

# NiSspa: software de código abierto para la caracterización de aislados de *Staphylococcus aureus* por *spa*-typing para epidemiología molecular

Carlos Alfredo Carmona-Gasca<sup>1\*</sup>, Luis Eduardo Talavera-Morales<sup>2\*</sup>, Francisco Ramón García-Rodríguez<sup>1\*</sup>, Ana Patricia Orozco-Ortiz<sup>2</sup>, Juan José Valdez-Alarcón<sup>1</sup>✉

<sup>1</sup>Centro Multidisciplinario de Estudios en Biotecnología Facultad de Medicina Veterinaria y Zootecnia Universidad Michoacana de San Nicolás de Hidalgo Km 9.5 Carretera Morelia-Zinapécuaro Posta Veterinaria, La Palma. C.P. 58893, Tarímbaro, Michoacán, México.

<sup>2</sup>Universidad Tecnológica de Morelia, Av. Vicepresidente Pino Suárez No. 750 CP.58200, Morelia Michoacán.

## Resumen

*Staphylococcus aureus* es un patógeno versátil capaz de causar infecciones en humanos y animales con diferentes tipos de patologías. El gen *spa* de *S. aureus* presenta secuencias cortas repetidas (SSR's) que se utilizan en la técnica de *spa*-typing para la caracterización molecular de aislados del patógeno con fines epidemiológicos. Las bases de datos y los diferentes tipos de software disponibles para *spa*-typing son poco amigables y presentan la información fragmentada. En este trabajo se presenta una aproximación al diseño de un software alternativo, NiSspa (Nicolaita-software para *spa*-typing), que relaciona de forma amigable la información de *spa*-typing con la información presente en las bases de datos globales de epidemiología molecular de *S. aureus*. El software se probó contra las herramientas disponibles en línea utilizando datos experimentales generados en el laboratorio.

**Palabras claves:** Proteína A, *spa*-tipos. SSR's, Secuencias Cortas Repetidas

## Abstract

*Staphylococcus aureus* is a versatile and pathogenic bacteria that is able to cause infections in humans and animals with different pathologies. The protein A gene (*spa*) of *S. aureus* presents short sequences repeat (SSR's) that are used for *spa*-typing in molecular epidemiology for the characterization of *S. aureus*. The databases and different available tools for *spa*-typing are unfriendly and present fragmented information. This paper presents an approach to design an alternative software, NiSspa (Nicolaita-software for *spa*-typing), that friendly links *spa*-typing information with the global databases of molecular epidemiology of *S. aureus*. The software was tested against the available online tools using data generated in the laboratory.

**Keywords:** Protein A, *spa*-types. SSR's, Single Sequence Repeat's

## Introducción

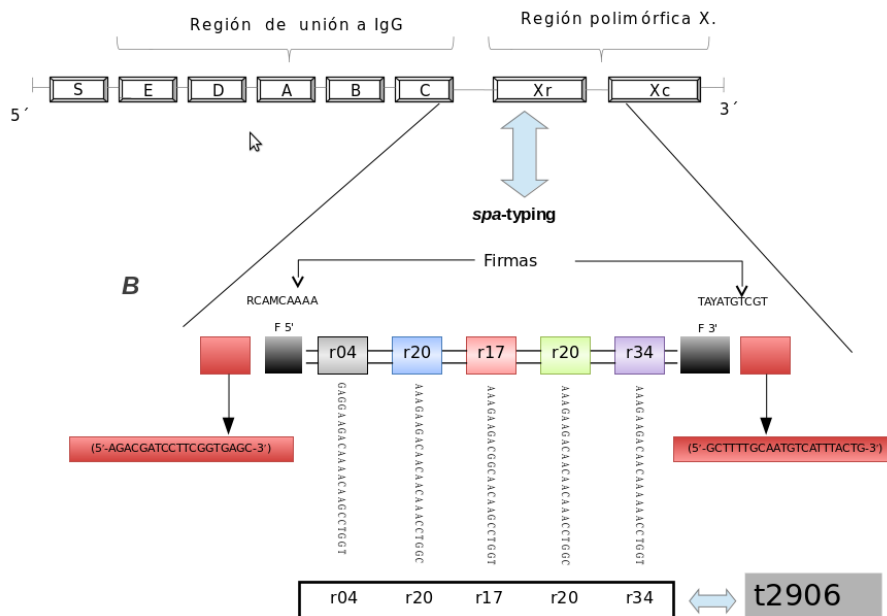
*Staphylococcus aureus* es un agente patógeno que causa diferentes tipos de infecciones en el hombre, en animales domésticos y salvajes. Entre las patologías más comunes en humanos se encuentran las infecciones de piel, tejidos blandos e intoxicaciones provocadas por enterotoxinas o la toxina de choque tóxico. Todas estas infecciones son Adquiridas en la Comunidad (denominadas AC), *S. aureus* es, además, uno de los patógenos más importante de infecciones Adquiridas en Hospitales (denominadas AH), en las que provoca osteomielitis, endocarditis infecciosa y septicemias (Bustos-Martínez *et al.*, 2006). En animales es capaz de causar piodermas, infecciones asociadas a intervenciones quirúrgicas y mastitis en los animales productores de leche (Graveland *et al.*, 2011). *S. aureus* es un patógeno exitoso gracias a que existen diferentes genotipos de y cada genotipo contiene variantes de un conjunto de factores de virulencia (Baizabal-Aguirre *et al.*, 2009). Uno de estos es la Proteína A (Spa), que

