

SITIOS CTCF COMO HERRAMIENTA PARA LA IDENTIFICACIÓN DE GENES ASOCIADOS A ENFERMEDAD

Comentario sobre el artículo: "Genome-wide CTCF distribution in vertebrates defines equivalent sites that aid the identification of disease-associated genes".

Publicado en Nature Structural & Molecular Biology 2011; 18: 708-714.

Olvera-Sumano Verónica,*,** Günther Jeannette*

*Subdirección de Enseñanza e Investigación. Hospital Regional de Alta Especialidad de Oaxaca.

**Facultad de Medicina y Cirugía. Universidad Regional del Sureste.

CORRESPONDENCIA/CORRESPONDENCE

Verónica Olvera Sumano

Subdirección de enseñanza

Hospital Regional de Alta Especialidad de Oaxaca
Aldama S/N San Bartolo Coyotepec Oaxaca
C.P. 71256

Tel. 50-18080 ext. 1095

veronica_o_s@hotmail.com

DETALLES DEL ARTÍCULO

Recibido el 01 de Junio de 2011.

Aceptado el 10 de Junio de 2011.

Rev Eviden Invest Clin 2010; 4 (2): 65-67.

CTCF sites like an identifying tool for gen associated diseases

Palabras clave: Genes, Enfermedad, ADN.

En el artículo se aborda el estudio de alteraciones genómicas relacionadas a enfermedades humanas ubicadas en secuencias reguladoras no codificantes del ADN, (catalogadas antes como ADN basura) por lo regular se localizan lejos de regiones promotoras; situación que plantea un reto para asociar estas mutaciones no codificantes o variantes de riesgo a sus genes diana. Para esto, fue analizada la distribución genómica de sitios CTCF (sitios de unión de la proteína CTCF) en células de humano, ratón y pollo; demostrando que

éstos ocupan posiciones sintéticas (en el mismo cromosoma) en genomas de vertebrados y una fracción considerable demostraron ser constitutivas además de evolutivamente conservadas en mamíferos. También así fueron analizadas las características de las secuencias relacionadas, en las que fue encontrada una asociación estadísticamente significativa ($p<0.001$) entre sitios CONSYN-CTCF (sitios conservados ratón-humano-pollo) y sitios E2F-1. Estos sitios CTCF de preferencia flanquean genes que codifican para factores

de transcripción, a menudo relacionados con enfermedades humanas y funcionan como bloqueadores de secuencias *enhancer* (potenciadoras) *in vivo*, sugiriendo que actúan como límites de genes conservados de manera evolutiva.

Una vez observado que estos sitios CTCF ocupan posiciones sintéticas en genomas de organismos vertebrados y que son evolutivamente conservados, este concepto fue aplicado más tarde para predecir y demostrar que de manera funcional las variantes polimórficas asociadas con esclerosis múltiple ubicadas en el gen ***EV15*** (*Ecotropic viral integration site 5 protein homolog*) afectan también al gen ***GFI1*** (*Growth factor independent 1 transcription repressor*) adyacente.

Estas variantes polimórficas (polimorfismos) relacionadas a esclerosis múltiple han sido identificadas en un bloque de desequilibrio de ligamiento, donde además se encuentran ***GFI1***, ***EV15***, ***RPL5*** y ***FAM69A***. El mapeo fino de estas regiones genómicas ha señalado a los polimorfismos del intrón 17 del gen ***EV15*** como la mayor parte de las probables variantes causales de asociación. Sin embargo, estos hallazgos no aclaran el rol funcional de esta región de riesgo ***EV15***. El análisis de sitios CTCF dentro de este bloque genético indica que el intrón 17 del gen ***EV15*** quizá pertenece a ***GFI1*** y no al dominio regulatorio de ***EV15***. Por último, fue demostrado que el incremento en la expresión de ***GFI1*** y no de ***EV15***, se asocia a un haplotipo (conjunto de alelos de múltiples loci que se transmiten juntos de padres a hijos) de riesgo de esclerosis múltiple. Por lo tanto, se concluye que ***GFI1*** y no ***EV15*** es el gen causal relacionado con esclerosis múltiple.¹

COMENTARIO

La esclerosis múltiple (OMIM 126200) es la afección neurológica progresiva más común en el adulto joven. Su prevalencia global es de 2-150 por 100 000 individuos. De acuerdo a su patogenia es una enfermedad autoinmune que lleva a la desmielinización de los axones.² En genética es un desorden complejo que resulta de una combinación de factores genéticos y no genéticos. Además de los antígenos HLA humanos ubicados en el locus de esclerosis múltiple, existen otros factores genéticos poco claros; por lo que la localización de estas variantes causales está aún pen-

diente de determinar.³ La ubicación de éstos podría ser en secuencias no codificantes, como secuencias reguladoras.

