertida. Esas personas no están tan preocupadas de lo ético que pueda ser tocar la secuencia del ADN, ni si tenemos un 100% de seguridad en lo que estamos proponiendo para revertir el defecto. El sufrimiento personal puede ser muy influyente en nuestra mirada sobre el precario equilibrio bien *vs* mal en este caso. De este grupo de personas escucho comunmente otras palabras. Esperanza. Búsqueda. Milagro. ¿Cuándo?

El mundo se enfrenta a una revolución tecnológica sin precedentes que puede permitirnos influir como nunca antes sobre nuestro propio destino biológico. A las posibilidades de extender la vida humana y de mejorarla, se suman también la de moldear antojadizamente nuestras características y las de nuestros descendientes. La secuenciación y la edición genómica, entre otras técnicas como la clonación y el uso de las células madre, abre puertas -o cajas de Pandora- a las cuales podemos entrar atropelladamente como solemos hacerlo o, por el contrario, resolver detenernos un instante para evaluar su conveniencia. Esto último ha ocurrido, afortunadamente, en el caso de la edición genómica¹ gracias a que la comunidad científica esta consciente de que la tecnología avanza más rápido que la discusión de sus implicancias. Manipular genéticamente al ser humano de manera heredable esta, por el momento, impedido por una moratoria voluntaria

¹ <http://www.nationalacademies.org/gene-editing/Gene-Edit-Summit/index.htm>

en casi todas las naciones². Con las demás especies no hay tal moratoria, sin embargo, pues las hemos estado manipulando genéticamente por varios miles de años sin mayores cuestionamientos. La misma tecnología de edición genómica permitirá acelerar esa marcha y abre nuevas opciones de manejo de especies productivas e incluso de manipulación de poblaciones naturales. A las tecnologías de manipulación se suman las de caracterización de los genomas y la generación de datos masivos sobre ellos, al punto que estamos *ad portas* de poder secuenciar a bajo costo el genoma de cada persona. Vendrá la medicina personalizada en que los individuos serán tratados por los médicos de manera específica y no genérica, como hasta ahora. Al beneficio evidente de este cambio se contrapone la inquietante posibilidad de que la información genética se use en nuestra contra. Sería imaginable que aumentaran las primas de seguros de salud de los portadores de ciertos “genes”, sobre todo aquellos relacionados con enfermedades o susceptibilidades³. Si se permite la edición genómica en humanos, se suma la posibilidad de que nos equivoquemos y que perturbemos un sistema delicadamente esculpido por millones de años de evolución.

Pero, ¿que entiende la sociedad sobre estas discusiones y cómo podría participar de ellas? Si bien, hay un ánimo inclusivo cuando los científicos nos reunimos a pensar sobre la ética científica, no somos más que una élite ilustrada (pero sesgada) que toma decisiones para el resto. La primera opción tiene que ser, indudablemente, educar e informar. Muchas veces, la ciencia se desprestigia innecesariamente al dejar correr falsas concepciones o mitos sobre el alcance e intención de las tecnologías. En otras, la ciencia ha sido, efectivamente, irresponsable al no dimensionar los efectos de su quehacer (la bomba atómica es un ejemplo doloroso). Al abrir una discusión, creo que es deber de los científicos no dictar cátedra ni salirse de su experticia y, en cambio, entregar todos los elementos de juicio que puedan usarse para la toma de decisiones por parte de la comunidad o nación. El objetivo de esta monografía, es desarrollar el tema de la genómica, con sus diversas aristas y consecuencias, en términos de poner al alcance del lector lo que entendemos hoy como genoma, su funcionamiento y su manipulación. Más importante, intentaremos proporcionar suficiente información como para entender cómo operan los genes a

² Se especifica en el acuerdo que la restricción se refiere a manipulaciones en embriones llevados a término o a la línea germinal (los gametos, espermios u óvulos) y, por ende, la transmisión de esas modificaciones a los descendientes. No se prohíbe intentar corregir genes en células somáticas (todo el resto de los tejidos) para curar enfermedades de origen genético.

³ Sería el colmo de la condición preexistente.

nivel molecular, lo cual le quita misterio y transparenta palabras tan cargadas como “mutante”, “transgénico”, “enfermedad genética”, “clonación”, “determinismo”, “terapia génica”, u otras.

Este no es un texto de biología molecular ni de genética. Referimos al lector a fuentes especializadas o a referencias primarias de la literatura científica para los detalles más técnicos o que requieren un conocimiento más profundo. Asumiremos un conocimiento de nivel escolar de la biología, ya que de otro modo, usaríamos muchas páginas en repasar términos y conceptos mayormente conocidos⁴. Ocasionalmente usaremos términos en inglés (tratando en lo posible de ofrecer una traducción fiel) debido a que hay palabras que se han inventado o cooptado específicamente para ser usadas en la disciplina; recordemos que el inglés es el idioma preferido por la ciencia mundial en la actualidad⁵. Lamento que las fuentes citadas estén en ese idioma, pero no hay alternativa.

Tampoco es este un libro de filosofía o que pretenda establecer verdades o dogmas sobre el que hacer. Humildemente asumimos, basados en evidencia que nos convence a la mayoría de los científicos, que la vida apareció en la tierra accidentalmente mediante procesos físico-químicos conocidos, que la evolución es el mecanismo por el cuál se diversificó esa vida -y nos incluimos en esa diversidad- y que los procesos biológicos están basados en fenómenos describibles, repetibles y verificables pues su sustrato es la materia. Sin perjuicio de lo anterior, hay aspectos de la genética y genómica que nos atañen directamente en nuestra dimensión moral. Trataremos de explicar desapasionadamente y sin proselitismo estos conceptos pero entregando datos objetivos que puedan servir a la discusión o a formar opinión.

Antropocéntricamente, hablaremos de preferencia de los humanos (y de nuestros parientes cercanos, los animales, por un sesgo disciplinario del autor) dado el interés que despierta el conocimiento de nosotros mismos. Pero todos los organismos vivos poseen genes y genomas que operan bajo los mismos principios; nuestro conocimiento se basa en un ensamblaje de lo aprendido de todos ellos, partiendo por las bacterias. En biología usamos “modelos”, especies que se prestan para la experimentación y cultivo o mantenimiento en el laboratorio. La bacteria intestinal *Escherichia coli* es la favorita de los microbiólogos, mientras que en eucariontes, tenemos un repertorio de especies de diversos *taxa* como la levadura de la cerveza *Saccharomyces cerevisiae*, la

⁴ Llegando a encontrarse con un tema ignorado o que ha olvidado, Wikipedia es una excelente fuente.

⁵ Hice un esfuerzo descomunal y hablo del material genético como ADN y ARN y no como DNA y RNA, mi tendencia natural.

maleza *Arabidopsis thaliana*, la mosca del vinagre *Drosophila melanogaster*, el pez cebra *Danio rerio*, el sapo africano *Xenopus laevis*, el pollo *Gallus gallus* y el ratón *Mus musculus*. Dada la importancia de estos modelos en la biología experimental, sus genomas fueron los primeros en ser secuenciados, junto con el de *Homo sapiens*. Lo que sabemos viene primordialmente de los genomas de esas especies.

Es importante destacar, finalmente, que muchos de los temas que recorreremos son recientes y están en proceso de definirse aún. Todos los días aparecen nuevos datos y hallazgos que remecen nuestras preconcepciones. Hemos aprendido que el único dogma válido en biología es que no existen los dogmas, pues estos son eventualmente derrumbados por excepciones o nuevos mecanismos previamente insospechados. Recordemos que la definición de ciencia lleva implícita la falseabilidad, el que sea siempre posible demostrar que estamos equivocados. Junto con ello, las tecnologías avanzan tan rápido que la esfera de lo factible experimentalmente crece de manera exponencial. Es probable que, al poco tiempo de publicarse, este libro quede obsoleto. Por el momento, trataremos de cubrir de manera sencilla estos temas, no libres de controversia, ya que es fundamental que la sociedad tome consciencia de lo que podemos, queremos y debemos hacer -y no hacer- con el nuevo conocimiento, el cuál inaugura la era genómica.

Introducción

Conocer los genomas nos permite ubicar a las especies en su contexto evolutivo y entender la diversidad biológica.

El gran escritor argentino Jorge Luis Borges (1899-1986) hizo un ejercicio teórico en *La Biblioteca de Babel* (1941) donde imaginaba un recinto que contuviera todos los posibles libros de 410 páginas escritos combinando aleatoriamente 25 caracteres (22 letras, punto, coma y espacio). Evidentemente, la biblioteca era inmensa e imposible de recorrer por los lectores pero, necesariamente, incluiría todas las obras, biografías, recetas y verdades imaginables (al igual que falsedades), perdidas en un vasto mar de volúmenes incomprensibles e inútiles producto de la combinatoria azarosa de los caracteres. Los bibliotecarios perdían la vida buscando las obras relevantes o tratando de hacer sentido de los textos que iban encontrando en su recorrido.

Uso esta analogía para describir lo que a primera vista es el genoma de un organismo biológico complejo como somos los humanos: una enorme secuencia de caracteres en que hay pequeños trozos de información interpretable, interrumpida por eternas secuencias de caracteres sin un orden aparente o legible. Hace algunos años, los biólogos se sentían frente a una verdadera Biblioteca de Babel cuando contemplaban la inmensidad del genoma humano, muchos llegando a pensar que era imposible hacer sentido de él o, peor aún, que no valía siquiera la pena hacer el esfuerzo por descifrar su secuencia, dado su tamaño (3×10^9 pares de bases, o letras)⁶. Sin embargo, tuvimos mejor suerte que los personajes de Borges porque, primero, se trata de un abecedario mucho más simple: sólo 4 caracteres (los nucleótidos que componen el ADN se abrevian G, A, T, C). Segundo, el idioma también es simple: tríos de estas cuatro letras se traducen en aminoácidos, el lenguaje de las proteínas, y este código (el Código Genético) lo conocemos desde los años 60. Tercero, el ingenio y la tecnología nos han permitido acceder al texto completo (la secuencia de las letras) y hemos entrado en una fase, mucho más desafiante, de comprensión de él. Tenemos las partes, pero somos conscientes que de la suma de ellas emergen

⁶ En genómica, usamos terminología análoga a la de capacidad computacional (cuya unidad es el *byte*): 1000 bases=1 kilobase (Kb); 1000Kb= 1 megabase (Mb); 1000Mb= 1 gigabase (Gb). El genoma humano tiene 3Gb.

propiedades de un nuevo orden que aún no somos capaces de describir siquiera. La Biología de Sistemas es una disciplina surgida justamente para examinar el accionar colectivo de genes, proteínas, metabolitos, etc. y tratar de explicar la emergencia de células, órganos, cuerpos, conductas, ecosistemas u otros niveles de organización. Hemos leído el “libro” pero nuestra comprensión lectora todavía no madura al punto de poder hacer sentido completo de éste.

Dejemos de lado, por el momento, la caja negra que abarca el salto entre genes y propiedades emergentes y examinemos aquello de lo cuál estamos más seguros. Secuenciar nuestro genoma nos permitió inequívocamente confirmar que somos producto de la evolución biológica. Es perfectamente reconocible nuestra ancestría, al encontrar en nosotros genes, o más bien, secuencias, altamente conservadas con otras especies. Los genes se copian en cada división celular y se transmiten a la descendencia fielmente. La robustez de esa heredabilidad es tan alta que podemos afirmar que especies que existieron hace cientos de millones de años tenían muchos de los mismos genes que hoy encontramos en nuestro genoma. ¿Cómo lo sabemos? Porque no somos los únicos descendientes actualmente vivos de esos antepasados y hemos secuenciado los genomas de otros de sus descendientes. Tenemos “primos” como los simios o parientes algo más lejanos como los demás mamíferos o los demás vertebrados, etc. En ellos encontramos versiones muy similares a nuestros genes, señal no sólo de nuestro emparentamiento, sino de la estabilidad de esas secuencias. Dado que el genoma es “la instrucción” para la construcción de nuestro cuerpo y características, la conservación de genes sólo viene a reafirmar lo que intuyó brillantemente Darwin: la relación y continuidad entre todas las especies desde que comenzó la vida. Algunos genes incluso se comparten entre todas las especies, desde bacterias hasta humanos⁷. Recordemos que el código genético (el diccionario para traducir desde nucleótidos a aminoácidos) es universal y que, por lo tanto, finalmente, todos los seres vivos venimos de un mismo antecesor original que usaba ese código. Cuando nos comparamos con otras especies, mientras más recientemente existió nuestro antepasado común, más similitud de secuencias encontramos entre los respectivos genomas. Por ejemplo, entre *Homo sapiens* y *Pan troglodytes* (el chimpancé) las secuencias genómicas muestran un 96% de identidad⁸. No hay otra especie viviente que se parezca más a nosotros. Estas diferencias se acumularon durante los últimos 5 a 7

⁷ El uso del “desde-hasta” no implica un rango de superioridad o de “mayor” evolución; simplemente, es un rango de diversidad o, a lo mucho, complejidad.

⁸ Varki & Altheide. 2009. <http://genome.cshlp.org/content/15/12/1746.full>

millones de años, tiempo hace el cuál vivió el primate antecesor de ambas especies. Pero, ¿que ocurrió durante la evolución de estas dos especies a partir de ese momento de divergencia? La respuesta a esta pregunta es la clave para entender tanto la evolución de las especies como la interpretación de la estructura de los genomas a través del tiempo. Darwin propuso los principios básicos del fenómeno, aún sin tener ninguna noción sobre cómo se heredaban los caracteres (lamentablemente, los trabajos pioneros de Mendel pasaron desapercibidos a sus contemporáneos). El primer principio es el de la variación. Si bien, la herencia es fundamental para preservar características ventajosas de una generación a otra y, sobre todo, mantener una estructura genética coherente con la vida, es imprescindible que el sistema de replicación de la información genética sea levemente imperfecta. Hay dos fuentes principales para esa imperfección. Una, son las mutaciones, que pueden introducirse estocásticamente ya sea durante la replicación misma del ADN o también durante la vida de los individuos por daño al material genético producto de agentes biológicos, químicos o físicos. Las mutaciones relevantes, en el caso de los eucariontes, deben ocurrir en la llamada línea germinal, o en los gametos, para que se transmitan a la siguiente generación. La segunda fuente de “imperfección” en los organismos que poseen reproducción sexual, ocurre durante el proceso de formación de los gametos, más precisamente, durante la meiosis. En esta etapa crítica, en que se selecciona la información genética que transmitiremos a nuestros descendientes, ocurre un proceso de recombinación entre los cromosomas homólogos que recibimos de cada uno de nuestros padres. Esta mezcla de los genes (o caracteres) permite que nuestros hijos hereden alelos (versiones ligeramente diferentes del mismo gen, de los cuales cada padre aportará uno de los dos que tiene) en una combinación absolutamente nueva⁹. Es decir, aunque no se introdujeran mutaciones en el ADN heredado, cada genoma es único debido a la particular secuencia de alelos que posee en sus cromosomas materno y paterno. Entonces, en cada nueva generación, incluso entre hermanos o mellizos, hay una diversidad nueva de genotipos con ventajas y desventajas frente al ambiente. Aquí entra a operar el segundo mecanismo darwinista: la selección de aquellos individuos mejor adaptados al ambiente particular del momento. Esto genera la reproducción diferencial de los individuos según cómo estén dotados para enfrentar esas condiciones; algunos alelos (o mutaciones) se transmiten a la siguiente generación y otros, simplemente, están destinados a perderse. Si este

⁹ También es en la meiosis donde pueden ocurrir recombinaciones fallidas que alteran la macroestructura de los cromosomas. Estos cambios no siempre son catastróficos y pueden ser fuente de innovación.

“movimiento” genético (que puede ser gradual o acelerado según las presiones selectivas que operen o la permisividad del ambiente) se acompaña de aislamiento de los grupos variantes (poblaciones) podemos llegar a la especiación, donde los grupos divergentes ya no intercambian genes y siguen su marcha genética de manera independiente. Así nos ocurrió en nuestro pasado cuando los homínidos se fragmentaron para eventualmente generar especies como *Homo sapiens* y *Pan troglodytes* y las ya extintas *Homo neanderthalensis* u *Homo habilis*. Los cambios debido a mutaciones se acumulan muy lentamente en los genes ya que normalmente son dañinos para la función de la proteína codificada por ellos. Así, las secuencias de los genes pueden permanecer inalteradas por millones de años y ser semejantes entre especies que divergieron. Llamamos a estos genes homólogos (por su semejanza o identidad de secuencias). Si determinamos que el gen que comparten dos especies fue heredado de un antepasado que tenía ese mismo gen, hablamos de genes ortólogos.

Una pregunta que se escucha con frecuencia es si el hombre sigue evolucionando en la actualidad. Para formularnos esa pregunta con rigor científico, tenemos que preguntarnos más bien si ambos principios darwinistas siguen operando en nosotros. La variación genética es indudable; basta mirar en detalle los genomas de dos personas cualquiera (incluyendo a los gemelos) y veremos la variación. La selección, en cambio, requiere reproducción diferencial para enriquecer alelos en la población. Quizás se estén seleccionando en contra alelos que predispongan a ciertas enfermedades (o a favor aquellos que nos protejan), pero recordemos que esa selección tiene que operar previo o durante la etapa reproductiva de la especie¹⁰. Por esta razón, podrían desaparecer de a poco enfermedades de origen genético que produzcan malformaciones congénitas, pero no enfermedades como el cáncer o el Alzheimer familiar, que típicamente aparecen en la edad post reproductiva. En la ecuación evolutiva del hombre hay que incorporar al menos dos elementos más: el aspecto social-cultural y la manipulación genética o genómica (ver secciones II y III). Estamos entrando en la era de la genómica individual (con la medicina personalizada en mente) lo que traerá la posibilidad de hacer un seguimiento en gran escala del acervo genético humano a través del tiempo. Esto permitirá rastrear el “progreso” evolutivo del hombre con lujo de detalle. Y no podemos olvidar aquella hipótesis alternativa a la

¹⁰ Lo que, por cierto, no ocurre, es el mito urbano de la desaparición paulatina del dedo pequeño del pie.

de Darwin, propuesta por Lamarck, que involucra la herencia de caracteres adquiridos¹¹. Sabemos ahora que sobre el ADN se incorpora información proveniente de la experiencia de vida de los individuos, información que ha adquirido una relevancia cada vez más reconocida en generar un fenotipo. Exploraremos el concepto del eipgenoma, información no codificada en el ADN sino que por sobre este, y explicaremos que el ADN funciona dentro de un contexto, el cuál modula y extiende sus posibilidades.

Genes y Genomas: el balance entre conservación y cambio.

¿Cree Ud. que somos fundamentalmente distintos a los chimpancés? Dejemos de lado por el momento la conciencia (*sentience*) y repitamos la pregunta. Uno tendería espontáneamente a decir que sí, obvio que somos diferentes. Ellos tienen pelo sobre el cuerpo, se cuelgan hábilmente de los árboles y se apoyan en sus manos cuando corren. Bien. Pero si enumeramos y comparamos todas las características anatómicas y fisiológicas, probablemente podemos encontrar, por lejos, más similitudes que diferencias. Especialmente si modifico la pregunta y voy más allá: ¿somos fundamentalmente diferentes a un pollo? ¿A una rana? ¿A un pez? Terminé con los vertebrados, ¿A una hormiga? ¿A una almeja? ¿A una lombriz? A estas alturas, el chimpancé se está pareciendo bastante más a nosotros. Ya dijimos que la similitud de secuencias con el chimpancé es de 96% y decrece a medida que nos “alejamos” evolutivamente. Pero ese porcentaje sigue siendo importante entre los vertebrados. Por ejemplo, un 70% de los genes en el pez cebra (*Danio rerio*) tienen un equivalente (gen ortólogo) en el genoma humano¹². Si quitamos la ausencia de pulmones, la puesta de huevos, las aletas y escamas, de nuevo nos quedamos con una larga lista de similitudes que sobrepasan con creces las diferencias. En algunos estadios de nuestro desarrollo embrionario, de hecho, nos parecemos más a los peces de lo que nos gustaría. Es en estos estadios donde operan genes que son extremadamente conservados entre todos los animales; compartimos un grupo de ellos con organismos tan disímiles como los insectos, los moluscos y las lombrices. Se desprende a partir de estos increíbles descubrimientos sobre nuestro

¹¹ Se ridiculizó a Lamarck por pensar que la jirafa, al estirar el cuello para alcanzar ramas más altas, transmitiría este “alargamiento” a su descendencia. Lamarck usó un mal ejemplo, pero hoy tenemos evidencia de que no estaba del todo perdido.

¹² Howe et al., 2013. <http://www.nature.com/nature/journal/v496/n7446/full/nature12111.html>

origen y nuestra íntima relación de parentesco con el resto de la naturaleza, la necesidad de una profunda humildad y respeto hacia lo viviente. No somos más complejos en mecanismos o contenido genético que otros organismos, sólo somos diferentes anatómicamente en detalles que deberíamos considerar superficiales. Incluso, si vamos a explorar el origen de lo que los humanos valoramos por sobre todas nuestras capacidades, la conciencia y el pensamiento racional, posiblemente no encontremos una base para ello en los genes *per se*. Los genes son meros ladrillos: con ellos puedo hacer una pared, una casa o una catedral. La diferencia, es entonces, qué hacemos con los ladrillos o cómo los usamos y combinamos en la construcción. Ese 4% de diferencia que tenemos con el chimpancé tiene más que ver con esa construcción que con los genes mismos, donde la diferencia es, en realidad, insignificante. ¿A que nos referimos con la “construcción” del organismo a partir del genoma? En la sección I de este libro intentaremos explicar cómo se expresan los genes reguladamente desde el genoma y describir cómo diferencias en elementos que no están insertos en los genes mismos (que podemos llamar regulatorios o estructurales) son la clave para entender la lógica del armado y funcionamiento de lo viviente. Hay un gran espacio genómico entre los genes, donde hay secuencias que son libres de variar y donde no podemos fácilmente reconocer elementos funcionales, ni menos, conservados. Aquí se concentran las diferencias entre especies e, incluso, entre individuos de una misma especie¹³. Más aún, muchas de las enfermedades genéticas humanas tienen su origen en mutaciones en estos lugares (*loci*) alejados de los genes, lo que hace muy compleja su interpretación y posibilidades de corrección. Finalmente, los biólogos nos hemos dado cuenta que la secuencia nucleotídica del ADN es una primera capa de información que se encuentra superimpuesta por sucesivas capas adicionales, mucho menos estables, pero esenciales para el funcionamiento de los genes. Esta superestructura, que colectivamente llamamos epigenética, tiene un origen histórico en cada célula (depende de estados anteriores) y se modifica por influencias externas (el ambiente). Es decir, es un elemento genómico que es único e irrepetible para cada individuo e incluso para cada célula de cada individuo. Es, entonces, la expresión del genoma lo que precisamos entender para conocer, de verdad, las instrucciones para la construcción del organismo. El genoma es el ingeniero, pero el epigenoma y la herencia histórica (externa al ADN), son el arquitecto¹⁴.

¹³ Entre cada individuo humano y cualquier otro hay un 0.1% de diferencias en secuencia genómica, variantes ubicadas, en su mayoría, en las regiones intergénicas.

¹⁴ Del griego *arkhitekton*, *arkhi-* ‘maestro’ + *tektōn* ‘constructor.’

Conciliando el determinismo y la plasticidad genética

La biología molecular y la genómica moderna han, primero, exacerbado y, luego, reconciliado la aparente dialéctica entre el conjunto de instrucciones codificadas en el ADN (*nature*), y la historia y circunstancias en las que se desenvuelve el organismo portador de esas instrucciones (*nurture*). Concordamos que actualmente esa es una falsa dicotomía. No se puede entender la biología fuera del contexto ambiental-histórico; hay una relación sinérgica y complementaria entre ambas. Por muchos años, a los biólogos moleculares se les atribuyó un determinismo ortodoxo (“eres lo que dice tu ADN”) con alguna justificación debido a la excitación que produjeron los descubrimientos de la segunda mitad del siglo XX. La estructura del ADN, el código genético y la secuenciación de los primeros genomas revelarían, supuestamente, nuestra naturaleza más profunda. El gen fue aislado, se estudió *in vitro*, se comprobó su poder como constructor del fenotipo. Se le atribuyó un “egoísmo”¹⁵ y teleología propias; los organismos seríamos simples máquinas reproductoras de genes los cuales están en constante -y salvaje- competición darwiniana entre ellos. En los últimos años, sin embargo, ha ocurrido un vuelco conceptual entre los mismos científicos dedicados a la biología molecular, dado en buena parte gracias a tecnologías que nos permiten estudiar los genes en su contexto real: estos están en un genoma que reside en un núcleo, dentro de una célula, dentro de un organismo vivo. La secuencia de letras que llamamos “genoma” es una abstracción genérica que representa el orden de los desoxiribonucleótidos en los cromosomas de un organismo. El verdadero genoma está empacado y enmascarado de una forma tal que no toda esa información está accesible en un momento dado. Hoy, tenemos formas de observar y manipular este estado de accesibilidad lo que nos abre una nueva puerta para describir el comportamiento del genoma como un todo, en términos de su expresión y estructura. Las conclusiones son sorprendentes e imponen un límite a lo que podemos saber de un individuo al secuenciar su genoma; podremos conocer, quizás, su predisposición genética a ciertas características (expresado como una probabilidad), y poco más. Ejemplo: mujeres con una mutación en el gen BRCA1 tienen un 55% de probabilidad de manifestar un cáncer de mama, comparado con un 12% en mujeres sin la

¹⁵ Dawkins, R. 1976. *The Selfish Gene*.

mutación¹⁶. Por lo tanto, la mutación en BRCA1 no asegura un cáncer a la portadora, ni evitarlo si no tiene la mutación. Que desarrolle o no la enfermedad depende del resto de su genoma (el cáncer resulta generalmente de una suma de mutaciones o variantes en varios genes) y de su vida previa (factores de riesgo ambientales como exposición a agentes tóxicos o mutagénicos, alimentación, historia reproductiva, etc.). Son poco comunes los fenotipos o patologías que son determinados por un gen; probablemente tenemos un fuerte sesgo para pensar lo contrario porque las enfermedades monogénicas son las más fáciles de estudiar desde el punto de vista genético. La enfermedad neurodegenerativa de Huntington resulta de mutaciones que son dominantes y 100% penetrantes (todos los portadores de la mutación desarrollarán sí o sí los síntomas). Sin embargo, personas con la misma mutación pueden manifestar la patología entre los 2 y los 70 años de edad¹⁷, nuevamente reflejando que aún los genes más determinantes están fuertemente modulados por los factores ambientales o los contextos genéticos.

