

# Implicaciones funcionales de los polimorfismos de un solo nucleótido (SNP) en genes codificantes de proteínas y no codificantes en enfermedades multifactoriales

Julián Ramírez-Bello\* y Mayra Jiménez-Morales

Unidad de Investigación en Enfermedades y Metabólicas, Hospital Juárez de México, Ciudad de México, México

## Resumen

Los SNP representan las variantes genéticas más comunes del genoma humano. Estas alteraciones de un solo nucleótido, localizadas tanto en genes codificantes como en no codificantes de proteínas, se dividen en neutras y funcionales. Las primeras no tienen impacto biológico, mientras, las segundas afectan diversos procesos y constantemente se asocian con riesgo en enfermedades multifactoriales. Los SNP funcionales de los promotores de ambos tipos de genes, denominados SNP reguladores (rSNP) y rSNP de microRNA (miR-rSNP), respectivamente, afectan la expresión génica. Los SNP funcionales ubicados en la estructura del RNA heterogéneo nuclear (exones e intrones), mRNA maduros (5' UTR, secuencia codificante y 3' UTR), y miRNA (pri, pre y miRNA maduros), se denominan SNP RNA estructurales (srSNP) y miR-srSNPs, respectivamente. Los primeros afectan el corte y empalme; los segundos, la traducción, la estabilidad, la secuencia de aminoácidos, la estructura y la función de las proteínas, y la interacción de mRNA/miRNA; y los terceros alteran la estructura, el procesamiento, la interacción miRNA/mRNA y su función. La evaluación del efecto biológico-funcional alélico de los SNP ha contribuido enormemente a un mejor entendimiento de las enfermedades multifactoriales. El objetivo de esta revisión es actualizar al lector acerca del efecto funcional de estas variantes localizadas en algunos genes codificantes y no codificantes de proteínas, y su relación con las enfermedades multifactoriales.

**PALABRAS CLAVE:** Polimorfismo de un solo nucleótido. RNA codificante. RNA no codificante.

## Abstract

Single nucleotide polymorphisms (SNPs) represent the most common type of variation in the human genome. The SNPs located in protein-coding and non-coding RNA genes are classified as neutral and functional. The neutral have no effect, while the functional affect different biological processes and continually confer risk for multifactorial diseases. Functional SNPs found in the promoters of protein-coding and non-coding RNA genes (microRNAs: miRNAs) termed regulatory SNP (rSNPs) and miRNAs rSNPs (miR-rSNPs), respectively, affect the gene expression. Functional SNPs located on the structure of the precursor mRNAs (exons and introns), mature mRNA (5' untranslated region [UTR], coding sequence, and 3' UTR), and primary, precursor, and mature miRNAs are termed structural RNA SNPs (srSNPs) and miR-srSNPs, respectively. The srSNPs affect the splicing (and alternative splicing), srSNPs affect the splicing (and alternative splicing), the translation, stability, amino acid sequence, structure, and function of proteins and interaction between mRNA/miRNAs. Finally, the miR-srSNPs affect the struc-

### Correspondencia:

\*Julián Ramírez-Bello  
Unidad de Investigación en Enfermedades Metabólicas y  
Endócrinas del Hospital Juárez de México  
Av. Politécnico Nacional, 5160  
Del. Gustavo A. Madero,  
C.P. 07760, Ciudad de México, México  
E-mail: dr.julian.ramirez.hjm@gmail.com

Fecha de recepción en versión modificada: 30-03-2016

Fecha de aceptación: 30-03-2016

ture, processing and interaction between miRNAs/mRNAs. Functional characterization of potentially harmful risk alleles of the SNPs located in protein-coding and non-coding RNA genes have contributed to an understanding of their functions in the complex diseases. The objective of this review is update the reader on the functional role of the SNPs located in protein-coding and non-coding RNA genes and their relationship with multifactorial diseases. (Gac Med Mex. 2017;153:238-50)

**Corresponding author:** Julián Ramírez Bello, dr.julian.ramirez.hjm@gmail.com

**KEY WORDS:** Single nucleotide polymorphism. Protein-coding gene. Non-coding RNA.

## Introducción

A principios del año 2000 se publicaron sendos artículos en dos de las revistas con mayor factor de impacto, *Nature* y *Science*, y ambos mostraron el número aproximado de nucleótidos del genoma humano y un elevado número de variantes genéticas comunes (más de un millón); estos últimos fueron llamados SNP (*Single Nucleotide Polymorphism*) o variantes de un solo nucleótido (SNV, *Single Nucleotide Variant*)<sup>1,2</sup>. Las características más importantes de los SNP se describen en seguida: a) se localizan en todo el genoma humano, en regiones intragénicas y extragénicas; b) representan las variantes genéticas más comunes; c) generalmente son bialélicas; d) son fácilmente evaluadas por métodos automatizados; y e) una gran parte de ellas tienen repercusiones directas en enfermedades humanas (Tabla 1). Desde el punto de vista funcional, es de vital importancia identificar el papel biológico de los nucleótidos de los SNP que tienen una frecuencia alélica menor (MAF, *Minor Allele Frequency*), localizados en diferentes genes<sup>3</sup>. Mediante estudios de asociación de gen candidato o del genoma completo (GWAS, *Genome-Wide Association Study*) han sido identificados genes involucrados en diversas enfermedades complejas (Fig. 1)<sup>4-6</sup>. Por su parte, los estudios de biología-genética molecular y bioquímicos han contribuido en la identificación del papel funcional *in vitro* o *in vivo* de los alelos

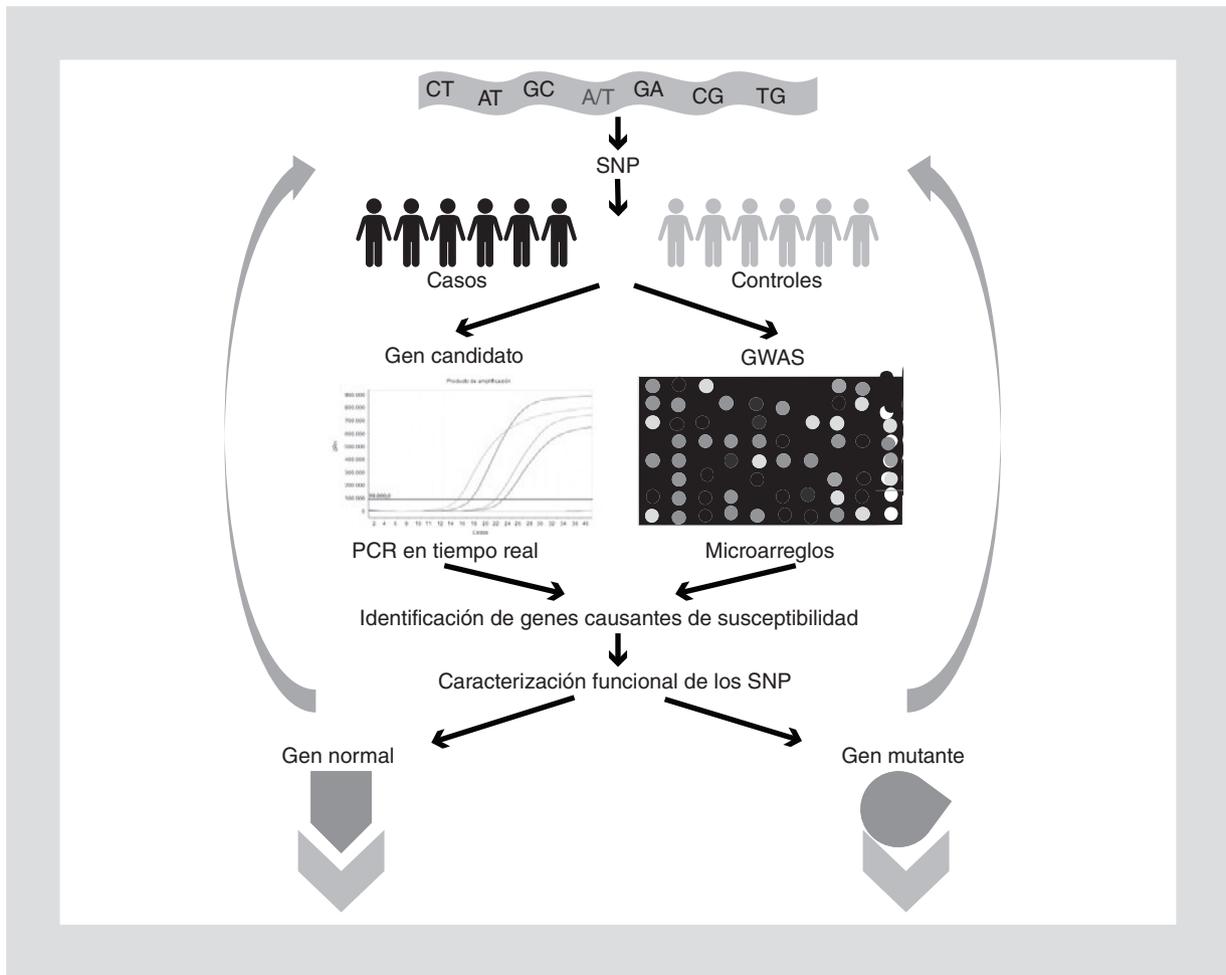
menos frecuentes de los SNP localizados en la estructura de genes codificantes de proteínas y no codificantes (Fig. 1)<sup>3,7-10</sup>. Dichos estudios han definido el papel funcional de los alelos de menor frecuencia en la expresión génica, el corte y empalme (y el alternativo), la traducción, la alteración de la estructura y de la función de las proteínas y de los microRNA (miRNA), la estabilidad de los RNA mensajeros (mRNA) y la interacción de miRNA y mRNA. Estos estudios, además, han contribuido a entender cómo los alelos menos comunes de los SNP confieren riesgo para desarrollar enfermedades multifactoriales humanas, y cómo afectan la respuesta a ciertos medicamentos<sup>3-10</sup>.

