Universidad Nacional Autónoma de México

Instituto de Investigaciones Biomédicas
Centro de Ciencias Genómicas

"Predicción de Residuos Funcionalmente Importantes en el dominio de unión al ligando de la familia de factores transcripcionales Crp/Fnr en bacterias"

Tesis que para obtener el grado de:

Licenciatura en Investigación Biomédica Básica

presenta:

Mónica Ivonne Peñaloza Spínola

Asesores de Tesis:

Dr. Ernesto Pérez Rueda
Dr. Julio Collado Vides

Cuernavaca, Morelos 2007

En memoria del abuelo

Alberto Spínola Spínola (1925 – 2006)

En recuerdo del tío abuelo

Ángelo Spínola Spínola (1923 – 2006)

Abuelo, me hubiera gustado tanto compartir

esto y más en vida contigo

A todas aquellas personas que con su compañía en diferentes momentos de mi vida,

me han dado tanto y enseñado otro mucho

Agradecimientos

Definitivamente, este trabajo no lo hubiera concluido en este tiempo y no hubiera logrado concretar los resultados y discusión sin la guía, discusiones, pláticas, ayuda en programación e incansable apoyo de Irma Lozada Chávez. Mil gracias amiga y colega de Ciencia.

En cuanto al desarrollo de los algoritmos y programas principales de este trabajo, le debo un gran agradecimiento a Hely y Bruno. Sin sus conocimientos y larga experiencia en estos menesteres, no lo hubiera logrado.

Por otra parte, la paciencia y compañía de mi familia (José Luís, Lourdes y Pamela) durante el lento y tortuoso desarrollo de esta tesis ha sido un gran apoyo. Sé que aunque no entendían por qué tardaba, nunca duraron que lo lograría (al menos eso quiero pensar). Finalmente pago mi deuda con ustedes.

También la compañía, apoyo y paciencia de uno de los mejores amigos que me ha dado la vida, Christián, ayudaron mucho a que no desfalleciera en mis intentos de terminar este trabajo.

A esta tesis misma, por haberme enseñado mucho sobre lo que se puede lograr al perseverar y luchar por la realización de las ideas.

Demás compañeros de laboratorio (sin orden especial: Vero, Irma M, Albert, Soco, Martín, Sarath, Romualdo, Víctor, Conchita, Juan, Arturo) mil gracias por todos los momentos que hemos compartido, pláticas, reflexiones y desarrollo de los proyectos bioinformáticos y de la vida.

Finalmente, gracias a mis tutores Julio Collado y Ernesto Pérez Rueda; a Julio por haberme dejado realizar este proyecto (sé que no le ve mucha relación con lo demás que hacemos en el laboratorio, pero espero convencerlo que tiene conexión y aplicación directa) y a Ernesto por las recomendaciones, guía, todas las opciones que buscaba para métodos y opciones de realización de esta idea y su paciencia para que terminara este proyecto.

Índice Página

0 Resumen

Actualmente, con la obtención de cientos de miles de secuencias de genes de una gran diversidad de organismos a través de los proyectos genómicos, se hace necesario utilizar estrategias alternativas a las experimentales que permitan avanzar rápidamente en la caracterización funcional de estas secuencias. Dentro de estas estrategias se encuentran las computacionales, que se valen del entendimiento de los procesos biológicos y conocimiento de sus componentes, modelandolos lo más real posible.

Con el uso de estas estrategias computacionales se puede, inicialmente, conocer el tipo de moléculas funcionales a las que codifican. Por ejemplo, RNAs, proteínas, entre otras. Posteriormente, se puede identificar las características propias de estos productos que los definen y los hacen específicos a sus funciones; por ejemplo: las proteínas están conformadas por cierto tipo de aminoácidos que les confieren una estructura específica, la capacidad de reconocer a sus sustratos, la capacidad de realizar las reacciones metabólicas que transforman a los sustratos, etc. Uno de los resultados finales con el uso de estas estrategias es que se logra conocer la función de los productos de las secuencias, sin la necesidad de hacer los experimentos exhaustivos individualmente para cada gen.

El trabajo que se presenta en esta tesis es el avance en la caracterización de una familia de proteínas que está involucrada en la expresión genética, la familia de factores transcripcionales Crp/Fnr. El objetivo principal radica en la predicción de residuos funcionalmente importantes (RFIs) en su dominio de unión al ligando.

Los resultados obtenidos aquí se basan en una estrategia que considera el contenido informacional, y su análisis estadístico para todas las posiciones del alineamiento de estas secuencias. Los resultados muestran diferentes RFIs para los grupos analizados, unos ya caracterizados experimentalmente y otros que aún no se caracterizan experimentalmente; estos últimos se proponen que podrían ser funcionalmente indispensables para los factores transcripcionales. Estos resultados han permitido identificar, dentro de las secuencias y estructuras, varios residuos que se proponen como importantes para las interacciones con los ligandos y/o entre subunidades, mantenimiento del plegamiento, y otras funciones que hacen común la respuesta a la regulación transcripcional de estos factores en diferentes organismos.

0.1 Summary

Currently, after obtaining hundreds of thousands of gene sequences from a great diversity of organisms through genomic projects, it has become necessary to use alternative strategies to the experimental ones, in order to quickly advance in the functional characterization of such amount of sequences. Within the alternative strategies are the computational ones, which considering the understanding of the biological processes and the knowledge of their components, model them the realest as possible.

With the use of these computational strategies it is possible, first, to get to know the type of functional molecules that they code for (e.g. RNAs, proteins, etc.). Later, it can be identified the properties that define and make specific, these products, for their functions (e.g. the type of amino acids that conform a protein confer a specific folding, a specific substrate recognition, the ability to do metabolic reactions to transform substrates, etc.). One of the final results of these strategies is to know the function of the sequence products, without doing exhaustive individual experiments.

The current work is an advance in the characterization of a protein family that is involved in the transcriptional regulation, the Crp/Fnr family. The main objective has been to predict the 'residues functionally important' (RFI) in their ligand binding domain.

The results here obtained are based on a strategy that considers the informational content of each position in the alignment of the domain sequences and its statistical analysis. There were identified different RFIs for the analyzed groups of sequences, ones of these RFIs are already characterized, but other RFIs remain to be experimentally characterized. These last RFIs are proposed to be functionally indispensable for the domain fold or structure, for different ligand and/or intersubunits interactions, or for other functions that allow the regulatory response of these transcription factors.

1 Introducción

1.1 Transcripción celular y regulación transcripcional

Todos los organismos tienen su información biológica codificada en tres tipos de moléculas básicas, DNA, RNA y proteínas. Cada tipo de molécula lleva a cabo funciones distintas y que se complementan entre sí. El DNA almacena la información genética que codifica para todas las funciones y características del organismo; el RNA es el transmisor y regulador de la información entre el DNA y las proteínas; y las proteínas son las ejecutoras de las funciones estructurales y de catálisis que la célula lleva a cabo, así como los vehículos de las reacciones metabólicas, las que dan estructura, transmiten señales, transportan, entre otras (Lewin 2000).

Durante la vida de la célula, se pasa de un tipo de molécula a otro para obtener diferentes tipos de moléculas funcionales a través de varios procesos biológicos que se representan en el Dogma Central de Biología Molecular (DCBM) (Crick 1970).

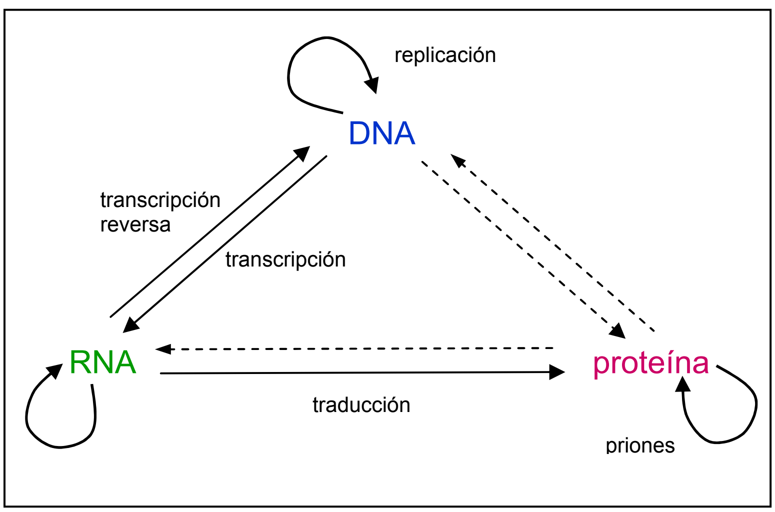


Figura 1. Dogma central de la biología molecular. Representa los nueve flujos de información biológica entre los tres diferentes tipos de moléculas básicas de la vida (DNA, RNA y proteínas), tanto los conocidos (líneas sólidas) como los que aún no se han descrito o inexistentes (líneas punteadas). Los procesos conocidos son: *Replicación DNA(DNA → DNA), Transcripción (DNA → RNA), Transcripción reversa (RNA → DNA), Traducción (RNA → proteína), Replicación RNA (RNA → RNA) y los recientemente descritos priones (proteína → proteína).*

Uno de los flujos del DCBM es la transcripción (figura 2), la cual se define como el proceso mediante el cual los genes, codificados en el DNA, son leídos por la enzima RNA polimerasa (y co-factores) para transcribir su código de DNA a RNA; y así poder ser procesados para obtener de ellos los productos funcionales para la célula (diferentes tipos de RNA, proteínas estructurales, enzimas, transportadores, etc.) (Lewin 2000).

Ya que el paso de información de DNA a RNA es aquel mediante el cual se obtienen los productos necesarios para el resto de los procesos celulares, es propio de las células regular este proceso para coordinar las diferentes tareas a realizar. Así, se entiende a la regulación transcripcional como "*la coordinación del proceso de la expresión genética controlando cuánto y cuándo un gen debe ser transcrito*." Esta regulación ocurre en diferentes etapas del proceso de transcripción y con diferentes mecanismos (etapas de transcripción y regulación, en figura 2). De ésta, de la que más se conoce es la regulación al inicio de la transcripción, comúnmente llamada 'regulación transcripcional'. Esta regulación es muy directa y eficiente, ya que permite el ahorro de energía, tiempo y recursos antes de empezar el proceso de transcripción, así como también hace la conexión directa con diversos tipos de señales del metabolismo (Lewin 2000).

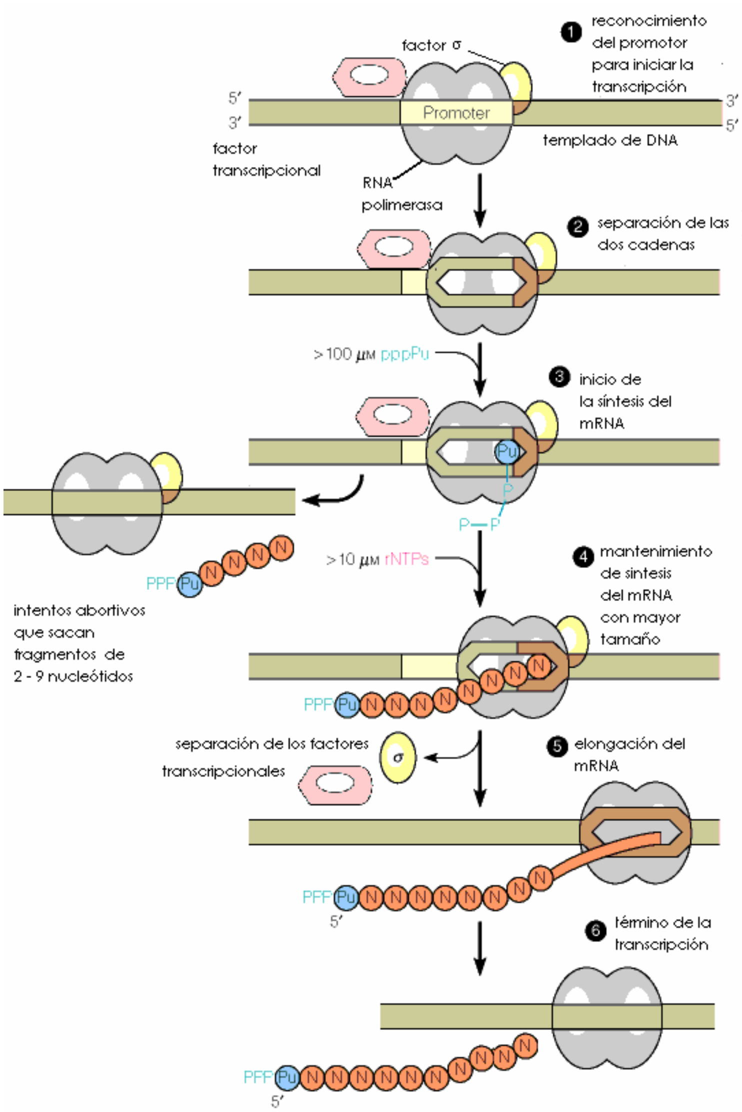


Figura 2. **Pasos del proceso de transcripción***:*
1) Reconocimiento del sitio de inicio de la transcripción, con ayuda de factores sigma y factores transcripcionales, 2) separación de las dos cadenas de DNA, 3) inicios cortos de transcripción con intentos abortivos y desprendimiento de la RNA-polimerasa de los factores, 4,5) elongación del transcrito, 6) termino de la transcripción
Etapas en las que ocurre la regulación transcripcional:
a) al inicio (etapa 1), b) posterior al inicio (etapa 2), c) durante la elongación (etapa 4), d) posterior al término (posterior a etapa 6).
Las formas en que ocurre la regulación son:
- inhibición/promoción a iniciar el proceso o para mantenerlo ocurriendo,
- disminución de la velocidad con que está ocurriendo,
- bloqueo del proceso (Lewin 2000; Browning and Busby 2004).