es la más abundante de la pared celular. Para su mejor estudio, Spa se divide en diferentes regiones funcionales (**Figura 1**): una región involucrada en su secreción (S), 5 dominios (E, D, A, B y C) de unión a inmunoglobulinas isotipo G (IgG) y una región denominada región X (**Figura 1A**). Esta se encuentra subdividida a su vez en una región constante (Xc), la cual es responsable de interactuar con la pared de *S. aureus* sirviendo como anclaje y la región (Xr), la cual está compuesta por un número variable secuencias cortas repetidas (SSR's) con un tamaño variable de 24 a 30 pares de bases (Shopsin *et al.*, 1999; Kim *et al.*, 2010) y que promueve procesos inflamatorios de una manera dependiente de la longitud de los SSR's (Garofalo *et al.*, 2012). La región Xr se encuentra delimitada por dos secuencias consenso que la flanquean y que se denominan secuencias "firma" (signature sequences). Las firmas en los extremos 5' (RCA MCA AAA) y 3' (TAY ATG TCG T), se utilizan así para identificar y delimitar la región de SSR's en Xr (**Figura 1B**).

La variabilidad en la longitud y en el polimorfismo de secuencias de las SSR's se ha aprovechado para utilizar a la región Xr como un marcador molecular en la caracterización de este patógeno, con la técnica llamada *spa*-typing (Harmsen *et al.*, 2003). En esta técnica se identifican el número y tipo de las SSR's y el orden en que se presentan. A la combinación de SSR's en la región Xr se le denomina *spa*-tipo y a la técnica se le conoce por su nombre en

✉ **Autor de correspondencia:** J.J. Valdez Alarcón. Centro Multidisciplinario de Estudios en Biotecnología Fac. De Medicina Veterinaria y Zootecnia Universidad Michoacana de San Nicolás de Hidalgo Km 9.5 Carr. Morelia-Zinapécuaro Posta Veterinaria, La Palma. C.P. 58893, Tarímbaro, Michoacán, México. Email: jjvaldez\_a@hotmail.com; jjvaldez@umich.mx Tel/Fax: +52(443)295-8029.

\*Estos autores contribuyeron de manera equivalente a este trabajo.



**Figura 1. a)** Estructura del gen *spa*. Las regiones del gen se describen en el texto. **b)** Organización estructural de la región Xr. Se muestran los oligonucleótidos spa-1113f y spa-1514r (rectángulos rojos) para la amplificación de la región Xr, la localización y secuencia de las firmas 5' (F 5') y 3' (F 3'), la secuencia (vertical) y designación de cada repetido y el *spa*-tipo asignado (t2906).

inglés *spa*-typing. Una de las ventajas de la secuenciación de la región Xr en la técnica de *spa*-typing es la portabilidad y sencillez con la que puede manejarse la información de las secuencias de nucleótidos que pueden ser representadas mediante un código de cuatro letras y manejarse en la forma de un archivo de texto, permitido el desarrollo de bases de datos globales disponibles en internet, para saber el *spa*-tipo, la secuencia de la región Xr de aislados obtenidos de diferentes fuentes y en diversas partes del mundo se compara con estas bases de datos. El sitio de internet para *spa*-typing es conocido como Ridom SpaServer (<http://spa.ridom.de/>). Este sitio recibe los archivos de salida de los equipos de secuenciación (electroferograma) y una persona, denominada curador, determina la calidad y confiabilidad del electroferograma y asigna un número de repetido (r) a cada SSR y al conjunto del número de repetidos le asigna un *spa*-tipo (t). El conjunto de *spa*-tipos reportados por los usuarios ha sido recopilado en esta base de datos internacional de libre acceso. Actualmente, se han reportado más de 561 repetidos distintos y 11,165 *spa*-tipos derivados de las combinaciones de repetidos cuyo número continúa en aumento.