## Clasificación funcional de los SNP

Los SNP funcionales se clasifican de acuerdo a la región donde se ubican y al efecto que ejercen sobre ella. Los SNP funcionales de los promotores de genes codificantes de proteínas y no codificantes se denominan SNP reguladores (rSNP) y SNP reguladores de los microRNA (miR-rSNP), respectivamente (Tabla 2); ambas variantes afectan la expresión génica (Fig. 2)<sup>3,11,12</sup>. Por otro lado, los SNP funcionales localizados en la estructura de los mRNA precursores (pre-mRNA) y los mRNA maduros se denominan SNP RNA estructurales (srSNP), mientras los srSNP de los microRNA se denominan miR-srSNP. Los srSNP afectan la traducción, el corte y empalme (y el alternativo), la estructura y la estabilidad

**Tabla 1. Características de los polimorfismos de un solo nucleótido**

Características	Descripción
Distribución	Un SNP se encuentra en promedio cada 250 pb; se han reportado cerca de 11 millones
Ubicación	Regiones intergénicas (o extragénicas) e intragénicas; en esta región se encuentran genes codificantes de proteínas y no codificantes.
Número de alelos	Generalmente son bialélicas, aunque pueden ser trialélicas y tetraalélicas
Efecto biológico	Neutras y funcionales
Evaluación de los SNP	Fácilmente genotipificados mediante tecnologías automatizadas
Usos en salud	Identificación de individuos genéticamente susceptibles para desarrollar enfermedades multifactoriales, gravedad, actividad y respuesta a medicamentos



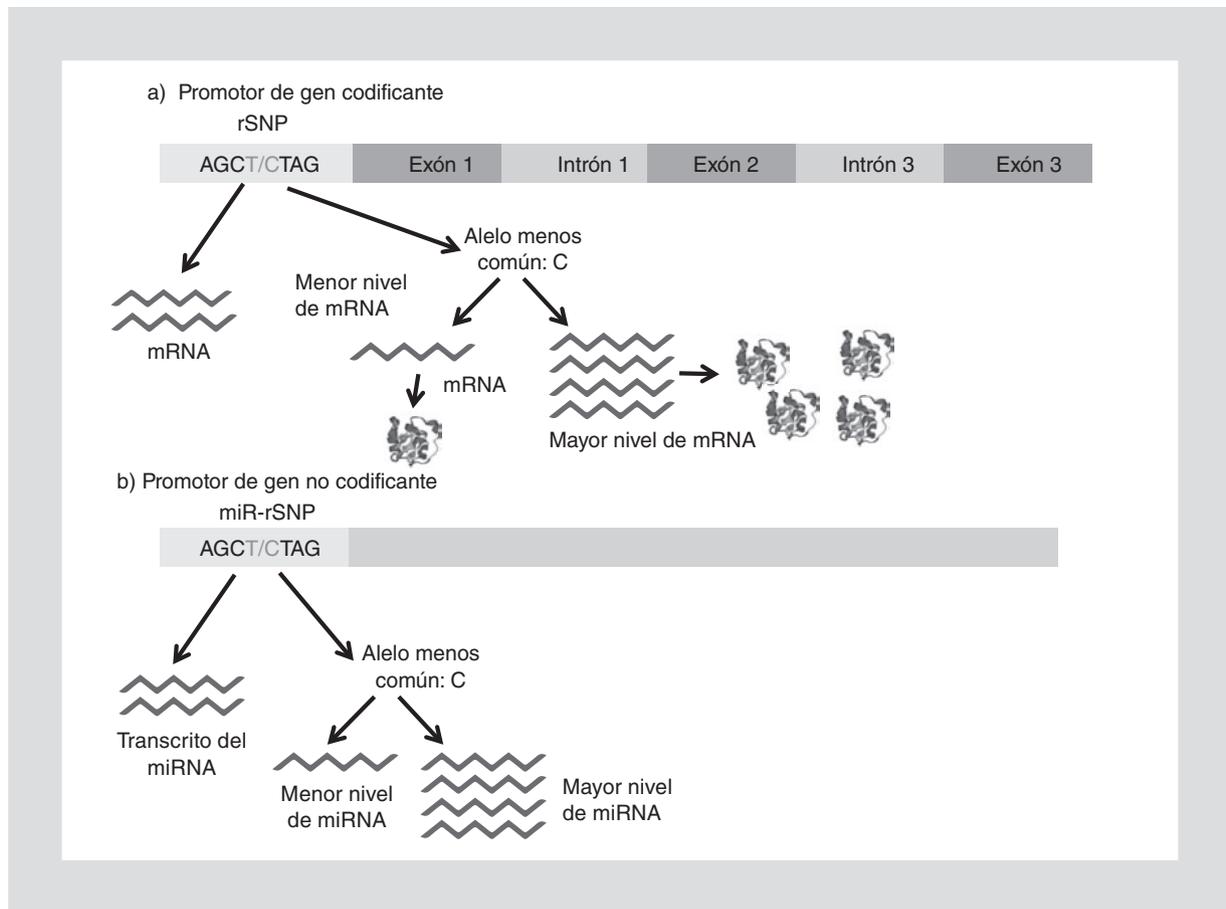
**Figura 1.** Caracterización funcional de los SNP. Los estudios de gen candidato o de asociación del genoma completo (GWAS) han contribuido enormemente en la identificación de loci involucrados con riesgo para desarrollar diversas enfermedades humanas multifactoriales. Los estudios de genética molecular y bioquímicos, a su vez, han contribuido en la identificación del efecto funcional de los alelos menos comunes de los SNP. Estos estudios pueden ir de gen candidato a identificación funcional o viceversa.

de los mRNA, la maduración, la función de las proteínas, y la unión entre miRNA y mRNA (Fig. 3 y Tabla 2)<sup>3,11</sup>. Mientras, los miR-srSNP afectan el corte y empalme, el procesamiento de transcritos grandes (miRNA primarios: pri-miRNA) y precursores (pre-miRNA), la unión de miRNA y mRNA, y su función (Fig. 4). Los SNP funcionales de la secuencia codificante se denominan sinónimos (sSNP) y no sinónimos (nsSNP) (Fig. 3). Los sSNP involucran un cambio de nucleótido y de codón (código genético degenerado: más de un codón da origen al mismo aminoácido), pero no de aminoácido, y aunque no cambian la secuencia de aminoácidos en las proteínas, pueden afectar ciertos rasgos (Tabla 2 y Fig. 3). A su vez, los nsSNP se subdividen en nsSNP sin sentido (*nonsense*) y de sentido erróneo (*missense*); los primeros generan un codón de paro y terminación prematura de la proteína, y los segundos generan un cambio de un aminoácido. Ambos pueden tener un efecto drástico,

pero con los últimos puede no ser grave si los aminoácidos reemplazados tienen similitud en la estructura química y las propiedades bioquímicas. Los dos tipos de variantes afectan la secuencia, la estructura y la función proteica (Tabla 2 y Fig. 3)<sup>3,13-16</sup>.

### Estructura y función de los promotores de genes codificantes y no codificantes de proteínas

El promotor de genes codificantes de proteínas y no codificantes regula de manera coordinada la expresión génica. En esta región se encuentran diversas secuencias que actúan en *cis*. En los promotores de genes codificantes, las secuencias *cis* se encuentran en el promotor basal o núcleo del promotor y en la región próxima al promotor basal<sup>9</sup>. Por otro lado, las secuencias consenso, estructura y organización de los promotores de genes



**Figura 2.** Efecto funcional de los rSNP y miR-rSNP. Los alelos menos comunes de los rSNP y miR-rSNP localizados en los promotores de genes codificantes de proteínas y no codificantes, respectivamente, llevan a una menor o mayor expresión génica, ya sea creando, destruyendo o modificando la afinidad de unión de diversas proteínas, como factores de transcripción.

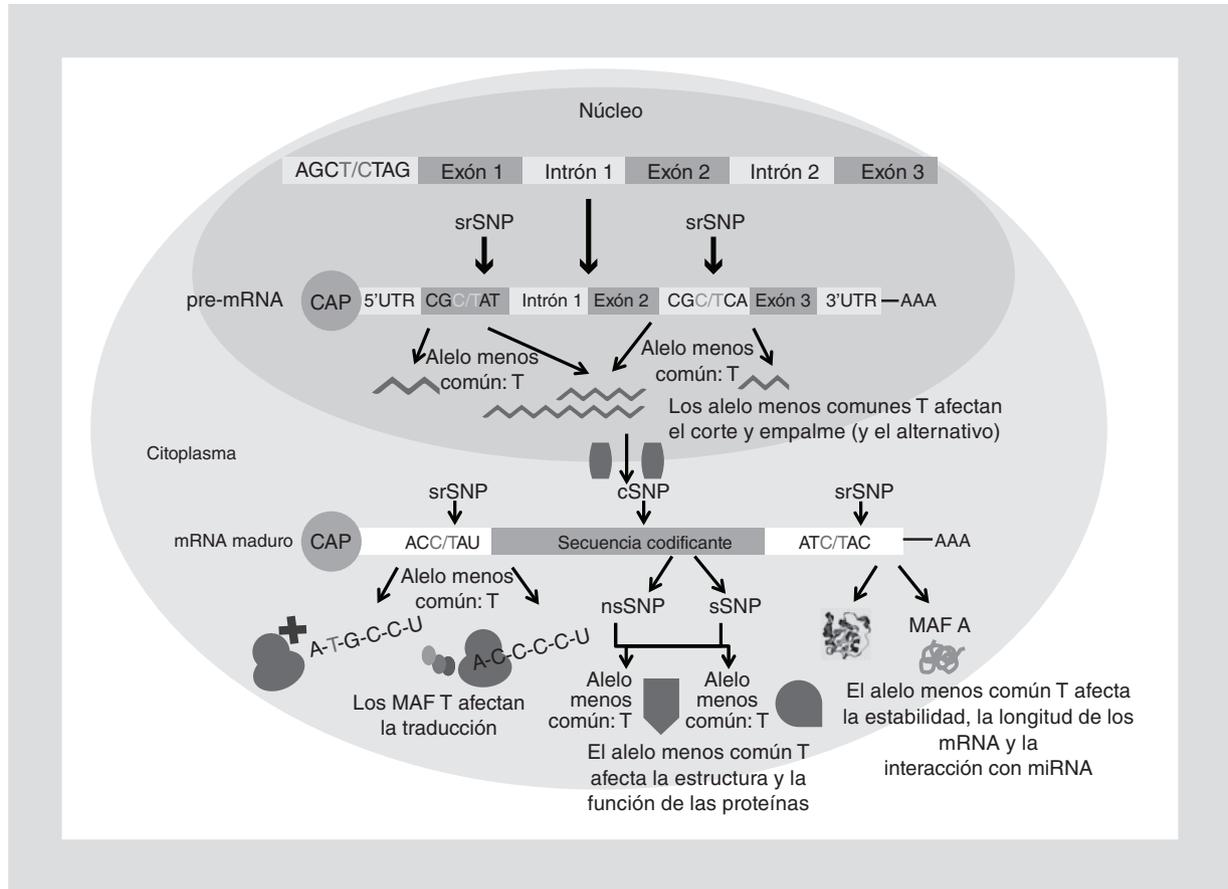
**Tabla 2. Clasificación funcional de los SNP**