La regulación transcripcional coordina la transcripción de los genes cuyos productos son necesarios para los procesos metabólicos que la célula está realizando en ese momento, mientras que también inhibe la transcripción de los genes cuyos productos son innecesarios en ese momento fisiológico (Lewin 2000; Browning and Busby 2004).

1.1.1 Elementos de la regulación transcripcional

La regulación transcripcional es llevada a cabo por proteínas que interaccionan entre ellas, con la RNA-polimerasa y/o con regiones del DNA, llamadas comúnmente reguladores transcripcionales y factores sigma (ver figura 3) (Lewin 2000; Browning and Busby 2004).

1.1.2 Los factores Sigma

Los factores sigma son una subunidad proteínica temporal de la RNA polimerasa (RNA-p) que aumentan la afinidad de la interacción RNA-p con la región del promotor del DNA en 10^7 veces (figura 3) (Perez-Rueda 1999) y promueven que la RNA-p abra las dos cadenas de DNA para que inicie la transcripción (Burgess and Anthony 2001). Para ello, los factores sigma tienen dos regiones, una donde reconocen a los promotores y otra donde interaccionan con la RNA-p y los factores transcripcionales (los cuales intervienen en la interacción de la RNA-p con el DNA). Sin los factores sigma, la RNA-p se une y despega del DNA en cualquier parte y no transcribe (Burgess, Travers et al. 1969).

A su vez, los factores sigma fungen como reguladores modulares de la transcripción, pues hay diferentes factores sigma para diferentes condiciones globales de la célula. De los factores y sus mecanismos que más se han estudiado en procariontes, se han caracterizado diecisiete factores sigma en *Bacillus subtilis* (Haldenwang 1995; Moreno-Campuzano, Janga et al. 2006) y siete en *Escherichia coli* (Gruber and Gross 2003).

1.1.3 Los Factores transcripcionales

En la etapa de reconocimiento del DNA por la RNA-p además intervienen las proteínas que son factores transcripcionales (FT), también llamados 'reguladores transcripcionales'. Estos factores controlan la transcripción de manera más fina que los factores sigma, modulan la actividad transcripcional de los diferentes procesos celulares en respuesta a señales celulares específicas (ver más adelante en pág. 7). Activan o reprimen la expresión de los genes de acuerdo a que tanto se requieren sus productos para los diferentes procesos celulares (figura 3) (Browning and Busby 2004).

Los FT mejor caracterizados en procariontes son los de *E. coli* (Perez-Rueda and Collado-Vides 2000) y *B. subtillis* (Moreno-Campuzano, Janga et al. 2006); mientras que para el resto de los procariontes los estudios se realizan por comparaciones con FTs de estos y otros organismos (Mironov, Koonin et al. 1999; Makarova, Mironov et al. 2001; Tan, Moreno-Hagelsieb et al. 2001; Rodionov, Vitreschak et al. 2004). Para *E. coli* se ha caracterizado aproximadamente el 30% de su red de regulación (Salgado, Gama-Castro et al. 2006); con esto se sabe que del total de sus genes cerca del siete por ciento codifican para FTs[1] (314 genes son FT) (Perez-Rueda and Collado-Vides 2000), de los cuales 163 han sido determinados experimentalmente y 151 computacionalmente (RegulonDB [http://regulondb.ccg.unam.mx/]). A su vez, del primer conjunto de FTs, se ha determinado que el 24% son activadores, 46% son represores y 20% son duales [activan/reprimen] (el resto son de efecto desconocido) (Perez-Rueda and Collado-Vides 2000).

[1] *El porcentaje de genes de un genoma que codifican para factores transcripcionales correlaciona con la observación de Pérez-Rueda et.al 2003, de que un organismo de ambientes y modos de vida versátiles tiene mayor cantidad de FTs (7-10% de sus genes, e.g. E. coli misma), mientras que un organismo que sólo está adecuado a uno sólo o pocos habitats tiene menos TFs (2-3%, e.g. H. pylori) (Perez-Rueda E and Collado-Vides J 2004).*

Tabla 1. **Ejemplos de factores transcripcionales de *E. coli*.**

Factor transcripcional	Efector	Efecto en la transcripción	Genes – tipo de procesos regulados
OmpR	Fosfato	Activa	Principales proteínas de membrana, en respuesta a cambios de osmolaridad
MalT	Maltotriosa, ATP	Activa	Metabolismo de maltosa
LacI	Alolactosa	Reprime	Metabolismo de lactosa
PhoP	Fosfato	Reprime	Respuesta a disponibilidad de Magnesio en el medio
TrpR	Triptofano	Reprime	Biosíntesis de triptófano
AraR	Arsenito	Reprime	Resistencia a arsénico
GalR	Galactosa	Dual	Metabolismo de galactosa
CRP	AMP-cíclico	Dual	Represión catabólica
Lrp	Leucina	Dual	Metabolismo de leucina

Fuente: RegulonDB (http://regulondb.ccg.unam.mx/index.html)

Los FTs realizan su función al interaccionar con el DNA en sus “sitios de unión” u “operadores”, ubicados alrededor del inicio de la transcripción (entre las posiciones -200 y +40 pares de bases alrededor del inicio de la transcripción) (Madan Babu and Teichmann 2003). El reconocimiento de los operadores por parte de los FT se da sin abrir las cadenas de DNA, uniéndose a ellas por medio de puentes de hidrógeno, y suele ocurrir en el surco mayor del DNA (donde están las bases que dan la especificidad de reconocimiento) (Harrison 1991; Suzuki and Yagi 1996).

Las señales celulares a las que responden los FT pueden ser: a) sustratos o productos de las mismas vías metabólicas cuyos genes regulan ellos mismos, por ejemplo metabolitos como: vitaminas, aminoácidos, azúcares, etc (Baker, Tomlinson et al. 2001; Danot 2001; Lewis 2005); b) cambios ambientales externos como cambios de pH, de osmolaridad, de oxigenación, etc. (Gunsalus and Park 1994); y c) diferentes topologías del DNA, tales como la curvatura o su superhelicidad (Yasuzawa, Hayashi et al. 1992).

Estas señales lo que hacen es cambiar la conformación o el estado químico (óxido-reducción) de los FT, que los inducen a unirse o despegarse del DNA para hacer su efecto de regulación. Actualmente se sabe que cerca de la mitad de estas señales para modular a los FT son moléculas pequeñas (Madan Babu and Teichmann 2003). A estas señales de los TFs se les llama comúnmente ligandos o efectores.

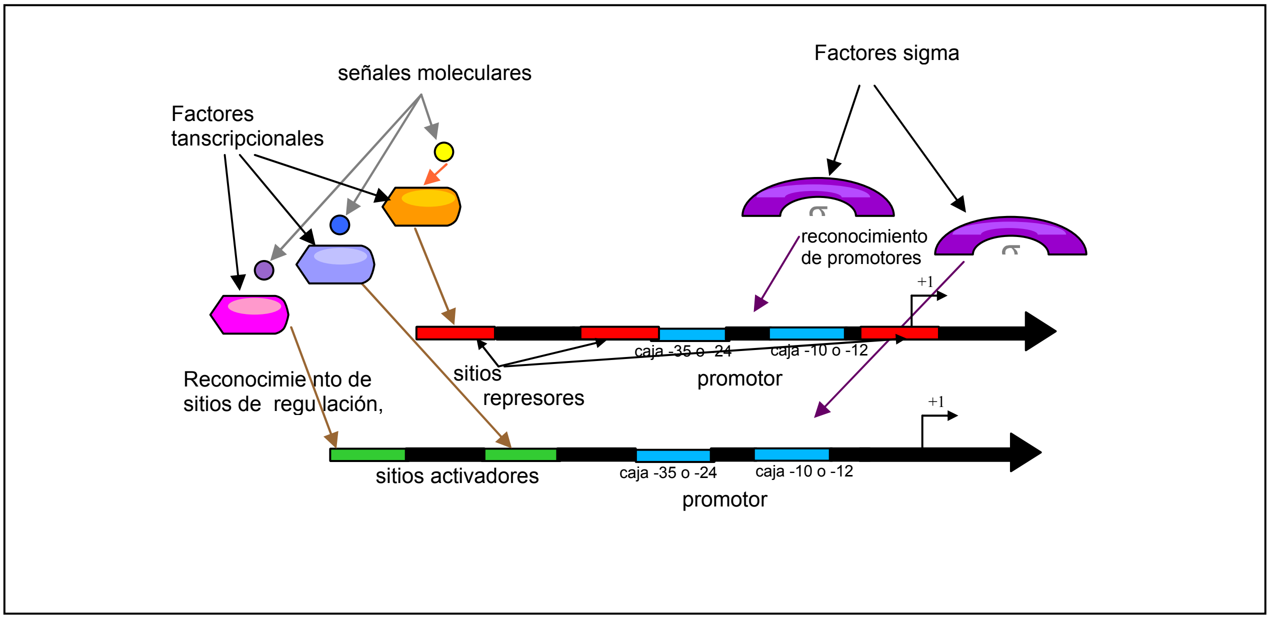


Figura 3. **Representación de ejemplo de los elementos de la regulación transcripcional junto a sus elementos en el DNA**. Los *factores sigma* interaccionan con las cajas de los *promotores* (localizadas en distancias centradas a -35 o -24 y -10 o -12 bases arriba del inicio de la transcripción +1) para promover que la RNA-polimerasa inicie la transcripción; mientras que los *factores transcripcionales*, en respuesta a diferentes *señales moleculares*, interaccionan con los *operadores* o *sitios de regulación* para facilitar o impedir que la RNA-polimerasa se una al DNA y/o inicie la transcripción.

a) Dominios que conforman a los factores transcripcionales

Los FT generalmente están conformados por dos dominios[2], uno de reconocimiento al DNA y el otro para unir al ligando o interaccionar con otras moléculas, incluyendo subunidades de la misma proteína u otras proteínas (Madan Babu and Teichmann 2003).

De los dominios de unión al DNA (DNA-binding domain, DBD) de los FT se tienen identificados cinco tipos de motivos (ver tabla 2), de los cuales el más común en bacterias es el tipo hélice-vuelta-hélice (helix-turn-helix, HTH) (Perez-Rueda and Collado-Vides 2001; Perez-Rueda, Collado-Vides et al. 2004). El HTH consiste de aproximadamente veinte residuos, se divide en dos α-hélices unidas por una vuelta de 3 o 5 residuos que forma entre ellas un ángulo de 120° (Brennan and Matthews 1989).

Tabla 2. **Estructuras de dominios de unión al DNA descritas para los factores transcripcionales.**

Dominio de unión a DNA	Características	Distribución
Hélice vuelta hélice (helix-turn-helix, HTH)	Formada por dos α-hélices de aproximadamente 10 residuos cada una, conectadas por una "vuelta" de alrededor 3 a 5 residuos de aminoácidos. La hélice C-terminal es la que da la especificidad del dominio uniéndose al surco mayor del DNA.	Procariontes, Eucariontes (modificada)
Dedos de Zinc	Consta de α-hélices y β-plegadas unidas por uno o más átomos de Zinc. Las α-hélices son las que interaccionan con el DNA.	Eucariontes, Arqueas
β-Plegada antiparalela	Conformada por tres o cuatro hojas β-plegadas, que interaccionan con el surco mayor del DNA.	Bacterias
Zipper de leucina	Formada por dos α-hélices que interaccionan entre sí a través de varios residuos de leucina.	Eucariontes
Hélice-asa-hélice (helix-loop-helix, HLH)	Consiste en una α-hélice corta conectada por un asa a una segunda α-hélice larga.	Eucariontes

Referencia: Harrison 1991.

[2] *En este trabajo, se toma la definición de dominio como la unidad mínima independiente estructural y funcionalmente (Lewin 2000).*
Tipos de dominios presentes en los FT: El 90% de los FT están conformados por un dominio de unión a DNA (DNA Binding Domain, DBD); aproximadamente el 75% de los FT tienen además un dominio extra donde interaccionan con su efector, entre sus subunidades o con otras proteínas (Dominio para Otras Interacciones, DOI); cerca del 12% de los FT tienen dos DOI; y alrededor del 3% tienen tres DOI (Madan Babu and Teichmann 2003).

De los dominios para otras interacciones (DOI) de los FT, tanto su función como estructura varían; su estudio y clasificación no es tan específico con los DBD, pero se pueden agrupar en cuatro tipos conocidos: a) de unión a moléculas pequeñas, b) tipo enzima, c) interacción con otras proteínas, d) receptor, y el resto siguen siendo desconocidos (Madan Babu and Teichmann 2003).