Para hacer menos precisa aún la información extraíble de un genoma, la secuencia misma y -obviamente- la epigenética, son inestables en el tiempo. Todas nuestras células somáticas (aquellas que no son gametos) están constantemente sufriendo mutaciones, pequeños cambios -en su gran mayoría, irrelevantes- por el bombardeo que diariamente sufrimos desde el exterior. Comer una hamburguesa al sol podría constituir un auto-atentado mutagénico por el daño oxidativo generado en nuestras células por los radicales libres en el alimento y por los rayos UV que rompen nuestro ADN. Hay un eficiente sistema de reparación que detecta y corrige las lesiones moleculares, pero unas pocas escapan del mecanismo de corrección y se acumulan a través del tiempo. De aquellas, una fracción puede afectar genes y, por lo tanto, mutarlos. Recordemos que tenemos dos copias de cada gen (por la diploidía de nuestras células, un cromosoma de cada padre) por lo que la pérdida de uno usualmente no resulta en un problema. En este caso, seremos heterocigotos, nuestras dos copias son diferentes (si son iguales, seríamos homocigotos). Hay dos escenarios terribles: el primero consiste en la ocurrencia de mutaciones que afecten ambas copias¹⁸. El segundo escenario, es que hayamos heredado una copia mutada (es decir, eramos *a priori* heterocigotos). La mala suerte de verdad, es que una “pérdida de

¹⁶ <https://www.cancer.gov/about-cancer/causes-prevention/genetics/brca-fact-sheet#q2>

¹⁷ Myers *et al.*, 2004. *Huntington's Disease Genetics*. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC534940/>

¹⁸ Algo improbable, pero posible, dada nuestra longevidad y tendencia a envenenarnos constantemente.

heterocigosis” ocurra y que perdamos la copia normal que nos quedaba de algún gen importante para regular el crecimiento o la división celular. Teóricamente, basta que en una de nuestras células se dé esta fatídica combinación para desencadenar divisiones descontroladas de ella y sus hijas: un tumor. Los tumores además, acumulan más y más mutaciones debido a que aumentamos las probabilidades de error en la síntesis de ADN y en la pérdida de los mecanismos de corrección, mutaciones que aumentan la tasa de división aún más -una especie de super-tumor- o adquieren la capacidad de migrar y establecerse en tejidos distintos al de origen (la metástasis). La secuenciación del ADN de células tumorales revela esa historia y, muchas veces, se pueden detectar las mutaciones (heredadas o adquiridas) que favorecieron este escenario. El cáncer es, finalmente, una enfermedad que tiene su origen en el genoma, donde debe existir un delicado balance entre preservar el acervo genético y mantener un control férreo de la expresión y función de los genes.

Volviendo al determinismo y a su descrédito. Sugeríamos que es falso que la secuencia de bases que tenemos en nuestro genoma sean “instrucciones” que nos pre-definen como individuos y que, irremediablemente, estemos destinados a ser como dictaminen esas instrucciones, una especie de Karma molecular. Sumemos más consideraciones.

Primero, la secuencia de bases esta lejos de ser la única instrucción que se usan nuestras células para funcionar y coordinarse cuando nos formamos. El desarrollo embrionario temprano es primariamente producto de información materna¹⁹. Aparte de todo el material celular materno, el genoma viene fuertemente impregnado de información epigenética de ambos padres que incide fuertemente sobre la expresión y función de los genes. Además, las células interactúan con -y responden a- el ambiente exterior a ellas y cambian su comportamiento acorde con las condiciones externas. Para un mamífero en gestación, por ejemplo, no es irrelevante el estado nutricional de la madre, independiente de los genes que posea. Un caso que ejemplifica lo que podemos llamar la plasticidad fenotípica (a un mismo genotipo, distinto fenotipo) ocurre en los experimentos de clonación. En la clonación, burlamos la reproducción sexual y nos saltamos la meiosis, generando un nuevo individuo a partir del genoma no recombinado de un organismo ya formado. En teoría, todos los individuos derivados de este “donante” deberían ser absolutamente

¹⁹ Cuando somos una sola célula, luego de la fecundación, casi toda ella esta constituida por materia proveniente de nuestra madre, a través del óvulo. El espermio, en cambio, contribuye el ADN paterno y poco más. Lo siento Señores.

idénticos (*ergo*, clones). Lo son pero, como mucho, en su secuencia de ADN. Ligeras variaciones en la exposición ambiental que sufren los clones durante su vida los harán distintos, tal vez imperceptiblemente. Aunque todos tienen los mismos genes, secuencias de ADN idénticas en sus cromosomas, los genes no son entes programados determinísticamente para funcionar como autómatas. El paso de genotipo a fenotipo requiere de la expresión de esos genes, la manifestación de ellos a través de la transcripción (ADN →ARN) y de la traducción (ARN → Proteína) lo cuál ocurre en coordinación con señales extracelulares (el ambiente). Analizaremos los mecanismos de expresión de los genes en la sección I. Podemos asegurar entonces, que el sueño (pesadilla, tal vez) de la clonación humana con el fin de reproducirse a uno mismo, es irrealizable. El “hijo” clonado podrá tener la misma secuencia de bases de ADN que el “padre”, pero jamás podrá repetirse la misma secuencia de eventos externos y azares que moldearon ese genoma y lo hicieron “expresarse” de una cierta forma. Hablaremos más sobre las consecuencias éticas de este tipo de iniciativas en la sección III.

Segundo, las mutaciones somáticas se acumulan durante toda nuestra vida. Células individuales pueden adquirir estas mutaciones estocásticamente y ser, por ende, distintas al resto y únicas, salvo que se dividan. Si esa mutación afecta la expresión o función de un gen, puede darle a esa célula una característica nueva, sea favorable o detrimental. Si el efecto es crecer más que sus vecinas (dividirse más veces) tenemos terreno fértil para un cáncer, como describimos anteriormente. Entonces, el núcleo somático que extraigamos del individuo deseoso de clonarse, será portador de un genoma distinto al que tuvo esa misma persona al nacer y, por lo tanto, el clon parte siendo distinto. Podría ser portador de mutaciones altamente peligrosas (por ejemplo, haber perdido una de las copias de un supresor de tumores) lo que haría al clon mucho más susceptible a enfermedades de origen genético. Podemos añadir que hay secuencias no-codificantes en el genoma (extensiones genómicas donde no hay genes) que contienen secuencias repetitivas y que cumplen un rol estructural (también detallado en la sección I). Es el caso de los telómeros, los extremos de los cromosomas, que naturalmente sufren un acortamiento con cada división celular. Es probable que la célula somática usada para extraer el núcleo a usar en la clonación tenga telómeros más cortos que los originales y que esto tenga una consecuencia negativa en el clon²⁰. Otras partes del ADN se mueven constantemente, como los transposones.

²⁰ El acortamiento de los telómeros es una señal de envejecimiento celular; es posible que el clon presente un envejecimiento prematuro.

Y otras más llegaron a nuestro genoma durante nuestra vida: son secuencias provenientes de virus que se integran y habitan en nuestro ADN²¹.

Tercero: ¿Saben, queridos lectores, de qué estamos hechos? Nuestro “organismo” en realidad es una simbiosis de millones de organismos; de hecho hay probablemente más microorganismos en una persona que células humanas. El clon más perfecto, crecido en condiciones idénticas al progenitor y con un genoma y epigenoma calcado a la letra, aún tendría que nacer, crecer y vivir con un repertorio de microorganismos asociados idéntico al de su antecesor, cosa realmente imposible de replicar. Hoy sabemos que la multitud de microorganismos que nos acompañan durante toda nuestra vida, son cruciales para nuestro desarrollo fisiológico y, posiblemente, hasta para nuestra conducta. No da lo mismo cuales bacterias viven en comunidad con nosotros y el tipo de especies y las proporciones relativas de ellas, son propias de cada individuo.

La genómica es una disciplina en constante evolución. Involucra mucho más que el desciframiento de secuencias y estructuras. Se hacen comparaciones entre especies y entre individuos de una especie. Se examina la expresión de los genomas y las influencias epigenéticas. Se desarrollan tecnologías para obtener secuencias de muestras cada vez más rebuscadas y para analizarlas computacionalmente. Y, recientemente, hemos entrado a modificar la secuencia del ADN y de su contexto, con poderosas técnicas de edición, mutagénesis y transgénesis. A la curiosidad que nos motiva a los científicos por entender el funcionamiento del genoma, una de las fronteras de la investigación en biología, se suman las posibles aplicaciones en salud humana, biotecnología y preservación de nuestro ambiente. Es decir, es un área del conocimiento rico en oportunidades para los nacientes científicos y tecnólogos pues tenemos mucho por avanzar y descubrir. Para los no-especialistas, representa un desafío meterse en el tema, por lo complejo y, a veces, árido que resulta. Pero creemos que es fundamental tener una apreciación mínima de algo que nos afectará a todos. Cuando nos compramos el último computador o teléfono celular, no leemos los manuales técnicos que describen los circuitos y el *software* que hay detrás, sólo los usamos y operamos sobre la elegante y simple pantalla que hace de amigable interfaz con el humano. En el caso de la genómica, no podemos ignorar lo que se podría llegar a hacer con nuestros propios cuerpos (o con el de nuestros hijos), con nuestros alimentos, con el entorno

²¹ El VIH, virus causante del SIDA, es un buen ejemplo. Infecta nuestras células inmunes y, a veces, permanece silente por años hasta que un día, se expresa (activa su transcripción) y se manifiesta la enfermedad.

natural. Hay decisiones políticas, éticas y personales que tomar. Es por ello que, en mi humilde consideración, la genómica tiene que ser un área en que todos los ciudadanos responsables de un país manejen conceptos mínimos y puedan entender los cómo, porqué y para que de lo que estamos haciendo los especialistas. Entendemos la inquietud que genera el tema y que se ha cargado de un aura de misterio o de malevolencia. A veces, se propone retroceder o prohibir el avance por el miedo o rechazo que genera. Pienso que, más que mala intención de aquellos detractores, es ignorancia por una lado y también la mezcla de un conjunto de temas en un gran saco de pseudo concepciones. Yo acepto la discrepancia y la voluntad de parar o limitar los avances científicos si se discuten con conocimiento y nos ponemos de acuerdo sensatamente en que nos conviene como humanidad. Pero para ello, hay que despejar los miedos y basarse sólo en hechos demostrados. La tecnología y el conocimiento pueden ser nuestros mejores aliados si nuestro objetivo es el bienestar y el progreso sostenible.

Examinaremos en las secciones que vienen, los elementos necesarios para entender, primero, como opera el genoma para lograr el delicado equilibrio entre la preservación de la secuencia de nucleótidos junto con la manifestación contextualizada de la información que contiene, la llamada expresión génica (parte I). Luego, revisaremos las tecnologías que nos han permitido conocer y, más recientemente, manipular los genomas (parte II). Finalmente, analizaremos adonde nos pueden llevar estos conocimientos y desarrollos tecnológicos y que decisiones tendremos que tomar al respecto (parte III).

I Genomas

Transcripción: los genes se manifiestan...a veces.

Una vez que Watson y Crick descifraran la estructura del ADN e intuyeran el mecanismo de replicación fiel de este²², y que Khorana, Nirenberg y Ochoa resolvieran el código genético, emergieron preguntas fundamentales respecto a cómo los genes codificados en el ADN se transformaban en las proteínas. Ambos pasos fueron resueltos de manera teórica primero y luego corroborados experimentalmente. Complejas máquinas celulares son capaces de transformar la secuencia de bases de una de las hebras del ADN en una copia idéntica de ARN mensajero (transcripción) que luego es tomado por una segunda máquina celular que lee tríos de bases para ensamblar una secuencia de aminoácidos (traducción). La bioquímica identificó a los actores, resolvió mecanismos y entendimos cómo los genes se manifestaban, es decir, cómo pasaban de ser mera información a ser agentes ejecutores de los fenómenos biológicos. Pero faltaba un elemento que se hizo rápidamente fundamental resolver. Los humanos, como el resto de los organismos multicelulares, somos enormemente complejos. Se sabía, con los experimentos de clonación de Sir John Gurdon²³, que todas las células de un organismo multicelular contienen una copia completa del genoma, por ende, cada célula contiene todos los genes, con el potencial de ser transcritos y traducidos (expresados). Pero, las células no expresan todos los genes. Muy por el contrario: cada tipo celular expresa un subconjunto de ellos, los que requiere para diferenciarse y adquirir su función específica. El problema de la expresión génica diferencial pasó a ser un objetivo científico crucial, abordado por cientos de investigadores en todo el mundo ya que respondería una de nuestras preguntas más importantes: ¿Cómo es que con la información contenida en nuestro ADN es posible construir una persona con todas sus diferentes partes? Tenemos millones de células y, al menos decenas si no cientos, de tipos de células y tejidos. Tenemos una cabeza arriba del cuerpo, dos brazos a los costados y dos piernas con las que

²² Watson J.D. and Crick F.H.C. Nature 171, 737-738 (1953)

²³ Gurdon, J. B. *The developmental capacity of nuclei taken from intestinal epithelium cells of feeding tadpoles*. J. Embryol. Exp. Morphol. 34, 93-112 (1962). Nótese que este experimento, esencialmente de clonación, antecedió -al mucho más famoso- de Ian Wilmut con la oveja Dolly por más de 30 años.

caminamos. No da lo mismo cuales células aparecen en cada región del organismo y algo las hace ser diferentes. Ejemplifiquemos con neuronas y células pancreáticas. Llamaremos a dos genes *neu* y *tro*, ambos presentes en nuestro genoma. En neuronas, se expresa *neu*: sufre transcripción para formar un ARN mensajero (ARNm) y luego este se traduce para producir la proteína NEU, la cuál es responsable de que las neuronas formen axones para conectarse con otras células²⁴. En células pancreáticas en cambio, *neu* no se transcribe por supuesto, ya que estas células no necesitan axones para conectarse con otras células. En cambio, sí fabrican enzimas digestivas, una de las cuales (TRO) esta codificada por el gen *tro* que se expresa sólo en este tipo celular. Afortunadamente, la neurona no expresa TRO ya que no esta encargada de producir enzimas digestivas. Entonces, si las neuronas y células pancreáticas tienen un genoma idéntico y que, por lo tanto, contiene ambos genes, ¿cómo “saben” los genes en una célula dada si tienen que expresarse o no? La respuesta tiene que ver con la historia de ambas células y cómo llegaron a ser lo que son, durante el desarrollo del organismo. Esa historia involucró la aparición de proteínas encargadas de activar la transcripción de genes, los factores de transcripción, muchos de los cuales son tejido-específicos. En nuestro ejemplo, *neu* es regulado por un factor transcripcional presente sólo en neuronas (llamémoslo, proNEU) mientras que *tro* es regulado por un factor transcripcional presente sólo en el páncreas (proTRO). Pero, de la respuesta emerge inmediatamente una nueva pregunta: ¿cómo llegaron a expresarse de manera tejido específica los genes *proneu* y *protro*? No hay problema, les digo, pues estos son regulados por otros activadores transcripcionales, también tejido específicos, pre-proNEU y pre-proTRO. Y así, en el más clásico argumento estilo huevo-gallina, podemos seguir *ad infinitum* sin entender el mecanismo que culmina con que dos células exhiban expresión diferencial de genes. Lo interesante de nuestro ejercicio, aparentemente fútil, es que estamos retrocediendo en el tiempo ya que cada vez que invocamos un factor previo, estamos recurriendo a un estado anterior de las células respectivas, o mejor dicho, a células precursoras o progenitoras de las neuronas y células pancreáticas. Forzando al máximo este retroceso ontogénico, llegaríamos eventualmente a la primera -y única- célula original: el huevo fecundado. En esta célula, el genoma esta “virgen”: esta todo por determinarse, que genes se expresan y cuales no. En muchos organismos, como nosotros, esta condición se mantiene por un tiempo, hasta que han ocurrido unas pocas divisiones celulares. De

²⁴ En nuestra nomenclatura, usaremos *cursivas* para referirnos a un gen (ADN o ARN) y mayúsculas para una proteína.

hecho, cuando en un embrión que tiene dos células, estas se separan, podemos generar dos individuos completos (gemelos con genoma idéntico). Estas primeras células son las llamadas totipotentes, o células madre, porque tienen la capacidad de diferenciarse a cualquier tipo celular y de ahí su potencial valor terapéutico como agentes regeneradores de todos los tejidos.

Volviendo a la expresión diferencial, quedó pendiente explicar cómo se generan las primeras diferencias en la transcripción de genes entre distintas células, fenómeno que es concomitante con la pérdida de su calidad de “madres” o totipotentes²⁵. Los mecanismos involucrados son varios, y van desde lo estocástico hasta las asimetrías inducidas por el ambiente, y también son diferentes en cada especie. Por ejemplo, el RNA mensajero o la misma proteína correspondiente a un regulador transcripcional puede ser heredado por una célula hija y no por la otra, luego que una célula madre se divida. Este regulador se puede distribuir asimétricamente mediante el transporte por el citoesqueleto (andamios intracelulares que movilizan organelos y macromoléculas) o ser activado por señales entre células, las cuales están en una constante conversación molecular para ponerse de acuerdo en qué hacer. El regulador transcripcional puede entonces iniciar una cascada de activaciones sucesivas que van transformando a las células descendientes en precursores de un tipo celular determinado, como pueden ser las neuronas. Evidentemente, la otra célula hija producto de la primera división asimétrica que describimos, tendrá otro destino eventual (ej., páncreas) y, por lo tanto, expresará otro conjunto de genes. Dado que el genoma (entendido como la secuencia de bases del ADN) se mantuvo inmutable²⁶ durante el proceso de diferenciación, lo que generó las diferencias entre las neuronas y las células pancreáticas fue su historia ontogenética (embrionaria), lo que instruyó a ese genoma a que genes expresar.

La mecánica de la transcripción no la revisaremos en detalle aquí, pero es necesario describir cómo los factores de transcripción, que son proteínas, funcionan de manera tan específica (por ejemplo, proNEU regulando a *neu*). Los factores de transcripción tienen estructuras moleculares dadas por su secuencia de aminoácidos que les permiten “encontrar” en el vasto genoma a sus genes blanco, uniéndose al ADN cercano a ellos. Son capaces de reconocer secuencias específicas ubicadas generalmente en la vecindad del sitio donde se une la maquinaria

²⁵ La llamada diferenciación celular.

²⁶ El término, admitidamente antropocéntrico, se aplica bien en este caso.

transcripcional y de atraer esa maquinaria para que empiece su trabajo de síntesis de mRNA. La interacción proteína-ADN ocurre por afinidad entre ambas moléculas, lo que redundará en una unión física semi-estable. Llamemos a la secuencia de ADN reconocida por proNEU, como el elemento E-neu con la secuencia TACATACA. Usando algunos trucos de la ingeniería genética, podemos aislar la secuencia E-neu y generar un trozo de ADN pequeño que es capaz de unirse específicamente a la proteína proNEU. También podemos tomar el trozo de ADN que contiene el gen *tro* (ese que se expresa en el páncreas y que tiene en su vecindad una secuencia E-tro que une proTRO). Con ellos, fabricamos un ADN recombinante en que juntamos el ADN que tiene la secuencia E-neu con el gen *tro*. Este gen artificial, si estuviera presente en el genoma sería capaz de unir la proteína proNEU (presente sólo en neuronas, como dijimos) y activar la transcripción de *tro*, para producir la proteína TRO. Pero, ¡horror! Si hacemos eso, y este gen artificial es introducido de veras en el genoma, ¡ocurriría la expresión de una enzima pancreática en una neurona! En efecto, hemos creado un “monstruo”, pero uno que nos enseña algo fundamental: la secuencia E-neu, que une al factor proNEU, tiene el control total de la expresión del gen que este en su vecindad, no importa cuál sea este. En otras palabras, en el genoma, hay secuencias de control transcripcional que operan para activar genes de manera específica, ya sea en un tiempo o en un lugar (tejido o tipo celular) particular. Mas aún, esas secuencias de control son modulares e intercambiables, y controlan indiscriminadamente al gen que se les ponga por delante. En inglés, estos elementos de control reciben el nombre de *enhancers*, cuya traducción más cercana es potenciadores y se ubican normalmente próximos al lugar en el ADN donde se inicia la transcripción del gen que regulan. Dada su importancia, se ha definido que estos elementos potenciadores son parte constituyente de los genes y ha habido en los últimos años un gran esfuerzo por encontrarlos. Esto no ha sido fácil. Primero, son secuencias pequeñas, de unas pocas bases de largo. Segundo, no existe un lugar específico relativo a los genes donde buscarlos; pueden estar cerca o lejos, de un lado o del otro del gen. Tercero, no existe conservación entre ellos ya que se permite cierta promiscuidad en la unión factor de transcripción-potenciador (en nuestro ejemplo, proNEU perfectamente puede unirse a las secuencias TACATACA, TAGATACA y TACACACA²⁷). Entonces, identificar los elementos potenciadores buscando las secuencias de unión en el genoma es como buscar una aguja en un pajar (un pajar de 3×10^9 letras

²⁷ Quizás, no elegí bien las secuencias de nucleótidos para usar de ejemplo...

en el caso humano). Y no es porque no sepamos buscar esa secuencia pequeña (8 bases) dentro del genoma. Los programas computacionales de búsqueda son altamente sofisticados y fueron la base para los buscadores de internet como *Google*²⁸. De hecho, una búsqueda detectaría cientos de “elementos” TACATACA, pero la gran mayoría de ellos serían falsos positivos (no serían auténticas secuencias de unión de proNEU). Hoy, tenemos formas de encontrar los potenciadores, no en base a su secuencia, sino porque hemos logrado situar, más bien, a los factores de transcripción y determinar dónde están unidos en el ADN. Si, en nuestro ejemplo, tomamos neuronas, fijamos al ADN las proteínas que están unidas a él, rompemos el ADN en fragmentos pequeños y aislamos una proteína como proNEU, arrastrará consigo un pedacito de ADN que contiene la secuencia TACATACA ubicada en la vecindad del gen *neu*. Este juego combinatorio entre elemento potenciador y gen, lo hemos aprovechado de diversas maneras tanto en investigación fundamental como para el desarrollo de la biotecnología.

¿Cuál es la relación entre el factor de transcripción y el gen que es regulado? Esto ha sido también un tema de intensa investigación y no ha sido fácil resolverlo. Se sabe que inmediatamente antes de la región codificante de un gen existe una pequeña secuencia llamada el promotor, que es muy similar entre todos los genes, donde se une la maquinaria de síntesis de ARN (la ARN Polimerasa). Hoy sabemos que, en los genes que están activamente en transcripción, los factores de transcripción unidos a los potenciadores están interactuando físicamente con la ARN Polimerasa en el promotor. Es decir, como hay una distancia a veces considerable entre potenciador y gen, el ADN hace un lazo y vuelve sobre sí mismo para formar una estructura circular. En el genoma, los genes activos forman parte de sectores donde estas interacciones son permitidas, mientras que los genes silenciados, se encuentran empaquetados en densas super-estructuras de cromatina (ADN junto con proteínas nucleares, tema que revisaremos a continuación). El estudio de la arquitectura nuclear se ha transformado en una de las áreas de estudio más intensas en los últimos años, pues entrega una visión global de cómo el genoma se estructura espacialmente y, combinado con las modificaciones epigenéticas, es indispensable para entender la regulación de la expresión génica.

Un aspecto importante en el contexto de la expresión, es que los genes no se expresan de una forma “todo o nada”. Es decir, su transcripción es regulada muy finamente, al punto de que

²⁸ Todos hemos ingresado en un buscador una frase de 30 caracteres para recibir, en milisegundos, decenas de páginas que contienen esa frase en una base de datos (secuencia de letras) mucho mayor que el genoma humano.

un poco más o menos de la dosis normal de ARN (o proteína) puede ser fatal para el organismo. Doy dos ejemplos de cómo esto es crítico. Una de las funciones celulares más importantes es el control de la proliferación, también llamado crecimiento celular por división. Este mecanismo de generación de nuevas células está exquisitamente regulado y hace que crezcamos de una manera ordenada y coordinada. Nuestros brazos crecen durante nuestro desarrollo pero luego paran de crecer; lo mismo nuestros órganos y tejidos. Si comemos de más, acumulamos reservas, pero nuestra dimensión intrínseca no cambia sustancialmente. Incluso, cuando regeneramos una estructura, lo hacemos para recomponer lo que se perdió, y el tejido en regeneración “sabe” cuando parar de crecer. Todo ello es posible porque las células tienen sistemas de control para regular cuantas veces dividirse y, sobre todo, cuando no hacerlo. A su vez, este control depende en buena medida, del control de la expresión de genes involucrados en división y crecimiento: sus niveles deben aumentar, disminuir o permanecer constantes dependiendo del contexto o necesidad. La cantidad de ARNm producido por cada gen está regulado finamente por la actividad de los factores transcripcionales, los cuales actúan en conjunto para potenciar o atenuar la síntesis de mensajero.

En segundo término, hay diversos mecanismos de modulación post-transcripcionales, que actúan una vez que el ARNm ha sido producido, para regular sus niveles. En general, son mecanismos de atenuación que producen la degradación de la molécula o que impiden su traducción a proteína. Volveremos a estos mecanismos más adelante.

Cromatina: el genoma eucariótico se encuentra empaquetado en una compleja estructura

Antes de conocerse que el ADN era el material genético, se observó en el núcleo de las células eucarióticas una estructura que poseía afinidad por diversas tinciones (colorantes) por lo que se la llamó cromatina. Eventualmente, se descubrió que el ADN es parte de la cromatina y que convive con una serie de proteínas que la empaquetan y protegen. Recordemos que el genoma humano, distribuido en 46 pedazos (cromosomas) mediría más de un metro si lo extendiéramos de punta a cabo, pero éste cabe en el núcleo que mide corrientemente un micrómetro (milésima de milímetro) de diámetro. El ADN se enrolla apretadamente sobre complejos proteicos con forma cilíndrica llamados nucleosomas que, a su vez, se amontonan en superestructuras compactas y

relativamente estables. Los nucleosomas se componen de 8 proteínas, dos de cada una de las 4 histonas, que tienen una alta afinidad por el ADN. A veces, los nucleosomas se ubican distanciadamente unos de otros, dejando pequeños trozos de ADN expuesto (desprovistos de la protección que ofrecen los nucleosomas) en una configuración que se asemeja a un collar de perlas. Este tipo de ADN esta en una condición de ser reconocido por los factores de transcripción y, por lo tanto, de que los genes correspondientes se expresen. Pero hoy sabemos que la mayor parte del genoma (y, por ende, los genes) usualmente se encuentra inaccesible, oculto en los nucleosomas y en las superestructuras formadas por complejos proteicos. En otras palabras, los genes son invisibles para la maquinaria transcripcional, que no tiene forma de encontrarlos. Como dijimos arriba, la transcripción sólo se inicia cuando esta presente un factor de transcripción que se une a los elementos potenciadores. Muchas veces, los factores de transcripción estan encargados de desenmascarar los genes, abrir la cromatina para que se expongan los genes que de otra manera estarían ocultos. Estos factores, que modifican y remodelan la cromatina, poseen actividad enzimática que altera las histonas de la cromatina en la vecindad del gen que será activado, reduciendo su afinidad por el ADN y permitiendo la llamada descondensación. Es decir, un gen expresado (que sufre transcripción) esta ubicado en una región de cromatina descondensada, en que los nucleosomas estan modificados, y esta asociado a complejos de factores transcripcionales que estimulan la transcripción. Esto lo distingue de otros genes, no expresados en esa célula, que se encuentran desprovistos de estas características y se encuentran dentro de la cromatina condensada. De hecho, una excelente forma de examinar el genoma y buscar genes expresados, es buscar regiones de cromatina abierta (descondensada) o buscar zonas donde las histonas de los nucleosomas presenten modificaciones que reducen su afinidad al ADN. Hoy tenemos métodos precisos para analizar el genoma completo y mapear estas zonas en un tejido o tipo celular dado. Si hiciéramos nuestro análisis en neuronas y células pancreáticas -aprovechando el ejemplo de la sección anterior- encontraríamos que en neuronas, el gen *neu* se ubica en cromatina descondensada y el gen *tro* en cromatina compacta. Lo contrario ocurriría en células pancreáticas.