La base de datos de Ridom SpaServer contiene además información epidemiológica de los aislados asociados

con los *spa*-tipos. Dicha información se encuentra de forma dispersa y redundante en 4 tablas con los datos epidemiológicos y moleculares relevantes como la combinación de repetidos, la frecuencia global de cada *spa*-tipo y su relación con el análisis de secuencias de *loci* múltiples o multi-*locus* sequence typing (MLST), otro método de caracterización molecular para *S. aureus* (Maiden 2006; Enright *et al.*, 2000). Esta técnica permite determinar el fondo genético compartido entre los aislados de *S. aureus* y agrupar a los mismos de acuerdo al número de alelos compartidos entre siete genes que se analizan. A la combinación de alelos de los siete genes se les denomina Secuencia-tipo (ST). A los grupos de similitud que comparten seis, cinco o cuatro de los siete alelos, se les denomina Complejos Clonales (CC). Ridom SpaServer proporciona una base de datos que contiene el país de origen, el tipo de muestra, y el alotipo del casete de resistencia a meticilina (SCC*mec*) de los aislados de *S. aureus* (IWG-SCC 2009).

De manera alterna, para designar el *spa*-tipo de forma rápida y sin tener que esperar a la evaluación del curador, existen herramientas en internet como SPA Searcher (PeiXe<software>; <http://seqtools.com/>) y Spa Typer/Identifier (<http://fortinbras.us/cgi-bin/spaTyper/spaTyper.pl>) destinados exclusivamente

para la identificación de los *spa*-tipos y de las firmas que flanquean a la región Xr. Estas dos herramientas dependen de la base de datos en la página de Ridom SpaServer. Sin embargo, no manejan los datos epidemiológicos presentes en Ridom SpaServer. A la fecha no existen programas de código abierto para el análisis de los *spa*-tipos y de la información epidemiológica relevante a los aislados, por lo que en el presente trabajo se creó, hasta donde sabemos, el primer programa bioinformático de código abierto para la búsqueda de *spa*-tipos que permita usar la base de datos internacional de *spa*-typing. El software se denominó NiSpa (Nicolaita-Software para *spa*-typing) y fue diseñado para identificar los *spa*-tipos en aislados de *S. aureus* de bovinos y humanos obtenidos del municipio de Tarímbaro, Michoacán. A diferencia de las herramientas de internet, este software presenta simultáneamente el análisis de los *spa*-tipos junto con los datos epidemiológicos sobresalientes, en forma ordenada y no redundante y en un solo formato de salida que es visualmente amigable.

## Materiales y métodos

### Construcción de las bases de datos

La base de datos que al momento de este trabajo contenía 554 repetidos y 10,646 *spa*-tipos, y la información epidemiológica contenida en cuatro tablas de la página Ridom SpaServer (<http://spa.ridom.de/>), se copió para conjuntarla en una hoja de Excel de Microsoft Office, en una sola tabla que incluyó todos los datos asociados a cada *spa*-tipo. Para tal fin se modificó la tabla de frecuencias reordenándola de acuerdo al orden ascendente de los *spa*-tipos. También se reconstruyeron las secuencias de los *spa*-tipos a partir de las secuencias de los números de repetidos, de manera que se generó una columna con las secuencias de cada *spa*-tipo. La base de datos final de Excel contiene todas las variables de forma no redundante en una sola tabla que contiene únicamente 7 variables: *spa*-tipo, frecuencia internacional, país de origen, información relevante, Secuencia-tipo (ST), combinación de repetidos y la secuencia completa del *spa*-tipo. Para evaluar de manera preliminar la nueva

base de datos, la tabla de variables generada en Excel se introdujo en un archivo del programa de Access de la suite de Microsoft Office, con su respectivo formulario, permitiendo después de varias modificaciones hacer una búsqueda de cualquier *spa*-tipo y de la información relevante.

Diseño y programación del software

Para el diseño del software se utilizó el programa Visual Basic, que es un conjunto de herramientas gratuito para programación de aplicaciones en lenguaje BASIC (Begginers All purpose Symbolic Instruction Code 1). Los códigos desarrollados para programar el software *NiSpa* se presentan a continuación.