1.2 Clasificación de los factores transcripcionales

En general, las proteínas se pueden clasificar y analizar basándose en su secuencia, estructura y función (Holm 1998). Esta clasificación genera agrupamientos, de los cuales a su vez, se generan inferencias evolutivas, funcionales, de localización celular, de conformaciones, de estructuras (Mount 2001).

El estudio de secuencias, del que se comienza cualquier análisis de genes o proteínas, permite hacer el resto de las inferencias (Wilson, Kreychman et al. 2000). El estudio de estructuras usualmente apoya el análisis de secuencia para hacer las inferencias funcionales y de localización celular (Li, Jaroszewski et al. 2002).

a) Análisis de secuencia

Este tipo de análisis se realiza con la comparación de secuencias de genes o proteínas. Para esto se han diseñado diferentes algoritmos, entre los más utilizados se encuentran FASTA (Pearson 2000), BLAST (psi-blast o blast) (Altschul, Gish et al. 1990) que hacen búsquedas de similitud de secuencias a partir del alineamiento de las mismas; y los Hidden Markov Models (HMMs) (Eddy 1998) que utilizan cadenas de Markov que toman en cuenta la memoria de las posiciones anterior y posterior de un perfil de secuencias. Las comparaciones pueden ser entre un par de secuencias (alineamientos pareados) o varias secuencias a la vez (alineamientos múltiples); y las identidades entre secuencias se consideran de dos maneras, a lo largo de toda la secuencia (global), o por regiones o dominios (local) (Mount 2001).

Las secuencias de proteínas se relacionan en grupos de acuerdo a su porcentaje de similitud. Si las secuencias son muy similares, desde el 100% hasta un 40% de similitud, es altamente probable que las proteínas conserven la misma conformación estructural y conforme más similares sean las proteínas será más probable que realicen las mismas funciones bioquímicas (Wilson, Kreychman et al. 2000); si la similitud de secuencia disminuye hasta un 25%, estructuralmente son muy parecidas pero sus funciones ya son variantes (Chothia and Lesk 1986; Wilson, Kreychman et al. 2000).

Actualmente las secuencias de genes y proteínas se pueden consultar y obtener de bases de datos como GeneBank (Benson, Karsch-Mizrachi et al. 2003), SWISSProt (Boeckmann, Bairoch et al. 2003) y PIR (Protein Information Resource) (Barker, Garavelli et al. 1999). Por otra parte, los agrupamientos de secuencias de proteínas en familias, basados en sus alineamientos y dominios se pueden consultar y obtener de la base de datos Pfam (Sonnhammer, Eddy et al. 1998). Pfam utiliza los algoritmos de BLAST para toda la secuencia, y HMM para los dominios.

b) Análisis estructural

En este otro tipo de análisis, las proteínas se relacionan por la similitud de sus estructuras secundarias y/o terciarias de sus dominios y de la proteína completa. Esta similitud se determina por el ordenamiento de los dominios y mediciones de distancias entre estructuras, principalmente. Este análisis también, se apoya en la similitud de secuencias (Li, Jaroszewski et al. 2002).

Actualmente, las estructuras tridimensionales se pueden consultar y obtener del repositorio Protein Data Bank (PDB) (Berman, Westbrook et al. 2000). Aparte, los agrupamientos y clasificaciones estructurales se tienen en varias bases de datos: SCOP, clasificación evolutiva y estructural (Murzin, Brenner et al. 1995); CATH, clasificación jerárquica de estructuras de dominios (Pearl, Bennett et al. 2003); y Superfamily, clasificación estructural de secuencias de proteínas en familias evolutivas (Gough 2002).

La clasificación de los FT se basa principalmente en la comparación de sus secuencias, divididos en dominios. A la fecha, hay dos grandes estudios para la clasificación de los

TF: uno que se basa en análisis de los dominios de unión al DNA (DBD), considerando una similitud de secuencia mínima de 25%, del cual deriva su clasificación en 75 familias evolutivas (Perez-Rueda and Collado-Vides 2000); y el otro que además de las secuencias de los DBD considera las estructuras secundarias y terciarias de estos dominios, y los clasifica en 11 familias estructurales (Madan Babu and Teichmann 2003).

1.3 Familias de reguladores transcripcionales

Los FT que se agrupan en familias relacionándose por similitud de secuencias, con un 25% de identidad mínimo, se observa que tienen: misma ubicación del DBD (región carboxilo o amino), poseen un grupo de arquitecturas comunes, mecanismos de acción comunes, tamaños de secuencia de entre los 150 hasta los 400 aminoácidos, regulación de procesos similares, e inclusive la tendencia a comportarse como activadores y/o represores (Perez-Rueda and Collado-Vides 2000; Madan Babu and Teichmann 2003). Algunos ejemplos de familias de TFs se muestran en la tabla 4.

Tabla 4. **Ejemplos de familias de FT, con las características comunes más relevantes**.

Familia	tipo de procesos regulados	efecto en la regulación	ubicación de su DBD
AraC/Xyls	Metabolismo de carbono, respuesta a estrés y patogénesis.	Activación/Represión	Carboxilo
LysR	Biosíntesis de aminoácidos	Activación/Represión	Amino
MerR	Estrés oxidativo, presencia de metales pesados y antibióticos	Activación	Amino
GalR/LacI	Metabolismo de azúcares alternativos.	Represión	Amino
GntR	Metabolismo de carbonos y ácidos grasos, biosíntesis de aminoácidos y varios más.	Represión	Amino

Fuente y referencias en: RegulonDB (http://regulondb.ccg.unam.mx/)

1.3.1 Familia de factores transcripcionales CRP/FNR

La familia Crp/Fnr[3] está representada por los dos reguladores globales que le dan nombre a la familia (Shaw, Rice et al. 1983). Crp es un FT de respuesta global que le permite a la célula contender con la falta de glucosa facilitando la activación de vías alternativas para alimentarse con otros carbohidratos (fenómeno conocido como represión catabólica) (Zubay, Schwartz et al. 1970); y Fnr es otro FT de respuesta global que le permite a la célula responder rápidamente a la ausencia de oxígeno re-dirigiendo su vía de respiración aeróbica a anaeróbica (Spiro and Guest 1990).

Debido a las importantes repercusiones que tienen los efectos de la regulación por Crp y Fnr, los integrantes de esta la familia han sido tema de diversos estudios desde su descubrimiento (Li, Wing et al. 1998; Green, Scott et al. 2001; Korner, Sofia et al. 2003). A la fecha, estos estudios se han enfocado mucho en la caracterización de las proteínas Crp y Fnr. Se han identificado nuevos homólogos de Crp, Fnr y de los nuevos miembros de la familia (Perez-Rueda and Collado-Vides 2000; Korner, Sofia et al. 2003). Se tiene la estructura tridimensional de Crp y de otros dos miembros de la familia, CooA y PrfA. Se sabe que los miembros de esta familia además de regular la represión catabólica y la respuesta a anaerobiosis, regulan otro tipo de procesos como metabolismo de aminoácidos, plegamiento de proteínas, producción de toxinas y *pilli*, respuestas a estrés oxidativo o por oxido Nítrico (NO), fijación de nitrógeno, fotosíntesis y factores de virulencia (Korner, Sofia et al. 2003). Son subdivididos en varios subgrupos: **ArcR**, **CooA**, **CprK**, **Dnr**, **FixK**, **Flp**, **FnrN**, **MalR**, **NnrR**, **NtcA**, **PrfA** y **YeiL**, propuestos por Körner *et. al* 2003 (Korner, Sofia et al. 2003). Para regular a los diferentes procesos los TFs de la familia reconocen diferentes ligandos (figura 4); sin embargo, aún no se caracterizan por completo estos ligandos a los que responde cada subgrupo, ni que determina la

---

[3] *Antecedentes de los estudios de la familia Crp/Fnr.*
A principios de los 70's se estudiaba el primer sistema de control positivo de la regulación, ***CRP*** *(**c**AMP **R**eceptor **P**rotein) de E. coli. CRP se estudiaba como un FT activador único en su tipo (Zubay et. al 1970) y como el responsable de la "represión catabólica". Al mismo tiempo se estudiaba al FT* ***FNR*** *(**F**umarate and **N**itrate-reductase **R**egulator) de E. coli como el primer sistema de control de respuesta a cambios en la concentración de oxígeno (Lambden and Guestl 1976). Años después (1983), se descubrió la homología de CRP con FNR (como parálogos) (Shaw, Rice et. al 1983), lo que derivó en considerarlos como una sola familia,* ***CRP/FNR****.*

especificidad para su reconocimiento.

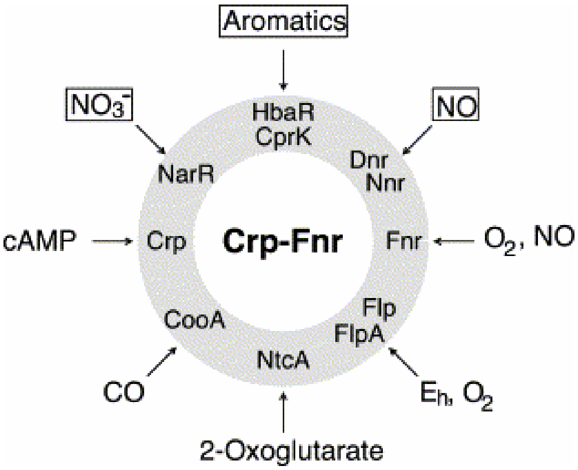


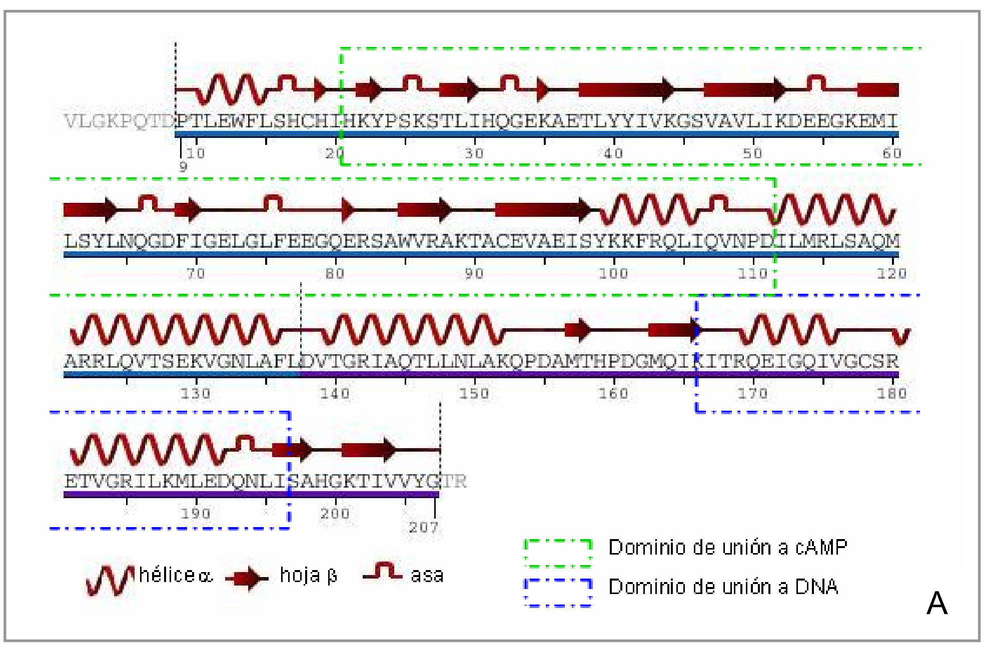
Figura 4. **Ligandos conocidos para los subgrupos de la familia Crp/Fnr**, definidos por Körner *et. al* 2003 (Figura tomada del mismo artículo).

Las características más relevantes de la familia Crp/Fnr son que:

- las proteínas son de tamaños similares (entre 230 – 250 aminoácidos) (Korner, Sofia et al. 2003);
- su DBD está ubicado en el carboxilo-terminal, es de aproximadamente 60 aminoácidos (a.a.) y su motivo es tipo hélice-vuelta-hélice alado, wHTH (wing-HTH, una variante de HTH un poco más abierto en la vuelta con lo que las hélices están un poco más separadas) (Perez-Rueda, Collado-Vides et al. 2004);
- su DOI está ubicado en la región amino, es de aproximadamente 170 a.a. y está clasificado estructuralmente como “cNMP-binding domain” (dominio de unión a un nucleótido mono-fosfatado cíclico) según Pfam (Sonnhammer, Eddy et al. 1998), aunque cada subgrupo reconoce diferentes ligandos (Korner, Sofia et al. 2003);
- tienden a reconocer operadores localizados entre las posiciones -40 o -44 arriba del inicio del gene; estos operadores son cajas palindrómicas perfectas o casi-perfectas de 5 o 6 nucleótidos separados por 4 u 8 nucleótidos; y los operadores aunque son de secuencias similares varían de acuerdo al FT que lo reconoce (Korner, Sofia et al. 2003).

a) Especificaciones estructurales de Crp

Crp de *E. coli* es una proteína de 210 aminoácidos, conformada por un DBD de tipo wHTH en la región de los residuos 166 al 197 y un dominio para unir AMPc (Adenina Mono-Fosfato cíclico) en la región de los residuos 21 al 112 (fig. 4a). La estructura secundaria de Crp combina α-hélices y hojas β-plegadas intercaladas (fig. 4.A). En la estructura terciaria se observa la región amino con una cavidad en el centro donde entra el AMPc y la región carboxilo con las dos α-hélices del HTH que contactan al DNA (fig. 4.C) (Lawson, Swigon et al. 2004). Crp normalmente se encuentra formando dimeros y sólo cuando tiene unida una molécula de AMPc en cada subunidad se modifica su conformación para unirse al DNA y llevar a cabo la regulación transcripcional (Garges and Adhya 1988) (fig. 4.B). A la fecha se han identificado varios residuos involucrados de diferentes formas en la función de la proteína, por mutagénesis dirigida (seleccionados por observaciones posteriores a la obtención y análisis de la estructura de Crp): Lys52 (K52), Asp53 (D53), Ser62 (S62), Glu72 (E72), Arg82 (R82), Tir99 (Y99), Leu124 (L124), Treo127 (T127) y Ser128 (S128) (Weber and Steitz 1987; Heyduk, Heyduk et al. 1992; Belduz, Lee et al. 1993; Lee, Glasgow et al. 1994; Baker, Tomlinson et al. 2001; Lin, Kovac et al. 2002; Gekko, Obu et al. 2004; Lawson, Swigon et al. 2004). En la figura 4.C se indican estos residuos dentro de la estructura de la proteína y su función se describe en Resultados.