Hay muchas modificaciones o “marcas” moleculares en la cromatina que nos señalan si los genes en esos lugares estan activos (con, al menos, la posibilidad inminente que se transcriban) o inactivos (en cromatina condensada). Como explicamos, las marcas se pueden encontrar en las histonas, pero también las hay en el ADN mismo. Es el caso de una modificación química de la

Citosina llamada metilación (se genera un metilcitosina o 5mC en parejas CG en el genoma). La metilación del ADN es una marca represiva o conducente al silenciamiento de los genes así modificados ya que esta modificación atrae proteínas que condensan la cromatina empacándola. Es una marca fácil de poner pero difícil de sacar (por las características de las metilasas y demetilasas) por lo que suele mantenerse por largos periodos en el ADN, usualmente hasta que el ADN se replica. Esto hace que la metilación sea una señal de represión de largo plazo, durando muchas veces por días o meses (en casos extremos, años) sin modificarse. Dado que las modificaciones del ADN y de las histonas en la cromatina son estables, no forman parte de la secuencia de ADN *per se*, y que influyen en la herencia (dado que se transmiten a células hijas o, en el caso de los gametos, a la siguiente generación), se han llamado colectivamente modificaciones epigenéticas (por sobre el genoma). La epigenética se dedica a estudiar los fenómenos que influyen en la herencia pero que son ajenas a la secuencia misma del ADN. Las modificaciones epigenéticas tienen varios orígenes, algunas dadas por la actividad de factores de transcripción (que vienen, finalmente, del genoma), pero, la mayor parte, son producto de la historia ontogénica de la célula o individuo o son impuestas por el ambiente. Como dijimos anteriormente, la epigenética es la razón por la cuál la secuencia del ADN contribuye sólo parcialmente con la información crítica para el desarrollo y funcionamiento de un organismo. Es decir, para un conocimiento completo de los factores involucrados en la herencia, necesitaríamos conocer la secuencia del ADN y el estado epigenético de la célula en cuestión. Para colmo, cada célula de un individuo complejo como una persona, difiere de las demás en su estructura epigenética (por mucho que la secuencia de ADN sea idéntica en todas sus células). Y además, la condición epigenética cambia con el tiempo. Como lo único práctico -por el momento- es examinar la secuencia de ADN (es lo único estable y medible con una simple muestra), nos conformaremos con esa parte de la información por ahora²⁹.

Explicamos que la cromatina es la forma molecular que adopta el material hereditario en la célula eucariótica. El empaquetamiento del ADN en esta superestructura permite el almacenaje de estas largas cadenas del orden de megabases (los cromosomas) en un pequeño espacio dentro del núcleo celular. Pero dentro de ese núcleo, los cromosomas no están segregados entre sí, sino

²⁹ De todas formas, esto no es poca cosa. Primero, la base sobre la cuál se arma la información epigenética es la secuencia. Segundo, el ADN tiene una influencia directa en el fenotipo por lo que conocerlo nos dice mucho sobre el potencial y riesgo genético de un individuo.

que conviven en lo que podríamos llamar un plato de spaghetti. Hay una madeja de moléculas entrelazadas en las cuales ciertas partes de la hebra se aproximan a partes distantes de la misma hebra o con otras hebras. Por ejemplo, si tuvieramos un plato de 46 fideos, podríamos imaginar que partes del fideo 4 se contactarían con partes del 6, 17 y 22, y con sigo mismo. En el genoma, esto no ocurre al azar y en general no ocurre entre cromosomas, sino dentro de un mismo cromosoma. Los contactos entre regiones distantes de la secuencia de ADN tienen un sentido estructural y, sobre todo, regulatorio, como explicamos en el capítulo anterior. Esta arquitectura permite contactos entre proteínas asociadas a elementos de control y otras proteínas asociadas a genes, lográndose una regulación específica a distancia. Existen técnicas que nos permiten examinar estas interacciones y revelar aquellas secuencias de control importantes para un gen. Gracias a este análisis, hemos descubierto que los elementos de control pueden estar muy distantes en el genoma de los genes regulados por ellos, habiendo, muchas veces, otros genes entre ambos. Aquellas zonas de un cromosoma que interactúan entre sí lo hacen de manera exclusiva, es decir, no hay contactos fuera de esa zona; la zona de exclusión esta aislada y recibe el nombre de barrio cromosómico o, más elegantemente, dominio topológicamente asociado. A cada lado del barrio, hay aisladores, secuencias que unen proteínas que segregan esa parte del genoma del resto. De esa manera, un potenciador sólo activa genes dentro de su barrio, evitando el caos regulatorio que ocurriría si empezara a contactar y activar a otros genes.

ADN intergénico: ¿basura evolutiva o mecanismo regulatorio y estructural?

Los humanos tenemos un genoma relativamente grande en tamaño comparado con la mayoría de las especies conocidas, lo que sugiere a primera vista que la cantidad de bases tiene que ver con la “complejidad”. Pero esta presunción se derrumba por dos razones. Primero, hay abundantes ejemplos de organismos como algunos protozoos, anfibios o plantas con genomas mucho más grandes que el nuestro. De hecho, en un mismo grupo de organismos, los tamaños genómicos pueden variar en órdenes de magnitud. Segundo, entre organismos evolutivamente distantes -y supuestamente diversos en complejidad-, no hay una diferencia significativa en el número de genes que contienen sus genomas. ¿Cómo reconciliamos estos hallazgos con las aparentes diferencias entre especies? ¿O es que esas diferencias son superficiales? Ya hablamos de

la relación de parentesco profundo y la homología que podemos encontrar entre las especies al comparar sus genes. También explicamos que las diferencias morfológicas y fisiológicas surgen de cambios a nivel regulatorio (principalmente transcripcional) que hacen uso de manera distinta de los mismos genes. Y recién describimos que tenemos que hacer una distinción entre los genes, constituidos por las secuencias que son transcritas (que pasan a ser ARN) junto con los elementos reguladores como los potenciadores, y las secuencias llamadas intergénicas donde, teóricamente, no hay nada. Hagamos un cálculo en una servilleta. Hay, aproximadamente³⁰, 25.000 genes en el genoma humano y el tamaño promedio de un gen es de 30Kb, por lo que tenemos “ocupados”, a lo más, 750.000Kb con genes. Dado el tamaño genómico de 3×10^9 bases o 3.000.000Kb, nos “sobra” la mayor parte del genoma. Esto, sin considerar que la parte útil o efectivamente codificante para un gen es mucho menor a 30Kb ya que los genes están segmentados en pequeños fragmentos codificantes (llamados exones) interrumpidos por largas extensiones de secuencia que, luego de transcritas, son removidas del ARN mensajero (los llamados intrones). Se calcula que sólo el 2% del genoma humano es efectivamente codificante. El 98% no codificante carece de secuencias que reconozcamos como funcionales y se le llegó a llamar “ADN basura”³¹, una paradoja ya que los sistemas biológicos no suelen derrochar recursos y energía sin razón. La hipótesis del ADN basura implica que, durante nuestro pasado evolutivo hemos acumulado secuencias inservibles por la incorporación de parásitos genómicos como los transposones y virus, y por duplicaciones de grandes segmentos del genoma, de cromosomas enteros o incluso del genoma completo³². Evidentemente, al existir estas secuencias masivas e inservibles, ellas no pueden perjudicar la viabilidad del organismo huésped para continuar existiendo, y pasan desapercibidas esperando su depuración natural. Sabemos que en las regiones intergénicas, se acumulan mutaciones espontáneas de manera mucho más rápida que dentro de las secuencias codificantes y son la fuente principal de variación entre individuos de una especie.

³⁰ <http://bionumbers.hms.harvard.edu/search.aspx?log=y&task=searchbytrmorg&trm=human+genome&time=2017%2f03%2f10+16%3a26%3a24.048>

³¹ El término *junk DNA* se ha transformado en una especie de cliché muchas veces mal usado. Puede verse una interesante discusión en <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4014423/>

³² En este caso, un organismo diploide (dos copias de cada cromosoma) pasa a ser tetraploide (cuatro), como ocurrió con algunos peces incluyendo el salmón. Hay un molusco octaploide (no, no es un pulpo). Y entre las plantas hay muchos ejemplos de multiploidía, como el gran pino californiano *Sequoia sempervirens* que es hexaploide. La poliploidización se puede inducir artificialmente para producir individuos estériles, como la sandía sin pepa, que es triploide.

Describimos anteriormente que al menos una parte de las secuencias intergénicas cumplen un rol estructural y regulatorio. Hay elementos de control de la expresión que se ubican lejos de los genes mismos; estos elementos pueden unir factores transcripcionales con diversos roles en la configuración de la cromatina, afectando barrios o sectores del genoma donde pueden ubicarse muchos genes. Cuando existen grupos de genes “co-regulados”, estos se mantienen unidos durante la evolución (ligados) ya que su separación por movimientos de ADN en el genoma³³ resultaría en un brusco cambio de expresión o en su no expresión. Hay grupos de genes que se mantienen unidos durante largos periodos evolutivos como los llamados genes homeóticos. Estos, encargados de definir el patrón de distribución de órganos y estructuras en el cuerpo (cabeza *vs* pies, dorso *vs* vientre) se encuentran en un complejo de genes presente desde los invertebrados a los humanos, abarcando más de 500 millones de años de evolución. Perder estos genes o alterar su expresión resulta en monstruosidades como es el caso de la mosca a la que le crecen patas en la cabeza³⁴ o el caso -menos catastrófico- de la polidactilia, en que las personas afectadas presentan seis dedos en la mano. Los complejos de genes como los homeóticos, están “protegidos” durante la evolución porque constituyen un conjunto de genes y elementos de control inseparables y que pueden abarcar un gran fragmento genómico de, a veces, cientos de kilobases. Dentro de un complejo de este tipo, los genes mismos (las secuencias que son transcritas) pueden representar un pequeño porcentaje de la secuencia siendo la mayor parte secuencia intergénica con sus elementos regulatorios. Dada la conservación estructural y funcional en los complejos génicos entre especies, los elementos regulatorios pueden ser detectados también por conservación de secuencias. Así, en un vasto mar de letras no codificantes, hemos detectado algunas cortas secuencias que son idénticas o similares entre especies separadas por millones de años de tiempo evolutivo. Una manera de probar si son elementos regulatorios, es aislarlos y probar su capacidad de regular la transcripción. Es más. Debido a su alto grado de conservación, se han tomado elementos regulatorios provenientes de

³³ Estos “movimientos” son en realidad rearrreglos de largos trozos del genoma por roturas en los cromosomas que son mal reparados, resultando a veces en la inversión de fragmentos o su reubicación en otro cromosoma.

³⁴ En la mosca *Drosophila melanogaster*, la mutación en el gen homeótico *Antennapedia* (literalmente, “patas en lugar de antenas”), afecta su patrón de expresión, tal que cambia desde la región torácica de la mosca (donde deben estar las patas) a la región cefálica (donde deberían estar las antenas).

humano y se han introducido en diversos animales (moscas, peces o ranas) y ¡funcionan!³⁵ Estos elementos ultraconservados nos han abierto una puerta para el estudio de los potenciadores y cómo ejercen su rol de controladores de la expresión, incluyendo la localización en ellos de marcas epigenéticas específicas. Además, nos han enseñado lo importante que es el contexto genómico en el que se ubican, lo que explica en parte la necesidad de contar con un “espacio” adecuado entre los genes: es una imposición dada por el mecanismo regulatorio. Claramente, no hay “basura” entre los genes.

Se ha develado en los últimos años un nuevo y misterioso fenómeno respecto a este tema: prácticamente el genoma completo se transcribe, azarosa y sutilmente, y que los ARNs producidos, que no codifican para proteínas, contribuyen a una regulación fina de los niveles de ARNm. Estos ARNs varían en tamaños y roles: están los largos no codificantes (*long non-coding RNAs*, *lncRNAs*) o los micro ARNs (*microRNAs*, *miRNAs*), para ir sumando a la ya nutrida sopa de letras. Los micro ARNs poseen secuencias que tienen complementareidad a ARNms y pueden inducir su degradación, siendo un mecanismo de silenciamiento y dosificación de la expresión. Evidentemente, los micro ARNs (que provienen de transcritos más largos que son procesados) son genes también pues su mutación genera un fenotipo. Cuando analizamos un genoma hoy en día, incluimos una búsqueda de los micro ARNs y de otros transcritos no codificantes y los mapeamos en la secuencia genómica. La cromatina, como conjunto de ADN y proteínas, es además un reservorio de maquinaria de síntesis, la cuál se ubica en una distribución amplia que es independiente del estado transcripcional de la célula. Nos hemos encontrado con sorpresas insospechadas al mirar la estructura y comportamiento molecular de la cromatina lo que sugiere que todavía no llegamos a encontrar -y menos aún comprender- todas las capas regulatorias existentes.

Los hallazgos recientes han hecho que vayamos rellenando cada vez más el genoma de “significado” y no me extrañaría que, eventualmente, haya un continuo de funcionalidad con poco espacio para secuencias “sobrantes” o “basura” genómica. Pero hay una parte del genoma - que en el caso de humanos llega a ser al menos la mitad de este- que presenta características únicas y vida propia. Son secuencias llamadas repetitivas, justamente porque son iteraciones -a veces interminables- de trozos de texto genético. Nos encontramos en el genoma humano con

³⁵ Por lógica, también deberían funcionar en el sentido contrario, aunque nadie en su sano juicio se animaría a probarlo.

largas extensiones de ATATATATATATATATATAT... o de CGCGCGCGCGCGCG³⁶. Estos repetidos más pequeños son, con toda probabilidad, estructurales y su largo, accidental. Sirven para espaciar los elementos de control y los sectores genómicos que deben mantener independencia regulatoria aunque, sin embargo, es probable que el tamaño exacto de los trechos de repeticiones no sean críticos. La secuencia CG además, es sustrato para una modificación directa del ADN (la metilación) que es importante en la represión transcripcional, como ya vimos. Pero las repeticiones más abundantes en términos del espacio que ocupan son secuencias que provienen de unos verdaderos parásitos genómicos, los llamaremos genéricamente transposones, aunque hay una variedad de ellos, con diversos orígenes y comportamientos. Los transposones fueron descubiertos en 1947 por Barbara McClintock³⁷ en el maíz. Ella se dio cuenta que en el maíz silvestre, que presenta un patrón de colores variegado (y hermoso), el patrón se origina por diferencias en el genoma de cada grano inducido por la activación de transposones³⁸. Estos elementos, pueden “saltar” de un lugar a otro en el genoma de manera azarosa alterando el ambiente regulatorio del sector donde aterrizan. ¿Cómo saltan? Los transposones contienen muchas veces una secuencia codificante para una o más proteínas que lo ayudan a hacerlo. Esta maquinaria, sintetizada por ellos mismos, reconoce los bordes del transposon, cortan el ADN a ambos lados, viajan con el transposón por el espacio nuclear en la célula, y llegan a un nuevo lugar en el genoma donde hay un nuevo corte e inserción del transposón. Es un verdadero *cut and paste*, tal como lo hacemos con una frase que cambiamos de lugar en un procesador de texto. Otros de estos parásitos (los retrotransposones) usan más bien la estrategia *copy and paste*: transcriben su secuencia formando un ARNm, el cuál codifica para una proteína que toma ese mismo ARNm y lo retrotranscribe (lo transforma de ARN en ADN) para luego integrar ese ADN en un nuevo lugar del genoma. Nótese que en este segundo caso hemos duplicado el elemento ya que el original se quedó en su lugar; en el caso anterior, sólo lo hemos movido. Es decir, el movimiento de un retroelemento genera un aumento en el tamaño del genoma, que puede llegar a ser significativo si ocurre con alta frecuencia. Los dos ejemplos que

³⁶ Si fuera un libro, sería como encontrarse de repente con “Bla, bla, bla...” cada cierto número de frases.

³⁷ Premio Nobel en Medicina o Fisiología 1983; la primera mujer que lo ganó sola. Científica excepcional, muy incomprendida en su época, que esencialmente descubrió la regulación génica (además de los transposones) mucho antes que lo propusieran Jacob y Monod, a quienes se les atribuye esa idea.

³⁸ Hoy sabemos que el genoma del maíz esta constituido en un 85% (!) por transposones.

acabo de mencionar describen procesos autónomos: los elementos contienen en su propio “genoma” todo lo necesario para moverse. Pero además, existen en el genoma huesped parásitos de los parásitos: elementos que no codifican para las proteínas movilizadoras y retrotranscriptoras sino que se aprovechan de las que generan los elementos autónomos. Es decir, saltan sólo cuando los otros están saltando. La realidad es que, en el genoma humano, la gran mayoría de los elementos de tipo transposón no codifican para proteínas en sus genomas sino que se aprovechan -cuando pueden- de la “nobleza” de los autónomos³⁹. ¿Se imaginaban Uds que dentro de cada célula de sus cuerpos, ocurre toda esta intriga, competencia, abuso y traición? Hay una verdadera telenovela molecular en nuestro genoma ya que convivimos con infraorganismos egoístas y potencialmente peligrosos. Imaginemos que un transposón (que se integra “caiga donde caiga” en el genoma, sin discriminación), se inserta en medio de un gen, o de un potenciador. En ese caso, perfectamente puede eliminar la función del gen haciéndolo inservible. Para nuestra fortuna, los transposones saltan poco; la frecuencia de movimiento es baja en relación a nuestra vida individual (sobre todo en los gametos) y su efecto sólo es notorio en tiempo evolutivo. Se calcula que, en humanos, hay un salto de retrotransposón cada 200 nacimientos. Hay además, un complejo sistema celular para mantener a raya los movimientos de los elementos transponibles. Nuestro genoma contiene secuencias que codifican para represores de la actividad de transposición modificando la cromatina en los lugares donde se encuentran para hacerlos inaccesibles a la maquinaria transcripcional. Por un lado, los ARNs largos no-codificantes promueven la metilación (silenciamiento) del ADN en las secuencias repetitivas. Por otro, existen factores celulares (proteínas) que son nuestras principales defensas en esta carrera armamentista que tenemos con los transposones. Una de las importantes se llama *p53*, que cumple varios roles incluyendo el de controlar la proliferación celular ya que es un antioncogen o supresor de tumores (su pérdida está asociada con cáncer). Justamente, en los tumores que han perdido la actividad de *p53*, una de las consecuencias es la “liberación” de los transposones que, al saltar agresivamente, generan más mutaciones y empeoran el fenotipo -y pronóstico- tumoral. No asustarse: como todo buen parásito, para tener éxito el transposón no debe perjudicar en demasía al hospedero. Lo peor que le puede pasar a un transposón es que se muera su huesped

³⁹ Un 17% del genoma humano está constituido de retrotransposones del tipo LINE o L1. Hay 500.000 copias de L1 de las cuales sólo 100 son funcionales (poseen maquinaria de transposición). El resto, son parásitos oportunistas de los LINEs funcionales.

ya que llegaría hasta ahí su propia existencia. Por lo tanto, no nos están afectando notoriamente; hay cosas mucho peores que nos acechan en este planeta.

La consecuencia de un exacerbado movimiento de transposones durante los últimos millones de años de evolución humana, han resultado en una amplificación relativamente repentina de nuestro genoma, la cual incluye al resto de los primates⁴⁰. Se calcula que el porcentaje total de elementos repetitivos en nuestro genoma alcanza un 50%⁴¹. Se ha especulado también, que parte de las diferencias que nos separan de nuestros peludos primos pueden deberse justamente a la influencia de los transposones en el genoma. Dado que las diferencias son regulatorias (ya explicamos que, a nivel de los genes, no las hay), tenemos que buscar la “humanidad” dentro de las secuencias intergénicas. Dijimos recién que los transposones autónomos tienen en su interior genes que codifican para proteínas que les permiten los saltos. Esto significa que también tienen sus propios elementos de control transcripcional (potenciadores, promotores). Un transposón que se inserta en la vecindad de un gen puede generar dos consecuencias: la mala es abolir o modificar negativamente su patrón de expresión causando un perjuicio al hospedero. La buena, la inserción confiere a un gen vecino un nuevo patrón de expresión que resulta ser ventajoso para la especie hospedera. Este fenómeno se ha visto con frecuencia en plantas. Por ejemplo, el arroz es una especie cuyo genoma está lleno de transposones activos, varios de los cuales han aterrizado en la vecindad de genes y les han modificado su expresión. Lo interesante de la historia del arroz es que algunos saltos han ocurrido en un tiempo tan reciente que tenemos variedades cultivadas que difieren según si han saltado o no el transposón. Observando las diferencias entre las variedades con y sin el salto del transposón podemos inferir que le ocurrió al gen afectado. En un ejemplo ilustrativo, la presencia del transposón cerca de un gen de tolerancia a la sequía le permite a la variedad que lo posee sobrevivir con menos agua que el resto de las variedades. Ocurrió una especie de avance tecnológico involuntario y que podemos aprovechar en nuestro beneficio. Al igual que con las mutaciones espontáneas, sin embargo, no siempre podemos darnos el lujo de que la lotería genética nos solucione los temas de alimentación, pues la probabilidad de que encontremos propiedades ventajosas por este mecanismo es baja.

⁴⁰ Para una revisión del tema: <https://www.nature.com/nrg/journal/v10/n10/full/nrg2640.html>

⁴¹ Cálculos más audaces indican que este número puede ser mayor y llegar a los dos tercios del genoma. <http://journals.plos.org/plosgenetics/article?id=10.1371/journal.pgen.1002384>

Hemos establecido entonces que los genomas se construyen durante la evolución a partir de eventos azarosos pero que, finalmente -y por definición- son compatibles con la sobrevivencia y el éxito de la especie. Incluso la parte del genoma humano que proviene de aquellos cohabitantes no invitados a nuestro genoma tendría una función (para ellos o nosotros) por lo que lo realmente superfluo es probablemente muy minoritario. La cosa es que por accidente o diseño⁴², el genoma humano tiene propiedades que emergieron junto con su crecimiento en tamaño y, aunque no podemos establecer una relación causal entre esta característica y lo que nos hace humanos, es intrigante pensar que nuestras capacidades -e innumerables falencias y defectos- provienen de la acción de pequeños entes que ni siquiera califican como organismos. ¿Cómo llegaron originalmente a nuestro genoma los transposones? Venimos heredándolos por millones de años de nuestros antecesores pero constantemente se incorporan nuevos elementos a nuestro genoma. La fuente más obvia son los retrovirus. Estos microorganismos, al infectar una célula, introducen su material genético consistente en ARN junto con una proteína llamada retrotranscriptasa que lo transforma en ADN y otra (integrasa) que lo integra al genoma hospedero. El mecanismo está codificado en el genoma del retrovirus y, una vez integrado, pasa a formar parte del genoma huésped. El virus normalmente generaría transcritos para propagarse a sí mismo, generando nuevos virus que saldrían de la célula para colonizar nuevas víctimas (otras células). Pero, si ocurre este fenómeno en una célula germinal (gameto), el retrovirus integrado pasaría a formar parte del genoma transmisible a la siguiente generación; el agente infeccioso ahora es un pasajero eterno del huésped (siempre que la descendencia futura del infortunado portador no haya sido perjudicada por esta “carga” genética). El ex-retrovirus eventualmente pierde los genes que le permiten salir de la célula, manteniendo el mecanismo de retrocopia e integración, convirtiéndose en un retrotransposon. Otras secuencias virales pueden también ser las causantes de la incorporación de algunos de los otros elementos repetitivos y/o transponibles. Una consecuencia interesante de todos estos movimientos de fragmentos genómicos es que a veces pueden acarrear no sólo al transposón sino al ADN aledaño, secuencias que pueden contener genes o fragmentos de genes. Un gen duplicado es una excelente oportunidad evolutiva: puede sufrir mutaciones que transformen su actividad en algo nuevo y aprovechable por la selección⁴³.

⁴² Me refiero (¡por favor!) a la elegante construcción de los organismos hecha por la selección natural.

⁴³ Llamamos a este proceso neofuncionalización, la creación de una nueva función biológica.

Otros genes nuevos pueden venir de las infecciones virales e incluso de otras especies: existe evidencia de la transmisión horizontal de ADN entre especies de bacterias y entre microorganismos (no sólo virus) y eucariontes superiores. La verdad, es que somos una amalgama de ADN de los orígenes más diversos: somos parte virus, parásito, mono, pez, esponja y bacteria. Recordemos que tenemos un antiguo genoma bacteriano presente en nuestras mitocondrias (una bacteria que hizo simbiosis con otro microorganismo generaron la primera célula eucarionte). ¿No es fantástico todo esto? Como para reconsiderar lo que definimos como el “yo”.

Ahora, que alcanzamos la comprensión de la estructura de los genomas y cómo funcionan (espero), pasemos a ver las técnicas que usamos para estudiarlos y, más importante, modificarlos. Es con esas técnicas que podemos introducir un quiebre en el curso natural de la evolución de los genomas y los organismos.

II Tecnologías

¿Cómo se secuencia un genoma?

El padre de la genética moderna, Thomas Hunt Morgan, abre su libro *La Base Física de la Herencia* (1919) con la siguiente perla: “El que los aspectos fundamentales de la herencia hayan resultado ser tan extraordinariamente simples despierta la esperanza de que entender la naturaleza sea completamente factible. Su publicitada insondabilidad ha sido, una vez más, revelada como una ilusión producto de nuestra ignorancia”. Morgan acertó meridianamente en lo último que dijo. Los científicos ocasionalmente caemos en la trampa de creer que hemos resuelto o comprendido a cabalidad un problema, sólo para darnos cuenta más tarde que hemos, simplemente, abierto nuevas preguntas y que somos en realidad más ignorantes que antes⁴⁴. Morgan se refería a que, habiendo descubierto el mecanismo hereditario básico (los cromosomas contienen los genes y la herencia sigue las leyes mendelianas) pensó que faltaba por dilucidar sólo la composición química del gen, que sería la unidad irreductible⁴⁵ de la herencia. Cien años después, la naturaleza no sólo nos sigue humillando, sino que seguimos frente a más y más preguntas, con el agravante de que la acumulación de datos ya sobrepasó los límites de la comprensión humana. La buena noticia, para los lectores más jóvenes por cierto, es que tenemos ciencia para rato. Revisemos entonces, donde estamos parados hoy (el 16 de abril de 2017) en el campo de la secuenciación de genomas y hasta donde hemos resuelto lo que le faltaba a Morgan. Repasaremos brevemente la historia de la secuenciación del ADN, para concentrarnos, más bien, en las tecnologías actualmente a nuestra disposición.

Desde mediados del siglo pasado sabíamos que el ADN de los genomas de todas las especies está constituido por las 4 bases (o desoxiribonucleótidos para ser preciso) A, G, T, C. Pero, más que la composición química de la molécula hereditaria (que, *per se*, no es información) la pregunta genética fundamental era, ¿en qué orden están esas bases en la secuencia lineal?⁴⁶

⁴⁴ Es como si, al ver la luz al final del túnel, nos damos cuenta al llegar que es sólo un farol y que el túnel sigue aún más allá. Sabía pero resignadamente, Sócrates concluyó: “Sólo sé que nada sé”.

⁴⁵ En ese tiempo, los biólogos aspiraban a ser como los físicos, quienes definían “partículas” fundamentales. Ahora, somos (un poco) más humildes.

⁴⁶ No es lo mismo “ $2a+2b+2r+e+i$ ” que “barbarie”.