-----

- FrmSpa.frm

```
Private Sub BtnBus_Click()
Dim DB As Database
Dim rs As Recordset
If Trim(TxtSeq.Text) <> "" Then
    sNomBd = App.Path & "\BDSpa.mdb"
    If gf_OpenBD(DB, sNomBd) = True Then
        Set rs = DB.OpenRecordset("tabla")
        c = 0
        Do While rs.EOF = False
            c = c + 1
            Cad = Trim(rs("secuencia completa").Value)
            cod = InStr(1, Trim(TxtSeq.Text), Cad)
            If cod <> 0 Then
                LblRep.Caption = rs("Repeat succession1").Value
                TxtTpo.Text = rs("Spa-type").Value
                TxtSeq2.Text = rs("secuencia completa").Value
                MLST.Caption = rs("MLST").Value & ""
                comment.Caption = rs("Comment").Value & ""
                Frequency.Caption = rs("Frequency").Value & ""
                origen.Caption = rs("Countries of origin").Value & ""
                ChkCin.Value = False
                If InStr(1, Trim(TxtSeq.Text), "GCACCAAAA")
                Then
                    ChkCin.Value = 1
                End If
                ChkTre.Value = False
                If InStr(1, Trim(TxtSeq.Text), "TACATGTCGT")
                Then
                    ChkTre.Value = 1
                End If
                'Izquierda
                ACAACAAAA.Value = False
                If InStr(1, Trim(TxtSeq.Text), "ACAACAAAA")
                Then
                    ACAACAAAA.Value = 1
                End If
                GCAACAAAA.Value = False
                If InStr(1, Trim(TxtSeq.Text), "GCAACAAAA")
                Then
```

```
GCAACAAAA.Value = 1
End If
ACACCAAAA.Value = False
If InStr(1, Trim(TxtSeq.Text), "ACACCAAAA")
Then
    ACACCAAAA.Value = 1
End If
GCACCAAAA.Value = False
If InStr(1, Trim(TxtSeq.Text), "GCACCAAAA")
Then
    GCACCAAAA.Value = 1
End If
'Derecha
TACATGTCGT.Value = False
If InStr(1, Trim(TxtSeq.Text), "TACATGTCGT")
Then
    TACATGTCGT.Value = 1
End If
TATATGTCGT.Value = False
If InStr(1, Trim(TxtSeq.Text), "TATATGTCGT")
Then
    TATATGTCGT.Value = 1
End If
MsgBox "Encontrado a partir de la posición " &
cod
End If
rs.MoveNext
Loop
DB.Close
End If
End Sub
Private Sub BtnCerrFrm_Click()
    FraDerAut.Visible = False
End Sub
Private Sub BtnSal_Click()
    Unload Me
End Sub
Private Sub CmdLim_Click()
    TxtSeq.Text = ""
    TxtSeq2.Text = ""
    TxtTpo.Text = ""
    ChkCin.Value = False
    ChkTre.Value = False
    ACAACAAAA.Value = False
    GCAACAAAA.Value = False
    ACACCAAAA.Value = False
    GCACCAAAA.Value = False
    TACATGTCGT.Value = False
    TATATGTCGT.Value = False
    LblRep.Caption = ""
    MLST.Caption = ""
    comment.Caption = ""
    Frequency.Caption = ""
```

```

    origen.Caption = ""
End Sub

Private Sub Form_Load()
End Sub

Private Sub TxtSeq_Change()
    Label5.Caption = "Longitud= " & Len(TxtSeq.Text)
End Sub

Private Sub TxtSeq2_Change()
    Label6.Caption = "Longitud= " & Len(TxtSeq2.Text)
End Sub

• Archivo Modspa.bas
Attribute VB_Name = "ModSpa"
Global sNomBd As String 'Nombre de la Base de Datos
Global BanPanErrBas As Boolean 'Bandera que indica que ya se detecto
el error de la BD

Function gf_OpenBD(dBDOpen As Database, NomBd As String) As
Boolean
    Dim rsOpen As Recordset
    If gf_ExiNomArc(NomBd) Then
        On Error GoTo ErrOpen
        Set dBDOpen = OpenDatabase(NomBd)
    Else
        GoTo ErrOpen
    End If
    gf_OpenBD = True
    BanPanErrBas = True
    Exit Function
ErrOpen:
    gf_OpenBD = False
    If Err = 3343 Then
        MsgBox "La base de Datos esta severamente dañada y no
        puede ser reparada", vbCritical, "Error " & Err 'Kill (sNomBd)
    End If
End Function

Function gf_ExiNomArc(NomArcBus As String) As Boolean
    gf_ExiNomArc = False
    On Error GoTo Fallo
    X = GetAttr(NomArcBus)
    gf_ExiNomArc = True
    Exit Function
Fallo:
    'If NomArcBus = sNomBd And BanPanErrBas = True Then

    If NomArcBus = sNomBd Then
        MsgBox "No existe la base de Datos"
        gf_ExiNomArc = False
        BanPanErrBas = False
    End If

    gf_ExiNomArc = False
End Function
-----

```

La carpeta de instalación contiene 30 archivos. Los tipos de archivos que contiene la carpeta con el programa son: El tipo de

archivo \*.DL se asocia con "animación" es el archivo que le da la imagen a la ventana del software, \*.LST identifica los archivos de datos de lista, \*.ex es un archivo fuente escrito y \*.tl es un operador menor el cual hace que el programa lea los algoritmos y pueda abrir sin errores. La instalación del software es ejecutada por el archivo contenido en la carpeta del programa con el nombre SETUP.EXE el cual permite la ejecución del software abriendo la ventana de instalación, donde se autoriza que el software sea corrido con un icono de seguimiento, para que pueda cargar los datos en una ventana. A su término creará un acceso directo para su ejecución.