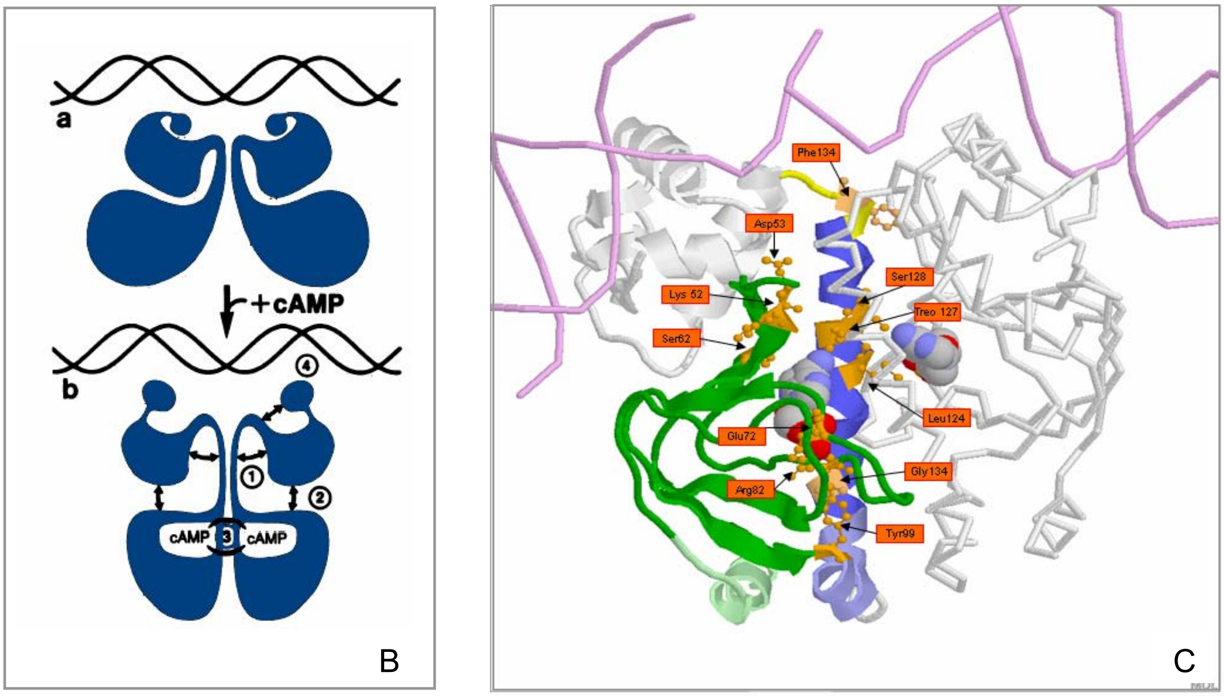
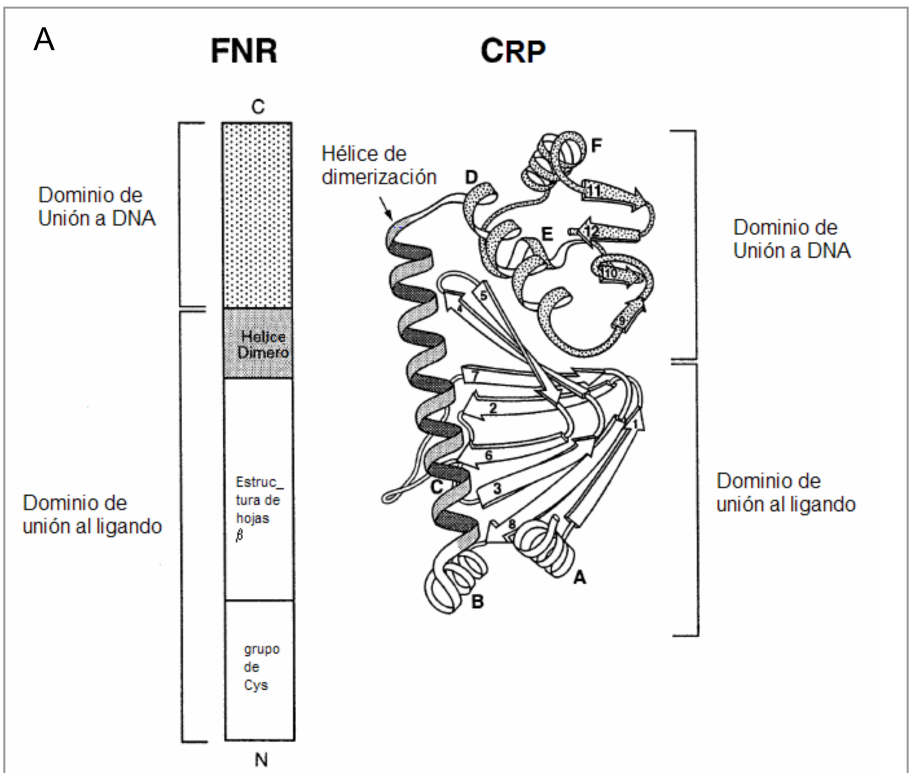


Figura 4. A) **Secuencia y estructura secundaria de CRP de *E. coli***, se indican las hélices α, hojas β y asas, junto con los dominios de unión a AMPc y a DNA (obtenida del Protein Data Bank, en la sección de "sequence details"). B) **Representación del cambio conformacional de CRP al unirse a AMPc**, resultado de esta unión se llevan a cabo 4 contactos con aminoácidos en diversas regiones de los dominios DUL y DBD, facilitando así la unión al DNA (figura tomada de Garge et. al 1988). C) **Estructura tridimensional de CRP** con los residuos probados experimentalmente con mutagénesis: Lys52 (K52), Asp53 (D53), Ser62 (S62), Glu72 (E72), Arg82 (R82), Tir99 (Y99), Leu124 (L124), Treo127 (T127) y Ser128 (S128). Cadena A monómero de Crp indicando: región N-terminal → verde claro, dominio de unión al ligando → verde, a-hélice de interacción inter-subunidades → azul, región bisagra → amarillo, dominio de unión al DNA →blanco; cadena B Crp en blanco; cadenas de DNA en lila. (figura generada con el software *Protein Explorer 2.8* [www.proteinexplorer.org]).

b) Especificaciones estructurales de Fnr

Fnr de *E. coli* es una proteína de 250 aminoácidos, conformada por al menos dos dominios estructurales: un DBD y uno de unión al ligando tipo cNMP. Fnr en vez de unir algún cNMP, más bien interacciona con el grupo prostético Hierro-Azufre 4Fe4S por medio de cuatro cisteínas ya identificadas (Cys22, 23, 26 y 122). Para Fnr aún no se ha obtenido el cristal de su estructura tridimensional, pero se han hecho comparaciones de su secuencia con la estructura de Crp para aproximar la ubicación las diferentes regiones de los dominios de la proteína (figura 5.A) (Kiley and Beinert 1999). En este trabajo se predijo la estructura secundaria de Fnr con *PsiPred 2.5* para poder ubicar espacialmente sus residuos funcionalmente importantes (Jones 1999; McGuffin, Bryson et al. 2000) (fig. 5.B). Se propone que Fnr sigue un arreglo estructural muy similar al de Crp, pero en la parte Amino-terminal FNR presenta una región adicional de ~35 a.a (figura 5.A y 5.B). Esta región es importante porque justo ahí se encuentran tres de las cuatro cisteínas necesarias para interaccionar con su grupo prostético 4Fe4S. La estructura de esta región se definió con estructura secundaria en forma de 'asa' con un bajo nivel de confianza (fig. 5.B), así que es muy probable que se acomode de forma distinta (pero para definir correctamente esta región hace falta información estructural de proteínas similares a su tipo).



P.J. Kiley, H. Beinert / FEMS Microbiology Reviews 22 (1999) 341–352

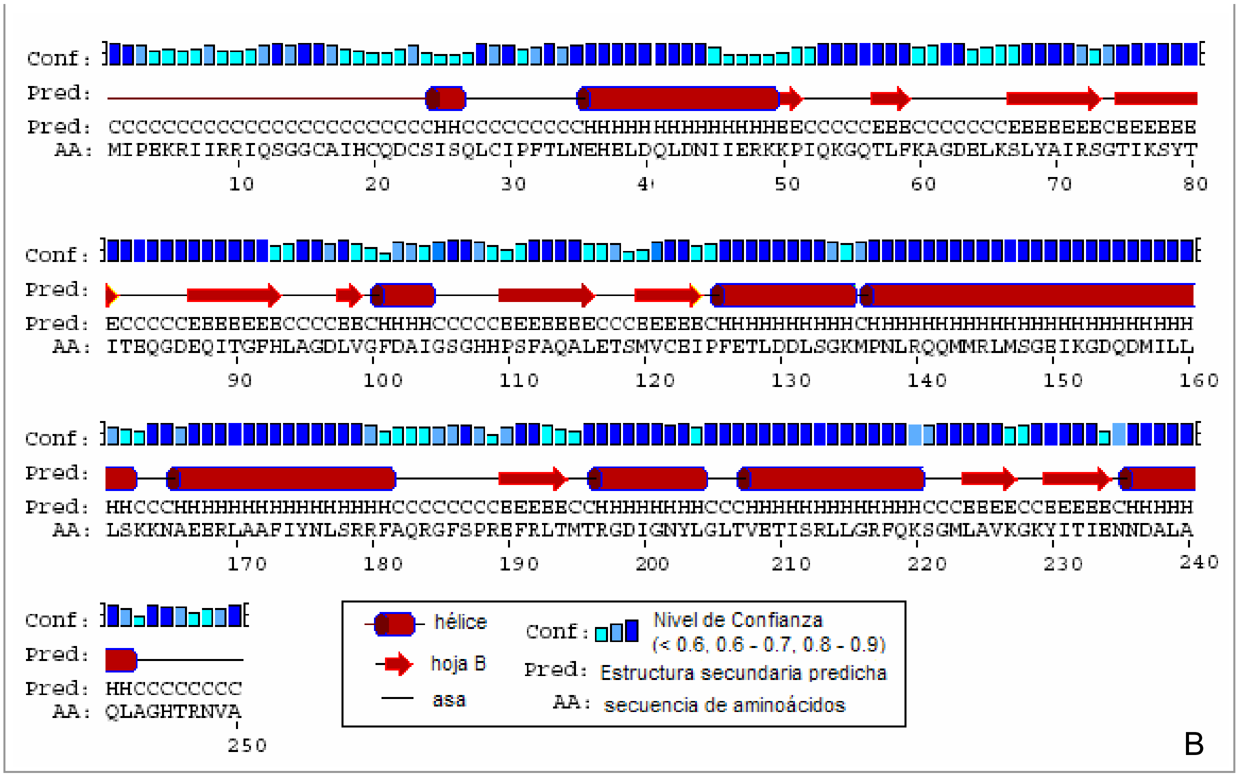


Figura 5. A) **Comparación esquemática de Fnr con la estructura de Crp.** En el esquema de Fnr se anotan las diferentes regiones con su correspondencia en la estructura de Crp en diferentes tonos de gris y blanco. Figura extraida de Kiley *et. al* 1999, sólo se tradujo el texto correspondiente. B) **Predicción de estructura secundaria de FNR de *E. coli***, realizada con PsiPred. En filas, la posición en la secuencia, la predicción de estructura y la secuencia de la proteína. El nivel de confianza está en barras azules: 0.9 – 0.8 azul fuerte, 0.7 – 0.6 azul y > 0.6 azul claro; el tipo de estructura: asa (C), hoja β (E) y α hélice (H).

El grupo prostético 4Fe4S siempre está unido a cada subunidad de Fnr provocando que Fnr se acople/desacople al DNA según su oxidación-reducción en respuesta a los niveles de Oxígeno (Lazazzera, Beinert et al. 1996; Khoroshilova, Popescu et al. 1997). Cuando aumenta mucho la concentración de oxígeno, el grupo prostético se desacopla de Fnr, y en cuanto la célula detecta esto degrada a estas moléculas de Fnr-apo (Fnr sin 4Fe4S) (Dibden and Green 2005). Fnr sólo se acopla como dímero cuando se une al DNA, el resto del tiempo está como monómero (Kiley and Beinert 1999).