Poco después de descubrirse la naturaleza estructural del ADN, se logró descifrar el código genético, el sistema que permite la traducción entre ambos idiomas: ARN y proteína. El mecanismo está basado en palabras de tres letras de ARN para cada aminoácido. Pero, dado que hay 64 combinaciones de tripletes con un alfabeto de cuatro letras y hay 20 aminoácidos para armar las proteínas, hay una redundancia de tripletes. Se acuñó la expresión “código genético degenerado” para referirse a que la mayoría de los aminoácidos son codificados por más de un triplete. Lo importante para nuestro relato, es que con esta información se dedujo de inmediato que, en los genomas, los genes que codifican proteínas estarían constituidos por secuencias organizadas en sucesivos tripletes de bases y que tendría que haber una secuencia de tripletes correlativa a la secuencia de aminoácidos para cada proteína codificada en el genoma⁴⁷.

La primera secuencia descrita no fue de ADN ni ARN sin embargo. Frederick Sanger, en 1955, logró secuenciar por primera vez una proteína, la insulina. Pasarían dos décadas para que, el mismo Sanger, en 1977, publicara un método para secuenciar el ADN⁴⁸. Con sus primeras y laboriosas metodologías logró secuenciar el primer genoma, el de un bacteriófago (virus de bacterias) llamado ϕ X174 con sus 5.386 nucleótidos. Luego vinieron genes de organismos más complejos y genomas de microorganismos. La metodología ideada por Sanger fue la base de todos los esfuerzos de secuenciación importantes durante los siguientes 20 años. Durante buena parte de la tesis doctoral de este autor (1988-1993), usé mi tiempo en correr largos geles de acrilamida (un polímero que permite la separación de moléculas por su tamaño) y la posterior lectura de las secuencias obtenidas en una autorradiografía (imagen en una película de rayos X). La secuenciación manual era tediosa e ineficiente, aunque de a poco se lograban ensamblar genes y trozos de genoma de unos pocos miles de bases. Usando la química de Sanger como principio, la secuenciación luego se automatizó y abarató, apareciendo las primeras máquinas para secuenciar a principios de la década de 1990. Para 1995, el grupo de J. Craig Venter publicó el primer genoma de un organismo de vida libre, la bacteria *Haemophilus influenzae*, una de las causantes de la molesta bronquitis del resfrío. Es en este momento cuando se empezó a hablar en serio sobre la posibilidad de acceder a la secuencia completa del humano y, por ende, de

⁴⁷ Francis Crick presentó esta hipótesis aún antes de conocerse el código genético: "On Protein Synthesis." The Symposia of the Society for Experimental Biology 12, (1958).

⁴⁸ Se ganó dos Premios Nobel: uno por cada uno de estos desarrollos (1958 y 1980). Uno de los centros de secuenciación más importantes del mundo es el *Sanger Centre*, en el Reino Unido.

cualquier organismo complejo. Un consorcio de países desarrollados decidió emprender el desafío con fondos públicos, estimándose un costo de varios miles de millones de dólares y 15 años; el líder del grupo fue, inicialmente, James Watson y luego, Francis Collins del Instituto de Salud Pública de EEUU. Se alzaron voces críticas respecto al proyecto argumentando que era un desperdicio de dinero mientras otros criticaban la metodología seleccionada. El mismo Venter había estado secuenciando los llamados ADN complementarios (*cDNAs*) -que son retrocopias de los ARNms- de humanos con la intención de patentar las secuencias e intentar su uso comercial. Armó una empresa con este fin (Celera Genomics) y obtuvo financiamiento privado para su esfuerzo de secuenciación. Dada esta presión, el consorcio público aceleró la marcha y ambos grupos llegaron por su respectivo lado a una versión preliminar⁴⁹ del genoma humano que se anunció conjuntamente a mediados del 2000 (5 años antes de la fecha estimada de término). El entonces Presidente de EEUU, Bill Clinton, acompañado del Primer Ministro del Reino Unido Tony Blair, y de Watson, Collins y Venter, anunció: “Con este nuevo y profundo conocimiento, la humanidad se encuentra a las puertas de alcanzar un formidable poder de sanación. La ciencia genómica tendrá un impacto auténtico en nuestras vidas y, más aún, en las vidas de nuestros descendientes. Revolucionará el diagnóstico, la prevención y el tratamiento de la mayoría, sino todas, las enfermedades humanas”.

Los dos grupos que simultáneamente alcanzaron la meta de secuenciar los 3 mil millones de bases del genoma humano usaron esencialmente la química de Sanger aunque con la ayuda de nuevas versiones fluorescentes (y no radioactivas) de las moléculas de marcaje (lo que ahorra el uso de los tediosos geles). Además, se automatizó el proceso, generándose máquinas que podían correr en paralelo muchas secuenciaciones y evacuar los datos a un computador que interpretaba las lecturas⁵⁰. En paralelo, y antes de iniciarse el proyecto genoma humano, Venter había ideado una estrategia que resultó ser la clave para secuenciar genomas de tamaño mediano a grande. En lugar de cortar el ADN en trozos y secuenciarlos de punta a cabo (como lo hizo

⁴⁹ Se publicaron dos artículos, uno por cada grupo. El del Consorcio Internacional: *Initial sequencing and analysis of the human genome*. Lander *et al.* *Nature* 409, 860-921 (2001). El de Celera Genomics: *The Sequence of the Human Genome*. Venter *et al.*, *Science* 291:1304 (2001). La secuencia inicialmente publicada no estaba completa, ya que faltaban las regiones altamente repetitivas en los extremos (telómeros) y centro (centrómeros) de cada cromosoma. Además había un alto número de interrupciones en los fragmentos ensamblados por lo que la continuidad de cada cromosoma era pobre para los estándares de hoy.

⁵⁰ Método desarrollado por Leroy Hood (*Anal. Biochem.* 154, 353. 1986) y que la empresa Applied Biosystems comercializó.

inicialmente el consorcio internacional), el usó la “técnica de la escopeta” (*shotgun sequencing*). Consistía en secuenciar azarosamente cientos de millones de pequeños fragmentos provenientes de un fragmento mayor del ADN genómico y luego, ensamblarlos computacionalmente. La verdadera dificultad radicaba en esto último ya que existía la duda de si sería posible hacer lecturas lo suficientemente masivas como para cubrir todo el fragmento seleccionado y si, luego de obtenidos los datos, se podría diseñar un software lo suficientemente poderoso para armar el puzzle. Venter, confiado que era posible, ofreció al equipo que estaba lentamente secuenciando el genoma de la mosca *Drosophila melanogaster*, terminar en un año el proyecto. Así lo hicieron⁵¹ y transformaron de inmediato la manera en que se estaba secuenciando el humano. Durante la década del 2000 al 2010, la secuenciación “masiva” introducida por Venter, Collins y sus colegas siguió su marcha y empezaron a acumularse las secuencias genómicas de organismos simples y complejos. Cada uno de estos resultados era un artículo en *Nature* o *Science*, y se generó toda una disciplina acompañada de una industria empeñada en hacer cada vez más accesibles las tecnologías de secuenciación.

Apareció, en ese tiempo, una segunda generación de secuenciadores (muchas, basadas conceptualmente aún en la estrategia de Sanger) que aumentaron en varios órdenes de magnitud su capacidad *versus* la de las máquinas disponibles y redujeron el costo, también en órdenes de magnitud. Las ingeniosas innovaciones implicaron, por ejemplo, secuenciar sobre una matriz sólida y tener cámaras digitales extremadamente sensibles y poderosas que podían medir *in situ* la secuenciación de cada molécula. Las metodologías nuevas recibieron varios nombres, incluyendo: secuenciación masivamente paralela (*massively parallel*), de alto rendimiento (*high throughput*) o, simplemente, de “siguiente generación” (*next generation sequencing, NGS*). La más usada de ellas en este instante es la tecnología Illumina que ofrece máquinas capaces de hacer cientos de millones de lecturas cortas con la estrategia *shotgun*, lo que es suficiente como para secuenciar, en una “corrida” de la máquina, un genoma humano completo. Las lecturas son realmente cortas en proporción: son de entre 150 y 300 bases, lo que comparado con el total a secuenciar (3 mil millones en el caso de humano), obviamente requiere muchas lecturas. ¿Cómo se deduce el genoma completo a partir de estos pedacitos? Son dos las estrategias. Primero, es importante contar con un genoma humano previamente secuenciado (lo que sí tenemos, desde el 2001) para

⁵¹ *The Genome Sequence of Drosophila melanogaster*. Adams *et al.* Science 287:2185-2195. 2001.

poder mapear cada segmento a su lugar. Es mucho más difícil armar un genoma *de novo* con esta tecnología, aunque es posible. Cuando secuenciamos un genoma humano hoy, simplemente lo armamos sobre el genoma de referencia existente en las bases de datos. Lo segundo, es que las lecturas cortas provienen de una muestra que tiene muchas copias del genoma, por lo que tendremos pedazos que se distribuyen azarosamente en la secuencia total y se sobrelapan. Es decir, cada sector del genoma es secuenciado muchas veces. A esto le llamamos la “profundidad” o cobertura de la secuenciación y corresponde a cuantas veces, en promedio, se secuenció cada base. Un análisis profundo y adecuado para ensamblar por completo un genoma humano requiere una cobertura de al menos 100X (se dice “cien por”), lo que indica que cada base fue secuenciada en promedio 100 veces. Si no nos importa mucho el ensamble y queremos un panorama más general, nos puede bastar una cobertura de 10X. Independiente de la cobertura, de todas formas los fragmentos pequeños son difíciles de ensamblar por que a veces son ambiguos. Explicamos antes que el genoma esta lleno de secuencias repetidas o muy parecidas entre sí. Si una lectura contiene una secuencia no única, no sabemos donde va. Ilustremos el problema con un pequeño texto (el “genoma”) que tenemos en varias copias y que luego secuenciamos con un método que nos da lecturas de sólo tres letras. Un posible resultado de las lecturas de tres letras que nos arroja el secuenciador podría ser el siguiente:

ama uem mel mea mim rqu ima ame apo mam ama ame lam ama ama uem eam orq ima map

No tengo espacio para poner miles de fragmentos, pero les dí la idea. Ahora, armados de un computador y software, deben armar la “frase” o genoma original⁵². Noten que tenemos tríos que se repiten (“ama”), lo que dificulta su ubicación. ¿Son lecturas provenientes del mismo lugar genómico o es que en el genoma ese trío se repite en varios lugares? En nuestro ejemplo, nos ayuda que en el texto combinamos las letras del abecedario; sería mucho más complejo usando sólo las 4 letras del ADN ya que la cantidad de fragmentos repetidos por azar sería enorme. Como discutimos en el capítulo anterior, el genoma humano (como casi todos) está lleno de secuencias repetidas imposibles de mapear. Lo que normalmente hacen los programas computacionales que se encuentran con esto, es simplemente dejar esas secuencias de lado y

⁵² Inténtelo. El “genoma” esta secuenciado con una cobertura de 2,5X y tiene un largo de 24 caracteres (sin contar espacios, los cuales hemos eliminado).

tratar de ensamblar lo que no es repetitivo. Un software potente y un buen bioinformático omiten esas secuencias y trabajan primero con el resto. De hecho, a veces no es necesario esforzarse más ya que los genes no son secuencias repetitivas, y los capturaremos en nuestro primer intento de ensamble. Si queremos tener un genoma mejor ensamblado, como para buscar variantes en el individuo que lo puedan predisponer a alguna enfermedad, tendremos que trabajar más con el ensamble o recurrir a otro tipo de secuenciación.

Desde hace un tiempo, las tecnologías de secuenciación corta están siendo reemplazadas por máquinas capaces de hacer lecturas mucho más largas (hasta decenas de kilobases). Esta incipiente tecnología (aún adolece de problemas de precisión) resuelve de inmediato el problema del ensamble ya que, si se encuentra con una secuencia repetitiva, casi siempre estará asociada a secuencias no-repetitivas y la sobreposición computacional de los pedazos leídos se ve enormemente simplificada. La forma en que se resolvió el tema del largo de las lecturas es usando sistemas que toman un fragmento de ADN (una molécula única, ojalá intacta) y la secuencian en toda su extensión. Una tecnología en uso ascendente es la creada por Pacific Biosciences (coloquialmente llamada ‘Pac-Bio’) que genera lecturas de varios kilobases (hasta 20) aunque tiene una tasa de error alta (10-15%). Es necesario aún, complementar corridas Pac-Bio con Illumina (que tiene una tasa de error de menos de 1%) para tener una secuencia precisa y bien ensamblada. Ha surgido muy recientemente una prometedora alternativa, de la empresa Oxford Nanopore Technologies (ONT), que, al igual que Pac-Bio, se basa en secuenciar moléculas únicas y generar lecturas largas de ellas, con un avance importante en la reducción de costo y en el tamaño de los aparatos. Las primeras pruebas de las máquinas de ONT parecen haber llegado al “número mágico” que se buscaba desde hace años, el de obtener un genoma humano completo por mil dólares (esto ya se había alcanzado con una secuenciación parcial del genoma). Desde su lanzamiento, los aparatos ONT han reducido la tasa de error de un inaceptable 20% a menos de un 10% y mejorando cada vez más. Una ventaja adicional de las máquinas de ONT, es su pequeño tamaño y la posibilidad de usarlas en terreno (fuera de un laboratorio), lo que las ha hecho muy versátiles en análisis de campo, como en lugares donde se presentan enfermedades infecciosas altamente transmisibles (ebola, dengue, virus de la influenza, etc). La empresa y pensadores futuristas han predicho que la genómica de bolsillo será común dentro de poco lo que permitirá -en el hogar- monitorear en tiempo real la eventual llegada de

patógenos⁵³, o en los aeropuertos el arribo de agentes infecciosos. El aparatito de ONT llamado MinIon (del tamaño de un teléfono celular) fue llevado al espacio y se probó su capacidad de secuenciar bacterias en ausencia de gravedad⁵⁴. Es evidente que dentro de poco, generará resultados de una calidad suficiente como para que cada persona pueda autosecuenciar su genoma y saber (con ayuda de programas amigables con el usuario que pueda mirar en su *smart phone*) sus características genéticas, riesgos y predisposiciones. La película de ciencia ficción “GATTACA” de 1997⁵⁵ predecía justamente este escenario, con todas las terribles consecuencias éticas, políticas y sociales imaginables. Ese mundo está aquí, a la vuelta de la esquina. Ni siquiera los científicos que trabajamos en esto lo esperábamos tan pronto. Es probable que Ud., lector/ lectora, tampoco.

Genomas, transcriptomas, metagenomas, microbiomas, metabolomas...las ciencias “ómicas”.

Un cálculo reciente⁵⁶, indica que hay más especies en la tierra que estrellas en la vía láctea. Del trillón (10^{12}) de las especies probables, conocemos alrededor de 5,6 millones ($5,6 \times 10^6$), la inmensa mayoría de ellas (y de las que quedan por describir) microorganismos. Es decir, llevamos menos del 0,001% del camino iniciado por Carolus Linnaeus, quien intentó clasificar las especies conocidas y ponerles nombre (*Homo sapiens*, etc.). Obviamente, si pensamos en animales y plantas, todo lo macroscópico (grande) lo hemos visto y catalogado. Nos faltan organismos multicelulares pequeños -como microinsectos- o algas unicelulares. También, todo lo que vive escondido aún en las junglas, desiertos y polos del planeta. Pero si se trata de bacterias, arqueas, virus y hongos, no tenemos idea de lo que hay ni en nuestro jardín. Claramente, no daremos abasto para describirlas una por una, como hasta ahora, y tendremos que inventar un sistema de clasificación

⁵³ Erlich, Yaniv (2015). "A vision for ubiquitous sequencing". *Genome Research*. 25 (10): 1411–1416. Este autor propone que el monitoreo casero se haga en el excusado (WC), lugar al que llegarían todos los agentes indeseables que pululan en el ambiente hogareño y que los secuenciadores serían uno más de los electrodomésticos esenciales. (Madre: “Hijo, ¿tiraste la cadena?”, Hijo: “Síiiii”, Madre: “¿secuenciasteeeee?”)

⁵⁴ NASA, agosto de 2016: “First DNA Sequencing in Space a Game Changer”. https://www.nasa.gov/mission_pages/station/research/news/dna_sequencing

⁵⁵ Columbia Pictures Corporation. <http://www.imdb.com/title/tt0119177/>

⁵⁶ Locey & Lennon. 2017. Scaling laws predict global microbial diversity. <http://www.pnas.org/content/113/21/5970.full>

nuevo como el que usan los astrónomos⁵⁷. Increíblemente, al trillón de especies por secuenciar tenemos que sumar unas cuantas que están extintas. Así es, el ADN en tejido fosilizado o preservado es recuperable y secuenciable por lo que ahora tenemos datos genómicos de especies como el mamut⁵⁸ y de antecesores de los humanos, como el *Homo neanderthalensis*⁵⁹.

Una nueva cohorte de biólogos, acompañados de sagaces ingenieros matemáticos e informáticos -sin los cuales no llegarían muy lejos-, se está dedicando en el presente a hacer este trabajo descriptivo y de exploración. A lo Darwin, que tuvo que subirse a una pequeña goleta y viajar por el mundo coleccionando especies, estos científicos buscan ahora descubrir lo que siempre ha estado ahí pero que no sabíamos cómo mirar: la diversidad genómica del planeta. Como lo desconocido es lo pequeño, debemos coleccionar muestras ambientales para examinar lo que pulula en la tierra, agua y aire de todos los ambientes. Tomar un balde de agua desde el mar y simplemente extraer todo el ADN ahí presente y secuenciarlo, permite tener una noción de cuáles organismos están presentes en la muestra, sin tener que llevarlos al laboratorio y cultivarlos de manera aislada. De hecho, la gran mayoría de los microorganismos no son cultivables por lo que jamás los conoceríamos sin esta estrategia de análisis *in situ*. Dado que en ese tipo de muestra hay una diversidad enorme de especies (y reinos), lo que obtenemos de secuenciar esta sopa es una mezcla de fragmentos provenientes de todas ellas, que luego hay que clasificar, separar, ordenar y ensamblar para obtener los genomas presentes. Es igual o más difícil, desde el punto de vista bioinformático, que trabajar con pequeños fragmentos de un genoma complejo, como el humano, pero el principio es el mismo (similar a la estrategia de la escopeta, *shotgun sequencing*). A esta aproximación se le otorgó el nombre de metagenómica, disciplina que procura describir y entender las comunidades de especies que conviven en un determinado espacio o ambiente. Probablemente, el primer experimento serio de este tipo lo lideró Craig Venter cuando secuenció el metagenoma del Mar de los Sargazos, encontrando en su muestra 1800 especies de microorganismos de los cuales 150 eran desconocidas⁶⁰. Se puede hacer metagenómica del mar,

⁵⁷ Dado que a ellos también se les acabaron los nombres para las estrellas hace rato (e.g., *Alpha Centauri*), ahora recurren a números de catálogo (e.g., NGC 3031 X6).

⁵⁸ [http://www.cell.com/cell-reports/abstract/S2211-1247\(15\)00639-7](http://www.cell.com/cell-reports/abstract/S2211-1247(15)00639-7)

⁵⁹ <http://science.sciencemag.org/content/328/5979/710>

⁶⁰ Venter et al. *Environmental genome shotgun sequencing of the Sargasso Sea*. Science. 2004 Apr 2;304(5667):66-74. <http://science.sciencemag.org/content/304/5667/66.long> También ver su charla *Ted* al respecto: https://www.ted.com/talks/craig_venter_on_dna_and_the_sea

suelos, aire, hielos, baños, transporte público, restaurantes, su intestino, su piel, desechos industriales y, muy pronto, del espacio y cuerpos extraterrestres. La exobiología recurrirá a esta estrategia para buscar vida fuera de la tierra, con el entendido que, si se encuentra ADN, es probable que sea una “contaminación” terrestre. Es muy posible que un análisis metagenómico de la luna revele ADN llevado por los astronautas que la visitaron, y capaz que hasta en Marte, los rovers hayan contaminado ese ambiente con ADN terrestre. Quien sabe, es posible que ya tengamos un ecosistema de organismos evolucionando en Marte por culpa nuestra⁶¹.

Metagenómica es un término amplio, que se aplica a la búsqueda de cualquier ADN, sea su origen microbiológico o de especies más complejas (animales, plantas). Por ejemplo, en la muestra de mar, aparecerán algas y animalículos (fito y zooplancton), junto con bacterias, arqueas y virus. En una muestra de suelos, pueden aparecer -junto con bacterias- pequeños insectos, gusanos, semillas, polen o hasta restos de ADN de grandes animales⁶². El metagenoma de un ecosistema es inclusivo de todas las especies, vivas o muertas. Si se apunta a sólo caracterizar los microbios o microorganismos, se habla del microbioma. Probablemente donde más se este analizando el microbioma sea en nosotros mismos. Como mencionamos anteriormente, nuestro cuerpo esta lleno de microorganismos, excluidos del interior de nuestro cuerpo por potentes barreras físicas y biológicas, lo que incluye al intestino, cuyo lumen corresponde físicamente al ambiente exterior. Allí reside una increíble comunidad de microorganismos que se benefician del abundante alimento disponible, cómodos 37°C y un ambiente estable que dura muchos años. No tenemos espacio aquí para referirnos en detalle a este mundo, pero señalemos que en el microbioma hay más células que todas las células humanas de nuestro cuerpo y una diversidad de especies más rica que en la mayoría de los ambientes exteriores. La mayor parte de estos organismos vive en simbiosis con nosotros, es decir, no nos perjudican y, muy por el contrario, a menudo nos benefician. Descomponen alimentos para los cuales no tenemos las enzimas digestivas para hacerlo, nos protegen de patógenos al producir antibióticos naturales y mantienen una homeostasis necesaria para nuestra fisiología. También

⁶¹ Las misiones espaciales tratan de ser “limpias” (estériles) y además el viaje espacial es muy tóxico para cualquier tipo de vida, por lo que es probable que encontremos trazas de ADN pero no organismos vivos. De todos modos, en la tierra hay organismos extremadamente resistentes a condiciones extremas, como las que hay en otros planetas. La exobiología de verdad, asumirá que la vida extraterrestre no esta basada en ADN y, si lo fuera, que nuestro ADN viene a su vez de esa fuente (la panspermia).

⁶² Pudieron haber muerto en el sector y haberse descompuesto, o, más probablemente, haber dejado heces.

educan a nuestro sistema inmune para reconocer las bacterias beneficiosas y distinguirlas de las patológicas. Tanto así, que el mundo moderno, con su excesiva limpieza, ha terminado por afectar el microbioma intestinal generándose problemas de salud inesperados. Se sabe que los bebés no producen las enzimas necesarias para degradar todos los azúcares presentes en la leche materna. Este trabajo lo hace una comunidad de bacterias que deben provenir -apenas nacido el infante- de la madre o del ambiente. La higiene -del mundo occidental sobre todo-, ha hecho que estas especies ya no se encuentren en todos los niños, lo que hace que su nutrición, su inmunidad y probablemente su desarrollo, sean deficientes. Por otro lado, a veces le tiramos bombas atómicas a nuestro intestino para solucionar una infección: los antibióticos. Este tratamiento genera una condición muchas veces de “borrón y cuenta nueva” en el microbioma intestinal ya que eliminamos indiscriminadamente a las bacterias “buenas y malas”. El daño colateral termina por generarnos periodos de malestar hasta que recuperamos nuestra flora y nos obliga a suplementar la dieta con los “probióticos” que no son otra cosa que bacterias que mandamos a recolonizar ese territorio ahora vírgen. En otro ejemplo, se cree que muchas alergias, en creciente aumento en el mundo, se deben a que nuestra inmunidad se confunde frecuentemente si no tiene acceso, desde la infancia, a un amplio repertorio de antígenos, moléculas muchas veces presentadas por microorganismos. Dada la importancia de las bacterias de la flora para nuestra salud y desarrollo, hay un esfuerzo importante de investigación para catalogar las especies prevalentes en el interior de las personas⁶³. Se sabe que cada persona es única, contiene un patrón específico de especies con proporciones diferentes entre individuos. Lo que se conserva entre un hospedero y otro, es el cúmulo de funciones que ejecutan las bacterias: el conjunto de habilidades metabólicas de la comunidad de microorganismos es relativamente constante. Siempre hay alguna especie que ocupa cada nicho disponible y, finalmente, todo lo que comemos es aprovechado al máximo⁶⁴. Tanta es la influencia de esta comunidad microbiana en nosotros, y

⁶³ *Human Microbiome Project*. (National Institutes of Health). <http://www.hmpdacc.org/>

⁶⁴ El cuerpo entero es un verdadero zoológico y no hay organismos sólo en nuestro intestino sino sobre todas las superficies (y cavidades) que nos cubren. A las bacterias, hongos, y virus posiblemente tengamos que agregar, en uno u otro caso, animales. Pulgas, garrapatas, piojos, ácaros, gusanos y protozoos parasíticos, el ocasional mosquito y quien sabe que otros pasajeros pueden llegar a ocupar nuestros cuerpos. ¡El humano es un metagenoma andante!

su rol en nuestro desarrollo e identidad, que ya se propone redefinir el concepto del “yo” tomando en cuenta al metaorganismo que en realidad somos⁶⁵.

Se han caracterizado también comunidades de organismos en diversos ambientes de la tierra. Expediciones globales que recorren los océanos y continentes, pretenden catalogar las especies habitantes de todas las latitudes, desde los polos, los desiertos, los volcanes y las profundidades abisales⁶⁶. Recordemos que el 70% del oxígeno de nuestra atmósfera se produce por pequeñas algas unicelulares, las cuales viven en un precario equilibrio ecológico que depende de temperaturas, corrientes marinas e irradiación lumínica. Claramente, el grado de polución o de intervención de un ambiente afecta este equilibrio, ni que hablar del cambio climático. En este sentido, el balance y proporción de las especies propias de cada ambiente están siendo monitoreadas por vía de la metagenómica como indicadores de estas perturbaciones. Es un tema de creciente interés también para la agricultura, ya que organismos presentes en el suelo donde hay cultivos, generan un ambiente esencial para la sobrevivencia de las plantas que allí habitan; recordemos que varias especies vegetales requieren organismos simbióticos para fijar nitrógeno y para captar ciertos metabolitos o elementos esenciales. Asimismo, se está recurriendo a conjuntos de microorganismos en procesos industriales (producción de proteínas, remediación ambiental o aprovechamiento de extracciones mineras, entre otros) donde el análisis tedioso de caracterizar las comunidades de bacterias está siendo reemplazado por un seguimiento metagenómico.