Los factores de transcripción represores CTCF (11-zinc finger protein o bien CCCTC-binding factor), contribuyen a la regulación de la expresión génica y a la organización de gran orden del genoma. Sus secuencias son evolutivamente conservadas y se encuentran distribuidas en forma amplia a lo largo del genoma de vertebrados y de *Drosophila melanogaster*. Sin embargo la función primaria de CTCF, parece no estar ligada de forma directa con su distribución genómica. Una buena proporción de sitios CTCF se encuentran evolutivamente conservados entre humano y ratón a nivel secuencial, y se ha planteado la posibilidad de que los sitios obligados CTCF se encuentren en posiciones sintéticas en diferentes vertebrados, ya que estas secuencias conservadas a lo largo de la evolución podrían servir para resolver asociaciones ambiguas con genes blanco, tal como es el caso de la esclerosis múltiple y los genes ***GFI1*** y ***EV15***.⁴

CONCLUSIONES

Los primeros resultados arrojados por el proyecto Genoma Humano demostraron que alrededor de 3-5% de la totalidad de secuencias del humano son codificantes;⁵ y por mucho tiempo la atención de los científicos fue centrada en estas regiones buscando mutaciones y variantes relacionadas a enfermedad; sin embargo, las regiones no codificantes aguardan funciones por demás interesantes del genoma como son las funciones de regulación, de las que depende la expresión o no expresión de un gen o grupo de genes, así también, por mucho tiempo fue asumido que estas regiones se encontraban a una corta distancia de la secuencia del gen. Sin embargo, hoy día las evidencias orientan a plantear una nueva opción en cuanto a la regulación génica.

Un importante porcentaje de sitios CTCF se encuentran evolutivamente conservados entre humano y ratón, lo que sugiere una fuerte conservación evolutiva en mamíferos; además estos sitios muestran actividad de bloqueo sobre secuencias *enhancer* que flanquean genes del desarrollo asociados a enfermedad en el humano; por lo que estas secuencias CTCF pueden

ser utilizadas como guía para identificar y asignar mutaciones que ocurren en regiones no codificantes a genes diana relacionados a patología humana.

Estos resultados sin duda alguna reorientarán el estudio de las enfermedades complejas, ya que si bien las regiones codificantes del ADN pueden resultar muy informativas todo parece indicar que son las regiones reguladoras las que contienen información valiosa so-

bre el buen o mal funcionamiento de aquellos genes que de manera estructural se observan inalterados.

El entendimiento de los mecanismos de expresión genética así como de sus variantes, contribuirán de forma sustancial a ampliar y mejorar el conocimiento sobre los mecanismos de etiopatogenia y permitirá ofrecer mejores opciones terapéuticas, elevando no sólo la esperanza sino la calidad de vida de estos pacientes.

REFERENCIAS

1.-MARTÍN D, PANTOJA C, FERNANDEZ-MIÑÁN A, ET AL. GENOME-WIDE CTCF DISTRIBUTION IN VERTEBRATES DEFINES EQUIVALENT SITES THAT AID IDENTIFICATION OF DISEASE-ASSOCIATED GENES. NATURE STRUCTURAL & MOLECULAR BIOLOGY 2011; 18: 708-714.

- 2.-HANDEL A, HANDUNNETHI L, GIOVANNONI G, ET AL. GENETIC AND ENVIRONMENTAL FACTORS AND THE DISTRIBUTION OF MULTIPLE SCLEROSIS IN EUROPE. EUR J. NEUROL 2010; 17:1210-1204.
- 3.-OKSENBERG JR, BARABZINI SE, SAWCER S, ET AL. THE GENETICS OF MULTIPLE SCLEROSIS: SNPs TO PATHWAYS TO PATHOGENESIS. NAT. REV. GENET . 2008; 9:516-526.
- 4.-ESSIEN K. CTCF BINDING SITE CLASSES EXHIBIT DISTINCT EVOLUTIONARY, GENOMIC, EPIGENOMIC AND TRANSCRIPTOMIC FEATURES. GENOME BIOLOGY. 2009; 10 (11): R131.
- 5.- THE INTERNATIONAL HUMAN GENOME SEQUENCING CONSORTIUM. INITIAL SEQUENCING AND ANALYSIS OF THE HUMAN GENOME. NATURE. 2001; 409: 860-921.