Con estos dos casos representativos de la familia, se expone que aunque sean de la misma familia e interaccionen con el DNA para regular la transcripción, estos factores reconocen ligandos diferentes, los residuos que intervienen en las interacciones, transmisión de señales, reconocimiento del ligando, acoplamiento con el DNA, etc. son distintos; así como también sus estados conformacionales son diferentes (uno puede estar como monómero, mientras que el otro está como dímero).

Por otra parte, se ha explicado también que cada subgrupo interacciona con ligandos diferentes y regulan procesos distintos (Korner, Sofia et al. 2003).

Como quedará explicado en la siguiente sección, los residuos que intervienen en las diferentes funciones y roles estructurales de los dominios están mucho más conservados a lo largo de la evolución respecto al resto de los residuos.

1.4 Análisis de residuos funcionalmente importantes

El proceso de caracterización de las proteínas, que comienza con los análisis de secuencia y conformación estructural, se sigue con la determinación de la función.

La función y la conformación de las proteínas son dos características que se ha visto que están directamente correlacionadas (James and Tawfik 2003). Si la proteína no se conforma de forma específica, no logra ser funcional. A lo largo de la evolución se ha mantenido la selección de ambas características conjuntamente (Saito, Sasai et al. 1997). En una aproximación teórica, una proteína funcional cumple dos características mínimas que se mantienen a lo largo de la evolución:

a) Sigue el principio de “capacidad de plegamiento” (foldability) que es aquel que sigue la proteína para conformarse espacialmente con el mínimo de energía (Saito, Sasai et al. 1997; James and Tawfik 2003), dicha propiedad es estable ante mutaciones, así como a variaciones ambientales (temperatura, solvente, etc.) (Mirny and Shakhnovich 1999).

b) Tiene un “núcleo de plegamiento” que está constituido por residuos que conforman un grupo espacialmente contiguo en la estructura (esto no implica que estén contiguos en la secuencia) (Mirny and Shakhnovich 1999). Este núcleo(s) es el primero en conformarse cuando la proteína se empieza a plegar; posterior a estos puntos espaciales, el resto de la proteína se pliega gastando el mínimo de energía (Mirny and Shakhnovich 1999).

Gran parte de los residuos que se conocen como ‘Residuos Funcionalmente Importantes’ (RFI) son los que generan y mantienen estas características (Mirny and Shakhnovich 1999). La diversidad y variación de estos RFIs derivan en funciones y acomodos conformacionales diferentes. Actualmente, hay dos estrategias que se siguen para lograr determinar a estos RFI, una experimental y otra computacional.

a) Análisis experimental. Los estudios se realizan experimental e individualmente para cada proteína que se requiere conocer a sus RFIs. Esta estrategia consiste en realizar ‘análisis mutagénicos’ (Hutchison, Phillips et al. 1978) que implican mutar los residuos que se ubican en el sitio de unión, de los que se tiene idea que pueden participar en la interacción o de cada uno de los residuos de la proteína, y después se evalúa como funciona la proteína cambiada, y de acuerdo a ello, se determina la relevancia del residuo mutado para la función (Hutchison, Phillips et al. 1978). (Mucha de esta información sobre mutaciones hechas en proteínas y los efectos que tienen en la función y/o estructura de las proteínas se encuentra disponible en la PMD, Protein Mutant Database, (Kawabata, Ota et al. 1999)).

b) Análisis computacional. Los RFIs se pueden identificar computacionalmente en familias de proteínas que sean homólogas[4] (identificables por su similitud de secuencia) o que tengan un mismo plegamiento (identificables por similitud estructural), debido a que la evolución de las funciones y conformaciones de las proteínas sigue ciertos comportamientos evolutivos. Por una parte, a las proteínas homólogas es que proviniendo de un ancestro común, aunque la similitud de sus secuencias sea de sólo 25% su estructura varía muy poco, dentro de 2Å (Chothia and Lesk 1986). Por otra, a las proteínas con el mismo plegamiento, pueden tener un "núcleo de plegamiento" similar (Mirny and Shakhnovich 1999). En cuanto a los RFIs, la conservación de los residuos tiende a correlacionar directamente con su importancia para la función (Zvelebil, Barton et al. 1987), y/o mantenimiento de la conformación estructural; pero pueden también conservarse por razones históricas (falta de tiempo para divergir) (Mirny and Shakhnovich 1999). Así que evaluando la conservación de los residuos y analizando su posible funcionalidad con diferentes medidas se pueden identificar a los RFIs en grupos de proteínas homólogas o con el mismo plegamiento, de las que se tiene información estructural de alguna(s) de ellas.

La estrategia computacional es la más utilizada, ya que es rápida y se puede realizar para muchas proteínas a la vez; aunque tiene algunos inconvenientes: a) depende inevitablemente de la diversidad y del número de secuencias con que se trabaje y b) del tipo de mediciones que se hagan para determinar la conservación/variabilidad de los residuos (Mirny and Shakhnovich 1999).

---

[4] *Homología se define como la relación de dos caracteres cualquiera que han descendido, usualmente por divergencia, de un caracter ancestral común (Fitch 2000).*
Ortología se define como la relación de dos caracteres homólogos posterior a la divergencia de su ancestro común, donde estos caracteres se encuentran en diferentes organismos que divergieron del ancestro común (Fitch 2000).
Paralogía se define como la relación de dos caracteres homólogos cualquiera que surgen de una duplicación y usualmente han permanecido en un mismo organismo (Fitch 2000).

c) Comparación de métodos computacionales

Los métodos computacionales siguen varios principios comunes, pero cada uno varía en ciertas consideraciones. Los que se han considerado más relevantes se explican en el siguiente cuadro:

Autores de diferentes métodos	Hipótesis	Criterios para identificación de residuos funcionales	Medidas de funcionalidad / especificidad de los residuos
Método de Traza Evolutiva (Lichtarge, Yao et al. 2003)	Las funciones comunes se retienen o evolucionan de una misma proteína ancestral, a las descendientes con cambios en o alrededor del mismo sitio estructural. Los cambios a nuevas funciones se observan como cambios de residuos que están implicados en la función.	1) Alineamiento múltiple de secuencias de una familia. 2) Construcción de árbol por sustituciones que separa diferencia de secuencias por porcentajes	La conservación es la medida del tipo y/o posición del residuo en el grupo de secuencias alineadas. Esta se observa como la presencia o ausencia del residuo.
(Johnson and Church 2000)	Una familia de proteínas con el mismo esqueleto (backbone) estructural que reconoce ligandos distintos, tiene residuos variantes alrededor de la cavidad del ligando. Estos residuos son críticos para el reconocimiento específico de cada ligando.	1) Familia de proteínas con el mismo *backbone* estructural. 2) Alineamientos múltiples de secuencias y de estructuras	Ubicación del residuo alrededor de la cavidad, con una distancia menor a 4.5 Å del ligando.
(Mirny and Shakhnovich 1999), (Mirny and Shakhnovich 1999)	Se asume que las secuencias parálogas y ortólogas[4] realizan las mismas funciones bioquímicas de forma general, pero los ortólogos tienen la misma especificidad. Proteínas evolutivamente relacionadas tienden a conservar el mismo plegamiento.	1) Alineamiento múltiple de secuencias, donde el porcentaje de identidad agrupa ortólogos y separa parálogos 2) Medición del Contenido Informacional de cada posición del alineamiento	Las posiciones con Contenido Informacional significativo estadísticamente corresponden a residuos funcionalmente importantes (ver en métodos)

De los métodos descritos en el cuadro anterior, el de Traza Evolutiva ha sido utilizado ampliamente en otros trabajos para identificar RFIs. En cada caso, las variantes del método están dadas en la manera de evaluar la significancia de los RFIs encontrados (Landgraf, Fischer et al. 1999; Innis, Shi et al. 2000; Armon, Graur et al. 2001; Landgraf, Xenarios et al. 2001; Lichtarge, Yao et al. 2003; Berezin, Glaser et al. 2004). Este método se limita a sólo identificar a los RFIs que interaccionan con los ligandos, además que estas proteínas deben reconocer a sus ligandos en las mismas regiones o sitios activos.
Alternativamente, el método de Johnson y Church (2000), es útil sólo para casos en que la familia de proteínas sigue exactamente el mismo plegamiento, que interaccionen con su ligando en el mismo sitio, tan sólo variando el ligando que reconocen. Igualmente, este método se centra sólo en la identificación de RFIs que interaccionan con el ligando a partir de proteínas con información estructural y/o experimental.
Por último, el método de Mirny y Gelfand (2002) propone que la identificación de RFIs no se debe centrar solo en aquellos residuos que unan a un ligando; además, que los RFIs no tienen que ubicarse en los mismos sitios para todos los miembros de la familia, es decir pueden localizarse en diferentes regiones del alineamiento, no se limita a que las proteínas sigan exactamente el mismo plegamiento, y la evaluación de los RFIs encontrados no depende de observaciones sobre una estructura o su distancia al ligando, sino de su significancia estadística.

Dado que los miembros de la familia Crp/Fnr reconocen diferentes ligandos y no lo unen en una misma región, se decidió implementar el método de Mirny y Gelfand 2002 para identificar los RFIs que podrian interaccionar con un potencial ligando.

2 Hipótesis

Las diferencias funcionales dentro de una familia de factores transcripcionales parálogos pueden identificarse computacionalmente como residuos específicos en diferentes posiciones del dominio de unión al ligando.

3 Material y métodos

3.1 Colección de secuencias y estructura tridimensional

En este trabajo se decidió utilizar la clasificación de familias de factores transcripcionales (FTs) de Pérez-Rueda *et. al* 2000, ya que la información disponible de estos es principalmente de secuencias completas y sus dominios.

Todas las secuencias de las proteínas usadas en este trabajo para la familia Crp/Fnr, se extrajeron de la base de datos Genebank del National Center for Biotechnology Information (NCBI) [http://www.ncbi.nlm.nih.gov/entrez/ (sección 'protein')] hasta el 2005. Estas secuencias pueden ser consultadas utilizando los identificadores GI (gene identifier) de esta misma base de datos (ver Anexo 1: lista de secuencias utilizadas).

La estructura tridimensional que se utilizó como referencia para la familia fue la de Crp de *E. coli* obtenida de la base de datos Protein Data Bank (PDB), cuyo identificador es 1J59 ([http://www.rcsb.org/pdb/] Bermann et. al 2000).

3.2 Detección de los miembros de la familia Crp/Fnr

Se obtuvieron 276 proteínas homólogas a las secuencias de Crp y Fnr de *E. coli* de la base de datos SwissProt ([http://expasy.org/sprot/] (Boeckmann, Bairoch et al. 2003) utilizando el programa *blastp* (E-value 10^{-6}) (Altschul, Gish et al. 1990). De los candidatos de la búsqueda anterior, todas aquellas secuencias que tuvieran el mismo dominio de unión al DNA de tipo Helix-Turn-Helix (SSF46785) y el dominio de unión al ligando cNMP (SSF51206), según la base de datos SuperFamily (Gough 2002), fueron considerados homólogos (ortólogos y parálogos) pertenecientes a la familia CRP/FNR. Se encontraron y eliminaron 9 secuencias que no cumplieron con estos requisitos mínimos.

Con todas estas secuencias se generó el alineamiento múltiple (multiple sequence alignment, MSA) de la familia con el programa *clustalx* (Thompson, Gibson et al. 1997). Con el alineamiento obtenido se generó su perfil de secuencia con una matriz de peso utilizando los programas *hmmbuild* y *hmmcalibrate* del paquete *HMMer 2.3* (Eddy

1998). Posteriormente, se mapeo el perfil de la familia Crp/Fnr a través de 200 genomas de bacterias y arqueas utilizando el programa *hmmsearch* de *HMMer 2.3* (E-value $\leq 10^{-3}$) (Eddy 1998). De esta búsqueda se obtuvieron 171 secuencias (Anexo 7.1: Tabla de secuencias).

Como el propósito de este trabajo fue estudiar el dominio de unión al ligando (DUL), se eliminó del alineamiento el dominio de unión al DNA de la región carboxilo y se dejo sólo el DUL de la región amino, de aproximadamente 150 residuos. Posteriormente, se eliminaron las secuencias del DUL que eran 100% idénticas entre ellas para evitar redundancia en el grupo de datos estudiados, utilizando el programa *CD-HIT* (Lin, Kovac et al. 2002); se eliminaron 37 secuencias. Con las 134 secuencias del DUL homólogas y no redundantes, finalmente se volvió a hacer el MSA con *clustalx* (valores por default, ver Anexo 7.2: alineamiento de secuencias estudiadas).