SNP funcionales	Ubicación	Función
rSNP y miR-rSNP	Promotor de genes codificantes de proteínas y no codificantes, respectivamente	Alteran la expresión génica
- srSNP	pre-mRNA y mRNA maduros	Alteran la traducción, la estabilidad, la longitud y la interacción de mRNA/miRNA
- miR-srSNP	pri, pre y miRNA maduros	Afectan la estructura, el procesamiento y la función de los miRNA
cSNP	Secuencia codificante	Afectan la estructura y la función o la actividad de las proteínas o enzimas
- sSNP		Los nsSNP sin sentido generan un codón de paro y terminación prematura de las proteínas
- nsSNP		Los nsSNPs de sentido erróneo generan cambios de aminoácidos
• Sin sentido		
• Sentido erróneo		

cSNP: SNP codificantes; miR-rSNP: SNP reguladores de microRNA; miR-srSNP: SNP RNA estructurales de microRNA; nsSNP: SNP no sinónimos; rSNP: SNP reguladores; srSNPs: SNP RNA estructurales; sSNP: SNP sinónimos.

no codificantes se conocen poco, en comparación con los promotores de los genes codificantes<sup>17</sup>. Algunos elementos reguladores encontrados en el promotor de ambos tipos de genes son la caja TATA, elementos que

reconocen al factor de transcripción general II B, la secuencia iniciadora y más de la mitad de los promotores de los miRNA están asociados con islas CpG<sup>17</sup>. Diversas secuencias *cis* localizadas en esta región regulan la



**Figura 3.** Efecto funcional de los srSNP localizados en los pre-mRNA y los mRNA maduros. Los alelos menos comunes se encuentran en los pre-mRNA localizados en intrones y exones de genes codificantes de proteínas afectan el corte y empalme (el alternativo y la eficiencia). A su vez, los srSNP localizados en los mRNA maduros afectan la traducción, la estabilidad, la vida media y la interacción de mRNA y miRNA; además, pueden afectar la secuencia codificante, la estructura y la función de las proteínas.

unión de diversos factores de transcripción, transactivadores, etc., que actúan en *trans*. La unión *cis-trans* regula la expresión génica de manera coordinada.

### rSNP y miR-rSNP e implicaciones biológicas en enfermedades complejas

Una gran proporción de rSNP han sido identificados en el genoma humano (35,000-47,000); sin embargo, el número de miR-rSNP localizados en los genes no codificantes no ha sido reportado<sup>18,19</sup>. El alelo C del rSNP -169T/C de *FCRL3*, el cual codifica para la proteína 3 parecida a la fracción cristalizante gamma, aumenta la expresión de *FCRL3* a través de la modificación de la afinidad de unión para el factor de transcripción NF- $\kappa$ B, y confiere susceptibilidad para desarrollar artritis reumatoide (AR) y lupus eritematoso sistémico (LES)<sup>20</sup>. Por otro lado, los rSNP -376G/A, -308G/A y -238G/A del gen *TNF- $\alpha$* , el cual codifica para el factor de necrosis tumoral alfa, llevan a una

mayor expresión de mRNA y proteína, y aumentan el riesgo de desarrollar diversas enfermedades crónicas inflamatorias y autoinmunitarias<sup>3,21</sup>. Los individuos con genotipos G/A y AA del rSNP -308G/A del *TNF- $\alpha$*  no responden a la terapia biológica anti-TNF $\alpha$ , en comparación con el genotipo común, G/G<sup>3,21</sup>. Por su parte, el rSNP -168G/A del gen *MHC2TA*, que codifica para el transactivador del complejo principal de histocompatibilidad de clase II, disminuye *in vitro* su expresión y confiere riesgo para desarrollar AR, esclerosis múltiple e infarto agudo al miocardio<sup>22</sup>. Otro ejemplo lo constituye el rSNP -125G/A del gen *BAX*, el cual codifica para la proteína proapoptótica *BAX* (*un gen supresor de tumor*): el alelo A lleva a una menor expresión del RNA/proteína supresora de tumor en comparación con el alelo G, y confiere riesgo para desarrollar leucemia linfoblástica crónica<sup>23</sup>. Otros rSNP localizados en diferentes genes alteran el grado de expresión génica y confieren susceptibilidad para el desarrollo de enfermedades complejas humanas<sup>24-26</sup>. Por otro lado,

pocos miR-rSNP han sido reportados afectar su expresión. El alelo G del miR-rSNP rs57095329A/G del *miR-NA-146a* lleva a una menor expresión del miR-146a, el cual regula negativamente varios genes de la vía del interferón (un marcador de gravedad y actividad en el LES), y la disminución del miR-146a confiere riesgo para desarrollar LES<sup>27</sup>. Por otro lado, el alelo C del miR-rSNP rs4938723C/T (se encuentra en la isla CpG) del gen *miR-34b/c* disminuye su expresión *in vitro* y en células tumorales renales, y aumenta el riesgo para desarrollar este tumor<sup>28</sup>. Otros miR-rSNP localizados en diferentes genes de miRNA afectan sus niveles de expresión y confieren susceptibilidad para el desarrollo de diversas enfermedades humanas<sup>29-31</sup>.

## Estructura de los mRNA

Dada la importancia de cada región de los mRNA, se describirán brevemente su estructura y función. Diferentes regiones forman a los mRNA precursores (pre-mRNA) y maduros, y cada región participa en diferentes eventos biológicos. Los pre-mRNA están formados por:

- Exones e intrones, y en su extremo 5' y 3' del transcrito se encuentran la estructura caperuza, o Cap, y cola de poliadeninas (poliA), respectivamente.
- Por su parte, los mRNA maduros o simplemente mRNA están formados por la región 5' no traducida (UTR, *Untranslated Region*).
- Secuencia codificante.
- Región 3' UTR, y en sus extremos 5' y 3' del transcrito se encuentran la Cap y la cola de poliA, respectivamente; los intrones ya fueron eliminados (Fig. 3).

## Estructura y función de los 5' UTR

En el extremo 5' de la región 5' UTR de los RNAm primarios y maduros se encuentra la Cap, y sobre esta región se localizan diversas secuencias *cis* que regulan la traducción; entre ellas la secuencia de entrada a los ribosomas (IRES, *Internal Ribosome Entry Site*), el codón de inicio de la traducción (AUG) y otras AUG alternativas localizadas hacia arriba. La presencia de más de una IRES, AUG y diversas estructuras secundarias (p. ej., tallo-asa) regulan importantemente la traducción de los mRNA. En esta región se unen diferentes proteínas que regulan la entrada de los mRNA a los ribosomas, el inicio-alargamiento de la traducción y frecuentemente la expresión génica (por la cercanía de esta región al promotor de los genes). En los

eucariontes, el inicio de la traducción depende de dos eventos: el inicio dependiente de Cap y la entrada de los mRNA a los ribosomas<sup>32-34</sup>. El inicio de la traducción comienza cuando el factor IF4E se une a la Cap, inmediatamente después se une eIF4F y se forma un complejo que consta de eIF4E, eIF4A y eIF4G. La eIF4A, una helicasa de RNA dependiente de ATP, desdobra la estructura secundaria de la región 5' UTR estimulada por la proteína de unión al RNA: eIF4B. Posteriormente, el mRNA se une a la subunidad pequeña 40S de los ribosomas; esta subunidad encuentra el AUG e inicia la síntesis de proteínas<sup>32-34</sup>.

## srSNP localizados en la región 5' UTR de genes codificantes de proteínas y sus implicaciones biológicas en enfermedades complejas

Diversos srSNP localizados en la región 5' UTR de varios genes han sido caracterizados a nivel funcional. El srSNP +112A/C del exón 1 del gen *UCP1*, el cual codifica para la proteína desacoplante 1, se ubica en un elemento de respuesta para insulina. Un estudio mostró que el alelo de menor frecuencia altera los niveles de mRNA y confiere riesgo para diabetes mellitus tipo 2<sup>35</sup>. El srSNP rs3813946A/G del gen *CR2*, que codifica para el receptor 2 del complemento, no solo afecta su expresión génica, sino que también altera la accesibilidad de algunas proteínas a la cromatina y confiere susceptibilidad para desarrollar LES<sup>36</sup>. Por otro lado, el srSNP rs751402C/T del gen *ERCC5*, el cual codifica para la proteína 5 del complemento y está involucrado en la reparación por escisión de nucleótidos, crea un sitio de marco de lectura abierto río arriba del original, y afecta la expresión de la proteína y su capacidad para ser sintetizada después de un daño al DNA. Los individuos que tienen la variante de riesgo presentan resistencia a medicamentos basados en platino<sup>37</sup>. Otros srSNP localizados en esta región afectan la traducción de los mRNA, o el grado de expresión génica de sus respectivos genes, y confieren susceptibilidad para desarrollar enfermedades multifactoriales humanas, como asma, resistencia a la insulina y trastornos psiquiátricos<sup>38-40</sup>.