El programa fue probado en 3 versiones del sistema operativos de la plataforma Windows: Windows XP Service Pack 1, 2 y 3, Windows vista (x86) y Windows 7 (x86). El programa no mostró errores al momento de su ejecución en ninguna de las plataformas utilizadas.

### Validacion; spa-typing de *S. aureus* de origen bovino

Se seleccionaron 23 aislados de *S. aureus* obtenidos en las localidades de Cotzío y Tájaro de los Izquierdo, municipio de Tarímbaro, Michoacán, México, pertenecientes a una colección del Laboratorio de Epidemiología Molecular y Biotecnología de Enfermedades Infecciosas del Centro Multidisciplinario de Estudios en Biotecnología (CMEB), en la Facultad de Veterinaria de la Universidad Michoacana de San Nicolás de Hidalgo (UMSNH). Los aislados se obtuvieron a partir de muestras de leche positivas a la Prueba de California para la detección de la mastitis bovina (Rice, 1997) en medios de cultivo selectivos y/o diferenciales (Bautista-Trujillo *et al.*, 2012) y se denominaron con las letras MRI (Mastitis Related Isolate) y un número secuencial. Se utilizó también la cepa de referencia de *S. aureus* NCTC 8325. Los aislados se desarrollaron en medio Infusión Cerebro Corazón (BHI) por 12 hrs a 37 °C, y se centrifugaron a 5,000 x g. El ADN fue obtenido con un protocolo modificado reportado previamente (Enright *et al.*, 2000), con 500 µl como volumen máximo y solución amortiguadora TEN<sub>30/1/100</sub> (30mM Tris-Hcl, 1 mM EDTA, 100mM NaCl pH 8), Proteínasa K (20 mg) y SDS (1% v/v). La estrategia de spa-typing se realizó según lo recomienda Ridom Spa Server. La amplificación de la región Xr del gen se trabajó en un volumen de 50 µl, conteniendo ADN purificado, 200 µM de dNTP's, 10 pmol de los oligonucleótidos spa-1113f (5'- TAA AGA CGA TCC TTC GGT GAG C -3') y spa-1514r (5'-CAG CAG TAG TGC CGT TTG CTT -3'), 1X de DreamTaq™ Buffer y 1.25 U de DreamTaq™ DNA polimerasa (Fermentas). Los ciclos térmicos fueron los siguientes: 95°C durante 5 min, seguidos de 35 ciclos a 94°C durante 45 s, 60°C durante 45 s y una extensión de 72°C durante 90 s y una extensión final de 72°C durante 10 min. La secuenciación se llevó a cabo utilizando los análogos de dideoxinucleótidos del kit Big Dye Terminator Cycle Sequencing Ready Reaction V 3.1 (Applied Biosystems). La polimerización se realizó en las siguientes condiciones: 1 ciclo por 5 min a 95°C; 25 ciclos de 0.5 min a 96°C, 0.5 min a 50°C y 4 min a 60°C; 1 ciclo sostenido de enfriamiento a 4°C. Los productos de polimerización se resolvieron en un secuenciador automático ABI Prism 310 Genetic Analyzer, utilizando un tiempo de inyección de 30 segundos, y la electroforesis a 15 kV, 50°C y 45 minutos de corrida;

en un capilar de 47 cm de longitud. La edición de los electroferogramas se realizó con el programa MEGA v5.0 (Tamura *et al.*, 2011), y la secuencia consenso de cada aislado fue introducida en el programa NiSpa. Los resultados fueron corroborados con las herramientas SPA Searcher (PeiXe<software>; <http://seqtools.com/>) y Spa Typer/Identifier (<http://fortinbras.us/cgi-bin/spaTyper/spaTyper.pl>).