3.3 Análisis evolutivo para la división en subgrupos

A las 134 secuencias del DUL se les hizo un análisis evolutivo para conocer su agrupamiento en subgrupos dentro de la familia utilizando el software para cálculos filogenéticos y graficación *MEGA* (Kumar, Tamura et al. 2004). Se seleccionó el modelo evolutivo para generar el árbol con el software *ProtTest* (REF). Se construyó su árbol filogenético siguiendo el método de Neighbor Joining utilizando el modelo JTT (parámetro α = 2.0) (Jones 1999) y se sometió a 100 replicas de bootstrap y una prueba de consensus para verificar que los agrupamientos fueran consistentes (ver Resultados: figura 1R, Anexo 7.3 : gráfica del árbol de la familia).

3.4 Identificación de residuos funcionalmente importantes

Ya que los residuos funcionalmente importantes (RFI) se comparten entre ortólogos (misma especificidad), pero no con parálogos (diferente especificidad), estos RFIs se pueden diferenciar entre estos grupos. En un alineamiento de secuencias de proteínas homólogas, las columnas donde se ubican estos residuos (RFI) tienen un *contenido informacional (CI)* diferenciable al del resto de las columnas. El CI es una medida que permite la asociación de la conservación de los residuos con la especificidad de la proteína (Mirny and Shakhnovich 1999).

Por lo tanto, la identificación de RFI se realizó calculando el CI de cada columna del alineamiento de la familia Crp/Fnr con la siguiente fórmula:

$$CI_i = \sum_{\substack{x=1,\ldots,20 \\ y=1,\ldots,Y}} f_i(x,y) \log \left(f_i(x,y) / f_i(x)\, f(y) \right) \qquad (1)$$

Donde,
x = *cada uno de los 20 aminoácidos (A,C,D,G,…,W,Y)*
y = *número de grupos en que se dividen las secuencias homólogas*
$f_i(x)$ = *frecuencia del residuo tipo x en la posición i del alineamiento múltiple*
$f(y)$ = *fracción de proteínas que pertenecen al grupo y*
$f_i(x,y)$ = *frecuencia del residuo x en la posición i dentro del grupo y*
$f_i(x,y) / f_i(x)\, f(y)$ = *razón de la frecuencia de un residuo en el grupo respecto al alineamiento multiple, considerando la cantidad de proteínas en el grupo*

Este cálculo se hizo con un programa en el lenguaje de programación *PERL*: ContInfo-aa-alignment-PRO.pl (Anexo 7.6: códigos de programas). El cálculo de CI de las columnas del alineamiento con *gaps* se consideró igual a cero ($CI_{gap} = 0$).

Dado que el *contenido informacional* se puede sesgar por una muestra pequeña o por la composición de aminoácidos (Mirny and Shakhnovich 1999), no se puede tomar sólo este valor para identificar a los determinantes de especificidad; sino que se debe calcular su significancia estadística (descrita abajo) y utilizar esta medida junto con el valor de CI-*reales* (los obtenidos para las secuencias de la familia Crp/Fnr) para predecir los RFIs.

3.5 Cálculo de significancia estadística

Para conocer el valor estadístico de los CI-*reales* de cada posición de los grupos del alineamiento, se utilizó la estrategia de comparación de cada valor de CI-*reales* contra CI-*referencias* de una población mayor de secuencias (que sigue una distribución Normal, Anexo 7.4), obteniendo su valor Z. Posteriormente, los CI-*reales* que tuvieron un valor Z ≥ 3 con un área ≤ 0.05 se consideraron *CI-reales* que corresponden a residuos funcionalmente importantes.

La población utilizada fue una generada con secuencias “pseudoaleatorias” que tienen la misma variabilidad intra-grupos (llamada aquí ρ) que la de los grupos del alineamiento (ver explicación en 3.5.b y discusión en resultados 4.3).

a) Obtención de secuencias pseudos-aleatorias

La generación de secuencias pseudoaleatorias, con variabilidad ρ (ver en punto 3.5.b) se hizo siguiendo los puntos 1 y 2 (explicados enseguida abajo) utilizando el programa en PERL Calc-seqs-PseudoRandom.pl (Anexo 7.6: códigos de programas). Se generaron igual número de secuencias como miembros de cada subgrupo y el proceso se repitió para la familia completa 10^3 veces.

1) Generación de secuencias parentales para cada subgrupo: Para cada grupo, se considera la distribución de aminoácidos en cada columna (que es $f_i(x)$) y con probabilidad P ($f_i(x)$) se hereda alguno de esos aminoácidos.

2) Generación de secuencias pseudo-aleatorias: Se toma la secuencia parental de cada grupo y la $f_i(x)$ de cada columna para ese grupo. De acuerdo con la probabilidad de ρ las secuencias pseudo-aleatorias heredan el mismo a.a. de la secuencia parental o cambian por algún otro a.a. de la $f_i(x)$ para esa columna.

b) Variabilidad intra-grupos (ρ)

La variabilidad intra-grupos (ρ) se utiliza para determinar la similitud dentro de los grupos de las secuencias pseudos-aleatorias y así dar una medida umbral para la herencia o cambio de aminoácido de las secuencias parentales a las nuevas secuencias pseudos-aleatorias. Si ρ = 1 implica que todas las secuencias generadas son idénticas, por el contrario, si ρ = 0 las secuencias son tan distintas como parálogos. La ρ se obtuvo posterior a la comparación de identidades por pares cada secuencia del alineamiento, utilizando el programa en PERL Calculo_rho.pl (Anexo 7.6: códigos de programas), y luego de analizarlas por agrupamientos (clustering) a través del Centroit Linkage Clustering, con una métrica de distancia "uncentered correlation" del software Cluster 3.0 (Eisen 1998) y con el programa en PERL ColorSeqRelation.pl (Anexo 7.6: códigos de programa) (ver figura 4R en Resultados). Para la familia de secuencias de Crp/Fnr se determinó utilizar una ρ = 0.75 para obtener una similitud intra-grupo cercana a las proteínas conocidas.

c) Obtención de los valores Z

Los valores Z de los *CI-reales* para cada posición del alineamiento se obtuvieron con el programa en PERL MedsStats.pl (Anexo 7.6: códigos de programas) utilizando la fórmula:

$$Z = \frac{x - \mu}{\sigma} \qquad (2)$$

Donde: *x = valor CI-reales a evaluar*
μ = media de la población de CI-referencias
σ = desviación estándar de la población de CI-referencias

3.6 Caracterización de estructura secundaria y terciaria

Para la familia Crp/Fnr se tienen reportadas varias estructuras de Crp de *E. coli* en la base de datos PDB, de las cuales se seleccionó una de las que tiene una mayor resolución, la 1J59 a 2.5Å. El mapeo de los residuos se hizo utilizando el software ProteinExplorer (www.proteinexplorer.org).

4 Resultados y discusión

Bajo la hipótesis de que las diferencias funcionales dentro de una familia de factores transcripcionales homólogos (ortólogos y parálogos) pueden identificarse computacionalmente como diferentes residuos específicos en diferentes posiciones del dominio de unión al ligando, la subdivisión de la familia en grupos podría hacer más sensible la identificación de los diferentes residuos funcionalmente importantes para cada uno de los grupos.

4.1. Identificación y agrupamiento de miembros de la familia Crp/Fnr

Para el estudio de residuos funcionalmente importantes (RFIs) del dominio de unión al ligando (DUL) de la familia Crp/Fnr se comenzó por la identificación de secuencias homólogas de CRP y FNR de *Escherichia coli* (CRP_{ECO} y FNR_{ECO}) a través de 200 genomas de Bacterias y Arqueas. De las 171 secuencias encontradas, se tomaron únicamente las regiones Amino donde se encuentran los Dominios de Unión al Ligando (DUL) de aproximadamente 150 a.a. (se eliminaron las regiones carboxilo de las secuencias). Posteriormente, se eliminaron 37 secuencias que eran redundantes (idénticas en los mismos organismos). Finalmente las 134 secuencias de los DUL se alienaron y se obtuvo su árbol filogenético como se explica en Métodos.

En la figura 1R se muestra el árbol filogenético de las 134 secuencias de los DUL. Lo primero a notar es que estas 134 secuencias se separan en dos grandes grupos correspondientes a los dos parálogos que definieron incialmente a la familia: los que se diferencian con Crp (región izquierda del árbol; marcados con etiqueta naranja “CRP”) y los que se diferencian con Fnr (región derecha del árbol; marcados con etiqueta roja “FNR”). En seguida se observa que a la vez estas 134 secuencias se separan en 9 grupos (diferenciados en recuadros de diferentes colores) que representan secuencias de 4 agrupamientos filogenéticos: *Proteobacteria* (alfa-α (G0: rojo, G8: azul), beta-β (G3: verde) y gama-γ (G1: amarillo, G2: naranja), *Firmicutes* (G4: verde seco, G5: azul claro), *Actinobacteria* (G6: rosa) y *Cianobacteria* (G7: lila) (ver también Anexo 3).

Es importante hacer notar que la subfamilia Fnr posee un mayor número de miembros subdivididos en seis grupos pertenecientes a γ (G2), β (G3) y α *Proteobacteria* (G0, G8), *Firmicutes* (G5) y *Cianobacteria* (G7); mientras que los miembros de la subfamilia

CRP son menos y se dividen en tres grupos pertenecientes a *γ-Proteobacteria* (G1), *Firmicutes* (G4) y *Actinobacteria* (G6). Esto no sólo indica una mayor divergencia de los miembros de la subfamilia Fnr, sino de una mayor diversidad funcional y reconocimiento a diferentes ligandos en los organismos aquí descritos (ver Körner *et. al* 2003 que lo han reportado previamente).

Además de los ortólogos a Crp y Fnr (grupos G1 y G2 respectivamente) se identificaron algunos homólogos a FixK y YeiL, que son otros dos miembros ya caracterizados de la familia (ver Introducción) (Körner *et. al* 2003). En el grupo G0 hay 9 secuencias anotadas como FixK de *α-Proteobacteria*, y en el grupo G4 hay cinco secuencias anotadas como YeiL de *γ-Proteobacteria*. Es importante remarcar que estas secuencias de YeiL se agruparon con el grupo 4 de *Firmicutes* (cuadro punteado amarillo, figura 1R).

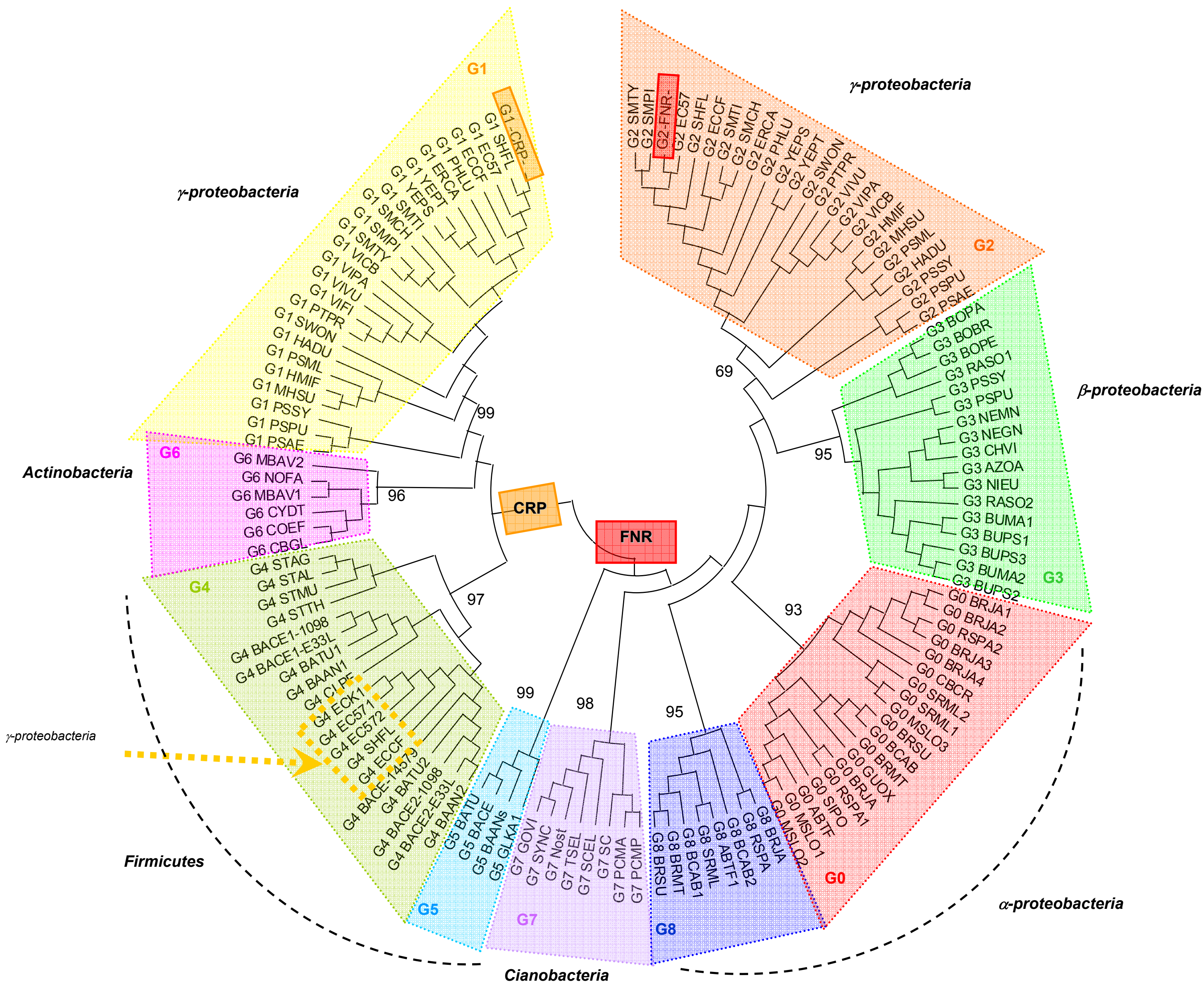


Figura 1R. **Representación circular del árbol de los 134 miembros identificados para la familia Crp/Fnr** divididos en 9 grupos, con valores bootstrap mayores a 65 (números negros en ramas que divergen a cada uno de los 9 grupos). Estos nueve grupos corresponden a cuatro grupos filogenéticos: *Proteobacteria* (α, β, γ), *Firmicutes*, *Actinobacteria* y *Cianobacteria* (nombres indicados al lado de los grupos por la parte externa del árbol circular). Los homólogos a Crp se separan en las ramas marcadas con etiqueta naranja "CRP" y se subdividen en tres grupos marcados recuadros amarillo (*Protebacteria* γ G1), rosa (Actino*bacteria* G6) y verde seco (*Firmicutes* G4). Los homólogos a Fnr se separan en las ramas marcadas con etiqueta roja "FNR" y se subdividen en seis grupos marcados en recuadros azul claro (*Firmicutes* G5), lila (*Cianobacteria* G7), azul y rojo (α-*Proteobacteria* G8 y G0), verde (β-*Proteobacteria* G3).