## Estructura y función de los intrones

Los intrones representan secuencias no codificantes localizadas entre exones de genes codificantes de proteínas o pseudogenes. Generalmente son grandes cuando se comparan con los exones, llegando a tener

una longitud de cientos a miles de nucleótidos. Un estudio reciente mostró que esas secuencias representan cerca del 24% de todo el genoma humano<sup>41-43</sup>. El corte y empalme es el mecanismo biológico por el cual los intrones son eliminados y unidos los exones. Datos recientes indican que el 70-98% de los genes codificantes de proteínas sufren este proceso<sup>3</sup>. Inicialmente se propuso que estas secuencias eran basura del DNA (dado que son eliminadas de los pre-mRNA), pero hoy sabemos que en ellos se encuentran genes de miRNA (solos o agrupados)<sup>44</sup> y diversas secuencias *cis* involucradas en el corte y empalme (y el alternativo), y en la eficiencia del corte y empalme, tales como sitios de ramificación y potenciadores o inhibidores del corte y empalme intrónicos (ISE, *Intronic Splicing Enhancer*, e ISS, *Intronic Splicing Silencer*, respectivamente), entre otras. El spliceosoma, que regula la eliminación de intrones y la unión de exones, está constituido por diversas ribonucleoproteínas pequeñas nucleares (snRNP, *Small Nuclear Ribonucleoprotein*, como U1, U2, U4, U5 y U6) y factores no relacionados con snRNP<sup>45,46</sup>. Un intrón típico contiene un sitio de corte y empalme en el extremo 5' (5 ss, *5 splice site*), una secuencia de ramificación (BPS, *Branch Point Sequence*), un tracto de polipurimidina (PYT, *Polypyrimidine Tract*) y otro ss en el extremo 3' del intrón (Fig. 3), además de contener comúnmente ISE o ISS<sup>3,45,47</sup>.

### srSNP localizados en intrones e importancia biológica en las enfermedades complejas

Los srSNP localizados en los intrones afectan el corte y empalme, el corte y empalme alternativo, y la eficiencia del corte y empalme. El srSNP rs9930761T/C, que cambia una timina por una citosina, localizado en el intrón 8 del gen *CEPT*, el cual codifica para una proteína que transfiere colesterol, se ubica en el sitio de ramificación, afecta la inclusión/exclusión del exón 9, está asociado con niveles alterados de lipoproteínas de baja densidad (LDL, *Low-Density Lipoprotein*) y confiere riesgo cardiovascular dependiente del sexo<sup>48</sup>. Por otro lado, tres srSNP (rs17026688C/T, rs17026651C/G y IVS8+48delG; el último, una delección del nucleótido G) localizados en el intrón 8 del gen *GADL1*, el cual codifica para la proteína 1 parecida a glutamato descarboxilasa, afectan fuertemente la terapia basada en litio en pacientes con enfermedad bipolar I. La variante IVS8+48delG afecta el corte y empalme de los exones 7 y 8 de *GADL1*, generando una isoforma corta debido a la eliminación de ambos exones en una línea celular neural, y este evento afecta

la respuesta al litio (fármaco empleado como primera línea de tratamiento en la enfermedad bipolar)<sup>49-51</sup>. Por otro lado, los srSNP rs2283265T/G y rs1076560T/G, localizados en los intrones 5 y 6, respectivamente, del gen *DRD2* que codifica para el receptor D2 de la dopamina, regulan la exclusión del exón 6 generando un transcrito corto. Estudios realizados en tejido de corteza prefrontal y estriado confirmaron que los portadores del haplotipo GG (proviene de ambos srSNP) generan un transcrito de *DRD2* corto, afectan la unión de las proteínas SRP55 y SRP40, las cuales regulan el corte y empalme, y alteran la actividad neuronal y la memoria<sup>52</sup>. Un ejemplo más lo representa el srSNP rs9406328A/G, localizado en el intrón 10 del gen *THBS2* que codifica para la trombospodina 2, proteína de matriz extracelular que regula los niveles de metaloproteasas involucradas en su remodelamiento. El srSNP rs9406328A/G se encuentra en el tracto de polipurimidina. Un estudio mostró que esta variante afecta la exclusión del exón 11 *in vivo* y confiere riesgo para desarrollar hernia de disco lumbar<sup>53</sup>. Otros srSNP localizados en los intrones de diferentes genes afectan el corte y empalme (y el alternativo), y la eficiencia del corte y empalme, y confieren riesgo para el desarrollo de diversas enfermedades multifactoriales humanas<sup>54-57</sup>.

### Estructura y función de los pre-mRNA

Los pre-mRNA (en sus extremos 5' y 3' tienen la Cap y la cola de poliA, respectivamente) están formados por exones e intrones, los exones forman las regiones 5' UTR, la secuencia codificante y la 3' UTR; por su parte, los intrones interrumpen la secuencia codificante. Los exones localizados en los pre-mRNA tienen un papel fundamental en la regulación del corte y empalme (y el alternativo), en la unión de diversas proteínas al RNAm, en la accesibilidad de proteínas que regulan el corte y empalme de los pre-mRNA, y en la formación de la estructura y la estabilidad de los pre-mRNA (Fig. 3). Diversas secuencias *cis* han sido identificadas al inicio y al final de los exones, tales como potenciadores e inhibidores del corte y empalme exónicos (ESE, *Exonic Splicing Enhancers*, y ESS, *Exonic Splicing Silencer*, respectivamente), secuencias que responden a la activación del corte y empalme, y sitios aceptores de corte y empalme, entre otros<sup>3,58-60</sup>. Estas secuencias son reconocidas por snRNP y otras proteínas, las cuales aumentan o inhiben la eficiencia del corte y empalme (y alternativo), generando diversas isoformas de mRNA, con inclusión o exclusión de exones, retención de intrones, etc., lo cual se traduce en proteínas con diferentes longitud y actividad<sup>3,58-60</sup>.

## srSNP localizados en los exones de los pre-RNA e importancia biológica en las enfermedades complejas

Los srSNP localizados en los exones de los pre-mRNA afectan el corte y empalme, el corte y empalme alternativo, y la eficiencia del corte y empalme. El srSNP C77G localizado en el exón 4 del gen *CD45*, el cual codifica para una proteína tirosina-fosfatasa, tipo receptor C, se encuentra en uno de los residuos del tracto de pirimidina. Este polimorfismo altera la función normal del ESS1, inhibe el corte y empalme, y afecta la inclusión del exón 4 en el transcrito de CD45. La eliminación de este exón lleva a la disminución de su actividad en las células T y contribuye en la hiperactividad del sistema inmunitario y en la susceptibilidad para desarrollar diversas enfermedades autoinmunitarias<sup>59</sup>. Por otro lado, el srSNP rs5883C/T, localizado en el exón 9 del gen *CETP*, afecta la secuencia de un ESE, y junto con el rs9930761T/C, localizado en el intrón 8 (el cual afecta el sitio de ramificación), alteran la inclusión del exón 9 y afectan en última instancia los valores de LDL y el riesgo cardiovascular<sup>48</sup>. A su vez, el srSNP rs767455A/G, localizado en el ss del extremo 3' del exón 1 del pre-mRNA de *TNFR1*, el cual codifica para el receptor 1 del factor de necrosis tumoral alfa, afecta la inclusión del exón 2; esta variante ha sido asociada con síndrome periódico asociado a TNFR<sup>61</sup>. Otros srSNP afectan el corte y empalme, la eficiencia del corte y empalme, el empalme alternativo, la estabilidad de la estructura de los pre-mRNA y la accesibilidad de diversas proteínas que regulan el corte y empalme, además de conferir riesgo para el desarrollo de diversas enfermedades humanas complejas (Tabla 3)<sup>62-64</sup>.

## Estructura y función de la secuencia codificante de los mRNA maduros

La secuencia codificante de los mRNA maduros está involucrada en la síntesis de las diferentes proteínas. En este proceso, el código genético utiliza 61 codones (código genético degenerado) para colocar 20 aminoácidos diferentes en las proteínas. Desde hace varios años se sabe que la secuencia codificante no solo codifica aminoácidos y sintetiza proteínas, sino que cumple otras funciones, como mantener estable la estructura secundaria o terciaria de los mRNA, a través de la unión de diversas proteínas de unión a RNA (RBP, *RNA-Binding Proteins*) involucradas en su plegamiento, las cuales además protegen al mRNA de la degradación<sup>65,66</sup>.

## sSNP y nsSNP en la secuencia codificante e importancia biológica en las enfermedades complejas

En teoría, los sSNP no tienen efecto en la célula, pero esto no es así, pues los sSNP localizados en los pre-mRNA y mRNA maduros pueden afectar el corte y empalme, la estructura, el plegamiento y la estabilidad de estos RNA, la respuesta a medicamentos y la interacción con diversas RBP (Fig. 3)<sup>13,65-69</sup>. Por otro lado, los nsSNP sin sentido y de sentido erróneo pueden afectar la estructura, el plegamiento, la estabilidad, la interacción con otras proteínas, la función y la actividad de diversas proteínas (o enzimas), y la respuesta a medicamentos<sup>15,69-72</sup>. Un sSNP localizado en el exón 26 del gen *MDR1*, el cual codifica para la glucoproteína p (P-gp, *Permeability Glycoprotein*), funciona como una bomba de eflujo que contribuye en la farmacocinética de diversos medicamentos, se ubica el sSNP 3435C/T (isoleucina/isoleucina) y no altera los niveles de mRNA y de proteína (los niveles son similares cuando se comparan ambos alelos); sin embargo, el cambio del codón común ATC por el raro ATT afecta la velocidad de traducción, el plegamiento, la función y la inserción de la P-gp a la membrana celular, y lleva a una menor efectividad para responder a ciertos medicamentos<sup>73,74</sup>. Por otro lado, el sSNP 971C/T, localizado en el gen *CDSN*, el cual codifica para la corneodesmosina, afecta la estabilidad de la estructura del mRNA (debido a que esta variante se encuentra en un motivo de estabilidad del mRNA), altera la unión de una proteína al mRNA y confiere riesgo para psoriasis (Tabla 3)<sup>75</sup>.