### Resultados

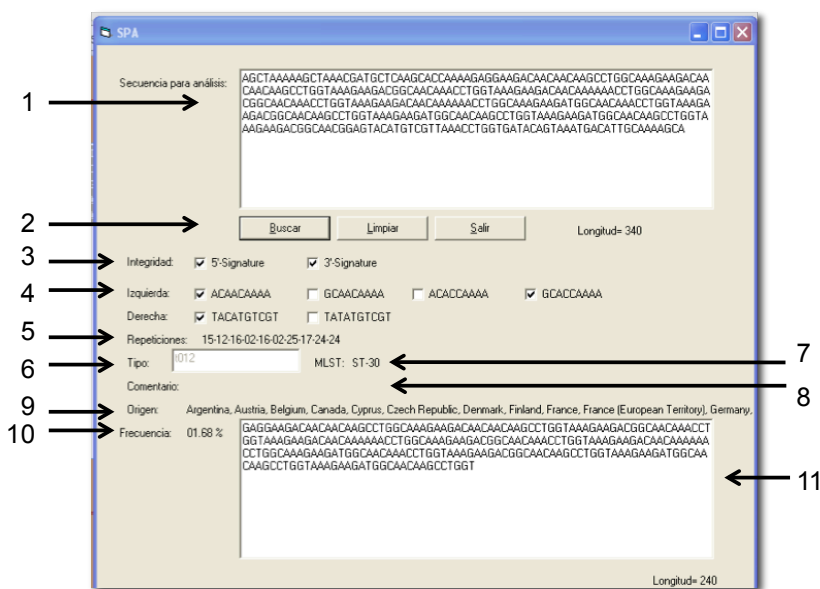
El software NiSpa procesa la información obtenida de la secuenciación de la región Xr para la búsqueda e identificación independiente de las secuencias firma 5' y 3' que flanquean la región de SSR's y la secuencia correspondiente a cada uno de los spa-tipos disponibles en la base de datos recuperada de Ridom SpaServer.

En la **Figura 2** se puede observar la ventana de análisis del software NiSpa. En ella se muestran los siguientes elementos: 1) ventana para insertar la secuencia a analizar; 2) botones para las acciones buscar/limpiar/salir, con los que se activa la búsqueda, la limpieza de la ventana para insertar una nueva secuencia o la salida del software; 3) cuadros de marca para indicar el hallazgo y la integridad de las secuencias firma 5' y 3'; 4) cuadros de marca para indicar las secuencias encontradas en las firmas 5' (Izquierda) y 3' (Derecha); 5) indicador de la secuencia de repetidos; 6) ventana que indica el spa-tipo encontrado, seguido de 7) la secuencia-tipo (MLST) asociada; 8) comentarios derivados de la base de datos de Ridom SpaServer; 9) origen reportado de los aislados; 10) frecuencia del spa-tipo encontrado; 11) secuencia delimitada que contiene los SSR's.

Para validar el software diseñado, se analizaron secuencias obtenidas en el Laboratorio de Epidemiología Molecular y Biotecnología de Enfermedades Infecciosas y secuencias de cepas de referencia. Las secuencias se analizaron en el software de NiSpa y se compararon con lo obtenido en SPA Searcher y Spa Typer/Analyzer. La **Tabla 1** se muestra los resultados de las secuencias que concordaron en los tres programas, las secuencias firma 5' y 3' identificadas, la secuencia de repetidos designada y el spa-tipo correspondiente. De la colección de aislados, no fue posible

**Tabla 1.** Caracterización de los aislados de *Staphylococcus aureus* de origen bovino utilizando NiSpa.

Cepa/ aislado	Firma 5'	Firma 3'	Secuencia de repetidos	spa-tipo
MRI-12	GCACCAAAA	TACATGTCGT	07-23-12-21-17-34-33-34	t224
MRI-13	GCACCAAAA	TACATGTCGT	11-19-12-21-17-34-24-34-22-25	t008
MRI-147	GCACCAAAA	TACATGTCGT	07-23	t605
MRI-148	GCACCAAAA	TACATGTCGT	07	t693
MRI-149	GCACCAAAA	TACATGTCGT	07-17-34-34-34-33-34	t1965
MRI-150	GCACCAAAA	TACATGTCGT	07-23-12-21-34	t4570
MRI-155	GCACCAAAA	TACATGTCGT	07-23-12-21-17-34	t189
MRI-163	GCACCAAAA	TACATGTCGT	07-23	t605
MRI-166	GCACCAAAA	TACATGTCGT	07-23-12-21-17-34-34-34-33-34	t267
MRI-171	GCACCAAAA	TACATGTCGT	07-23-12-21-17-34-34-34-33-34	t267
MRI-172	GCACCAAAA	TACATGTCGT	07-23-12-21-17-34-34-34-33-34	t267
MRI-22	GCACCAAAA	TACATGTCGT	07-23-12-21-17-34-33-34	t224
MRI-3	GCACCAAAA	TACATGTCGT	11-19-12-21-17-34-24-34-22-25	t008
MRI-32	GCACCAAAA	TACATGTCGT	07-17-34-34-34-33-34	t1965
MRI-34	GCACCAAAA	TACATGTCGT	07-17-34-34-34-33-34	t1965
MRI-35	GCACCAAAA	TACATGTCGT	26-16-23-31-29-29-17-17-25-16-28	t8265
MRI-36	GCACCAAAA	TACATGTCGT	07-17-34-34-34-33-34	t1965
NCTC-8325	GCACCAAAA	TACATGTCGT	11-19-12-12-21-17-34-24-34-22-25	t211