Los genomas y metagenomas entonces, nos revelan lo que potencialmente puede producir un organismo o comunidad de organismos por el colectivo de genes que contiene. A veces, sin embargo, nos interesa no sólo ver lo que se encuentra presente como ADN, sino los productos de ese o esos genomas. En otras palabras, que cosa están produciendo a nivel de transcritos (ARN), proteínas o productos del metabolismo. Campos de estudio dedicados a estos análisis han sido desarrollados y se agrupan coloquialmente como las ciencias “ómicas”, para analogarlas a la genómica. La transcriptómica, corresponde al análisis de los ARNs producidos por un organismo, células o tejidos. Es una secuenciación masiva, no del genoma, sino de los genes de ese genoma que se están expresando en un momento o contexto dado y es muy útil, como

⁶⁵ Rees et al., 2018. *How the microbiome challenges our concept of self*. *PLoS Biology* 16(2):e2005358. <http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.2005358>

⁶⁶ Ejemplos son el *Tara Oceans Expedition* (<https://oceans.taraexpeditions.org/en/m/about-tara/>) y el *Earth Biogenome Project* (<https://www.earthbiogenome.org/>).

primera aproximación, para conocer un genoma. Como captura sólo los genes expresados, el transcriptoma es mucho menos complejo que un genoma, y nos da una idea de las secuencias de las regiones codificantes (transcritas del genoma) sin preocuparnos por el 99% que es no-codificante, repetitivo, etc. Sin embargo, dado que se secuencian RNA, podemos detectar sólo aquello que se expresa en la muestra que tomamos y no es representativo de todos los genes (recordemos que no todos ellos se expresan en un tejido/momento dado). En este caso, existe la alternativa de secuenciar el llamado exoma, que es la secuenciación de todos los exones, o secuencias que potencialmente pueden llegar a ser transcritas y producidas como un ARNm. Dado que es más barato y simple de ensamblar comparado con un análisis del genoma completo, en muchos casos se recurre a la secuenciación de exoma para analizar los genomas de personas que puedan ser portadoras de enfermedades genéticas o cánceres; hay diversos proyectos dedicados a recopilar esta información en humanos⁶⁷.

El destino final de la transcripción, salvo en los casos en que el ARN sea el ejecutor de la función, es producir una proteína. Al igual que podemos mirar los genes en el genoma y los transcritos del transcriptoma, podemos evaluar las proteínas de una muestra lo que constituye el proteoma. Este análisis es completamente distinto a los que hemos descrito hasta ahora, ya que no estaremos mirando ácidos nucleicos sino secuencias de aminoácidos y ensamblando proteínas. Por lo tanto, la química -y la estrategia- involucrada es diferente. Lo que es parecido es que el análisis global del proteoma nos da una visión mucho más concreta de lo que la célula u organismo está haciendo realmente, pues son las proteínas las que están llevando a cabo las instrucciones del genoma. Nuevamente, la proteómica, se ha constituido como una disciplina en sí, con cultores y especialistas, con sofisticados equipos de análisis, programas bioinformáticos y métodos cuantitativos para comparar muestras. No satisfechos con todo lo anterior, hay un grupo de científicos dedicados a ir más allá todavía, al metaboloma, que es el conjunto de metabolitos (productos usados y generados por la fisiología celular) presentes en una muestra biológica o incluso en una muestra ambiental donde hayan organismos vivos. Un estudio que intenta describir un ecosistema o un organismo en el sentido más completo, tendría que considerar una mirada global a todos los niveles mencionados, pasando por todas las “ómicas”, para tener un reflejo apropiado de las características íntimas de ese sujeto de estudio. Ahora, cada uno de estos

⁶⁷ Por ejemplo: <https://www.nhlbi.nih.gov/research/resources/genetics-genomics/mendelian>

análisis genera una monumental cantidad de datos, que deben ser almacenados y procesados para su entendimiento. Y es aquí donde los biólogos llegan a su límite disciplinar: nos falta entrenamiento en sistemas informáticos y de diseño de programas que nos ayuden a interpretar lo que hemos colectado. Si no hubiéramos pedido ayuda externa, estaríamos aún en la edad de piedra de la “ómica”.

Accediendo a los datos: la Bioinformática

Si alguna vez han usado el programa de internet *Google Maps* o alguno similar que contenga mapas detallados del mundo, habrán hecho el ejercicio de partir con una mirada global (el planeta completo o el continente si se quiere) y luego un zoom tras otro para ir acercándose al punto de interés, que puede ser nuestro hogar. Eventualmente, al llegar al límite de resolución de las fotografías aéreas, el programa “baja” y nos muestra una foto de la puerta de la casa desde la calle⁶⁸. Ya que podemos llegar prácticamente a cualquier casa del mundo, tendríamos que darnos cuenta que, detrás de esta herramienta, hay una inmensa base de datos que contiene mapas globales y detallados de todo el planeta⁶⁹. Es decir, hay datos masivos que están organizados para que sean visualizables amigablemente por cualquier persona, aún aquella sin ningún entrenamiento en el funcionamiento de los *softwares* involucrados ni menos de los computadores. Hoy, podemos hacer lo mismo con muchos genomas, incluido el de humano. Hay programas interactivos para hacer esto en internet de una manera genérica y simple donde podemos partir mirando un cromosoma -el cuál contiene cientos de millones de nucleótidos o cientos de megabases- y podemos irnos acercando hasta ver los nucleótidos individuales. Esto es útil para entender la magnitud de la información almacenada en ese pequeño espacio físico. Los biólogos usamos una herramienta para esto que llamamos *Genome Browsers* (o navegadores del genoma) que nos permiten ver, de manera gráfica, desde la estructura global de un cromosoma hasta los

⁶⁸ Supongo que un día, mis pequeños hermanos, el máximo aumento nos mostrará a nosotros mismos, mirando en el computador -en ese momento- el *Google Maps*.

⁶⁹ El usuario tiene acceso a una cantidad de información similar a la que tenemos al frente con el genoma humano: hay alrededor de 1.5×10^9 edificios o casas en el mundo mientras que la secuencia de bases del genoma humano contiene 3×10^9 nucleótidos.

nucleótidos individuales. Si entran a uno de estos sitios y miran un ejemplo⁷⁰, podrán observar una región genómica llena de complejas anotaciones y simbologías. Con algo de entrenamiento, los biólogos podemos interpretar la información así presentada lo que nos permite conocer la distribución de genes, la posición de los elementos regulatorios, la estructura génica, la presencia de elementos repetitivos, las modificaciones epigenéticas más comunes, una lista de genes similares (ortólogos) en otras especies, mutaciones que originan enfermedades, etc. En estos navegadores del genoma, es posible moverse en ambas direcciones a lo largo del cromosoma o hacer *zoom* hacia afuera (alejarse) o hacia adentro (acercarse). Observar la estructura de un gen humano típico puede requerir un nivel de resolución que exponga en la pantalla unos 30Kb (3×10^4 bases) de secuencia. Recordemos que el genoma humano completo contiene más de diez mil segmentos de este tamaño, lo que da una idea de la dimensión de la información contenida en esta base de datos. Ahora, piensen en todas las especies secuenciadas, unas con genomas tan grandes como este y otras, la mayoría, más pequeños. A esto, sumemos toda la información adicional a la secuencia que podemos encontrar asociada a ella, incluyendo la información epigenética. Esta cantidad de información, que proviene de una multitud de experimentos de secuenciación que arrojaron miles de millones de trozos de secuencias de As, Gs, Cs y Ts, tuvo que ser organizada y presentada de una manera amigable e interpretable para los mortales como nosotros. Este trabajo lo hace la bioinformática, disciplina que ordena y da sentido biológico a lo que es, inicialmente, un caos de datos. Es una subdisciplina de la informática y también del llamado *data science*, o manejo de grandes volúmenes de datos que los humanos no somos capaces de manejar ni entender sin procesamiento previo. Hay varias aplicaciones similares en otros campos del saber. Ya hablamos de *Google Maps*, que es un mapa virtual del mundo con una resolución de algunos metros cuadrados. El mismo *Google* tiene un buscador de “palabras” de todo lo que esta público en la red, una base de datos probablemente más grande que la de todos los genomas conocidos. Datos masivos también son captados por la comunidad médica (fichas de pacientes con datos clínicos), por los astrónomos (datos de las observaciones de los telescopios y radio-observatorios que se acumulan cada noche), de meteorología (datos que ayudan a mejorar las predicciones; quizás es el sistema de datos más caótico y con mayor número de variables), la

⁷⁰ https://genome.ucsc.edu/cgi-bin/hgTracks?db=hg38&lastVirtModeType=default&lastVirtModeExtraState=&virtModeType=default&virtMode=0&nonVirtPosition=&position=chr9%3A133252000-133280861&hgside=582961307_yPoewWbkPnGxxsa0JrrMfjYPnc38

economía (probablemente el menos predecible por el factor humano), etc. Este siglo ha sido llamado el de la información, y estas colecciones de datos son un reflejo de lo que estamos acumulando y que tenemos que poder usar profesionalmente o como simples ciudadanos.

Pero la bioinformática es mucho más que ordenar y hacer accesibles los datos. Es la disciplina que toma los resultados de los experimentos de secuenciación (miles de millones de pequeños fragmentos) y los ensambla (emulando el “amoroso” ejercicio que hicimos anteriormente). Este proceso es más complejo de lo que parece. Además de tener que resolver dónde colocar los fragmentos en el mapa (los cuales presentan ambigüedad debido a su tamaño) también considera los errores de secuenciación, la variación entre individuos de una especie, la existencia de dos cromosomas no necesariamente idénticos y posibles contaminaciones⁷¹. Todo ello implica algoritmos de alta complejidad que requieren usualmente el desarrollo de un software propio y el establecimiento de *pipelines* (encadenamientos de programas que son corridos secuencialmente) de análisis reiterativo de los datos. Como los programas manejan volúmenes gigantescos de información, se usan supercomputadores con numerosos *cores* (núcleos de procesadores) corriendo en paralelo. Claramente, se requiere un entrenamiento profundo en ingeniería informática, lo que implica que se debe contar con científicos especializados en este tema para llevar a cabo un proyecto de genómica. Más aún, para enfrentar los desafíos que presenta el manejo coherente de la información genómica, no bastan los “computines” (no puede resistir) sino hacen falta ingenieros matemáticos que diseñen las estrategias lógicas que deben seguir los programadores. Estamos frente a una tarea interdisciplinaria, como es el caso en muchas de las ciencias de frontera de hoy. Y si bien, los biólogos no se entrenan como ingenieros, también han sabido ser creativos en el diseño de tecnologías de manipulación de los genomas. Programan, no con unos y ceros, sino con As, Cs, Ts y Gs.

Tecnologías de manipulación de genes

En 1975, cuando asomaron las primeras técnicas de ADN recombinante, los pioneros de las tecnologías llamaron a una reunión urgente para discutir la necesidad de imponer una moratoria

⁷¹ Es habitual que en las secuenciaciones de muestras humanas, aparezcan secuencias de *E. coli* u otras bacterias y a veces trazas de ADN humano del experimentador, no del sujeto de interés, que se “colaron” en la muestra.

o, derechamente, una prohibición para su uso indiscriminado⁷². Estos métodos, permitían cortar y pegar fragmentos de ADN de cualquier origen, gracias a que todas las especies llevamos nuestra información genética codificada con los mismos ladrillos moleculares. Es decir, es perfectamente posible colocar un pedazo de secuencia de bacteria junto con una proveniente de humano generando una molécula híbrida y única, siendo la nueva molécula completamente artificial⁷³. Es importante destacar que no hay ADN humano ni ADN bacteriano, sólo hay ADN. Es distinto que la secuencia (la información) sean de origen humano o bacteriano. Esta universalidad química y estructural permiten que un organismo no distinga el origen del ADN que contiene (si es propio o ajeno) desde el punto de vista de su replicación y mantención en la célula (veremos luego algunas excepciones a esta aseveración). En esos primeros experimentos, se introdujeron secuencias codificantes para proteínas humanas en pequeños “cromosomas” bacterianos, los cuales fueron replicados y mantenidos por las bacterias, y heredadas por todas sus hijas. En un cultivo de estas bacterias “transformadas”, cada bacteria puede llevar muchas copias del ADN híbrido produciéndose su amplificación exponencial; a esto llamamos “clonar” en la jerga técnica ya que involucra generar miles de millones de réplicas del mismo ADN. ¿Que diferencia le hace a la bacteria llevar en su interior un ADN con secuencias humanas? Ninguna⁷⁴. La bacteria, en cada división celular, reproduce fielmente todo su ADN sin discriminación, incluyendo esas secuencias humanas. Pero, me dirán Uds., “¿es una secuencia humana! ¿No producirá una proteína humana?” Depende, les contesto. Recordemos que la producción de una proteína a partir de un gen requiere transcripción y traducción. La bacteria normalmente hace esto con sus genes, pues contienen promotores y potenciadores para la transcripción y señales para la traducción específicas para esa bacteria. Si, al gen humano le añadimos estas secuencias de control (que tienen que ser compatibles con los sistemas bacterianos de síntesis), efectivamente se producirá la proteína humana. El sistema es tan susceptible de responder a un diseño lógico que, apenas desarrollados los métodos de ADN recombinante, se empezaron a producir proteínas humanas “recombinantes”, sintetizadas por bacterias y, luego, por otros sistemas celulares incluyendo los eucarióticos. La hormona del crecimiento, el

⁷² *The Asilomar Conference on Recombinant DNA* organizado por Paul Berg, quien hace un excelente recuento en *Nature*: <https://www.nature.com/nature/journal/v455/n7211/full/455290a.html>

⁷³ Por eso se llama ADN recombinante, porque se “combinan” ADNs de diversas fuentes.

⁷⁴ Se aplica el refrán “El burrito de San Vicente, lleva carga y no la siente”

interferón, la insulina son algunos ejemplos de los cientos de productos para uso en la medicina que empezaron a sintetizarse de esta manera. Nació la biotecnología, el aprovechamiento de sistemas biológicos con fines productivos, todo gracias a la posibilidad de manipular el ADN y de generar células portadoras de estos “genomas” híbridos. La resistencia a estos avances primero vino desde dentro de la comunidad de científicos, debido al potencial mal uso de la tecnología. Por ejemplo, era posible con relativa facilidad, generar toxinas o armas biológicas produciendo proteínas para este fin. Peor aún, es posible con ello generar organismos modificados para ser ultra patogénicos. Aquellos investigadores que trabajaban con oncogenes (genes que inducen cáncer), temían que al introducir los oncogenes en bacterias, estas podrían “transmitir” el cáncer si infectaban una persona. Se imaginó también, la posible creación de un virus de la influenza modificada para ser letal, lo que implicaba una pandemia mundial propagada simplemente via el estornudo. Estamos seguros que -lamentablemente- en laboratorios secretos distribuidos por el mundo, existen diversas armas biológicas producidas con ADN recombinante. Pero el resultado de la aplicación sensata de salvaguardas y restricciones de seguridad en los laboratorios académicos e industriales (probablemente, medidas exageradas en casi todos los casos) han evitado hasta ahora cualquier incidente producido por este tipo de investigación.

Una vez que se demostró que el ADN creado artificialmente en el laboratorio podía ser usado para introducir información genética en bacterias, se empezó a pensar en eucariontes, en organismos superiores y en el humano. Una barrera técnica consistía en que el ADN eucariótico se encuentra almacenado en un núcleo (que la bacteria no tiene) y empacado en cromatina, como describimos antes. Se logró realizar la llamada transfección, en que se logró incorporar ADN exógeno a células animales y vegetales en cultivo y que estas lo incorporaran en sus cromosomas⁷⁵. El ADN exógeno se integra habitualmente en cualquier lugar del genoma, sin discriminar dónde ni en cuantas copias, lo que implica que muchas veces hay que repetir el experimento para encontrar integraciones que funcionen bien y que no perturben un gen endógeno⁷⁶. De nuevo, si ese ADN codifica para una proteína, y ponemos la secuencia

⁷⁵ Las células vegetales tienen una pared celular por lo que no pueden ser fácilmente transfectadas ni microinyectadas. Con ellas, se usa una técnica de bombardeo de partículas metálicas cargadas de ADN que depositan ese material en el interior de la célula, transformándola. También es posible aprovechar microorganismos simbiotes que tienen la habilidad de integrar su ADN en el genoma de algunas plantas.

⁷⁶ Si un ADN externo se integra justo donde hay un gen (improbable en genomas grandes como el humano), puede mutarlo y abolir su función.

codificante acompañada de sus elementos regulatorios para la transcripción, etc., produciríamos la proteína a voluntad en esas células. Si la célula usada está inmortalizada (crece en cultivo continuamente y puede mantenerse para siempre si se desea), tenemos una cepa o línea celular productora a nuestra disposición. El paso siguiente fue más audaz, ya que se deseaba saber si era posible hacer esto con un organismo multicelular, como un mamífero por ejemplo. Los mamíferos, por un lado, son complejos de manipular experimentalmente, ya que su desarrollo embrionario ocurre *in utero*, lugar de difícil acceso. Quizás hubiera sido más fácil empezar con animales más simples, como insectos, peces o ranas, donde hoy es rutinario hacerlo. Pero por otro lado, el ratón es un modelo ampliamente usado debido a que se supone que será más parecido a nosotros que animales no-mamíferos. De tal modo, en 1980, Ralph Brinster y Richard Palmiter decidieron intentar introducir un ADN externo en un ratón mediante su microinyección en un embrión de una sola célula, el cigoto recién fecundado. La idea era, que si se integraba el ADN como lo hacía en las células transfectadas, todas las células del ratón resultante contendrían ese ADN. El proceso implicaba la tortuosa tarea de sacar los oviductos de una hembra que se hubiera cruzado recientemente con un macho, rescatar los cigotos (huevos) fecundados, microinyectarlos con el ADN, devolverlos al interior del útero de otra hembra (la primera tuvo que ser sacrificada) y esperar que alguno se implante y sobreviva para ver el resultado. Como el esfuerzo era enorme, decidieron hacer el primer intento no con cualquier ADN, sino con uno que inmediatamente demostrara la utilidad de la tecnología. Para ello, generaron un ADN artificial (lo que en la jerga llamamos “constructo”) que contenía la secuencia que codifica para la hormona del crecimiento de la rata (una especie emparentada, cuya hormona funciona en el ratón). Como explicamos, si queremos que la proteína se exprese, debemos anteponer a la secuencia codificante los elementos de control transcripcional. A los investigadores se les ocurrió que no era buena idea que el ratón expresara la hormona de crecimiento descontroladamente ni constitutivamente, por lo que antepusieron al gen de la hormona un potenciador que responde a la presencia de metales y que se activa sólo cuando el ratón consume una cantidad elevada de metal (en este caso, zinc). Obviamente, el ratón, como todos los animales, no contiene en su genoma el gen de la hormona del crecimiento regulada por metales, sino por otras señales propias que le permiten crecer en la medida adecuada en cierta etapa de su vida. Pero, por encima de ello, el ratón modificado tendría esta segunda copia de la hormona en su genoma, la cuál presumiblemente no se expresaría a menos que el ratón estuviera expuesto a zinc. Palmiter y

Brinster lograron inyectar los cigotos, implantarlos, y esperaron el par de semanas a que nacieran las crías desde la madre postiza a la cuál se le implantaron los embriones. En efecto, el ADN implantado en el genoma del ratón estaba presente en algunas de las crías sobrevivientes, y estas nacieron y maduraron hasta la adultez. Crucialmente, cuando estos ratones se aparearon con otros, el ADN exógeno fue transmitido a la siguiente generación, en proporciones mendelianas, lo que demostró el éxito del proceso. Había nacido el primer organismo multicelular transgénico. Por definición, los transgénicos son organismos que portan en su genoma ADN introducido exógenamente y que transmiten ese ADN fielmente a su descendencia. El ratón transgénico era absolutamente normal e indistinguible de sus hermanos no-transgénicos. Al igual que las bacterias o células transformadas, el ratón contenía en su genoma un poco más de ADN que el habitual, pero esto no le significa ninguna carga o perjuicio ya que no lo distingue de su propio material genético. El experimento clave ahora, era ver si ese ADN podía ser funcional. Como contenía una secuencia codificante (la hormona del crecimiento) y elementos de control (activación en respuesta a metales), el transgen tenía el potencial de conferirle al ratón una nueva característica (fenotipo) basado en ese nuevo gen (genotipo). Palmiter y Brinster tomaron los ratones transgénicos y a la mitad les dieron de beber agua con zinc durante las primeras semanas de vida. El resultado fue sorprendente y apareció en la tapa de la revista *Nature* donde publicaron su trabajo⁷⁷. El tratamiento con zinc indujo la expresión de la hormona y los ratones crecieron exacerbadamente, alcanzando el doble del tamaño de sus hermanos no tratados. La implicancia biotecnológica y médica fue tremenda. Podía pensarse en introducir características deseables en animales y plantas en un sólo paso. Bastaba con tener identificado el gen asociado a la característica y poner un elemento de control transcripcional que permitiera regular antojadizamente la expresión del transgen. Si pudiera introducirse en humanos, se pensó en corregir defectos genéticos o en “vacunas” genéticas que nos protejan de enfermedades infecciosas, entre otras aplicaciones.

La transgénesis en especies productivas es hoy ya parte de nuestra vida. Discutiremos más adelante las implicancias, virtudes y peligros con esta y otras tecnologías, pero queremos dejar en claro ahora lo que son y no son los transgénicos. El genoma contiene miles de millones de pares de bases, secuencias interminables de los 4 nucleótidos A, G, T y C. Un transgen esta compuesto

⁷⁷ Palmiter RD, *et al.* Dramatic growth of mice that develop from eggs microinjected with metallothionein-growth hormone fusion genes. *Nature*. 1982 Dec 16;300(5893):611-5.

por una secuencia de algunos miles de pares de bases de las mismas 4 “letras”. La combinación de ambos en un organismo transgénico no representa un cambio en la composición química del material genético (sigue siendo ADN) y, *per se*, no es un cambio “notado” por el organismo en cuestión. Sí, es diferente al organismo “silvestre”, pero esto no representa una diferencia conceptual relevante pues todos los organismos de una especie son diferentes entre sí a nivel de su secuencia de ADN. Aparte de todas las variantes puntuales en el genoma de las personas, tenemos también inserciones de ADN únicas producto de virus que nos han infectado, transposones que han saltado o movimientos de ADN espontáneos. Los conceptos de “silvestre” y “mutante” son, en realidad, engañosos, porque somos todos mutantes con respecto a cualquier otro ser humano, lo mismo en cualquier especie. Pensemos en el ratón transgénico recién descrito: si no se le daba de beber zinc, era indistinguible de los ratones no transgénicos. La diferencia radica en algo más importante que el ADN mismo: la información que codifica ese ADN. La secuencia lo es todo: si el organismo transgénico posee un ADN cualquiera, no codificante, sin elementos de control, etc., y, por lo tanto, no se le introdujo información, es un transgénico irrelevante y no debería ser tratado como algo genéticamente especial ni distinto del “silvestre”. Sigamos pensando sólo en el ADN y evaluemos la diferencia entre lo que es información codificada en él vs su composición química. Todos los organismos están compuestos de moléculas diversas (proteínas, lípidos, azúcares, ácidos nucleicos) de los cuales el ADN es el más consistente y constante entre ellos. Si me entregan un ADN degradado, una mezcla de nucleótidos A, G, T y C, no puedo decir de donde viene, pues se perdió la información que codificaba⁷⁸. De hecho, esto ocurre en nuestro intestino cuando comemos. Mucho de lo que comemos esta (o estaba) vivo y por lo tanto, contiene ADN. Una hamburguesa de su local de comida rápida preferida contiene ADN de vacuno (espero), lechuga, tomate, gallina (de la mayonesa si es hecha de huevo), pepino, tal vez repollo o mostaza⁷⁹. Las macromoléculas del alimento son degradadas por las enzimas digestivas y las usamos para armar nuestras propias macromoléculas. El ADN de vaca o de lechuga es descompuesto y usamos todas esas As, Gs, Cs y Ts para sintetizar nuestro propio ADN cada vez que lo replicamos y una nueva célula nuestra se

⁷⁸ Imaginen que tomo las páginas de El Quijote y con una tijera lo recorto en miles de pedacitos, cada uno con una letra del texto, y se los doy en una bolsa. ¿Sigue siendo El Quijote?

⁷⁹ A mis alumnos, que acostumbran comer en cualquier carrito de la calle, les digo que probablemente consumen también bastante ADN de *E. coli* en cada mordisco.

forma. El “ADN” de la vaca o del tomate no entran a nuestro cuerpo o, mejor dicho, la información genética no entra aunque sí lo hacen los constituyentes básicos del material hereditario. En este sentido, si el animal o plantas consumidas fueran transgénicas, el ADN del transgen tiene el mismo destino. Recordemos esta consideración para más adelante cuando nos preguntemos que peligros pueden tener -o no- los organismos transgénicos. Pero convengamos en que la categorización de un organismo como transgénico, o eufemísticamente, organismo genéticamente modificado (en inglés, *Genetically Modified Organism, GMO*), es una definición carente de sentido biológico. Todas las especies son “genéticamente modificadas” ya que constantemente su genoma esta sufriendo modificaciones, tanto evolutivamente como en su vida individual. Dejaré de lado mi queja por la imprecisión del lenguaje y asumiré que la etiqueta *GMO* se refiere a un organismo cuyo genoma ha sido manipulado artificialmente por el hombre, aunque esto no considera los siglos de manipulación artificial de los cultivos y animales de producción o compañía (siendo ellos, también, artificiales), sino la modificación genómica hecha en un paso y con la tecnología de la transgénesis. De acuerdo. Entonces, si veo la etiqueta *GMO* en un producto en el supermercado⁸⁰, debo entender que es transgénico para algo. ¿Para que? No aparece por ningún lado. ¿Es una hormona? ¿Es una resistencia al frío? ¿Es un gen de tolerancia a un herbicida? ¿Es una papa con más ojos, una zanahoria más anaranjada, un plátano más dulce? Nunca le entregan al consumidor el único dato que sería relevante: aquella que describe la modificación genética e indica que información se le adicionó a esa especie. Mucho menos que pruebas se hicieron para descartar que el producto genere alguna consecuencia negativa al ser consumido. Es decir, la etiqueta es inútil y tendenciosa: proclama el uso de una tecnología sospechosa para la elaboración y sugiere su rechazo *a priori*⁸¹. Volveremos sobre esto en la parte III.

CRISPR/Cas y la edición de genomas: la revolución que se viene

⁸⁰ En EEUU, hay un avance: se indica claramente en las etiquetas de los productos cuando estos no son genéticamente modificados (*non-GMO*). Mi queja lingüística sigue válida, sin embargo, ya que no existe casi ninguna especie *non-GMO*.

⁸¹ En las etiquetas no dice “vegetal regado con aguas servidas” o “usamos un pesticida altamente tóxico para impedir las infestaciones de gusanos en nuestro producto”, ambas que me asustarían bastante más que “GMO”.

Si consideramos que las dos revoluciones biotecnológicas importantes hasta ahora son el ADN recombinante y la transgénesis, la tercera será la de mayor impacto sobre nuestra condición humana, al menos potencialmente. Hasta ahora, carecíamos de una tecnología que nos permitiera intervenir de manera segura y efectiva en el genoma humano. En animales, plantas y microorganismos nos atrevemos, por nuestra asumida y temeraria complacencia con la explotación de otras especies⁸². Pero, si hemos desarrollado la genómica y hemos secuenciado al humano con el pretexto de curarlo de todos los males genéticos (Clinton *dixit*), ¿cómo intervenimos ese genoma? Una de las esperanzas era la transgénesis. Si logramos introducir ADN al genoma del ratón, ¿no lo podríamos hacer con un humano? Perfectamente podríamos: tomamos óvulos recién fecundados, los microinyectamos con el ADN de interés, los reimplantamos en la dama (en este caso, usaríamos a a misma persona tanto para tomar los óvulos como para reimplantarlos) y en nueve meses tenemos al bebé transgénico. Esto se puede hacer hoy, pero tiene varios inconvenientes. Primero, estaríamos agregando información genética, no corrigiéndola; serían contados los casos en que esto fuera útil. Segundo, la transgénesis es imperfecta en un sentido crucial: no podemos controlar en que lugar del genoma se integra el transgen. Existe la probabilidad que el ADN introducido se integre donde hay un gen endógeno preexistente, mutándolo. Sería peor el remedio que la enfermedad. Tercero, estaríamos curando a la siguiente generación, no a la que ya padece la enfermedad. Si bien es altruista pensar en reducir la carga genética perjudicial a nuestros descendientes, no soluciona el problema de la persona que necesita ser curada ahora ya. Para esto último, tendríamos que pensar en un método que cambie el genoma de un paciente, de una persona ya existente, en que todas sus células contienen el “error” genético. Esto es mucho más difícil, pues tendríamos que microinyectar o transformar todas sus células, o al menos, las del tejido afectado para lograr la “cura”. Este es el desafío de la terapia génica, ciencia que intenta reparar genomas en células somáticas (no en las germinales, que involucran a la siguiente generación). La esperanza de la terapia génica es desarrollar métodos que permitan entregar un transgen⁸³ a un alto número de células para tener un efecto. Hay un tejido que se presta mejor que otros para esta aproximación:

⁸² Explotación que hacemos por necesidad. Nótese que no me refiero a la crueldad carnívora pues incluyo a las especies no-animales; somos igualmente arrogantes con todas las especies.