Por otro lado, varios nsSNP localizados en la secuencia codificante de *DNase 1* y *DNase1L3*, que codifican para la desoxinucleasa 1 y la proteína 3 parecida a la DNasa 1, disminuyen su actividad, confiriendo susceptibilidad al LES, debido a que ambas DNasas no eliminan los restos de DNA (de nucleosomas), llegando a generar autoanticuerpos contra el DNA<sup>76</sup>. Por otro lado, el nsSNP rs5744174C/T, que cambia leucina por fenilalanina en el gen *TLR5*, el cual codifica para el receptor 5 parecido a *toll*, aumenta la producción de la quimiocina 20 en respuesta a la flagelina y confiere riesgo para desarrollar enfermedad de Crohn<sup>77</sup>. El alelo T del nsSNP 1858C/T del gen *PTPN22*, el cual codifica para la proteína LYP, un regulador negativo de la señalización de linfocitos T/B mediada por el receptor de células T (TCR, *T Cell Receptor*), cambia una arginina por triptófano en el

Tabla 3. Ejemplo de SNP funcionales involucrados con riesgo para desarrollar diversas enfermedades humanas multifactoriales

Tipo de SNP	Gen	Posición del gen	Alteración molecular y patología involucrada <sup>(ref.)</sup>
rSNP	<i>FCRL3</i>	-169T/C	Afecta la expresión génica y confiere susceptibilidad para LES y AR <sup>20</sup>
rSNP	<i>TNF-α</i>	-376G/A -308G/A -238G/A	Afecta la expresión génica y confiere riesgo para malaria cerebral, AR y LES <sup>21</sup>
miR-rSNP	<i>miR-146a</i>	Promotor	Afecta la expresión génica y confiere riesgo para LES <sup>27</sup>
miR-rSNP	<i>miR-34b/c</i>	Promotor	Afecta la expresión génica y confiere riesgo para cáncer renal <sup>28</sup>
srSNP	<i>CR2</i>	5' UTR	Altera la expresión génica y la accesibilidad para proteínas que regulan la expresión, confiere riesgo para LES <sup>36</sup>
srSNP	<i>ERCC5</i>	5' UTR	Crea un marco de lectura adicional, afecta la expresión y la síntesis de la proteína, causa resistencia a agentes basados en cisplatino <sup>37</sup>
srSNP	<i>CETP</i>	Intrón 8	Causa alteración en el corte y empalme, y se asocia con enfermedades cardiovasculares <sup>48</sup>
srSNP	<i>GADL1</i>	Intrón 8	Afecta el corte y empalme alternativo, y altera la respuesta a la terapia basada en litio en la enfermedad bipolar <sup>49</sup>
srSNP	<i>SLC6A4</i>	3' UTR	Afecta la expresión y la estabilidad del mRNA, y se asocia con una mayor intensidad para ingerir alcohol <sup>69</sup>
srSNP	<i>TNFR2</i>	3' UTR	Altera la estabilidad del mRNA y se asocia con obesidad y con resistencia a la insulina <sup>90</sup>
srSNP	<i>miR-146a</i>	Tallo del pre-miR-146a G/C	Afecta la integridad de la estructura y el procesamiento del miRNA, y confiere susceptibilidad a cáncer de próstata, carcinoma hepatocelular y sepsis <sup>100</sup>
srSNP	<i>miR-196a-2</i>	Tallo del pre-miR-196a-2 C/T	Afecta el procesamiento del miRNA y confiere riesgo para cáncer colorrectal <sup>101</sup>
sSNP	<i>MDR1</i>	3435C/T exón 26	Afecta la velocidad de plegamiento y la función de P-gp en la membrana de las células <sup>73</sup>
nsSNP	<i>TLR5</i>	Exón	Induce una mayor cantidad de CCL20 en respuesta a la flagelina y confiere riesgo para enfermedad de Crohn <sup>77</sup>
nsSNP	<i>PTPN22</i>	Exón 14	Altera la unión Csk-LYP y ya no puede inactivar a células del sistema inmunitario, lo cual genera riesgo para AR <sup>78-80</sup>

AR: artritis reumatoide; CCL20: Chemokine C-C Ligand 20; CETP: Cholesteryl Ester Transfer Protein; CR2: Complement Receptor 2; ERCC5: Excisión Repair Cross-Complementation group 5; FCRL3: Fc receptor-like protein 3; GADL1: Glutamate decarboxylase-like 1; LES: lupus eritematoso sistémico; MDR1: Multidrug Resistance protein 1; miR-146a: microRNA146a; miR-196a-2: microRNA 196a-2; miR-34b/c: microRNA 34b/c; miR-rSNP: SNP reguladores de microRNA; nsSNP: SNP no sinónimos; PTPN22: Tyrosine-protein phosphatase non-receptor type 22; rSNP: SNP reguladores; SLC6A4: Solute Carrier family 6; srSNP: SNP RNA estructurales; sSNP: SNP sinónimos; TLR5: Toll-Like Receptor 5; TNFR2: Tumor Necrosis Factor Receptor 2; TNF-α: Tumor Necrosis Factor α.

codón 620 (R620W) y rompe la interacción con la proteína CSK (*C-Src Kinase*); ambas (con LYP) regulan negativamente la activación de los linfocitos B/T, llevando a una menor actividad de LYP y una mayor activación de células T/B, confiriendo susceptibilidad a la AR<sup>78-80</sup>. Otros sSNP y nsSNP afectan la interacción con otras proteínas, la actividad enzimática, etc., y confieren susceptibilidad a diversas enfermedades humanas (Tabla 3)<sup>15,65,69,81</sup>.

### Estructura y función de la región 3' UTR de los mRNA maduros

La presencia de secuencias *cis* y diversas estructuras localizadas en la región 3' UTR de los mRNA maduros regulan la expresión génica postranscripcional, debido en gran parte a la interacción con factores *trans*, que incluyen RBP y miRNA. Esta región está involucrada en la estabilidad, la localización y la exportación de los mRNA hacia el citoplasma, y en la

interacción de mRNA y miRNA<sup>82</sup>. La longitud de 3' UTR varía considerablemente y depende del tejido, del órgano o de la condición patológica; por ejemplo, un mismo RNAm que codifica para la misma proteína puede tener una 3' UTR con longitud variable en diferentes tipos celulares, debido a los sitios alternativos de poliadenilación localizados en esta región<sup>83,84</sup>. Isoformas cortas de la región 3' UTR se asocian con una mayor estabilidad, mientras que las largas se asocian con menor estabilidad, y las isoformas cortas escapan de los mecanismos que regulan la degradación de los RNAm debido a la unión con diferentes miRNA, los cuales suprimen la traducción de los mRNA; así, las 3' UTR largas presentan un mayor número de sitios de unión a miRNA. Otras secuencias localizadas en esta región regulan el decaimiento de la vida media de los mRNA<sup>83-87</sup>.