**Figura 2.** Imagen de salida de la pantalla de resultados de NiSpa. La numeración se refiere en el texto.

adjudicar spa-tipos a 5 aislados, ya que en la combinación de repetidos hubo 2 aislados (MRI-171 y MRI-172) que presentan repetidos (designados con una X) los cuales indican la no concordancia con los SSR's registrados en la página web Ridom SpaServer y que por lo tanto, puedan ser SSR's nuevos, propios de los

aislados de bovinos, que deban registrarse. Para otros 3 aislados (MRI-14, MRI-23 y MRI-152), la secuencia de los repetidos no está registrada en la base de datos. La combinación de repetidos para cada aislado no identificado con NiSpa son: MRI-14 (23-12-21-34), MRI-23 (23-12-21-17-34-33-34), MRI-152 (20-17),

MRI-171 (07-23-12-X-34-34-33-34) y MRI-172 (07-X-13-34-34-34-33-34). Las firmas oficiales 5' son 4 distintas: R(C o G)CA M(A o C)CA AAA habiendo otras 2 firmas 3': TAY(C o G) ATG TCG T en las reglas oficiales de Ridom SpaServer. En nuestros aislados se encontró una nueva firma 5' la cual es: GCACCAAAT, donde cambia la última A por una T, en la cepa MRI-171.

## Discusión

El análisis de spa-typing depende primordialmente de las herramientas de internet. NiSspa ocupa la base de datos de Ridom SpaServer (<http://spa.ridom.de/>) que, aunque concentra la información de los spa-tipos no permite al usuario realizar comparaciones. Como alternativa existen dos herramientas en la red que dependen de Ridom SpaServer que son: SPA Searcher (SeqTools) (<http://seqtools.com/>) y Spa Typer/Analyzer (<http://fortinbras.us/cgi-bin/spaTyper/spaTyper.pl>) con las que se corroboraron los spa-tipo de los aislados, las firmas que lo flanquea y las secuencias de repetidos. Un aporte fundamental de Ridom SpaServer es que incluye datos epidemiológicos de los aislados y una relación de los perfiles genéticos obtenidos por las distintas técnicas utilizadas en *S. aureus* como Multi-locus Sequencing Typing (MLST) (<http://saureus.mlst.net/>), la identificación por la secuencia parcial del gen ARN ribosomal 16S (SRV's) (<http://www.ridom-rdna.de/>), SCCmec typing (<http://www.staphylococcus.net/>) y de perfiles de macro-restricción en electroforesis de campos pulsados (PFGE; <http://www.harmony-microbe.net/microtyping.htm>). SPA Searcher y Spa Typer/Analyzer son herramienta que solamente manejan el análisis de secuencia de repetidos pero no ligan la base de datos de Ridom SpaServer con su información epidemiológica, al análisis de la secuencia de la región Xr. NiSspa aporta en una sola ventana el spa-tipo de la secuencia problema y toda su información epidemiológica relacionada encontrada en la base de datos internacional.

Para realizar la comparación de la información generada con la información contenida en las bases de datos es necesario utilizar programas bioinformáticos especializados. En la actualidad existe un aumento en la tendencia a crear programas de uso libre. Existen pocos programas especializados para el análisis global de datos de epidemiología molecular, como los spa-tipos en *S. aureus*. Hasta donde sabemos ninguno de ellos es de código abierto. NiSspa será el primer programa distribuido y desarrollado libremente cuya función principal será la detección del spa-tipo a partir de una secuencia curada de la región Xr del gen spa de *S. aureus*. NiSspa puede mostrar en una sola ventana la información epidemiológica relacionada, y permite la identificación de spa-tipos pudiendo trabajar sin necesidad de una conexión continua al internet.