Con este árbol se pudo observar que el DUL ha sido suficiente para diferenciar las proteínas homólogas a la familia Crp/Fnr de acuerdo a los diferentes grupos de *Bacterias* mencionados; a la vez, esta separación de los DUL en diferentes grupos es indicativa de diferentes posibles funciones.

4.2. Identificación de residuos funcionalmente importantes

Ya que con la aproximación del 'Contenido Informacional' (CI) de los residuos de secuencias homólogas se pueden identificar posiciones en el alienamiento que tienen mayor carga de información, que implica una importancia para la función (Mirny and Shakhnovich 1999), se procedió a hacer el cálculo del CI para las ~180 posiciones del alienamiento de los DUL de los 9 grupos de la familia Crp/Fnr para identificar sus Residuos Funcionalmente Importantes (RFI). El cálculo de CI con la fórmula 1 explicada en Métodos (3.4) e implementada en el programa en el lenguaje de programación PERL 'ContInfo-aa-alignment-PRO.pl' (ver Anexo 7.6.1).

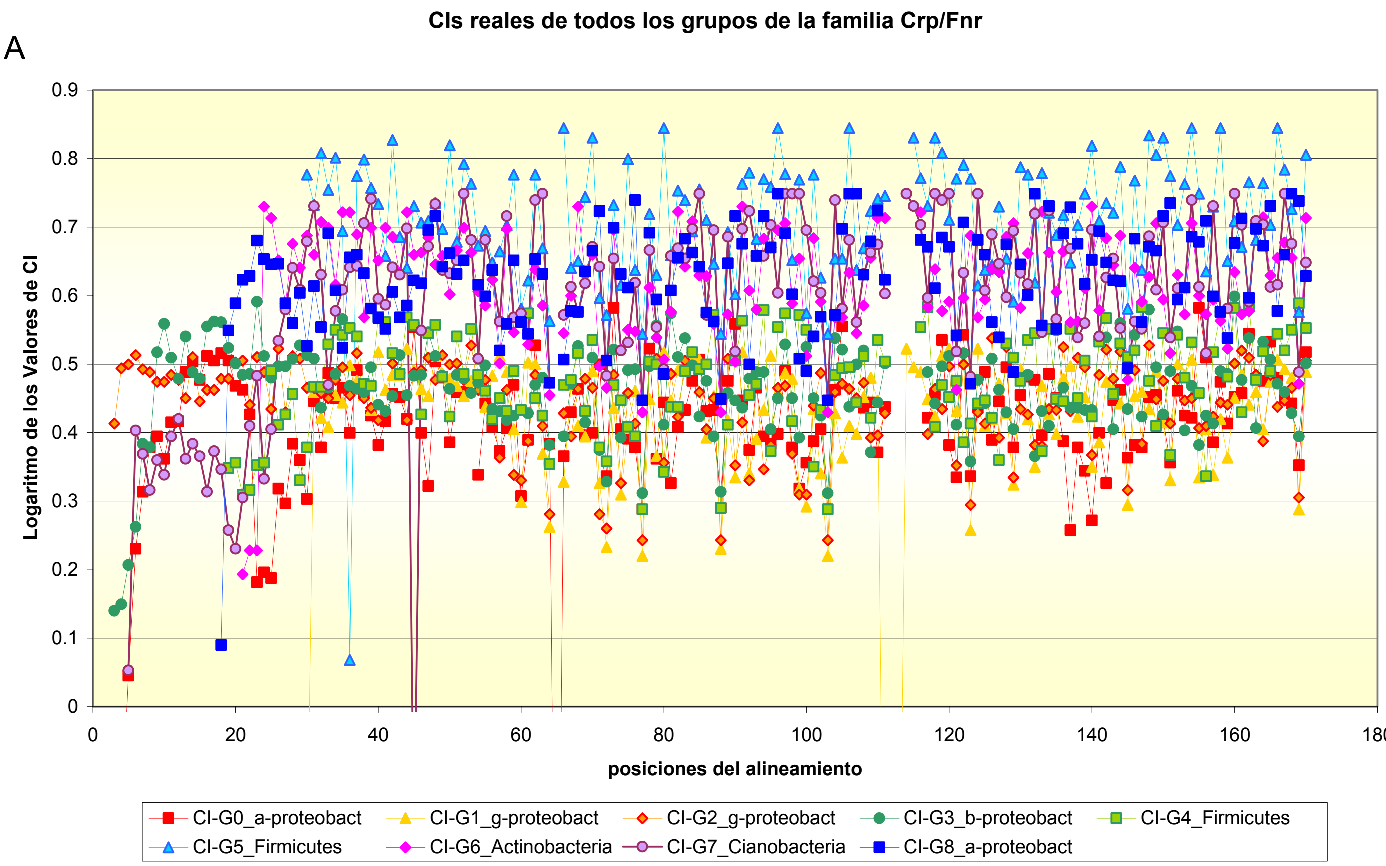
La distribución de los 9 grupos de secuencias DUL en 4 grupos filogenéticos distintos le dan diversidad (heterogeneidad) al universo de secuencias estudiadas. Por otra parte, el número de secuencias que constituyen a cada grupo no es homogéneo debido a la disponibilidad sesgada de secuencias en las bases de datos. En los grupos de *Actinobacteria* hay 9 secuencias [G6], de *Cianobacteria* 8 [G7], de *Firmicutes* 24 [G4, G5] y de *Proteobacteria* 93 (α tiene 27 [G0, G8], β tiene 17 [G3] y γ tiene 49 [G1, G2]).

Es importante notar que el grupo de *Proteobacteria* es el más representado.

En la figura 2R se grafican los valores logarítmicos del CI para cada una de las ~180 posiciones que conforman el alineamiento de secuencias DUL de los 9 grupos de la familia.

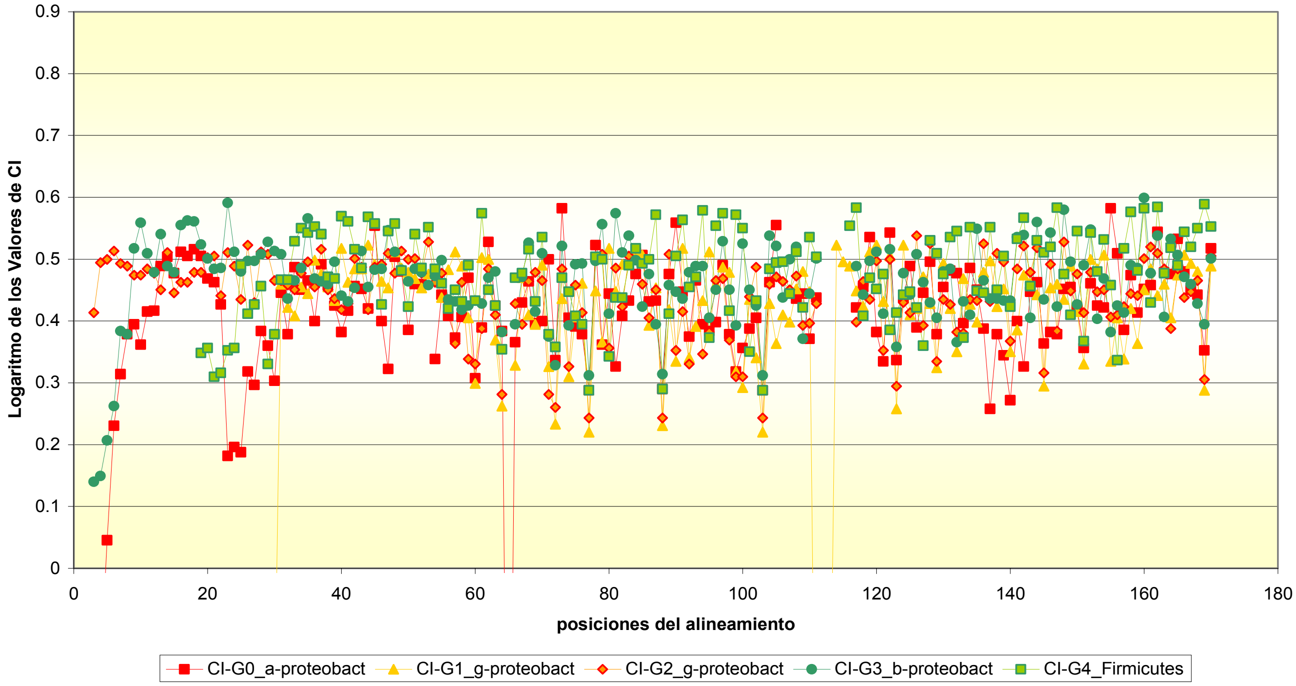
Lo primero que se puede notar es que los valores de CI se separan en dos tendencias distintas, una de valores altos y otra de valores medios (ver 2R-B y 2R-C). En la tendencia de valores altos, gráfica 2R-B, se encuentran los grupos de *Actinobacteria* [G6], *Cianobacteria* [G7], uno de *Firmicutes* [G5] y uno de α-*Proteobacteria* [G8]; que coinciden en ser grupos con menos de 10 secuencias y que tres de los cuatro grupos no son *Proteobacteria*. En la tendencia de valores medios, gráfica 2R-C, se encuentra el otro grupo de α-*Proteobacteria* [G0] junto con los de γ-*Proteobacteria* [G1, G2], β-*Proteobacteria* [G3] y otro de *Firmicutes* [G4]; que coinciden ser grupos con 17 o más secuencias y cuatro de los cinco grupos son *Proteobacteria*.

A



CIs reales de los grupos *Proteobacteria* (α, β, γ) y *Firmicutes* de la familia Crp/Fnr

B



CIs reales de los grupos de *Firmicutes, Actinobacteria, Cianobacteria* y α*-proteobacteria* de la familia Crp/Fnr

C

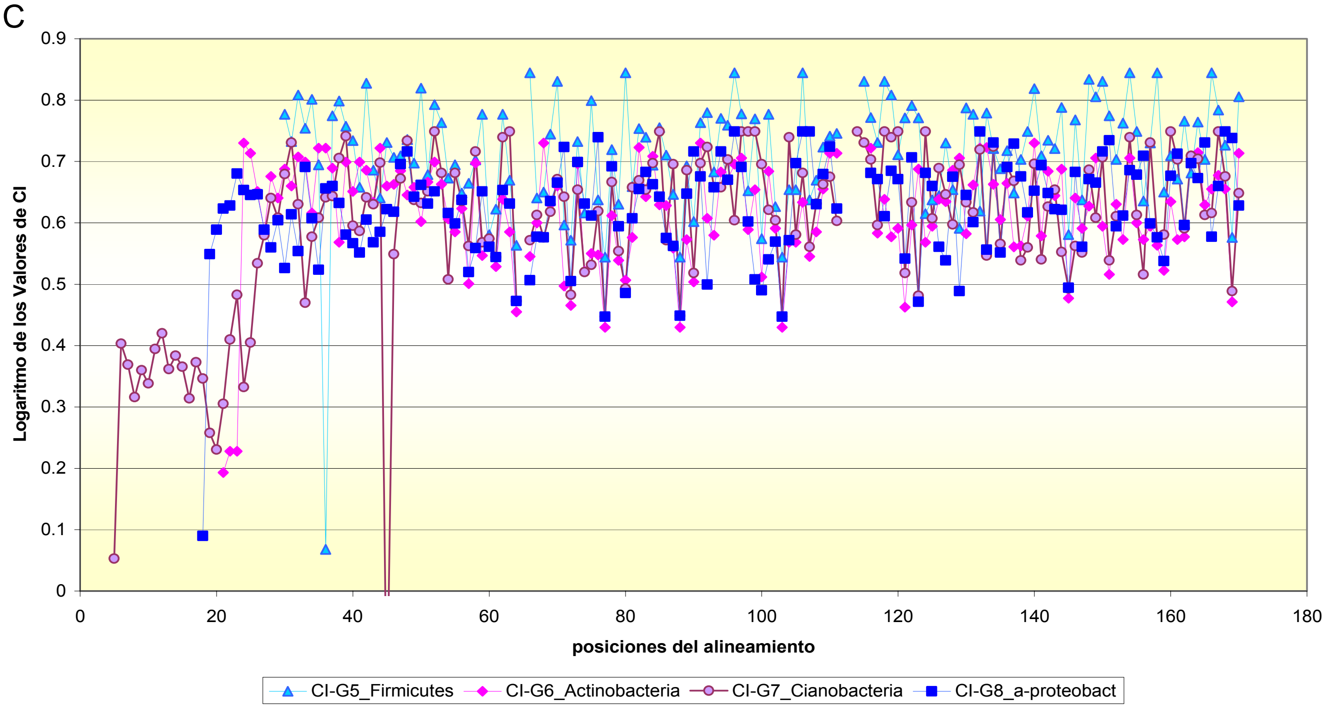


Figura 2R. A) **Logaritmos del Contenido informacional de los 9 grupos de la familia Crp/Fnr**. B) **Logaritmos del Contenido Informacional de cinco grupos con valores altos**. C) **Logaritmos del Contenido Informacional de cuatro grupos con valores medios**. Nota: El valor de CI de cada posición es discreto e independiente, pero se decidió trazar una línea que une a los puntos de cada grupo para facilitar su observación.