### srSNP localizados en la región 3' UTR e importancia biológica en las enfermedades complejas

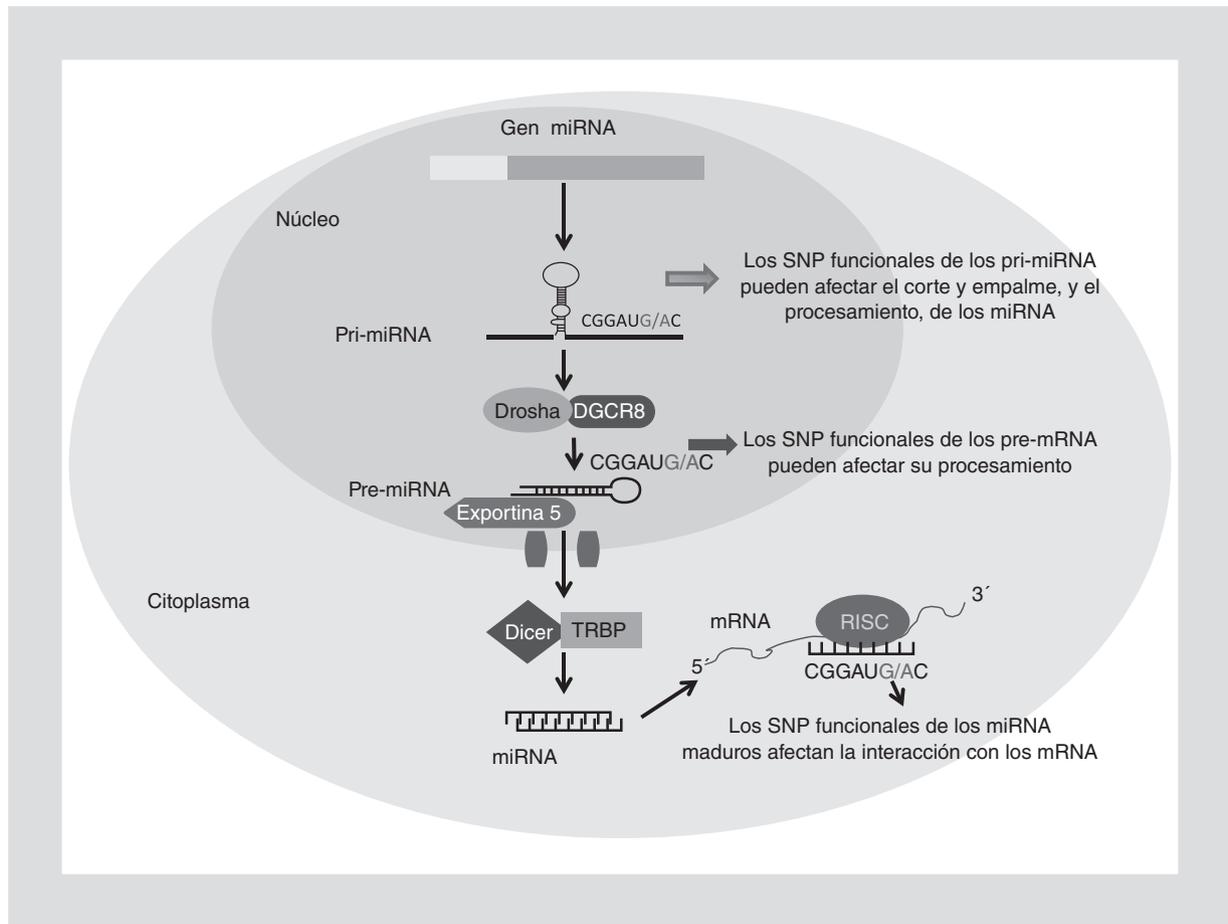
El srSNP rs3735590C/T del gen *PON1*, el cual codifica para la proteína paraoxonasa 1, afecta la unión del miR-616, el alelo C, reduce la afinidad de unión del miR-616 y aumenta la expresión del gen *PON1*. Esta variante confiere riesgo para desarrollar accidente cerebrovascular isquémico y rasgos subclínicos de aterosclerosis<sup>88</sup>. Por otro lado, el alelo G del srSNP rs1042173T/G, localizado cerca de un sitio de poliadenilación y del sitio de unión el miR-135 en la región 3' UTR del gen *SLC6A4*, el cual codifica para el transportador de serotonina (este gen se expresa en el cerebro de los humanos), aumenta los niveles de mRNA y proteína en comparación con el alelo T. Esta variante, además, mostró una asociación con la intensidad para beber alcohol. Algunos autores han hipotetizado que esta variante puede afectar la estabilidad del mRNA<sup>89</sup>. Tres srSNP localizados en la región 3' UTR del gen *TNFR2*, el cual codifica para el receptor 2 del factor de necrosis tumoral, afectan la unión de proteínas al mRNA, su vida media y su estabilidad. Estas variantes, además, han mostrado asociación con diversas enfermedades humanas complejas, como obesidad, incremento de leptina y resistencia a la insulina, entre otras (Tabla 3)<sup>90</sup>. Otras variantes que afectan la estabilidad de los mRNA y la unión con diversos miRNA, y confieren susceptibilidad a diversas enfermedades humanas, han sido descritas por varios autores (Tabla 3)<sup>91-93</sup>.

### Estructura y función de los genes que producen RNA no codificantes. Ejemplo de los miRNA

Los datos de la secuenciación del genoma humano mostraron que solo el 1.5% codifica proteínas; el resto de los transcritos provienen de genes no codificantes e incluye a los miRNA, tRNA, largos no codificantes, y rRNA, entre otros<sup>94-97</sup>. En esta sección se hablará específicamente de la estructura y la función de los miRNA. Los genes de *miRNA* se distribuyen prácticamente en todos los cromosomas, en el Y solo se han identificado un par de ellos. Dada su amplia distribución y tamaño, estos se pueden ubicar en regiones intergénicas, en intrones y en exones<sup>44,98</sup>. La mayoría de los pri-miRNA son sintetizados por la RNA polimerasa II, y de manera similar a los mRNA, contienen en sus extremos 5' y 3' la Cap y la cola de poliA, respectivamente<sup>99</sup>. Los miRNA maduros miden 18-22 nucleótidos y provienen de pre-miRNA de aproximadamente 70 nucleótidos, los cuales son generados de pri-miRNA más grandes<sup>97</sup>. Los pri-miRNA son escindidos por la RNasa Droscha, generando los pre-miRNA en el núcleo, los cuales se unen a la exportina 5 y viajan al citoplasma, en donde la RNasa III Dicer genera miRNA maduros dúplex de 22 nucleótidos. Posteriormente, una proteína argonauta selecciona la hebra guía, mientras la hebra antiguía es degradada por el complejo RISC (*RNA Induced Silencing Complex*), y el producto final es el miRNA, el cual se encuentra listo para unirse a sus diferentes mRNA (Fig. 4)<sup>97-99</sup>. Diversos estudios han mostrado que la región semilla de los miRNA (nucleótido 2-7) es determinante en la selección de mRNA<sup>97-99</sup>. Los miRNA se unen sobre todo a la región 3' UTR de los mRNA, y su principal función es inhibir la expresión génica a nivel postranscripcional, inhibiendo la síntesis proteica<sup>99</sup> (Fig. 4).

### srSNP localizados en la región 3' UTR e importancia biológica en las enfermedades complejas

Variantes genéticas localizadas en la estructura de los pri, pre o miRNA maduros afectan su procesamiento, actividad y función (Fig. 4). El miR-srSNP funcional rs2910164G/C, localizado en la región del tallo del *pre-miR-146a*, resulta en el cambio de C:U en lugar de G:U, afectando su integridad, procesamiento y forma madura del miR-146a. Esta variante, además, ha sido asociada con diversas enfermedades humanas, tales como cáncer de próstata, carcinoma hepatocelular y



**Figura 4.** Efecto funcional de los miR-srSNP localizados en pri-miRNA, pre-miRNA y miRNA maduros. Los MAF de SNP localizados en pri-miRNA, pre-miRNA y miRNA maduros afectan el corte y empalme (cuando se originan varios miRNA a partir de un solo transcrito), el procesamiento, la estructura, la interacción de miRNA y mRNA, y finalmente su función.

sepsis grave, entre otras<sup>100</sup>. Otro SNP funcional localizado en la región del tallo del *pre-miR-196a-2* es el rs11614913C/T. El alelo C de este SNP afecta su estructura, y estudios funcionales mostraron que el alelo C aumenta los niveles de *miR-192a-2* cuando se comparó con el alelo T en tejido tumoral de pulmón. El alelo menos común C se asoció con un aumento de la expresión de miR-196a-2 maduro. Estos resultados indican que el srSNP rs11614913C/T afecta el procesamiento del pre-miRNA a su forma madura. Estudios previos han mostrado que altos niveles de miR-196a promueven la migración y la invasión de células de cáncer colorrectal<sup>101</sup>. Por otro lado, un SNP funcional localizado en el *pre-miR-34a* reduce su expresión y promueve la migración y la proliferación de células de osteosarcoma<sup>102</sup>. Otros estudios han reportado cómo este tipo de SNP altera los niveles de expresión, el procesamiento o la actividad, confiriendo susceptibilidad a desarrollar diversas enfermedades multifactoriales humanas (Tabla 3)<sup>103,104</sup>.

## Conclusiones

Diversos estudios han mostrado la importancia funcional de los rSNP, los miR-rSNP, los srSNP y los miR-srSNP localizados en genes codificantes de proteínas y no codificantes en la fisiopatología de diversas enfermedades multifactoriales, afectando la expresión *génica*, el corte y empalme, la estabilidad (y la estructura), el procesamiento de los mRNA y miRNA, y la interacción de miRNA/mRNA. A su vez, los sSNP y nsSNP afectan la estructura, la función o la actividad de proteínas o enzimas implicadas en la respuesta a medicamentos. Comprender el efecto biológico de los alelos de los SNP en diferentes genes asociados con diversas enfermedades permitirá definir correctamente su influencia en la susceptibilidad, la gravedad y la actividad de las diferentes enfermedades multifactoriales humanas, y además contribuirá a identificar individuos que responderán o no a ciertos medicamentos o terapias biológicas.

## Agradecimientos

Se agradecen las facilidades proporcionadas por el Hospital Juárez de México.

## Conflictos de intereses

Los autores declaran no tener conflictos de intereses respecto a esta revisión.