Aunque NiSspa se operó en tres versiones de Windows, el programa está diseñado para trabajar en las versiones de Windows XP (x86) con Service Pack 3, Windows Vista (x86) con Service Pack 2 -, Windows 7 (x86), Windows Server 2003 (x86) con Service Pack 2, Windows Server 2003 R2 (x86), Windows Server 2008 (x86) con Service Pack 2, Windows Server 2008 R2 y la arquitectura de 32-Bit (x86). Los requerimientos del hardware son mínimos en cuanto a velocidad del procesador, memoria RAM

mínima de 512 MB, y solamente requiere 4 MB de espacio en disco duro. En su etapa de desarrollo actual, NiSspa solamente es capaz de identificar de manera independiente las secuencias firma y las secuencias de repetidos en los spa-tipos. Actualmente, se están desarrollando los códigos para incrementar las capacidades del software de manera que: 1) realice actualización automática de las bases de datos derivadas de la página <http://spa.ridom.de/>; 2) busque de manera independiente los SSR's individuales y 3) identifique las características de los posibles SSR's nuevos.

## Referencias

- Baizabal-Aguirre VM, J Oviedo-Boyso, JJ Valdez-Alarcón, A Bravo-Patiño y M Cajero-Juárez** (2009) Caracterización funcional de la interacción entre estructuras conservadas de *Staphylococcus aureus* y receptores celulares. En Bustos-Jaimes I, C Castañeda-Patlan, E Rendon-Huerta, H Reyes Vivas, Romero-Alvares I (eds.), *Mensaje Bioquímico Vol. XXIII*. Pp. 181-199.
- Bautista-Trujillo GU, JL Solorio-Rivera, I Rentería-Solórzano, SI Carranza-Germán, JA Bustos-Martínez, RI Arteaga-Garibay, VM Baizabal-Aguirre, M Cajero-Juárez, A Bravo-Patiño y JJ Valdez-Alarcón** (2012) Performance of cultura media for the isolation and identification of *Staphylococcus aureus* from bovine mastitis. *Journal of Medical Microbiology*. (En prensa) doi:10.1099/jmm.0.046284-0.
- Bustos-Martínez JA, A Hamdan-Partida y M Gutiérrez-Cárdenas** (2006) *Staphylococcus aureus*: la reemergencia de un patógeno en la comunidad. *Revista Biomédica* 17: 287-305.
- Enright MC, NPJ Day, CE Davies, SJ Peacock y BG Spratt** (2000) Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *Journal of Clinical Microbiology* 38: 1008-1015.
- Graveland H, B Duim, E van Duijkeren, D Heederik, y JA Wagenaar, JA** (2011) Livestock-associated methicillin-resistant *Staphylococcus aureus* in animals and humans. *International Journal of Medical Microbiology* 30: 630-634.
- Garofalo A, C Gai, S Lattar, N Gardella, M Mollerach, BC Kahl, K Becker, AS Prince, DO Sordelli, MI Gómez** (2012) The length of *Staphylococcus aureus* protein A polymorphic region regulates inflammation: impact on acute and chronic infection. *Journal of Infectious Diseases*. 206: 81-90.
- Harmsen D, H Claus, W Witte, J Rothganger, D Turnwald y U Vogel** (2003) Typing of methicillin-resistant *Staphylococcus aureus* in a university hospital setting by using a novel software for spa repeat determination and database management. *Journal of Clinical Microbiology* 41: 5442-5448.
- IWG-SCC, International Working Group on Staphylococcal Cassette Chromosome Elements** (2009) Classification of Staphylococcal Chromosomal Cassette mec (SCCmec): Guidelines for reporting novel SCCmec elements. *Antimicrobial Agents and Chemotherapy* 53:4961-4967.
- Kim HK, AG Cheng, H Kim, DM Missiakas y O Shneewind** (2010) Non-toxicogenic protein A vaccine for methicillin-resistant *Staphylococcus aureus* infections in mice. *Journal of Experimental Medicine* 207:1863-1870.
- Maiden MCJ** (2006) Multilocus sequence typing of bacteria. *Annual Review of Microbiology* 60: 561-588.
- Rice DN** (1997) *Using the California mastitis test (CMT) to detect*

*subclinical mastitis*. Bulletin G81-556A. Nebguide Publication. University of Nebraska.

**Shopsin B, M Gomez, SO Montgomery, DH Smith, M Waddington, DE Dodge, DA Bost, M Riehman, S Naidich y BN Krieswirth** (1999) Evaluation of protein A gen polymorphic region DNA sequencing for typing of *Staphylococcus aureus* stains. *Journal of*

*Clinical Microbiology* 37:3556-3563.

**Tamura K, D Peterson, N Peterson, G Stecher, M Nei y S Kumar** (2011) MEGA 5: molecular evolutionary genetic analysis using maximum likelihood, evolutionary distances and maximum parsimony methods. *Molecular Biology and Evolution* 28: 2731-2739.