⁸³ Podemos intentar introducir ADN pero que este no se integre, para reducir el riesgo de la interrupción de un gen endógeno. Pero esto no tendría un efecto permanente ya que el ADN no integrado usualmente termina desapareciendo de las células.

la sangre. Es el único tejido que podemos sacar del cuerpo (por un rato), manipular las células, y devolverlas al cuerpo modificadas. A esto llamamos la terapia génica *ex vivo*. Hicimos un primer intento hace algunos años y falló, lamentablemente. Se trató de curar una enfermedad inmune (aquella de los niños burbuja) con un ADN incorporado en un retrovirus (que representa una forma más eficiente de integración de ADN exógeno). Se introdujo el retrovirus en las células inmunes, pero el ADN se integró -en varios de los pacientes- en la vecindad de un gen endógeno, que resultó ser un oncogen. Estos pacientes desarrollaron una leucemia (porque se empezó a expresar el oncogen inducido por la integración del transgen en su vecindad), lo que agregó un cáncer a la deficiencia inmune previa. El negativo resultado hizo detener por muchos años las pruebas clínicas con terapia génica⁸⁴. Lo que nos falta es un sistema que permita introducir un ADN exógeno en un lugar preciso del genoma, y mejor aún, que haga la corrección del gen endógeno y no incorpore nuevo material, evitando los efectos no deseados del proceso. Bueno, es justamente lo que se acaba de desarrollar -o casi- en varios laboratorios del mundo. Se trata de la metodología llamada CRISPR/Cas, o CRISPR (pronunciado “crisper”), nombre proveniente de *clustered regularly interspaced short palindromic repeats*. La historia de cómo se generó esta herramienta que promete ser potentísima en muchas áreas de la biología es fascinante y recomiendo leer algunos libros sobre ello⁸⁵. Como tantos descubrimientos en ciencias (paso el aviso) este salió fortuitamente de investigación fundamental en biología. Resulta que algunas bacterias presentan un sistema de defensa contra la entrada de ADN exógeno que es altamente sofisticado, una especie de inmunidad contra los virus que frecuentemente las infectan. Estas bacterias inventaron un mecanismo para tomar ese ADN ajeno, extraer un pequeño fragmento, y guardarlo en su genoma. Lo almacenan en una parte especial de su genoma que esta marcada con secuencias iteradas muchas veces y que le sirven para distinguir lo propio de lo ajeno. Esas secuencias iteradas son cortas y palindrómicas⁸⁶, y se encuentran flanqueando una serie de pequeños fragmentos representativos del ADN externo que han ido coleccionando en el tiempo,

⁸⁴ Ver discusión en Williams & Baum, 2003. *Science* 302:400-401. <http://science.sciencemag.org/content/302/5644/400.full>

⁸⁵ Pueden buscar usando la palabra “crisper” en la sección libros de Amazon para ver lo que hay. Recomiendo en este instante (2017) el de Jennifer Doudna, una de las descubridoras del fenómeno: *A Crack in Creation: Gene Editing and the Unthinkable Power to Control Evolution*.

⁸⁶ Un palíndromo es un texto que se lee igual en ambas direcciones: “Anita lava la tina”. En el caso del ADN, se leen las hebras complementarias: GAATTC es un palíndromo pues la hebra complementaria leída en la dirección derecha a izquierda (5'-3') es la misma secuencia.

obtenidas de infecciones previas de virus. Estas secuencias son transcritas (ADN—> ARN) y patrullan la célula como ARNs pequeños junto con una enzima llamada Cas (proveniente de la *CRISPR Associated Sequence*), buscando una secuencia complementaria que delate al virus invasor. Cuando la encuentran (la reconocen porque la secuencia es complementaria ya que provino de una infección previa del mismo ADN), la enzima Cas degrada el ADN del virus. La elegante solución de la bacteria para defenderse de ADNs ajenos inmediatamente generó la expectativa de que la maquinaria involucrada (los pequeños ARNs repetidos/palindrómicos junto con las secuencias complementarias al blanco y la enzima Cas) podrían funcionar fuera de contexto, como lo hacen prácticamente todos los sistemas biológicos, cuando se trata de ácidos nucleicos. En otras palabras, la idea era tratar de que el pequeño ARN junto con la enzima Cas pudieran romper dirigidamente ADN dentro de cualquier célula, incluso las eucarióticas. Si podemos diseñar nuestras propias secuencias de ARN, podríamos llevar a la enzima Cas a donde nos plazca, en cualquier genoma. La prueba de principio de que el sistema era funcional lo hicieron Jennifer Doudna y Emmanuelle Charpentier en una fructífera colaboración que de seguro las llevará al Premio Nobel. Inmediatamente, muchos laboratorios comenzaron a hacer pruebas en células eucarióticas y eventualmente en modelos animales y plantas. Entre ellos, los grupos de Feng Zhang, Rudolph Jaenisch y George Church lograron hacer que funcionara la edición de manera muy eficiente en ratones y otros mamíferos, pavimentando el paso a pruebas en células y luego, embriones humanos. ¿Cómo lo hace esta maquinaria bacteriana para funcionar en una célula humana? Es simple: se diseña un ARN pequeño que tenga secuencias complementarias a alguna secuencia del ADN humano, y se introduce en la célula junto con la enzima Cas. Ambas moléculas se unen, llegan al sitio del genoma seleccionado por la complementareidad con el ARN pequeño y Cas causa una degradación o corte en el ADN genómico en ese lugar. La célula, que ve interrumpido su ADN, intenta una reparación para recomponer la continuidad de su secuencia genómica. Sin embargo, esta reparación es imperfecta: la maquinaria de reparación no siempre restituye la secuencia original sino que usualmente agrega o remueve unas pocas bases. Por lo tanto, se introduce una mutación. Por primera vez, usando un sistema muy simple (se introduce a la célula nada más que un pequeño ARN y la enzima Cas) se pudo mutar un sitio deseado en el genoma, con gran especificidad por ese sitio. En la tecnología CRISPR/Cas, que hoy usamos rutinariamente (solo 5 años después de su descubrimiento), nombramos al ARN pequeño como ARN guía, o ARNg (*guide RNA*, *gRNA*) y la versión de la enzima usada es la Cas9.

Ha causado gran preocupación la rapidez con que se han hecho pruebas en células humanas y, luego, en embriones humanos. Tanto en China como en EEUU, se han introducido mutaciones o hecho correcciones en el genoma de embriones humanos muy tempranos, en el estado de blástula. Estos embriones no fueron llevados a término, pero se demostró que la posibilidad de manipular el genoma humano a voluntad y generar personas con esos cambios, esta a la mano. La pregunta ahora es: ¿que hacer con esta poderosa tecnología? La edición genómica en embriones humanos es posible pero no deseable, por ahora. ¿Porque? Como dijimos antes, tendríamos que preocuparnos primero, de las personas existentes que sufren de enfermedades genéticas que no hemos podido resolver. Es cierto que muchas de esas personas quieren tener hijos y no quieren transmitirles sus “defectuosos” cromosomas⁸⁷. Pero, podemos ofrecer otras estrategias para hacer trampa en la ruleta genética: podemos seleccionar embriones producidos “naturalmente” por esa persona (y su pareja) que no sean portadores del gen mutante. Esto consiste en realizar fecundación *in vitro* para crear un grupo de embriones (tal como se hace en los procedimientos de reproducción asistida), hacer un test genético de ellos (se saca una célula de las varias que tiene el embrión temprano, lo que no afecta su desarrollo) y se implantan en la madre aquellos que dan negativo para la mutación. Hay médicos especializados en consejería genética, que nos orientan en este tipo de decisiones reproductivas. Esto es, obviamente, mucho más seguro y conveniente que la manipulación del genoma por edición con CRIPSR/Cas. Pero hay casos en que sería deseable hacer un cambio permanente en el genoma. Por ejemplo, se podría pensar en “vacunas” genómicas, cambios en el ADN que nos permitan sobrevivir infecciones futuras, como la del VIH. De hecho, ya sabemos cuales cambios son aquellos. Muchas de las personas que sobreviven el SIDA lo hacen porque tienen ligeras variaciones de secuencia en genes importantes para la defensa antiviral que los hacen resistentes (ver sección III). ¿No sería bueno “mutar” a toda la población futura para borrar este mal de la faz de la tierra? ¿O dotar a todos los futuros humanos de resistencia al dengue, malaria, hanta, ebola, etc., etc.? En principio, podemos imaginar esta aplicación de la edición genómica y no será difícil ponerla en práctica, junto con varias otras soluciones en el área de salud humana. Hay

⁸⁷ Toda persona viva porta en sus células alrededor de 50 mutaciones letales o causantes de enfermedades catastróficas. Al estar estas, casi siempre, en heterocigosis (una copia de las dos esta afectada, la otra siendo silvestre), no nos damos cuenta. El problema aparece cuando dos personas portadoras de mutaciones en el mismo gen, tienen la mala suerte de aportar cada uno un gameto con la versión mutante cuando forman un hijo; este, siendo homocigoto, presentará el fenotipo.

dos temas que trataremos en el siguiente capítulo que hacen que esta aproximación no sea deseable, al menos hasta que no lo pensemos un poco mejor. Pero si salimos de los humanos, ¿que nos impide hacer cambios en el genoma de otras especies a nuestro regalado gusto? Pues nada. Sumándose a los organismos transgénicos en agricultura ya existen hoy los organismos “editados”. Se viene una explosión de animales y plantas de consumo que serán editados para su mejoramiento; es crucial que nos demos el tiempo, como sociedad, para entender el proceso de edición y sus consecuencias. Pasemos a esto último.

III Consecuencias

Alcances de la manipulación genética de las especies.

La literatura de ciencia ficción esta llena de experimentos genéticos de los más mundanos a los más radicales, de los plausibles a los absurdos. Y a veces, lo que es claramente ficción al momento de la escritura, se aproxima con el tiempo a la realidad haciendo que admiremos la capacidad predictiva del autor. Es el caso de *Jurassic Park*⁸⁸, de Michael Crichton. Crichton tiene formación científica por lo que sus libros abordan temas firmemente anclados en la ciencia posible, salpicados de una dosis de fantasía. ¿Será posible hacer resucitar los dinosaurios? La idea no es totalmente descabellada pero tendrían que darse varios supuestos como para darle credibilidad hoy. En el libro (y la película correspondiente), entendían que no tendríamos acceso a la secuencia del genoma de un dinosaurio; el ADN tiende a descomponerse con el tiempo y en los restos de vida antigua -con suerte- encontramos pequeños fragmentos. En el cuento, el ADN de dinosaurio lo extraían de ámbar, donde quedaron atrapados insectos jurásicos que habían chupado sangre de dinosaurio. Los fragmentos obtenidos se empalmaron a ADN de reptiles o anfibios actuales y el genoma mosaico resultante se introdujo en huevos de lagarto (suponemos vía algún tipo de transgénesis), naciendo unos *Tyrannosaurus* que, en la parte más entretenida de la saga, terminaban devorando a los turistas del parque y sembrando el terror. El clásico tratamiento de este tipo de aventuras (lo llamaré el modelo Frankenstein) es una historia en que los científicos, por ambición y arrogancia, llevan una tecnología demasiado lejos y terminan generando un “monstruo” que nos alecciona sobre lo peligroso que es jugar con la naturaleza. En el inconsciente colectivo, se ha formado una impresión de que hacer estos experimentos conlleva un peligro: que estos se escapen del control de los científicos y se produzca una calamidad. La genética es particularmente susceptible a esta impresión, aunque sostengo que no todos los experimentos genéticos son temidos por la población. Desde hace miles de años, los humanos estamos haciendo notables experimentos genéticos con animales y plantas. La vaca

⁸⁸ https://www.amazon.com/Jurassic-Park-Novel-Michael-Crichton/dp/0345538986/ref=sr_1_1?ie=UTF8&qid=1503925551&sr=8-1&keywords=jurassic+park+book

lechera, el perro chihuahua⁸⁹ y el brocoli, son ejemplos de exitosas transformaciones genéticas completamente guiadas por humanos (no son especies naturales ni surgieron en la tierra mediante la selección natural). Desde que cultivamos y mantenemos especies para nuestro beneficio, hemos estado haciendo selección genética, buscando combinar propiedades útiles o atractivas. En el fondo, hemos creado genomas que jamás habrían existido sin nuestra intervención, la mayoría de las veces, sin tener idea alguna de lo que esos genomas contienen. La inocente abuelita que cada año selecciona las flores más bellas de su jardín y guarda esas semillas para plantar al año siguiente (en desmedro de las otras), esta perpetrando una flagrante intervención genética. Se calcula que el *Homo genetistus* existe desde hace unos 10.000 años, desde que dejamos de ser recolectores y empezamos a domesticar la naturaleza. Me atrevo a afirmar que ni el más recalcitrante vegano y purista de lo “natural” podría sobrevivir mucho tiempo prescindiendo del consumo o uso de organismos que fueron modificados genéticamente previo a 1970. Desde esa fecha, lo que ha cambiado es la velocidad⁹⁰. Si quiero generar una vaca que produzca más leche, puedo hacer lo de nuestros antepasados y criar vacas por generaciones o, en cambio, modificar un gen que incida directamente sobre este fenotipo y tener, en unos pocos años, la misma vaca *champion*. Evidentemente, hay modificaciones genéticas que nunca podríamos obtener sin la tecnología del ADN recombinante o de la transgénesis o que, simplemente, tomarían décadas o siglos obtener. Si somos honestos, la sensación de vértigo cuando pensamos en los procesos transformadores es la génesis del rechazo, más que el hecho en sí. ¿Porque? Pienso que hay ejemplos de “intervención” o de experimentación humana -no sólo en genética- en que un resultado malo o perjudicial se retiene en el inconsciente colectivo y son citados como casos que aleccionan a la sociedad y le imponen límites. Nos gusta ver que a las mentes más brillantes también les ocurren desgracias y que son humillados como cualquier mortal a causa de un error o por no haber considerado todas las consecuencias de su acto. Esa falibilidad reconforta y nos empareja como humanos en un gran colectivo que es profundamente ignorante de los secretos más íntimos de la naturaleza. Reconozco como científico que, en lo más

⁸⁹ Algunos tipos de canes podrían perfectamente ser considerados monstruosos; jamás sobrevivirían fuera de la protección del hogar.

⁹⁰ Algo válido para todas las tecnologías. Lo que nos asusta no son los cambios, sino lo brusco de ellos.

profundo, me siento ignorante e inseguro en muchas de mis hipótesis, ideas y principios⁹¹. De verdad, la naturaleza nos sorprende a cada rato con variables insospechadas, excepciones a las reglas y una enrabante capacidad de enrostrarnos nuestras limitaciones y falsas concepciones. Pero, a pesar de ello, la prueba está en que avanzamos más de lo que retrocedemos. Algo estamos haciendo bien. No haré aquí una defensa corporativa de mi profesión y de nuestras apuestas porque la evidencia del beneficio que ha traído la ciencia está al alcance de cualquiera y negarlo sería retrógrado e irracional. Lo lógico es seguir por este camino y pavimentarlo de buenas intenciones. La investigación científica tiene dos justificaciones: el avance del conocimiento *per se* y el mejoramiento de la calidad de vida y dignidad humanas⁹². Ambas razones son suficientes para sostener el progreso que buscamos.

Veamos los pro y contra de la intervención genética, primero en organismos no-humanos. En plantas y animales, las aplicaciones son múltiples. La producción eficiente y limpia es una obvía. También cosechar un mayor volumen de grano por hectárea usando menos fertilizantes o agua o aprovechar terrenos de otra manera estériles. ¿Quién podría quejarse de ello? Tener animales de producción que crezcan más rápido, sin enfermedades o pérdidas de viabilidad, que consuman menos alimento y requieran menos medicamentos, ¿acaso no sería desable? Voy a suponer que hasta el más fundamentalista de los opositores a la modificación genética tendría que estar de acuerdo en lo beneficioso que resultaría esto si no hubieran efectos colaterales no deseados a la tecnología. Al menos, no he escuchado argumentos serios de oposición que se basen solamente en un anti-progresismo ortodoxo. Más que tomar los argumentos del bando de los anti-GMO, haré mi propia crítica, que se sobrepone parcialmente con la de esos activistas, pero que tiene una diferencia fundamental: no haré una oposición global de la modificación genética (de la cual, por definición, no puedo estar a favor ni en contra⁹³) sino que pretendo que se evalúen méritos y defectos a cada caso de modificación en particular. Agrupo los efectos negativos de la modificación genética en tres campos amplios: la salud humana, la preservación del medio ambiente y la justicia económica/social. Las dos primeras son del ámbito de las

⁹¹ La definición de ciencia incorpora *de facto* la falseabilidad. Al verdadero científico le encanta que le demuestren que lo que piensa es falso...bueno, de vez en cuando nos gusta.

⁹² A veces se intenta contraponer esos dos objetivos, separando la investigación “fundamental” de la “aplicada”; aquí no hacemos esa distinción pues operan en la práctica sobre el mismo principio: el método científico.

⁹³ Hacerlo, sería como decir “estoy en contra de la radiación gamma”. Muy distinto es declarar “estoy en contra del uso de la radiación ionizante producida por un arma nuclear” o “estoy en contra de las bombas atómicas”.

ciencias biológicas; la tercera corresponde al campo de las ciencias sociales. La primera preocupación tiene que ver con que los consumidores de esos organismos modificados seamos las personas⁹⁴. ¿Hay algún peligro a nuestra salud asociado al consumo de productos GMO? De manera genérica, no. Ya vimos que el ADN, venga de donde venga, es el mismo en todas las especies. Comer plantas y animales implica comer ADN todo el tiempo lo que no es peligroso. Por lo tanto, un organismo cuyo ADN ha sido modificado no representa un peligro para la salud humana al ser consumido, punto. Insisto: el ADN no es peligroso, venga de donde venga. Lo que podría ser peligroso es, evidentemente, la información que codifica ese ADN. Si una cosa quisiera que quedara claro de este libro es que el ADN es una molécula de información y que esa información es la clave de la vida. Si estamos en presencia de un organismo GMO al cuál se le ha incorporado ADN pero no información, simplemente no es verdaderamente GMO. Como dijimos en el capítulo anterior, la categorización de producto GMO (y por consiguiente NO-GMO) es irrelevante y no deberíamos prestarle atención. Por el contrario, un consumidor informado estaría mucho más interesado en saber cuál modificación genética tiene el organismo en cuestión y conocer las pruebas de toxicidad, inmunogenicidad y seguridad alimentaria que se han hecho en ese organismo. Lamentablemente, no conocemos esas pruebas para ningún producto de consumo, sea o no GMO. Ni sería práctico obtenerlas. “Aquí tiene un tomate casera”. “Ah. ¿Fue tratado con pesticidas? ¿Cuales? ¿Cuanto tiempo pasó almacenado? ¿En que condiciones? ¿Tuvo siempre refrigeración? ¿Quién lo manipuló?” Lo que me gustaría señalar con este punto es que los consumidores, quienes no tenemos tiempo de investigar todos estos temas antes de partir al supermercado o feria, hacemos un voto de confianza con el sistema público-privado que nos rige. Los proveedores actúan bajo ciertos principios y normas⁹⁵ al igual que las instituciones que resguardan la salud de la población. Hay un acto de fe pública en este ámbito y nos sometemos al criterio de autoridades y reguladores, confiando en que nos cuidarán de envenenamientos, infestaciones y epidemias. Si poseen evidencia que la presencia de la bacteria *Salmonella* en los huevos de gallina es nociva, y prohíben su venta cuando es detectada, ¿porque no pensar que tendrán un criterio similar con un producto GMO? Como individuo, yo no puedo

⁹⁴ Esto puede ser indirecto ya que podríamos consumir productos de un animal que no ha sido modificado, pero que fue alimentado a su vez con productos GMO.

⁹⁵ Sí ya sé. Me dirán que el único principio que los guía es el de las leyes de mercado. Aún así, es de toda su conveniencia mantener la más alta calidad sanitaria para tener un negocio sustentable.

asegurar seguridad sanitaria de ningún producto alimenticio; lo tiene que hacer una autoridad responsable. En el caso de los GMOs, yo pediría como ciudadano que se hagan -por parte del estado- las pruebas necesarias para garantizarme que el producto no será perjudicial a mi salud y que se usen los criterios más rigurosos de que se dispongan. Definitivamente, puedo imaginar muchos productos GMO perjudiciales, incluso, quizás ya los hay. Si se expresa la hormona del crecimiento en los vacunos, para que estos alcancen un tamaño mayor al “natural”, no quisiera exponerme a la asimilación en mi cuerpo de esa hormona, que no es ADN sino una proteína producida por el ADN incorporado. Si una planta expresa una proteína de resistencia a un herbicida (por ejemplo, al glifosato, herbicida desarrollado en los '70 en la empresa Monsanto), quisiera saber si esa proteína es inocua para mi organismo⁹⁶. Y para ser más precavido: me gustaría que se probara, en el organismo transgénico que expresa una proteína introducida, si hay alteración en la expresión de sus propias proteínas, las endógenas. Se ha observado que la sobre-expresión de un producto génico, a veces genera un desbalance en la expresión de otros genes alterando sus niveles. Nuevamente, el tema se resuelve con pruebas adecuadas de toxicología e inocuidad.

Pasemos al segundo punto que, en mi opinión, es el más serio en cuanto a la precaución que hay que tener con los organismos productivos genéticamente modificados. Por razones prácticas, los cultivos de animales y plantas se realizan en los ecosistemas que comparten junto a todo tipo de especies “naturales”. Un campo de maíz es un pedazo de tierra donde podría haber, en ausencia de *Homo sapiens*, un bosque. Bueno, el hombre sobre la tierra modifica su entorno, no hay alternativa para su sobrevivencia moderna⁹⁷. El maíz, como ya explicamos, corresponde a una variante seleccionada artificialmente (domesticada) de la especie natural originaria del continente americano llamada teosinte. Hace 9.000 años, los habitantes de México central hicieron un experimento genético extraordinario, transformando un pasto con unos pocos granos en su espiga, en una planta que hoy es la fuente del 21% de la nutrición humana planetaria. Si bien hubo selección genética, se mantuvo de igual modo, en las variedades cultivadas inicialmente, una alta diversidad genética. De esta diversidad genética (genomas compatibles

⁹⁶ Ojo, no confundamos la toxicidad de la proteína de resistencia al herbicida con la toxicidad del herbicida aplicado en esas plantas, son cosas distintas y con pruebas distintas.

⁹⁷ Los más radicales propondrán que volvamos a ser cazadores-recolectores, único modo de vida en que actuamos como una especie más y nuestra intervención del entorno es mínima.

pero portadores de diferencias funcionales en genes) surgen distintas variedades de maiz, tolerantes a distintos climas, con diversos rendimientos y resistencias a condiciones locales. El hallazgo de variedades “mejores” genera inmediatamente un proceso de selección artificial: la predilección por el cultivo masivo de esas variedades. La consecuencia, obvia, es el desmedro de las otras variedades al punto de su desaparición. Es decir, partir de variedades y seleccionar la mejor para su propagación masiva, perjudica la misma variación que le dió origen. Al comienzo de este libro hablamos de evolución, y de lo crucial que resulta para la selección natural la presencia de variación en las poblaciones. Si reemplazamos variación por uniformidad, la selección no tiene sustrato para actuar y se abre la posibilidad de una consecuencia fatídica: la extinción. ¿Porque? Muy simple: las condiciones ambientales no son uniformes; el cambio permanente de medio ambiente (físico y biológico) genera una constante presión a las especies para acomodarse, un proceso adaptativo que incide directamente en sus genomas. Si la selección es muy estricta y rápida, por ejemplo, si una de las variantes genéticas en particular presenta una sobrevivencia muy superior al resto de ellas, se genera un cuello de botella genético. Este fenómeno natural, genera una vulnerabilidad a la especie durante el tiempo que demore en restaurarse la variación ya que un repentino cambio ambiental puede devastarla y hacerla desaparecer. En el caso de las plantas y animales de producción agropecuaria, nuestro instinto económico nos ha hecho tender hacia la uniformidad genética, desperdiciando la riqueza que nos ofrece la variabilidad natural. Si a esto sumamos una transgénesis que constantemente selecciona, corremos el riesgo de devastadoras consecuencias si surge un cambio no tolerado por la variedad predominante⁹⁸. Entonces, por el lado ambiental, el primer problema que enfrentamos con los transgénicos es el de la acelerada disminución de la diversidad genética de las especies que explotamos⁹⁹.

Por otro lado, existe el riesgo de que las modificaciones genéticas se transmitan a poblaciones naturales de la misma especie o de especies relacionadas, donde pueda ocurrir

⁹⁸ Podemos hacer una analogía -aunque en el sentido contrario- con los antibióticos y las bacterias. El abuso de los antibióticos ha favorecido la propagación de cepas que, gracias a la variación genética presente en las poblaciones microbianas, adquieren resistencia. Si las bacterias fueran genéticamente uniformes, podríamos barrer con ellas despiadadamente.

⁹⁹ Un interesante esfuerzo mitigador de esta amenaza se está haciendo en el banco mundial de semillas, un repositorio internacional ubicado en una caverna en las heladas montañas de Noruega, donde se almacenan reservas de todas las plantas cultivables que son colectadas en el mundo. Ver más en: <https://www.regjeringen.no/en/topics/food-fisheries-and-agriculture/jordbruk/svalbard-global-seed-vault/id462220/>

transmisión genética horizontal. Este fenómeno es natural y se ha observado en microorganismos y, en menor medida, en organismos superiores, sobre todo en plantas, aunque ocurre a una muy baja frecuencia¹⁰⁰. Sin embargo, el fenómeno puede exacerbarse cuando hay selección artificial y la transmisión de genes que confieren resistencia. Un individuo que adquiriera el gen de resistencia, inmediatamente posee una fuerte ventaja competitiva sobre sus congéneres, lo que puede hacer que su descendencia prospere en desmedro del resto. Es por ello, que la autorización para hacer cultivos transgénicos necesariamente debería tender hacia el aislamiento reproductivo de esas variedades, para impedir la transmisión horizontal. Una consideración adicional al de la transferencia del transgen, es el efecto de la fuerte selección que se impone a las especies sensibles al agente de selección, por ejemplo, un herbicida sobre las malezas. Se ha visto ya, que muchas malezas han adquirido resistencia a los herbicidas sin ser transgénicos; lo lograron acelerando su evolución y buscando en su propio genoma una solución. Hay una maleza que logró amplificar el número de copias del gen que es blanco del glifosato, haciéndola resistir a ese producto químico; lo hizo activando elementos repetidos en el genoma (transposones) y generando muchas copias del gen blanco. No nos sorprendamos cuando la naturaleza contraataca: tiene un arsenal inagotable de armas para defenderse de nuestra ineptitud.