## Bibliografía

- Sachidanandam R, Weissman D, Schmidt SC, et al. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature*. 2001;409:928-33.
- Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science*. 2001;291:1304-51.
- Ramírez Bello J, Vargas Alarcón G, Tovilla-Zárate C, Fragoso JM. Single nucleotide polymorphisms (SNPs): functional implications of regulatory SNP (rSNP) and structural RNA (srSNPs) in complex diseases. *Gac Med Mex*. 2013;149:220-8.
- Bønnelykke K, Sparks R, Waage J, Milner JD. Genetics of allergy and allergic sensitization: common variants, rare mutations. *Curr Opin Immunol*. 2015;36:115-26.
- Zeng P, Zhao Y, Qian C, et al. Statistical analysis for genome-wide association study. *J Biomed Res*. 2015;29:285-97.
- Kato N. Insights into the genetic basis of type 2 diabetes. *J Diabetes Investig*. 2013;4:233-44.
- Mekinian A, Tamouza R, Pavy S, et al. Functional study of TNF- $\alpha$  promoter polymorphisms: literature review and meta-analysis. *Eur Cytokine Netw*. 2011;22:88-102.
- Fiorillo E, Orrù V, Stanford SM, et al. Autoimmune-associated PTPN22 R620W variant reduces phosphorylation of lymphoid phosphatase on an inhibitory tyrosine residue. *J Biol Chem*. 2010;285:26506-18.
- Giancchetti E, Palombi M, Fierabracci A. The putative role of the C1858T polymorphism of protein tyrosine phosphatase PTPN22 gene in autoimmunity. *Autoimmun Rev*. 2013;12:717-25.
- Mishra, PJ, Mishra, PJ, Banerjee D, Bertino JR. MiRSNPs or MiR-polymorphisms, new players in microRNA mediated regulation of the cell: introducing microRNA pharmacogenomics. *Cell Cycle*. 2008;7:853-8.
- Sadee W, Wang D, Papp AC, et al. Pharmacogenomics of the RNA world: structural RNA polymorphisms in the drug therapy. *Clin Pharmacol Ther*. 2011;89:355-65.
- Obsteter J, Dovc P, Kunej T. Genetic variability of microRNA regulome in human. *Mol Genet Genomic Med*. 2015;3:30-39.
- Remensky V, Bork P, Sunyaev S. Human non-synonymous SNPs: server and survey. *Nucleic Acids Res*. 2002;30:3894-900.
- Chen R, Davydov EV, Sirota M, Butte AJ. Non-synonymous and synonymous coding SNPs show similar likelihood and effect size of human disease association. *PLoS One*. 2010;5:e13574.
- Haraksingh RR, Snyder MP. Impact of variation in the human genome on gene regulation. *J Mol Biol*. 2013;425:3970-7.
- Hull J, Campino S, Rowlands K, et al. Identification of common genetic variation that modulates alternative splicing. *PLoS Genet*. 2007;3:e99.
- Baer C, Claus R, Plass C. Genome-wide epigenetic regulation of miRNAs in cancer. 2013;73:473-7.
- Guo Y, Jamison DC. The distribution of SNPs in human gene regulatory regions. *BMC Genomics*. 2005;6:140.
- Kim BC, Kim WY, Park D, Chung WH, Shin KS, Bhak J. SNP@Promoter: a database of human SNPs (single nucleotide polymorphisms) within the putative promoter regions. *BMC Bioinformatics*. 2008;(9 Suppl):S2.
- Kochi Y, Yamada R, Szuki A, et al. A functional variant in FCRL3, encoding Fc receptor-like 3, is associated with rheumatoid arthritis and several autoimmunities. *Nat Genet*. 2005;37:478-85.
- Fragoso JM, Vargas Alarcón G, Jiménez Morales S, Reyes Hernández OD, Ramírez Bello J. Tumor necrosis factor alpha (TNF- $\alpha$ ) in autoimmune diseases (AIDs): molecular and biology and genetics. *Gac Med Mex*. 2014;150:334-44.
- Swanberg M, Lidman O, Padyukov L, et al. MHC2TA is associated with differential MHC molecule expression and susceptibility to rheumatoid arthritis, multiple sclerosis and myocardial infarction. *Nat Genet*. 2005;37:486-94.
- Moshynska O, Moshynskyy I, Misra V, Saxena A. G125A single-nucleotide polymorphism in the human BAX promoter affects gene expression. *Oncogene*. 2005;24:2042-9.
- Frey UH, Hauner H, Jöckel KH, Manthey I, Brockmeyer N, Siffert W. A novel promoter polymorphism in the human gene GNAS affects binding of transcription factor upstream stimulatory factor 1, Galphas protein expression and body weight regulation. *Pharmacogenet Genomics*. 2008;18:141-51.
- Lovewell TR, McDonagh AJ, Messenger AG, Azzouz M, Tazi-Ahni R. The AIRE -230Y polymorphism affects AIRE transcriptional activity: potential influence on AIRE function in the thymus. *PLoS One*. 2015;10:e0127476.
- Cui L, Gao Y, Xie Y, et al. An ADAM10 promoter polymorphism is a functional variant in severe sepsis patients and confers susceptibility to the development of sepsis. *Crit Care*. 2015;19:73.
- Luo X, Yang W, Ye DQ, et al. A functional variant in microRNA-146a promoter modulates its expression and confers disease risk for systemic lupus erythematosus. *PLoS One*. 2011;7:e1002128.
- Zhang S, Qian J, Cao Q, et al. A potentially functional polymorphism in the promoter region of miR-34b/c is associated with renal cell cancer risk in a Chinese population. *Mutagenesis*. 2014;29:149-54.
- Xu M, Qiang F, Gao Y, et al. Evaluation of a novel functional single-nucleotide polymorphism (rs35010275G>C) in MIR196A2 promoter region as a risk factor of gastric cancer in a Chinese population. *Medicine (Balt)*. 2014;93:e173.
- Li P, Yan H, Zhang H, et al. A functional polymorphism in MIR196A2 is associated with risk and progression of nasopharyngeal carcinoma in the Chinese population. *Genet Test Mol Biomarkers*. 2014;18:149-55.
- Cui L, Li Y, Ma G, et al. A functional polymorphism in the promoter region of microRNA-146a is associated with the risk of Alzheimer disease and the rate of cognitive decline in patients. *PLoS One*. 2014;9:ee89019.
- Wilkie GS, Dickson KS, Gray NK. Regulation of mRNA translation by 5'- and 3'-UTR-binding factors. *Trends Biochem Sci*. 2003;28:182-8.
- Mignone F, Gissi C, Liuni S, Pesole G. Untranslated regions of mRNAs. *Genome Biol*. 2002;3:REVIEWS0004.
- Van der Velden AW, Thomas AA. The role of the 5' untranslated region of mRNA in translation regulation during development. *Int J Biochem Cell Biol*. 1999;31:87-106.
- Mori H, Okazawa H, Iwamoto K, Maeda E, Hashiramoto M, Kasuda M. A polymorphism in the 5' untranslated region and a Met229 > Leu variant in exon 5 of the human UCP1 gene are associated with susceptibility to type II diabetes mellitus. *Diabetologia*. 2001;44:373-6.
- Cruikshank MN, Karimi M, Mason RL, et al. Translational effects of a lupus-associated polymorphism in the 5' untranslated region (UTR) of human complement receptor 2 (CR2/CD21). *Mol Immunol*. 2012;52:165-72.
- Somers J, Wilson LA, Kilday JP, et al. A common polymorphism in the 5' UTR of ERCC5 creates an upstream ORF that confers resistance to platinum-based chemotherapy. *Genes Dev*. 2015;29:1891-6.
- Ridderstråle M, Carsson E, Klannemark M, et al. FOXC2 mRNA expression and a 5' untranslated region polymorphism of the gene are associated with insulin resistance. *Diabetes*. 2002;51:3554-60.
- Holt RJ, Vandiedonck C, Willis-Owen SA, et al. A functional AT/G polymorphism in the 5'-untranslated region of SETDB2 in the IgE locus on human chromosome 13q14. *Genes Immun*. 2015;16:488-94.
- Duan J, Sanders AR, Molen JE, et al. Polymorphisms in the 5'-untranslated region of the human serotonin receptor 1B (HTR1B) gene affect gene expression. *Mol Psychiatry*. 2003;8:901-10.
- Grzybowska EA. Human intronless genes: functional groups, associated diseases, evolution, and mRNA processing in absence of splicing. *Biochem Biophys Res Commun*. 2012;424:1-6.
- Yenerall P, Zhou L. Identifying the mechanisms of intron gain: progress and trends. *Biol Direct*. 2012;7:29.
- Rogozin IB, Carmel L, Csuros M, Koonin EV. Origin and evolution of spliceosomal introns. *Biol Direct*. 2012;7:11.
- Ul Hussain M. Micro-RNAs (miRNAs): genomic organization, biogenesis and mode of action. *Cell Tissue Res*. 2012;349:405-13.
- Mattick JS, Gagen MJ. The evolution of controlled multitasked gene networks: the role of introns and other noncoding RNAs in the development of complex organisms. *Mol Biol Evol*. 2001;18:1611-30.
- Jin Y, Yang Y, Zhang P. New insights into RNA secondary structure in the alternative splicing of pre-mRNAs. *RNA Biol*. 2011;8:450-457.
- Zhang L, Li X, Zhao R. Structural analysis of the pre-mRNA splicing machinery. *Protein Sci*. 2013;22:677-92.
- Papp AC, Pinsonneault JK, Wang D, et al. Cholesteryl ester transfer protein (CETP) polymorphisms affect mRNA splicing, HDL levels, and sex-dependent cardiovascular risk. *PLoS One*. 2012;7:e31930.
- Chen CH, Lee CS, Lee MT, et al. Variant GADL1 and response to lithium therapy in bipolar I disorder. *N Engl J Med*. 2014;370:119-28.
- Baldessarini RJ, Tondo L. Does lithium treatment still work? Evidence of stable responses over three decades. *Arch Gen Psychiatry*. 2000;57:187-90.
- Rybakowski JK. Lithium in neuropsychiatry: a 2010 update. *World J Biol Psychiatry*. 2011;12:340-8.
- Zhang Y, Bertolino A, Fazio L, et al. Polymorphisms in human dopamine D2 receptor gene affect gene expression, splicing, and neuronal activity during working memory. *Proc Natl Acad Sci USA*. 2007;104:20552-7.
- Hirose Y, Chiba K, Karasugi T, et al. A functional polymorphism in THBS2 that affects alternative splicing and MMP binding is associated with lumbar-disc herniation. *Am J Hum Genet*. 2008;82:1122-9.