Finalmente, nos queda referirnos al tema de los impactos sociales y económicos de la transgénesis en organismos productivos. Este problema es análogo al que nos enfrentamos como sociedad en medicina, con fármacos y procedimientos. El sector privado invierte cuantiosos recursos en la investigación y pruebas necesarias para llevar un fármaco al mercado. La recompensa viene por el lado de la propiedad intelectual e industrial ya que usualmente, el compuesto o tratamiento es patentado y la empresa tiene, por un número de años, la venta exclusiva. En simple, es un monopolio acordado. De igual manera, las empresas del área de producción animal y vegetal, patentan variedades y organismos genéticamente modificados, pues invirtieron tiempo y recursos en crearlos. No nos compete pronunciarnos sobre la justicia de este esquema productivo y los países tienen -o deberían tener- mecanismos para proteger a los ciudadanos de los potenciales abusos. Proponemos que quizás una buena defensa sea incrementar la inversión en biotecnología para generar múltiples opciones o variedades eficientes

¹⁰⁰ El caso extremo de transmisión horizontal son los organelos de las células eucarióticas, los que son descendientes de organismos procariontes de vida libre que hicieron simbiosis con otros microorganismos.

que compitan entre ellas, la promoción de esfuerzos públicos en esta dirección¹⁰¹, y la cautela rigurosa de la magnitud de los monocultivos.

Dados los puntos que he tocado, ¿qué es lo que nos podría asustar de los transgénicos? A mi modo de ver, los temas 1 (riesgo a la salud humana por su consumo) y 3 (abuso económico de los actores productivos) son perfectamente solucionables con regulación. Hasta cierto punto, el 2 (riesgo genético-ambiental) lo es también, aunque es el que personalmente me preocupa más porque es aquel donde tenemos un conocimiento incompleto y muy poca experiencia histórica¹⁰². Nuestro maestro en este campo tendría que ser la evolución, diseñador que es mucho más “inteligente” que nosotros por la sencilla razón de que dispone de millones de años para hacer sus experimentos y esconde habilmente sus fallas. Es decir, hay que tener paciencia y disponer de los tiempos necesarios para las evaluaciones ambientales de largo plazo.

Hoy el tema de los organismos transgénicos/GMO se aborda de manera cautelosa por los gobiernos, por las empresas y por las organizaciones internacionales. Los científicos, hasta ahora, han permanecido demasiado silentes, más inseguros de cómo enfrentar comunicacionalmente el tema que contrarios a la tecnología. Un grupo numeroso de Premios Nobel han decidido sacar la voz y se han organizado para defender el uso de tecnologías GMO en agricultura¹⁰³. Los argumentos son fuertes, llegando a sugerir que la campaña en contra de los GMOs equivale a un crimen de lesa humanidad. En los países subdesarrollados de Asia y Africa, mueren alrededor de 2 millones de personas al año por deficiencia de Vitamina A. Esta vitamina no esta presente en suficiente cantidad en sus dietas, consistentes principalmente de arroz. Sin embargo, esta disponible una variedad de arroz transgénico, el llamado *golden rice*, que incorpora alta cantidad de Vitamina A en el grano; bastaría con que se plantara ese arroz en los lugares de carencia para reducir las muertes. Otros ejemplos son vacunas orales que pueden administrarse con plátanos transgénicos o maíz que es resistente a plagas; especies que estan prohibidas en varios países de Africa por presiones del primer mundo y que podrían solucionar algunas de las graves crisis

¹⁰¹ Conozco propuestas legislativas en EEUU donde se propone que la investigación hecha con fondos públicos, aún en la empresa privada, no pueda ser usada para fines privados; esto se enfoca a la industria farmacéutica, que se aprovecha del fuerte subsidio a la investigación biomédica en ese país y que luego aplica sobreprecios a las drogas que desarrollan producto de esa investigación.

¹⁰² Si bien, en los 40 años de historia de la ingeniería genética en alimentos no ha habido ningún problema serio, de todas formas es un tiempo corto en relación a los tiempos evolutivos o planetarios.

¹⁰³ Ver la página “*Support Precision Agriculture*”, http://supportprecisionagriculture.org/view-signatures_rjr.html

sanitarias y alimentarias que padecen. Es decir, a las precauciones sensatas que debemos imponer a los organismos GMOs de consumo, se opone la inmensa necesidad de ellos. Es demasiado cómodo opinar desde el lado más rico del mundo sobre que deben o no deben hacer las naciones pobres que enfrentan urgencias y crisis humanitarias devastadoras.

Transgénicos 2.0

Como vimos antes, la edición genética con la tecnología CRISPR/Cas presenta una importante diferencia con la transgénesis: en lugar de incorporar un ADN foráneo al genoma de la especie modificada, se hace un cambio en el propio ADN de esa especie. Es decir, no hay transferencia de información sino un cambio en la información previamente existente. Esa misma diferencia le impone limitaciones al alcance del método ya que debemos buscar formas para que la pérdida o modificación de alguna secuencia endógena altere la función génica en una dirección conveniente desde el punto de vista productivo. Quizás el mejor ejemplo actual sea el trigo sin gluten. Mediante CRISPR/Cas, a una variedad de trigo se le mutaron los genes que codifican para las proteínas que conforman el gluten, haciéndolo consumible para los celíacos¹⁰⁴. En este caso, era deseable quitar una propiedad a la planta. Por el contrario, como no podemos darle al organismo modificado una función génica que le es completamente ajena (sin usar transgénesis), los biotecnólogos que quieran usar CRISPR en este escenario deberán ponerse creativos. Ejemplo: se desea generar tolerancia en una planta de cultivo a una alta salinidad del suelo. Si usáramos transgénesis, incorporaríamos un gen que provenga de una especie que ya es tolerante a la sal y que confiera esta propiedad cuando transferido a la especie destinataria. Con CRISPR en cambio, tendríamos que silenciar genes propios de la planta que la hacen susceptible a la sal, si es que los hay. También esta la opción, de editar (o crear) secuencias regulatorias para que estimulen la expresión de genes que ya tiene la planta pero que no esta usando para la tolerancia a la sal. En este caso, no introducimos genes externos sino que modificamos la expresión de los genes “naturales”, generándose un organismo “mutante” pero no “transgénico”. En rigor, es genéticamente modificado (GMO), como todos los organismos de cultivo que han pasado por cualquier tipo de selección artificial, pero no debería llevar la impronta “GMO”

¹⁰⁴ Sánchez-León, S., et al. (2017). *Low-gluten, nontransgenic wheat engineered with CRISPR/Cas9*. *Plant Biotechnol. J.*
doi: <https://doi.org/10.1111/pbi.12837>

como es entendida por las regulaciones actuales, ya que no es transgénico, única modificación que exige tal nomenclatura (por ahora). Estoy consciente que esta distinción semántica no pasa de ser un ejercicio retórico de mi parte ya que el tema de los GMOs esta tan políticamente cargado que de seguro se intentará juntar los organismos editados (¿CRISPados?) con los transgénicos en un mismo malévolo saco de indeseables y peligrosos seres. No obstante, al ser una tecnología aún poco usada y conocida, los científicos tenemos ahora la oportunidad de dar la pelea para que se haga una discusión más seria del tema de fondo: ¿es válido modificar genéticamente organismos para mejorar la condición humana? Si somos capaces de educar a las sociedades respecto de algo que es fundamental para nuestra seguridad alimentaria en un contexto de progreso y sustentabilidad, tal vez se tomen decisiones racionales y científicamente sólidas. Si la conclusión, sensata e informada, dicta que deberíamos establecer restricciones, normativas estrictas o incluso moratorias, yo las apoyaría gustoso. Al mismo tiempo, me opondré a esas restricciones si se basan en mitos, pseudociencia o un dogmático antagonismo al progreso.

Echaré mano de mis pocos pelos para agregar a la confusa sopa bioética y mencionaré una aplicación de la tecnología de edición genómica con CRISPR que muchos no intuíamos. Y esta, en mi opinión, merece bastante más cuidado en su aplicación que la modificación de organismos con fines productivos. Resulta que la tecnología CRISPR permite crear potentes pseudo-organismos moleculares o parásitos genómicos con propiedades similares a las de un virus o transposón. Una estrategia, propuesta de manera teórica hace unos 15 años¹⁰⁵, consiste en integrar vía transgénesis en un organismo de reproducción sexual la enzima Cas9 y un ARNg (el ARN que guía a la nucleasa Cas9 a cortar el ADN en un lugar específico del genoma) dirigido contra una secuencia endógena del organismo portador. Si esta secuencia blanco esta en el cromosoma homólogo y en la misma posición genómica al lugar de la integración del ARNg/Cas9, la secuencia será cortada y el sistema de reparación natural de la célula usará al cromosoma intacto para la restauración, introduciendo una nueva copia del ARNg/Cas9 en ese lugar. El transgen que contiene el “casette” de edición entonces, quedará en ambos cromosomas pasando a todos los descendientes del individuo, copiándose en ellos de igual forma hasta que, en algunas generaciones, toda la población -o especie- porta el transgen. Esta estrategia, llamada *gene drive* (lo podríamos traducir como campaña o cruzada génica), fue propuesta por el grupo del Dr.

¹⁰⁵ Burt A. Proc. Biol. Sci. 270, 921.928 (2003).

George Church¹⁰⁶, quien ha llevado la tecnología CRISPR hacia aplicaciones poderosas, para algunos, apocalípticas. Han podido demostrar, en el laboratorio por ahora, que efectivamente es posible propagar aceleradamente un transgen de este tipo (lo llamaremos egoista) en una población de animales o plantas. Lo mismo ocurriría invariablemente en una población natural: si liberamos un organismo con este tipo de transgen egoista, se propagaría hasta abarcar la población completa, incluso aunque ser portador represente una leve desventaja selectiva (disminución del fitness reproductivo, por ejemplo). Es una herramienta que puede ser usada para llevar a una especie -en un tiempo brevísimo- a la extinción. ¿Porque haríamos aquello? Los bienintencionados científicos tenemos en la mira a varias especies que nos causan daño. Uno muy obvio es el mosquito, sobre todo, aquellas especies que son vectores de agentes infecciosos altamente dañinos como los del dengue, fiebre amarilla o malaria. Hemos estado, durante siglos ya, tratando de deshacernos de varias pestes y plagas con un arsenal que va desde el matamoscas casero hasta el DDT rociado desde aviones¹⁰⁷. La metodología del *gene drive* lograría el mismo resultado sin un esfuerzo mayor que el de soltar algunos mosquitos transgénicos. Antes de proceder, valdría la pena preguntarse si echaríamos de menos a los mosquitos si desaparecieran; me refiero a las consecuencias ecológicas de su ausencia ya que no creo que añoremos las molestas picaduras ni las horas nocturnas dedicadas a cazar a un infame zumbador. Hay ecólogos haciéndose la pregunta como parte de una evaluación seria de esta iniciativa y es posible que, dentro de poco, se autoricen pruebas piloto en lugares aislados (islas). En rigor, poco importa la respuesta ya que los proponentes del *gene drive* ya tienen una mejor alternativa al exterminio de la especie vector: con CRISPR podríamos alterar la genética de los mosquitos para que no transmitan más las enfermedades que queremos abolir. Antes de que nos deshagamos del repelente, esperemos ver cómo funciona esta alternativa, la que requiere algunas pruebas de laboratorio para evaluar su eficacia.

Los humanos hemos, irresponsable aunque inadvertidamente, llevado a la extinción a varias especies en los últimos siglos. Deliberada y conscientemente lo hemos hecho pocas veces, la

¹⁰⁶ Esvelt KM, Smidler AL, Catteruccia F, Church GM. Concerning RNA-guided gene drives for the alteration of wild populations. *Elife*. 2014 Jul 17;3.

¹⁰⁷ En 1958, Mao Tse Tung ordenó a los chinos exterminar las “cuatro pestes”: mosquitos, moscas, ratas y gorriones. Los gorriones estaban acusados de consumir los granos de las cosechas agrícolas; el resultado de su cuasi-exterminio fue la proliferación de las langostas (apetecidas por los gorriones), que diezmaron las cosechas de los años siguientes contribuyendo a la gran hambruna china de 1959-1961 en que murieron millones de personas.

más famosa siendo la erradicación del virus de la viruela. La herramienta del CRISPR nos permite hacer esto de manera más eficiente con algunas especies (de reproducción sexual), lo que nos otorga una nueva omnipotencia sobre nuestro planeta y el curso de la evolución sobre ella. En términos bioéticos, nos enfrentamos a dos consecuencias. Primero, como cualquier tecnología revolucionaria, ofrece aplicaciones destructivas, malignas o criminales, todas ellas de acceso fácil incluso para países o instituciones sin ninguna sofisticación¹⁰⁸. Segundo, estaríamos interfiriendo con un balance natural de una manera bastante más violenta (por lo instantáneo del efecto) que lo que normalmente hacemos con el cambio global por polución u otras consecuencias antropogénicas. Aún, el efecto del impacto de un asteroide en la desaparición de los dinosaurios fue paulatino (millones de años) comparado con lo que podríamos lograr con un *gene drive*. Veremos en seguida cómo podríamos usar CRISPRs en humanos, tema con su propia bioética, pero un *gene drive* que elimine versiones mutadas de genes o secuencias que nos predisponen a enfermedades es perfectamente factible. Evidentemente, es una solución a largo plazo para esas enfermedades, ya que su eliminación tardaría siglos por nuestro lento ciclo generacional.

Pienso que todo ciudadano en edad de sufragar debería tener claro lo que son y como funcionan las tecnologías de lectura, recombinación y edición de secuencias de ADN. Más que los aspectos técnicos, es importante informarse sobre lo que implica cada una de ellas para nuestra sociedad. Entender las ramificaciones biológicas y manejar conceptualmente los desarrollos tecnológicos permiten desmitificar la ciencia y nos hace enfrentar preguntas complejas racionalmente. Como titula un clásico panfleto evangélico, a veces me dan ganas de gritar: “¡Despertad!”. Me inquietan la desinformación y la ignorancia usadas en la toma de decisiones, una conducta que no es sólo culpa de nuestros gobiernos sino también de nosotros, los científicos, que somos incapaces de exponer de manera simple y concreta las evidencias que ayudarían a tomar mejores decisiones. La genómica, la transgénesis, la edición con CRISPR no son demonios; son herramientas que pueden ser bien o mal usadas y ese uso tiene que provenir de un consenso social. Y en el caso de nuestro propio genoma, tendremos pronto que tomar decisiones personales respecto a cuanto y cómo intervenir en él.

¹⁰⁸ Se llegó a proponer que quizás era conveniente no publicar la metodología para construir los *gene drive* para que no cayera en manos perversas (*Gene drives raise dual-use concerns. Science* 29 Aug 2014. Vol. 345, Issue 6200, pp. 1010-1011).

Gracias a la vida, que me ha dado TAAATCGCGTG

Imaginemos que un caballero viene ante mi, me entrega un *pen drive* y me dice: “Señor experto genómico: ahí esta la secuencia del genoma de mi hijo que esta por nacer; quiero que me diga con un 100% de certeza si será un genio o un bobo”. Después de un periodo de análisis de los datos, de reflexión, y de mucho respirar hondo, podría responderle: “Bueno Don Diógenes, no puedo darle ninguna certeza sobre lo que me pregunta ya que la inteligencia (o la idiotez) es un carácter altamente complejo que depende en gran medida de cómo Ud. eduque a su hijo. Lo que sí puedo decirle, es que no hemos encontrado mutaciones letales ni deletéreas en los genes de su hijo, por lo que esperamos que nazca sano y lleno de potencial como para llegar a ser muy inteligente. O sea, no eluda su responsabilidad y cuídalo”. Con este diálogo imaginario pretendo sincerar los alcances de la secuenciación del genoma de una persona, diferenciando lo que vamos a poder asegurar de lo que escapará de nuestro control. De a poco hemos ido morigerando el valor definitorio de la secuencia y entendiendo su rol en la complejidad genética de los organismos¹⁰⁹.

Esperamos haberlos convencido en nuestro relato previo de que la secuencia desnuda del ADN no describe completamente lo que “es” un organismo/individuo. Los genes pueden predisponer, pero no determinar. ¿Para que nos molestamos en tener la secuencia del genoma entonces? Porque esa secuencia describe el potencial genético, las posibilidades y limitaciones del individuo portador, existiendo por encima de ello un fenotipo plástico de amplio rango que es contexto dependiente. Vimos que la epigenética, la información contextual, es la que genera los “permisos” en el genoma, el cuál no tiene libertad de expresión, literalmente. Estas dos capas de información tienen roles separados si se quiere. La primera, la secuencia, es antigua y de cambio lento (lo hace en tiempo evolutivo) y la segunda, la epigenética, es dinámica, abierta al cambio en el tiempo individual, y va construyendo la narrativa de la vida de cada uno de nosotros. Esta capa es la que le da oportunidad al libre albedrío (y la responsabilidad individual), sin mencionar la cura de enfermedades, a pesar de la genética. Aliviados, ahora que sabemos que no venimos preconfigurados de nacimiento, veamos que podemos hacer con la secuencia de nuestro genoma, el genoma humano.

¹⁰⁹ El Premio Nobel Walter Gilbert anunció en un congreso en 1986 que “La secuencia completa es el santo grial de la genética humana”. Todos nos hemos dejado llevar por el entusiasmo alguna vez, ¿no?

Si fuéramos precisos, no se puede hablar de el genoma humano sino que tenemos que empezar a referirnos al genoma de un humano. El primer “genoma humano”, en realidad, se obtuvo a partir de varios individuos y, por ende, no era el de ninguno de ellos sino una combinación o promedio de secuencias; aquella es la que se encuentra en la base de datos a la que todos nos referimos cuando buscamos información sobre genes humanos. El primer genoma secuenciado de un individuo humano fue el de Craig Venter, quien secuenció su propio genoma. Ahora, tenemos con quien compararnos. Pero Venter es un señor caucásico que no necesariamente nos representa a todos los humanos; sería bueno contar con una variedad de personas para tener una muestra diversa de genomas humanos. Hace algunos años, se armó el programa 1000 genomas¹¹⁰, que se planteó la meta (alcanzada con creces ya) de secuenciar 1000 personas y describir la variación entre ellas. Hoy, tenemos ADN secuenciado de hombres y mujeres de todo el mundo, de una gran mayoría de las etnias, de individuos sanos y enfermos, jóvenes y viejos, vivos y muertos¹¹¹. Al momento de escribir este texto, me encuentro yo mismo en un proyecto de secuenciación masivo de personas de mi país¹¹². En Europa ya pasaron a los proyectos de 100.000 genomas y en China, a secuenciar millones de personas. ¿Para que sirve este esfuerzo? Muy pronto, comenzará la avalancha de secuenciaciones de los genomas de individuos que por razones médicas o personales decidan conocer esta información. La secuencia en bruto será fácil de obtener (vendrá en un pen drive, como ilustramos recién con nuestro encuentro imaginario con Don Diógenes), pero la interpretación de la información requerirá un análisis por parte de un experto, que a su vez necesita secuencias de referencia. Necesitamos saber *a priori* que es un genoma “normal” o, mejor, que secuencias son propias de personas sin enfermedades de origen genético y vice versa. Y como esto varía entre las poblaciones humanas, la comparación tiene que hacerse con una colección de genomas cercanos genéticamente. La comparación de un genoma particular con la “referencia” es un proceso informático complejo pues requiere distinguir las variaciones normales entre individuos de aquellas que son conducentes a enfermedad. La experticia requerida hoy para esto pasará luego a automatizarse y complejos programas harán este análisis.

¹¹⁰ *The One Thousand Genomes Project*. <http://www.internationalgenome.org/>

¹¹¹ Ya hablamos del genoma de *Homo neanderthlensis*. Se ha mirado el ADN de los restos de momias, personas que murieron a causa de las grandes pandemias y de personajes históricos.

¹¹² Proyecto 1000 genomas chilenos: www.1000genomas.cl.

Siguiendo con nuestra aspiración de contar con el genoma perfectamente secuenciado de una persona, nos encontramos con otra dificultad: el ADN que extraemos para secuenciar normalmente viene de un grupo de células (si la persona esta viva, lo que hacemos es extraer sangre ya que algunos leucocitos o glóbulos blancos contienen ADN). Pero, lo más probable, es que en ese grupo de células no hayan genomas idénticos entre todas ellas. Todo el tiempo, están ocurriendo mutaciones somáticas, aquellas que afectan las células de nuestro cuerpo que no transmiten material genético a la descendencia (en otras palabras, todas las células menos los espermios y óvulos). Cada vez que nos exponemos al sol u otras radiaciones, ingerimos agentes tóxicos o mutagénicos (el smog de Santiago de Chile siendo de los más potentes, con seguridad) o, simplemente, envejecemos: se introducen, lenta pero inexorablemente, cambios en nuestro ADN que se acumulan en cada célula. Se ha calculado que cada segundo de nuestra vida, ocurren en el conjunto de nuestras células alrededor de un millón de mutaciones. Es decir, si lográramos obtener los genomas de células individuales de una persona, encontraríamos que contienen entre ellas diferencias reconocibles en la secuencia de bases. Dado que esto es técnicamente complejo (porque cada célula contiene sólo una copia de cada uno de los pares de cromosomas y en los pasos de extracción y secuenciación se pierde material) y porque no es relevante para comprender el funcionamiento del genoma, nos conformamos con lo que es, en realidad, un “promedio” del genoma personal o individual. Además, es altamente probable que todas esas diferencias acumuladas en el genoma de las células de un individuo sean irrelevantes: son cambios que ocurren con mayor probabilidad en regiones no-codificantes (entre genes) y por lo tanto, sin consecuencia en las funciones celulares. Ocasionalmente, para nuestro infortunio, puede ocurrir en una célula de nuestro cuerpo, una mutación que afecte un gen. De nuevo, esto probablemente pase desapercibido por nosotros por varias razones. 1. Porque generalmente portamos dos copias funcionales de cada gen, uno en cada cromosoma homólogo (el paterno y el materno) y con una copia normal o “buena”, la célula andará muy bien (lo mismo se puede decir de nosotros, como individuos, cuando heredamos una copia de cada gen de nuestros padres: con una buena, basta). 2. Porque si la mutación afecta un gen esencial (es la única copia “buena” que queda) y la célula muere, tampoco nos daremos cuenta. Millones de nuestras células mueren diariamente, como las de la piel o del intestino, y simplemente las reponemos. Aunque sea una célula no reemplazable, como una neurona, también las perdemos diariamente con las consecuencias perfectamente conocidas por los que tenemos cierta edad: una ligera pero

constante pérdida de capacidad cognitiva. A veces, sin embargo, ocurre una mutación que puede ser catastrófica: puede inactivar un gen de control del crecimiento celular o aumentar la actividad de la proteína codificada por el gen mutado transformándola en gatilladora de un cáncer.

Dicho lo anterior, podemos secuenciar nuestro genoma, podemos detectar diferencias entre nuestra secuencia y la “normal” que sean relevantes para nuestra salud y, secuenciando a nuestros familiares, reconocer aquellas mutaciones heredadas vs. las que aparecieron por primera vez en nosotros (mutaciones somáticas). Veremos luego cómo la ingeniería de genomas podría abordar la corrección de estos errores pero debemos volver sobre la relevancia -o no- de toda esa información sobre nuestros genes que tenemos ahora a nuestra disposición en el mítico pen drive. Majaderamente, hemos insistido en relativizar la secuencia de bases como un actor en la definición de la individualidad humana. El ADN es la molécula de la herencia pero no de la esencia ni de la trascendencia¹¹³. Podemos clasificar los genes en tres tipos: 1, los genes esenciales que definen directamente y de manera exclusiva un fenotipo. Ya hablamos de enfermedades que provienen de mutaciones en este tipo de genes y son los mejores candidatos a ser reparados con un sistema de edición como CRISPR. 2, Los genes que inciden directa y exclusivamente en un fenotipo pero que no son esenciales (si se alteran o faltan no hay un perjuicio notorio). Un clásico ejemplo, empleado para enseñar genética en los colegios, es el de la discriminación gustativa de la feniltiocarbamida (sabe amargo o a nada, dependiendo de la versión del receptor gustativo que tengamos). Es decir, son versiones de genes que determinan características que no son muy relevantes¹¹⁴. 3, los genes que contribuyen en una medida menor a un fenotipo (sea este esencial o no) lo que los hace, generalmente, dispensables pues su efecto esta tamponado por la contribución de muchos otros genes. La gran mayoría de las características se manifiestan de esta manera, incluidas las más complejas en humanos, aquellas relacionadas con inteligencia y personalidad. Repitiendo una vez más, hay una contribución medible, pero minoritaria, de genes a la inteligencia¹¹⁵, dado que otra, buena parte, proviene de condiciones ambientales y epigenéticas. Es decir, para modificar este tipo de características, habría que hacer ingeniería

¹¹³ Y eso que tengo en mi closet una camiseta que proclama “*DNA is life*”.

¹¹⁴ Al menos, no son materia de vida o muerte. Para alguna gente, el color de los ojos, de la piel o la estatura son relevantes y determinantes en su vida.

¹¹⁵ Término que apenas me atrevo a usar dado que ni siquiera somos capaces de definirlo apropiadamente.

sobre muchos genes a la vez, la mayoría de los cuales tiene roles en diversos otros procesos lo que hace impracticable su manipulación. O al menos poco recomendable.

He hecho la distinción anterior para separar aquellos genes sobre los cuales tendríamos alguna posibilidad de éxito (y justificación ética y clínica) para manipular o editar (clase 1 definida anteriormente) en contraste con aquellos cuya manipulación es cuestionable éticamente o que son, simplemente, inabordables con la tecnología actual (clases 2 y 3). La genética, en alguna medida, y la genómica de manera más significativa, son todavía ciencias en evolución y podemos incluso decir que están en su infancia. Salvo en contados casos, ignoramos demasiado como para aventurarnos a definir blancos para “corregir” o “mejorar” el genoma humano. Es probable que, con el avance del conocimiento, podamos afinar la puntería e incrementar las posibilidades de resolver problemas genéticos complejos aún en aquellos casos donde hay múltiples genes involucrados o sutilezas en sus funciones que son reflejo de influencias externas que podremos manejar. De hecho, es probable que muchas soluciones (terapéuticas y otras) pasen por hacer cambios reversibles, no en la secuencia del ADN, sino en los mecanismos regulatorios o epigenéticos. Encender y apagar genes, controlando las modificaciones en la cromatina donde se encuentran, puede ser incluso más poderoso que introducir mutaciones ya que podemos decidir cuando y dónde activar el cambio y, si no nos gusta, revertir ese cambio. Mejor aún, la descendencia no se vería afectada ya que no modificaríamos la secuencia de bases ni tocaríamos la línea germinal. Sumamos a ello, la reciente descripción del uso de CRISPR para editar ARNm¹¹⁶, lo que permite hacer correcciones transitorias, de nuevo, sin tocar los genes. Lo que falta para lograr esto, aparte de mejorar la especificidad, es encontrar una buena forma de dirigir la edición a las células o tejidos apropiados en una persona adulta. Pensamos que esto, eventualmente, se solucionará pues hay ideas bien encaminadas. La nanotecnología es quizás la más promisoría ya que se requiere llevar una pequeña carga molecular hacia las células blanco de una manera masiva y eficiente. Pequeños nano-robots¹¹⁷ podrían cumplir la tarea perfectamente y tengo la impresión que los colegas físicos avanzan aceleradamente en esa dirección.