54. Wang D, Poi MJ, Sun X, Gaedigk A, Leeder JS, Sadee W. Common CYP2D6 polymorphisms affecting alternative splicing and transcription: long-range haplotypes with two regulatory variants modulate CYP2D6 activity. *Hum Mol Genet.* 2014;23:268-78.
55. Soemedi R, Vega H, Belmont JM, Ramachandran S, Fairbrother WG. Genetic variation and RNA binding proteins: tools and techniques to detect functional polymorphisms. *Adv Exp Med Biol.* 2014;825:227-66.
56. Morrison FS, Locke JM, Wood AR, et al. The splice site variant rs11078928 may be associated with a genotype-dependent alteration in expression of GSDMB transcripts. *BMC Genomics.* 2013;14:627.
57. Hecker M, Fitzner B, Blaschke J, Blaschke P, Zettl UK. Susceptibility variants in the CD58 gene locus point to a role of microRNA-548ac in the pathogenesis of multiple sclerosis. *Mutat Res Rev Mutat Res.* 2015;763:161-7.
58. de Souza JE, Ramalho RF, Galante PA, Meyer D, de Souza SJ. Alternative splicing and genetic diversity: silencers are more frequently modified by SNVs associated with alternative exon/intron borders. *Nucleic Acid Res.* 2011;39:4942-8.
59. Motta-Mena LB, Smith SA, Mallory MJ, Jackson J, Wang J, Lynch KW. A disease-associated polymorphism alters splicing of the human CD45 phosphatase gene by disrupting combinatorial repression by heterogeneous nuclear ribonucleoproteins (hnRNPs). *J Biol Chem.* 2011;286:20043-53.
60. Tazi J, Bakkour N, Stamm S. Alternative splicing and disease. *Biochim Biophys Acta.* 2009;1792:15-26.
61. Rittore C, Sánchez E, Soler S, et al. Identification of a new exon 2-skipped TNFR1 transcript: regulation by three functional polymorphisms of the TNFR1-associated periodic syndrome (TRAPS) gene. *Ann Rheum Dis.* 2014;73:290-7.
62. Bruun GH, Doktor TK, Andresen BS. A synonymous polymorphic variation in ACADM exon 11 affects splicing efficiency and may affect fatty acid oxidation. *Mol Genet Metab.* 2013;110:122-8.
63. Burkhardt R, Kenny EE, Lowe JK, et al. Common SNPs in HMGR in Micronesians and whites associated with LDL-cholesterol levels affect alternative splicing of exon 13. *Arterioscler Thromb Vasc Biol.* 2008;28:2078-84.
64. Kurmangaliyev YZ, Sutormin RA, Naumenko SA, Bazykin GA, Gelfand MS. Functional implications of splicing polymorphisms in the human genome. *Hum Mol Genet.* 2013;22:3449-5.
65. Shabalina SA, Spiridonov NA, Kashina A. Sound of silence: synonymous nucleotides as a key to biological regulation and complexity. *Nucleic Acids Res.* 2013;41:2073-94.
66. Tuller T, Zur H. Multiple roles of the coding sequence 5' end in gene expression regulation. *Nucleic Acids Res.* 2015;43:13-28.
67. Zhao Z, Fu YX, Hewett-Emmett D, Boerwinkle E. Investigating single nucleotide polymorphism (SNP) density in the human genome and its implication for molecular evolution. *Gene.* 2003;312:207-13.
68. Hunt R, Sauna ZE, Ambudkar SV, Gottesman MM, Kimchi-Sarfaty C. Silent (synonymous) SNPs: should we care about them? *Methods Mol Biol.* 2009;578:23-39.
69. Sauna ZE, Kimchi-Sarfaty C, Ambudkar SV, Gottesman MM. Silent polymorphisms speak: how they affect pharmacogenomics and the treatment of cancer. *Cancer Res.* 2007;67:9609-12.
70. Waldman YY, Tuller T, Keinan A, Ruppin E. Selection for translation efficiency on synonymous polymorphisms in recent human evolution. *Genome Biol Evol.* 2011;3:749-61.
71. Yates CM, Sternberg MJ. Proteins and domains vary in their tolerance of non-synonymous single nucleotide polymorphisms (nsSNPs). *J Mol Biol.* 2013;425:1274-86.
72. Stitzel NO, Tseng YY, Pervouchine D, Goddeau D, Kasif S, Liang J. Structural location of disease-associated single-nucleotide polymorphisms. *J Mol Biol.* 2003;327:1021-30.
73. Kimchi-Sarfaty C, Oh JM, Kim IW, et al. A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science.* 2007;315:525-8.
74. Komar AA. Silent SNPs: impact on gene function and phenotype. *Pharmacogenomics.* 2007;8:1075-80.
75. Capon F, Allen MH, Ameen M, et al. A synonymous SNP of the corneodesmosin gene leads to increased mRNA stability and demonstrates association with psoriasis across diverse ethnic groups. *Hum Mol Genet.* 2004;13:2361-8.
76. Ueki M, Kimura-Kataoka K, Takeshita H, et al. Evaluation of all non-synonymous single nucleotide polymorphisms (SNPs) in the genes encoding human deoxiribonuclease I and I-like 3 as a functional SNP potentially implicated in autoimmunity. *FEBS.* 2014;281:376-90.
77. Sheridan J, Mack DR, Amre DK, et al. A non-synonymous coding variant (L616F) in the TLR5 gene is potentially associated with Crohn's disease and influences responses to bacterial flagellin. *PLoS One.* 2013;8:e61326.
78. Burn GL, Svensson L, Sanchez-Blanco C, Saini M, Cope AP. Why is PTPN22 a good candidate susceptibility gene for autoimmune disease? *FEBS Lett.* 2011;585:3689-98.
79. Elshazli R, Settini A. Association of PTPN22 rs2476601 and STAT4 rs7574865 polymorphisms with rheumatoid arthritis: a meta-analysis update. *Immunobiology.* 2015;220:1012-24.
80. Xuan C, Lun LM, Zhao JX, et al. PTPN22 gene polymorphism (C1858T) is associated with susceptibility to type 1 diabetes: a meta-analysis of 19,495 cases and 25,341 controls. *Ann Hum Genet.* 2013;77:191-203.
81. Yates CM, Sternberg MJ. The effects of non-synonymous single nucleotide polymorphisms (nsSNPs) on protein-protein interactions. *J Mol Biol.* 2013;425:3949-63.
82. Michalova E, Vojtesek B, Hrstka R. Impaired pre-mRNA processing and altered architecture of 3' untranslated regions contribute to the development of human disorders. *Int J Mol Sci.* 2013;14:15681-94.
83. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB. Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science.* 2008;320:1643-7.
84. Mayr C, Bartel DP. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell.* 2009;138:673-84.
85. Miura P, Shenker S, Andreu-Agullo C, Westholm JO, Lai EC. Widespread and extensive lengthening of 3' UTRs in the mammalian brain. *Genome Res.* 2013;23:812-25.
86. Wabg L, Yi R. 3'UTRs take a long shot in the brain. *Bioessays.* 2014;36:39-45.
87. Matoulkova E, Michalova E, Vojtesek B, Hrstka R. The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA Biol.* 2012;9:563-76.
88. Liu ME, Liao YC, Lin RT, et al. A functional polymorphism of PON1 interferes with microRNA binding to increase the risk of ischemic stroke and carotid atherosclerosis. *Atherosclerosis.* 2013;228:161-7.
89. Seneviratne C, Huang W, Ait-Daoud N, Li MD, Johnson BA. Characterization of a functional polymorphism in the 3' UTR of SLC6A4 and its association with drinking intensity. *Alcohol Clin Exp Res.* 2009;33:332-9.
90. Pugal I, Lainez B, Fernández-Real JM, et al. A polymorphism in the 3' untranslated region of the gene for tumor necrosis factor receptor 2 modulates reporter gene expression. *Endocrinology.* 2005;146:2210-20.
91. Akdeli N, Riemann K, Westphal J, Hess J, Siffert W, Bachmann HS. A 3' UTR polymorphism modulates mRNA of the oncogene and drug target Polo-like kinase 1. *Mol Cancer.* 2014;13:87.
92. Miller CL, Haas U, Diaz R, et al. Coronary heart disease-associated variation in TCF21 disrupts a miR-224 binding site and miRNA-mediated regulation. *PLoS Genet.* 2014;10:e1004263.
93. Richardson K, Louie-Gao Q, Arnett DK, et al. The PLIN4 variant rs8887 modulates obesity related phenotypes in humans through creation of a novel miR-522 seed site. *PLoS One.* 2011;6:317944.
94. Mattick JS, Makunin IV. Non-coding RNA. *Hum Mol Genet.* 2006;15:R17-29.
95. Hüttenhofer A, Schattner P, Polacek N. Non-coding RNAs: hope and hype? *Trends Genet.* 2005;21:289-97.
96. Kagevama Y, Kondo T, Hashimoto Y. Coding vs non-coding: translatability of short ORFs found in putative non-coding transcripts. *Biochimie.* 2011;93:1981-6.
97. Mattick JS, Makunin IV. Small regulatory RNAs in mammals. *Hum Mol Genet.* 2005;14:R121-32.
98. Rodríguez A, Griffiths Jones S, Ashurst JL, Bradley A. Identification of mammalian microRNA host genes and transcription units. *Genome Res.* 2004;14:1902-10.
99. Cai X, Hagedorn CH, Cullen BR. Human microRNAs are processed from capped, polyadenylated transcripts that can also function as mRNAs. *RNA.* 2004;10:1957-66.
100. Shao Y, Li J, Cai Y, et al. The functional polymorphisms of miR-146a are associated with susceptibility to severe sepsis in the Chinese population. *Mediators Inflamm.* 2014;2014:916202.
101. Wang N, Li Y, Zhu LI, et al. A functional polymorphism rs11614913 in microRNA-192a2 is associated with an increased risk of colorectal cancer although not with tumor stage and grade. *Biomed Rep.* 2013;1:737-42.
102. Lv H, Pei J, Liu H, Wang H, Liu J. A polymorphism site in the pre-miR-34a coding region reduces miR-34a expression and promotes osteosarcoma cell proliferation and migration. *Mol Med Rep.* 2014;10:2912-6.
103. Dai ZJ, Shao YP, Wang XJ, et al. Five common functional polymorphisms in microRNAs (rs2910164, rs2292832, rs11614913, rs3746444, rs895819) and the susceptibility to breast cancer: evidence from 8361 cancer cases and 8504 controls. *Curr Pharm Des.* 2015;21:1455-63.
104. Slaby O, Bienertova-Vasku J, Svoboda M, Vyzula R. Genetic polymorphisms and microRNAs: new direction in molecular epidemiology of solid cancer. *J Cell Mol Med.* 2012;16:8-21.