¹¹⁶ *RNA editing with CRISPR-Cas13*. D. B. T. Cox et al., *Science* 10.1126/science.aag0180 (2017).

¹¹⁷ Se me ocurre, en inglés, *CRISPRites*.

Dado que pronosticamos la factibilidad de hacer modificaciones seguras en el ADN o en la capa epigenética de células somáticas de cualquier tejido de un individuo, podemos asumir que se podrá cruzar fácilmente la barrera de las modificaciones clínicamente necesarias (la persona esta definida como “enferma”) para llegar a introducir cambios estéticos o funcionales en caracteres que no inciden en la sobrevivencia o calidad de vida mínima. Hablamos de potencia física, capacidad intelectual, destrezas, aspecto exterior, incluso, personalidad. ¿Que tan abordables son esos caracteres? ¿Cuanto demoraría el cambio fenotípico en manifestarse una vez introducido el cambio genético? ¿Se puede reescribir la historia personal y hacer un cambio en alguien con “criterio formado”? Son preguntas que no podemos responder certeramente aún, pero pienso que, siendo algunos fenotipos más plásticos que otros, serán los primeros en ser sujetos de prueba. El atleta que ingiere hormonas para mejorar su potencia muscular, podría perfectamente reemplazar el doping por una modificación en la expresión de sus propias hormonas (testosterona, hormona del crecimiento, miostatina, etc.). Ese es un ejemplo de un “mejoramiento” claramente no esencial e imagino que el Comité Olímpico lo consideraría trampa, igual que el doping¹¹⁸. La obesidad, la calvicie, la intolerancia a ciertos alimentos, son también reparables por vía de edición genética ya que los genes involucrados son bien conocidos. Estos problemas son enfrentados seriamente por la medicina y merecen una solución si esta es segura. Y luego vienen los caracteres de mayor complejidad genética, donde sería necesario hacer muchos cambios para ver un efecto, siendo necesarias modificaciones tanto en secuencias como en el estado epigenético. Hablamos de personalidad o inteligencia, donde se podría lograr un efecto marginal en personas adultas debido a que la red neuronal construida en el desarrollo de esa persona ya esta establecida. Pero sí, a pesar de las dificultades técnicas aún por superar, es imaginable una intervención de este tipo. Veremos a continuación lo relacionado con enfermedades y trastornos que pueden y deben ser abordados por la edición de genomas y discutiremos si, para algunas de ellas, vale la pena considerar modificaciones en la línea germinal (afectando las futuras generaciones, no sólo al “paciente”). Pero entendamos que habrá una inmensa presión sobre los científicos y tecnólogos para las aplicaciones menos cuestionables éticamente y sobre las cuales probablemente haya también menos restricciones regulatorias, al ser materia de opción personal. Le doy unos 20 años al desarrollo de acceso masivo a estas

¹¹⁸ Habrá oportunidades laborales infinitas para los científicos genómicos ya que se generará toda una industria para detectar cambios introducidos en el genoma.

opciones por lo que lamento informarme, y a todos los de edad similar a la mía o mayor, que probablemente no llegemos a usarlas. Pero para el resto de Uds., es hora que vayan pensando en las opciones que se abren y si vale la pena asumir el riesgo que implica.

Alquimia genómica: ¿un quiebre en nuestra evolución?

Hace 5 millones de años, en un primate, ocurrió un pequeño cambio genómico que permite que Ud. este leyendo este libro. En el linaje que derivó hacia los Neandertales, los Denisovanos y los humanos actuales, un gen llamado *ARHGAP11* se duplicó. Del gen original, que detectamos todavía inalterado en los chimpancés, se formaron dos copias, *ARHGAP11A* y *ARHGAP11B*, en esta pequeña rama de los mamíferos, que son una pequeña rama de los animales, a su vez una pequeña rama de los eucariontes, etc. Este insignificante evento en el cosmos de la vida tuvo su consecuencia: la copia B del gen, al quedar liberada para tomar nuevas funciones (la copia A continuó haciendo lo que siempre hizo), experimentó un leve cambio de secuencia lo que generó la novedosa propiedad de promover la proliferación de progenitores basales del cerebro¹¹⁹. Esto, a su vez, permite la expansión de la neocorteza y los pliegues cerebrales. En simple, el gen *ARHGAP11B* se hizo cargo de agrandarnos el cerebro y de poner a nuestra disposición como especie, la posibilidad de incrementar nuestras capacidades cognitivas. En las ciencias genómicas, nos encontramos a menudo con estos eventos de duplicación, o creación de genes, que nacen en la evolución para luego perderse rápidamente al estar repitiendo innecesariamente la función del gen original. A veces, incluso, vemos los restos de esos genes duplicados (los llamamos pseudogenes) que son testimonio de posibilidades evolutivas desaprovechadas. El biólogo y ensayista Stephen Jay Gould, al observar que los experimentos evolutivos que fracasan son mucho más frecuentes que los exitosos, concluyó que la vida que conocemos es producto del azar y que, si echáramos a correr nuevamente el proceso desde el principio, de ninguna manera repetiríamos el resultado. Este “relojero ciego” es capaz de producir artefactos aparentemente perfectos, que dan la impresión de haber sido creados a propósito para su función. Como dijimos muy al principio, esta engañosa apariencia de diseño inteligente se debe únicamente a tres elementos: variación genética, selección natural y

¹¹⁹ *A single splice site mutation in human-specific ARHGAP11B causes basal progenitor amplification. Science Advances* 2016. Vol. 2, no. 12; <http://advances.sciencemag.org/content/2/12/e1601941.full>

muuuucho tiempo. Volviendo al gen *ARHGAP11B*, los portadores este sesudo regalo evolutivo lo aprovecharon bien: bajaron de los árboles, dominaron el planeta, acomodaron la evolución de las demás especies a su antojo y llegaron a desarrollar un método para dirigir la propia evolución a voluntad. La descubridora del CRISPR, Jenniffer Doudna, citada anteriormente, no en vano ha llamado a esta tecnología “una grieta en la creación”. ¿Seremos realmente capaces de saltarnos el proceso Darwiniano y decidir racionalmente que secuencia de ADN queremos tener los humanos de aquí para adelante? Antes de intentar responder esta pregunta, veamos que lecciones nos da la evolución que, mal que mal, nos trajo hasta aquí. Primero, nuestra creatividad en el diseño de genomas tiene un límite, el cual esta impuesto por lo que ya creó la naturaleza. A un simio-científico de seis millones de años atrás, dispuesto a usar CRISPR, no se le habría ocurrido duplicar el gen *ARHGAP11* para hacerse a él, o a sus parientes, más inteligentes. El gen original tiene una función totalmente ajena al desarrollo cerebral por lo que no había manera de predecir esa nueva función. Tampoco podemos inventar genes que hagan cosas que la naturaleza no ha creado, porque no tendríamos idea de donde partir. Sería como proponernos diseñar un gen que nos permitiera teletransportarnos de un lugar a otro, a lo *Star Trek*.

Dado que el relojero ciego no trabaja con un plan sino con el azar, veamos cuán creativa se ha puesto la naturaleza con nuestro genoma con algunos ejemplos¹²⁰. Apareció en EEUU hace algunas décadas una Señora -la llamaremos Mrs. Gena Blation- con una peligrosa enfermedad autoinmune (se llama síndrome *WHIM*) que no tiene tratamiento. Se trata de una enfermedad genética en que aparece una mutación dominante en un gen expresado en glóbulos blancos, haciéndolos estar en un permanente estado inflamatorio y de “ataque”. Gena sobrevivió con muchos padecimientos por un tiempo y luego se le perdió la pista, para aparecer mucho después en una clínica con dos hijas que presentaban la misma terrible enfermedad (al ser un alelo dominante, es más probable la transmisión de la enfermedad y basta que uno de los padres lo posea). Lo novedoso del cuento es que, en ese momento, Gena ya no presentaba ningún síntoma. La curiosidad de los médicos y genetistas ante el “milagro” los hizo indagar en el fenómeno y se encontraron con la sorpresa que, en su genoma, se había producido una mutación espontánea que eliminó el gen de la mutación dominante, dejando sólo el gen silvestre. Para aún mayor fortuna, la mutación espontánea había ocurrido en una célula de la médula ósea, lugar desde el

¹²⁰ Un compendio de ellos se describe en: Hirschhorn R. *In vivo reversion to normal of inherited mutations in humans. Journal of Medical Genetics*. 2003;40(10):721-728. doi:10.1136/jmg.40.10.721.

cuál se originan y reponen las células sanguíneas durante toda nuestra vida. De hecho, las demás células de Mrs. Blation seguían teniendo la mutación dominante, lo que no la afectaba porque el gen no tiene función en ellas. Es decir, ocurrió un cambio genético preciso en la célula precisa para “curar” una enfermedad genética potencialmente letal. Imaginen ahora que, enfrentados a un paciente con el síndrome *WHIM* que presenta esa misma mutación dominante, Uds le explican que se ha desarrollado una herramienta para entrar al genoma de una célula madre de su médula ósea y eliminar el gen alterado para revertir la enfermedad, tal como ocurrió con Mrs. Blation. “¿¡Y QUE ESTAN ESPERANDO!????” no tardaría en escucharse. En efecto, debemos asumir la perspectiva del que padece la crueldad del azar genético y acelerar tanto el desarrollo de los métodos (responsabilidad de los científicos) como de un marco legal y de salud pública adecuado (materia del poder político). Ojalá acompañado de políticas que pongan estos procedimientos al alcance de la mayor cantidad de potenciales pacientes posible.

Un segundo ejemplo que viene al caso, que ilustra cómo la edición genómica llevada a cabo por la naturaleza, jugando azarosamente con nuestros nucleótidos, se da en el caso de síndrome de inmunodeficiencia adquirida (SIDA) humana causada por el virus del VIH. Aparecieron personas infectadas con el virus que misteriosamente sobrevivían y no desarrollaban el SIDA. Se descubrió que estas personas presentaban una mutación natural de un gen que codifica para una proteína expresada en linfocitos (células inmunes) y que se ubica en la superficie celular. Es esta proteína, llamada CCR5, la que usa el virus del VIH para ingresar a las células inmunes y, eventualmente, matarlas, lo que causa la inmunodeficiencia. La versión mutante de CCR5 no se ensambla correctamente y no llega a la superficie celular en los linfocitos por lo que sus portadores (alrededor del 1% de la población caucásica es mutante para ambos alelos del gen) son inmunes al VIH. Se nos presenta entonces una perfecta oportunidad para generar una terapia génica para frenar el SIDA: es cosa de hacer edición genómica para alterar -o abolir- el gen *CCR5*¹²¹. De hecho, hay una empresa que ya ha iniciado pruebas clínicas con esa precisa estrategia y sus resultados preliminares son alentadores¹²². Uno se pregunta porque nuestro acervo genético incluye un gen cuya única función aparente es hacernos susceptibles al temido

¹²¹ Bastaría hacer edición en células progenitoras de la médula ósea, las cuales dan origen a los linfocitos, las células infectadas por el VIH.

¹²² *Gene Editing of CCR5 in Autologous CD4 T Cells of Persons Infected with HIV*. Tebas et al., 2014. New England Journal of Medicine. 370: 901–10.

VIH; las personas mutantes para *CCR5* no parecen tener ningún defecto. La clave se desprende de la frecuencia de ese alelo mutante en la población: no existe en absoluto entre los africanos y aparece primordialmente entre los habitantes del noreste Europeo. Aparentemente, el gen *CCR5* -en su versión silvestre o funcional- es esencial para la defensa inmune contra las infecciones de otros virus, entre ellos el Virus del Nilo Occidental (VNO), enfermedad mortífera transmitida por el mosquito del género *Culex*. Dada la prevalencia en África del VNO, la selección natural tiende a seleccionar fuertemente a favor de la mantención del gen *CCR5*, presión inexistente en las latitudes mayores (como Europa) donde no está presente el mosquito portador¹²³.

De tal modo, el estudio de la evolución humana y el análisis de la diversidad genética de la especie nos han servido en bandeja un par de buenos puntos de partida para hacer modificaciones en nuestro ADN. En otras palabras, la naturaleza nos regala información útil para combatir enfermedades devastadoras. Pero es importante notar que las mutaciones espontáneas descritas no se han propagado por selección dado que -como ilustramos recién- predisponen desfavorablemente a otras enfermedades; esencialmente una situación pierde-pierde (o empatata-empata). Lo que tenemos que lograr es hacer un poco de ingeniería de proteínas (que, finalmente, será ingeniería de genomas) para diseñar moléculas que incorporen los cambios deseados sin un efecto deletéreo, algo que de seguro podemos lograr con suficiente experimentación. Además, no tenemos porque ir cambiando residuos gen a gen; podemos juntar versiones de genes que funcionan bien juntos o hasta pensar en el diseño del genoma completo para abarcar el máximo posible de arreglos de una vez. Esta locura está a nuestro alcance hoy con microorganismos, donde los genes no son tan numerosos y se conocen mejor sus funciones exactas. Craig Venter (sí, el mismo que apuró las tecnologías de secuenciación de genomas y que se autosecuenció), ha estado empeñado en crear vida desde cero, armando un genoma bacteriano en el computador, primero, y luego sintetizando el ADN respectivo. Ha creado un microorganismo artificial, que contiene una combinación de genes inexistente en la naturaleza y que debería comportarse de una manera previsible de acuerdo al diseño¹²⁴. Este es un ejemplo

¹²³ *CCR5 Deficiency Increases Risk of Symptomatic West Nile Virus Infection*. Glass et al., 2006. *Journal of Experimental Medicine* 203: 35–40.

¹²⁴ *Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome*. Gibson et al., 2010. *Science* 329: 52-56. DOI: 10.1126/science.1190719. En este organismo sintético, Venter introdujo en el genoma una serie de “marcas de agua” o filigranas, entre ellas, un código para traducir entre nucleótidos y el alfabeto, los nombres de los investigadores involucrados, algunas citas y un correo electrónico secreto al cuál hay que escribir si uno lo descubre.

extremo de la Biología Sintética, con la que construimos módulos vivos a partir del conocimiento de las partes que lo componen, principalmente genes. Evidentemente, este camino nos llevará a crear otras formas de vida artificial, que eventualmente incluirá organismos más complejos. Venter está particularmente interesado en crear organismos que produzcan biocombustibles, lo que podría independizarnos del uso de combustibles fósiles. Si Ud. está preocupado por los transgénicos y mutantes, olvídelo. Pronto, esas tecnologías parecerán tan inocentes como cultivar flores en el jardín ya que tendremos que acordar un sistema de supervisión y seguridad biológica de una sofisticación insospechada para nuestras actuales normativas. Probablemente, los policías genómicos del futuro tengan una pistolita para leer (secuenciar) el código de barras presente en el genoma de estas especies artificiales. Muchas de ellas trabajarán para nosotros en procesos industriales o en medicina. En teoría, podríamos replicar algunas funciones “superiores” o crear algoritmos moleculares que tomen decisiones o que optimicen procesos, de cierta manera, inteligentes.

Ya que me lo pregunta, en mi modesta opinión, la creación o diseño de vida humana mediante herramientas biotecnológicas no será ni práctico, ni barato, ni necesario. Añadido, por otro lado, que sí será bueno recrear partes de humanos con el objeto de reparar tejidos y órganos dañados. Poder fabricarme un segundo corazón para cuando el actual falle, probablemente me permitiría llegar a los 150 años. Hoy, podemos reprogramar células para que se diferencien en la dirección que queremos y, dentro de poco, conduciremos a esas células a fabricar órganos *in vitro* (en cultivo). Dado que podemos partir con células propias, se desvanece el peligro del rechazo y se abre la posibilidad de introducir cambios en el genoma de ese órgano para eliminar mutaciones o corregir versiones de genes que prolonguen aún más la vida de esas células u órgano. Habrán órganos más difíciles de reparar o que no son transplantables como el cerebro (creo), y en ese caso, las terapias basadas en células madre que puedan colonizar los tejidos afectados serán el camino a seguir. Nuevamente, mientras estén fuera del cuerpo, las células pueden ser modificadas en su genoma para reparar mutaciones o dirigir el desarrollo de esas células a destinos específicos.

Lo anterior tiene justificación en el marco de la salud y bienestar de las personas nacidas con el complemento genético “natural”, aquel heredado de sus padres. Pero hemos hablado de la posibilidad que CRISPR u otras tecnologías de edición genómica puedan aplicarse en la modificación del genoma transmitido generacionalmente, en otras palabras, a nuestros

descendientes. ¿Porqué no evitarle a nuestros hijos enfermedades o susceptibilidades que afecten su salud si tenemos las herramientas para hacerlo? Racionalmente, suena sensato y altruista. En el curso de la evolución y la historia humana, hemos adquirido una creciente potestad para decidir sobre la existencia de nuestros hijos (podemos optar conscientemente por concebir o no). Incluso, hemos llegado al punto de implementar tecnologías para seleccionar (a favor o en contra) cuales de nuestras procreaciones llevaremos a término¹²⁵. Pero nunca antes hemos tenido la oportunidad, luego de decidir sobre su existencia, de tener un rol en el diseño de la vida que estamos engendrando, en aspectos relevantes para la sobrevivencia y calidad de vida futuras y, eventualmente, en aspectos no esenciales (características físicas y mentales). Durante décadas, hemos tenido una discusión, con posiciones irreconciliables, sobre los derechos del humano por nacer, definiéndose el humano en un rango que va desde el momento de la procreación (“que vengan los hijos que Dios nos mande”) hasta un feto relativamente avanzado. Pero en lo que todos estamos de acuerdo, es que el niño nacido adquiere derechos humanos fundamentales, que le son inalienables e irrenunciables, opinen lo que opinen los padres. De tal modo, si los padres u otros, toman una decisión sobre el genoma de esa persona, estan potencialmente vulnerando un derecho humano fundamental. Podemos concebir ese derecho como aquel que protege al individuo de la manipulación de su genoma por terceros (no hubo una decisión propia) sea esta previo o durante su existencia. Hay un terreno gris aquí dado que ya tomamos muchas decisiones por nuestros hijos sin su consentimiento (vacunas, alimentación, educación, etc.). Sin embargo, la intervención genética afecta al hijo y a todos sus descendientes, tal vez irreversiblemente, y además alteramos el genoma de la población y especie. Por primera vez habría un cambio genético en nuestra especie cuyo origen es totalmente artificial y que ha sido conscientemente introducido, prescindiendo de la selección natural o deriva génica a la que estamos sometidos como entes biológicos. Dado que no podemos hacer experimentos con humanos para probar la eficacia y seguridad de las hipotéticas mejoras, ni podemos probar la mayoría de esos cambios en modelos animales, estaríamos además operando a ciegas respecto al resultado.

...

Durante cientos de millones de años, los genomas han evolucionado para sostener la vida sobre la tierra y lo han hecho mediante una mezcla precisa de estabilidad y cambio, lo suficiente

¹²⁵ Los tests para detectar trisomías o para elegir el sexo de los embriones a implantar son rutinarios en las clínicas de fertilidad asistida.

de cada una para preservar lo que funciona y eliminar lo inadecuado. Demasiada estabilidad sería fatal: no existiría la posibilidad de innovación y la vida no habría prosperado como lo ha hecho. El mismo proceso evolutivo exige la aparición de una dosis de desorden, que llamamos variación, para tener un repertorio de soluciones eventuales para los desafíos que impone la sobrevivencia en un ambiente cambiante. Darwin reconoció en el proceso una crueldad e indiferencia chocantes; la naturaleza, según él, estaba teñida de rojo, marcada por garra y diente, dominada por la competencia feroz por sobrevivir, en que la mayoría de los ensayos evolutivos perjudican en vez de dotar. La combinatoria de eventos que lleva a un aumento de complejidad o a mejor adaptación son altamente improbables y, como explicamos antes, difíciles (o imposibles) de diseñar racionalmente. ¿Que confianza podemos tener, entonces, de obtener el efecto deseado al introducir un cambio en el genoma humano? La genética nos ayuda en una fracción de los casos, mostrándonos numerosos ejemplos de cambios perjudiciales, donde el simplemente revertirlos debería restaurar la “normalidad”. Una persona que padece de una enfermedad genética manifiesta un dolor humano que ahora tenemos la capacidad de aminorar o eliminar. Con seguridad, embarcarnos en este camino implica riesgos y errores, pero contaremos con la decisión del paciente cuya alternativa es inaceptable. Distinto es hacer un cambio permanente y heredable. Los biólogos sabemos que en la naturaleza nada es “para siempre”; un cambio en apariencia bien diseñado puede, más tarde, aparecer vergonzosamente arrogante, contraproducente y hasta negligente. No estamos en condiciones de meternos en el genoma humano futuro y, por ello, apoyo la moratoria actualmente en vigencia que lo impide.

Dicho lo anterior, reduciremos fuertemente la incertidumbre de las tecnologías de edición y manipulación de los genomas si profundizamos nuestro conocimiento de los mecanismos que operan en él. Hay dos fronteras que debemos cruzar: 1, el entendimiento de las interacciones entre genes y sus contribuciones relativas al fenotipo y 2, el funcionamiento de los genes en el contexto estructural del genoma, lo que incluye todas las capas epigenéticas que modulan la expresión de los genes. La genética molecular no terminó con la secuenciación de los genomas, recién empieza. Desafío e invito a los aspirantes a científicos a trabajar para desenmascarar la lógica de los genomas. Y a los demás, a seguir de cerca este apasionante desarrollo tecnológico.

Agradecimientos

Me sostengo precariamente sobre los hombros de varios gigantes que me anteceden en el campo disciplinar, en mi institución y en mi árbol genealógico (contribuciones, por cierto, epigenéticas y genómicas). Es primera vez que me enfrento al honor y desafío que significa escribir sobre ciencia para el lector no especializado. Traté de hacerlo pensando en las obras de los que considero máximos exponentes de este arte: Sagan, Gould y Asimov, entre otros. Aún cuando el estilo, aproximación al tema y profundidad son una elección personal, me sirvió mucho ver cómo los colegas de profesión se enfrentan a escribir para el respetable público sobre temas complejos dejando de lado la erudición y estrechez del mundillo académico. Descender los peldaños del marfileso edificio universitario requiere, disciplina, abstracción, paz y revisar una y otra vez el lenguaje para evitar el mareo que suele inducir la especialización o, en mi caso, el excesivo uso de la jerga técnica. Logré apartarme por un año de las tareas domésticas del catedrático haciendo un periodo sabático en Davis, pequeña villa californiana, donde logré ensamblar relativamente impertérrito esto que Ud. lee. A ese lugar fui a ponerme al día en aspectos descuidados de mi área de trabajo y a absorber temas que nunca tuve la oportunidad de aprender a fondo, entre ellas la genómica. Sobre todo, poder mirar las preguntas científicas de nuevo, quizás buscando una mirada fresca para acercarse a ellas. El beneficio de salida temporal me la dió mi institución (y el Decano de mi Facultad, Prof. Víctor Cifuentes), el proyecto que dirijo (el Centro de Regulación del Genoma, CRG, financiado por el programa FONDAP de CONICYT) y mi anfitrión, el Prof. Li-En Jao, de la Universidad de California. Largas y formativas discusiones sostengo semanalmente con mis colegas investigadores del CRG, Martín Montecino, Alejandro Maass, Mauricio González, Ariel Orellana y Rodrigo Gutiérrez. Agradezco el riguroso trabajo de corrección de Catherine Connelly, quien, en el ámbito de una lengua no materna, capturó numerosos errores en mi escrito preliminar; aquellos que queden son de mi responsabilidad. Aprecio el apoyo imprescindible para el espíritu explorador de nuestro país que entrega CONICYT con sus diversos programas; cuidémoslo, aunque cambie de nombre o forma. Agradezco a mis actuales proyectos: FONDECYT (1180606) y FONDAP (15090007).

Obviamente, la familia debe redoblar su tolerancia cuando se le ocurre al Señor Científico ponerse a escribir un libro, como si fueran pocas las horas que ya dedica a la tarea de mantenerse

activo y vigente. Aprovechamos el slogan deportivo, *mis dos amores*, para balancear los énfasis y tiempos.

¿Porque escribir un libro de divulgación? Yo tengo un contrato con la Universidad de Chile, y por consecuencia, con el Estado y pueblo de Chile, en que se me paga un salario y se me otorgan fondos de investigación para hacer avanzar el conocimiento y enseñar. Lo hago. Pero pienso que este camino exitante y maravilloso se vuelve, a veces, solitario y autorreferente. Más que justificar el financiamiento que he recibido y “devolver” algo a la nación, me gustaría más que mis compatriotas se entusiasmen y vibren como yo con los descubrimientos científicos. No les puedo pedir que lean mis artículos técnicos o que vengan a mis clases. Además, lo que yo hago es una gota en el oceano de los avances científicos que diariamente ocurren. En cambio, sí puedo digerir los temas complejos y vastos que he tenido la oportunidad de aprender y regurgitarlos con una perspectiva más amigable que aquel usado entre colegas. Si en alguno de mis lectores se despierta la sensación de asombro o se gatilla el cosquilleo de la curiosidad, me siento satisfecho. Mejor aún, si en algún joven logro sembrar el germen de la vocación por la indagación y una vida constructiva, de servicio al conocimiento universal, he logrado lo máximo a que puede aspirar un investigador: la continuidad de la búsqueda. Finalmente, me basta que los temas aquí tratados se discutan y se piensen. Divulgar también es provocar y estimular.

Sobre el Autor

Miguel L. Allende Connelly (1962, Nueva York). Licenciado en Ciencias Biológicas en la Pontificia Universidad Católica de Chile (1987) y Doctor en Biología Molecular de la University of Pennsylvania (1994). Realizó estudios postdoctorales en genética molecular en el Massachusetts Institute of Technology (1994-1997). Profesor Titular de la Facultad de Ciencias de la Universidad de Chile desde 2000 y Director del Centro de Regulación del Genoma (CRG) desde 2010. En 2016 fue elegido Miembro Correspondiente de la Academia Chilena de Ciencias.

Entre sus contribuciones científicas se destacan sus 75 artículos en revistas de corriente principal y la formación de más de 50 científicos jóvenes, entre estudiantes de pre- y postgrado y postdoctorados. Su investigación se ha centrado en el estudio de los mecanismos moleculares asociados al desarrollo embrionario, a la regeneración de tejidos y a la inmunidad, usando para todos sus proyectos el modelo del pez cebra (*Danio rerio*). Hoy, ha ampliado sus intereses científicos a la genómica de organismos sudamericanos, en especial los peces adaptados a ambientes extremos. Ha participado en numerosas actividades de divulgación, aprovechando el uso de los embriones de pez como herramienta de enseñanza para profesores secundarios y escolares.

Recientemente, impulsó la iniciativa 1000 Genomas Chile, cuyo objetivo es obtener datos de secuencia genómica de 1000 chilenos y 1000 especies habitantes del país. El proyecto es liderado por cinco centros de excelencia, incluido el CRG.

Portales de interés:

Laboratorio Allende: www.zebrafish.cl

Centro de Regulación del Genoma: www.genomacrg.cl

Proyecto 1000 Genomas: www.1000genomas.cl