Dos de las propiedades de la métrica de CI (ver Métodos) son que CI es igual a 0 si y sólo si x y y son estadísticamente independientes, y de que un valor grande de CI indica una fuerte asociación entre x y y (donde x es uno de los 20 a.a. y y es un grupo dentro del alineamiento); así que al observar estas dos tendencias de CI de las secuencias DUI estudiadas en la figura 2R, se quiso conocer de cual de las dos variables (x → tipo de aminoácidos, y → número de secuencias en cada grupo) consideradas en el cálculo del CI están influyendo más para tener estas diferencias:

a) Si se debe a que el número de miembros que conforman a los grupos es diferente (CI depende más de y);

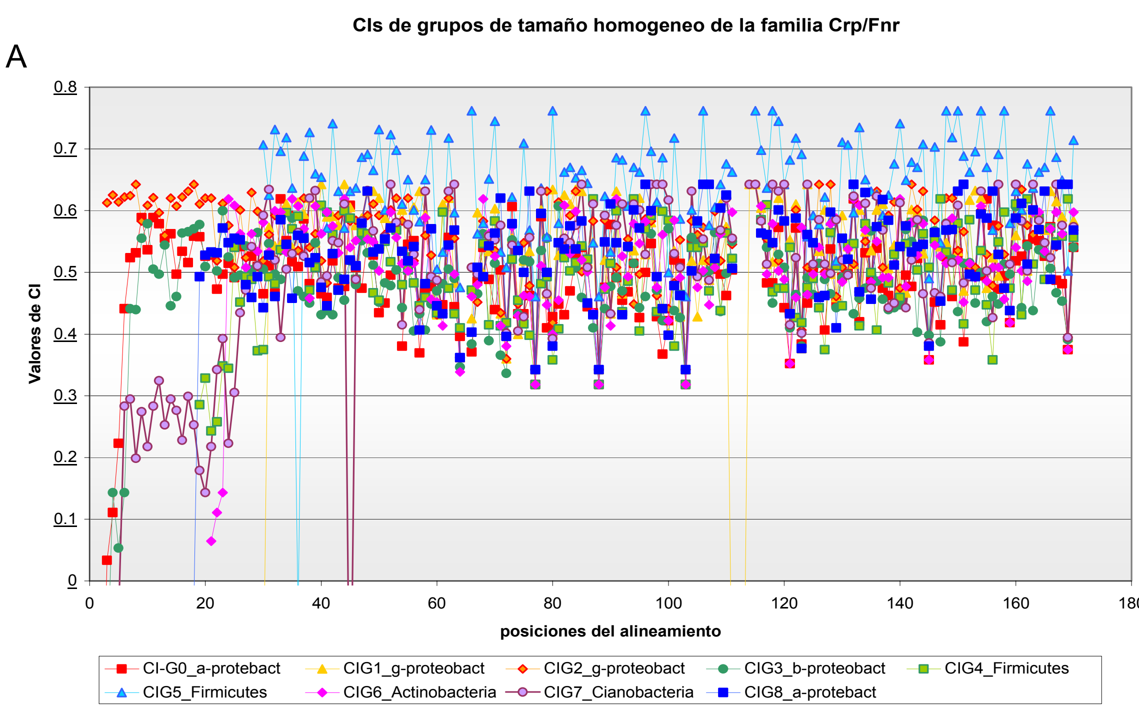
b) Si se debe a la diferente composición de aminoácidos en los grupos filogenéticos en el alineamiento (CI depende más de x).

Para la opción (a) se realizó otro análisis de CI para la misma cantidad de grupos de la familia, pero sin que su número de miembros variara (figura 3R-A). Para esto, se dejó el mismo número de miembros de los grupos que tienen 9 o menos secuencias y para los grupos con más de 17 miembros, se decidió que tuvieran al menos 9 secuencias seleccionadas al azar (ver tabla 1R). En la tabla 1R se muestra el número de secuencias que conforman a los grupos en la Familia Crp/Fnr (como están en el alineamiento) y el de los grupos con casi el mismo número de miembros seleccionados al azar.

Tabla 1R. **Número de secuencias en los grupos de la familia Crp/Fnr** como están en el alineamiento y de los grupos que se procuraron tuvieran el mismo número de miembros seleccionados al azar.

Grupo	**Número de secuencias de los grupos de la familia (observados)**	**Redistribución de miembros de grupos (azar)**
Actinobacteria G6	9	9
Cianobacteria G7	8	8
α-proteobacteria G8	8	8
Firmicutes G5	5	5
Firmicutes G4	19	9
α-proteobacteria G0	19	9
β-proteobacteria G3	17	9
γ-proteobacteria G1	25	9
γ-proteobacteria G2	24	9
Total	134	75

Para la opción (b) se realizó otro análisis de CI para la misma cantidad de grupos de la familia, pero la composición de aminoácidos de cada posición fue cambiada de forma aleatoria (figura 3R-B). El cambio de composición de aminoácidos se hizo con el programa en el lenguaje de programación PERL Shuffling-alin-PRO.pl (ver Anexo 7.6.3).



B

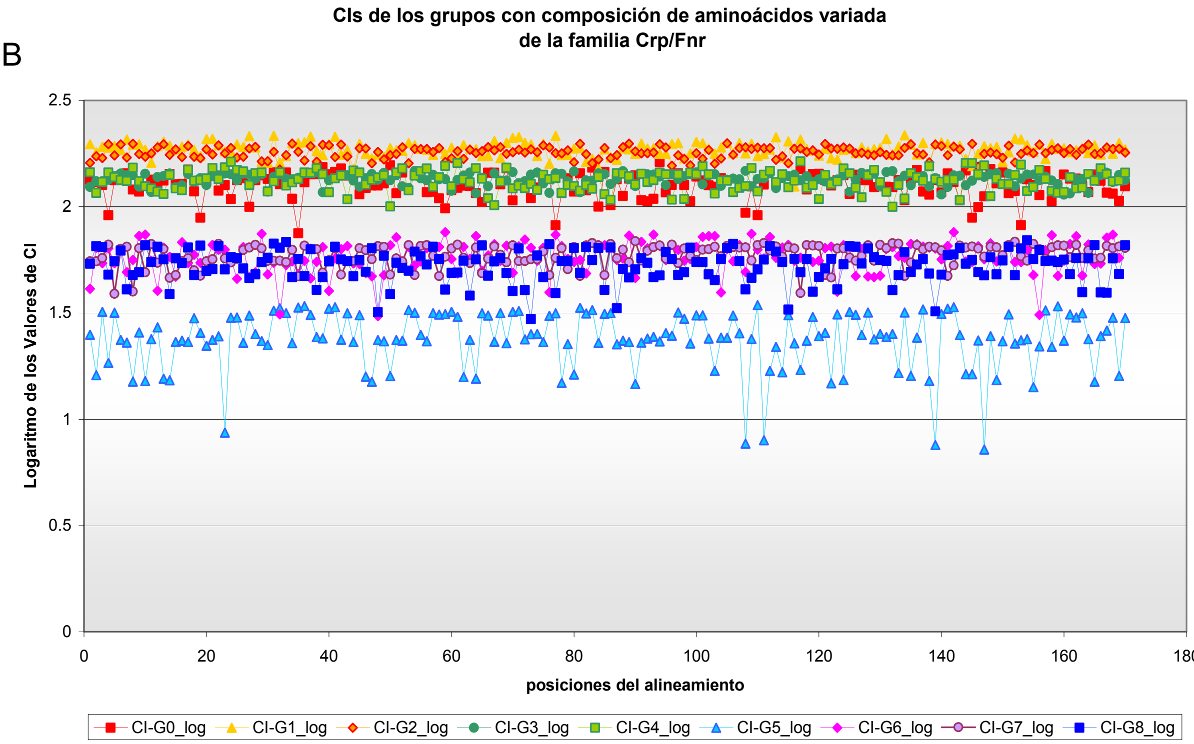


Figura 3R. A) **Logaritmos de *CI* calculados para 9 grupos con el mismo número de secuencias.** B) **Logaritmos de *CI* calculados para 9 grupos con composición de aminoácidos variada**. Nota: El valor de *CI* de cada posición es discreto e independiente, pero se decidió trazar una línea que une a los puntos de cada posición de los nueve grupos para facilitar su observación.

En la figura 3R-A, donde los grupos tienen aproximadamente el mismo número de secuencias, se observa que los valores de *CI* de todos los grupos son similares, y además que en general sus valores aumenta con una media aproximada de 0.5. Con esto se ve que los valores de *CI* dependen del número de secuencias en los grupos (*CI* depende más de y).

En la figura 3R-B, donde la composición de aminoácidos de los grupos está cambiada, se observa que aunque los valores de *CI* han cambiado, se siguen separando por el número de secuencias que tienen los diferentes grupos y no reflejan las relaciones filogenéticas de los diferentes grupos (*CI* depende más de y).

En el acomodo de las secuencias encontradas para la familia Crp/Fnr se observa claramente que el grupo 8 de α-*Proteobacteria* que sólo tiene 8 miembros se encuentra con los grupos de valores altos (figura 2R-B), mientras que el resto de los grupos de *Proteobacteria* (G0, G1, G2 y G3) que tienen al menos 17 miembros presentan valores de CI medios (figura 2R-C); y uno de los grupos de *Firmicutes* con 19 secuencias (G4) tiene un valor también medio de *CI* (figura 2R-C) y el otro grupo de *Firmicutes* con 5 secuencias (G5) tiene un valor alto (figura 2R-B).

Con estos dos ensayos se demuestra que las diferencias en los valores de *CI* dependen del número de secuencias en los grupos (y) de las secuencias estudiadas. También se muestra que los valores de *CI* no se ponderan siendo mayores o menores a causa del número de secuencias o de la composición de aminoácidos. Por lo tanto, para poder diferenciar los valores de *CI* que se relacionan con un alta conservación (funcionalidad) de los residuos se requirió hacer una evaluación estadística.

4.3. Identificación de la significancia estadística del *CI*

Ya que el valor de *CI* no es suficiente indicador por si solo de la asociación del residuo con su funcionalidad, se procedió entonces a calcular su significancia estadística comparándolo con el *CI* de secuencias que representen un universo "azaroso" de la familia Crp/Fnr. Este universo se conforma de secuencias pseudo-aleatorias que son variables pero mantienen la misma composición de aminoácidos de la familia, lo cual permite que sigan siendo secuencias similares. Si el universo de secuencias más bien se conformara de secuencias totalmente aleatorias, se estarían incluyendo secuencias cuyas funciones y RFIs son totalmente ajenas a las de la familia y no se podrían detectar *CI*s de RFIs comunes; este universo se alejaría del modelo biológico real estudiado, la familia de Crp/Fnr.

Para determinar la variabilidad de secuencias para el estudio se graficaron las identidades entre las 134 secuencias DUL de la familia Crp/Fnr (gráfica 4R). En la gráfica 4R, se muestran sobre la diagonal en rojo las identidades que indican secuencias similares de 100 a 95%, y alrededor de la diagonal se indican identidades menores: en naranja de 85 a 75%, en naranja claro de 74 a 70%, amarillo de 69-60%, y de verde-azul-negro las identidades menores de 60 hasta 0%. Con esta gráfica se puede ver que el universo para explorar las secuencias las pseudos-aleatorias se encuentran entre 60 y 90%. Para mayoría de los grupos el intervalo de 80 a 90% (naranja cercano a rojo) no existe. Tomar una $\rho \geq 0.9$ (que equivale a 90%) implica generar secuencias casi idénticas a las reales, y tomar una $\rho \leq 0.6$ implica considerar un intervalo ya demasiado alejado del modelo biológico real estudiado. Por esto, se utilizó una métrica de variabilidad intra-grupos de $\rho \leq 0.75$.

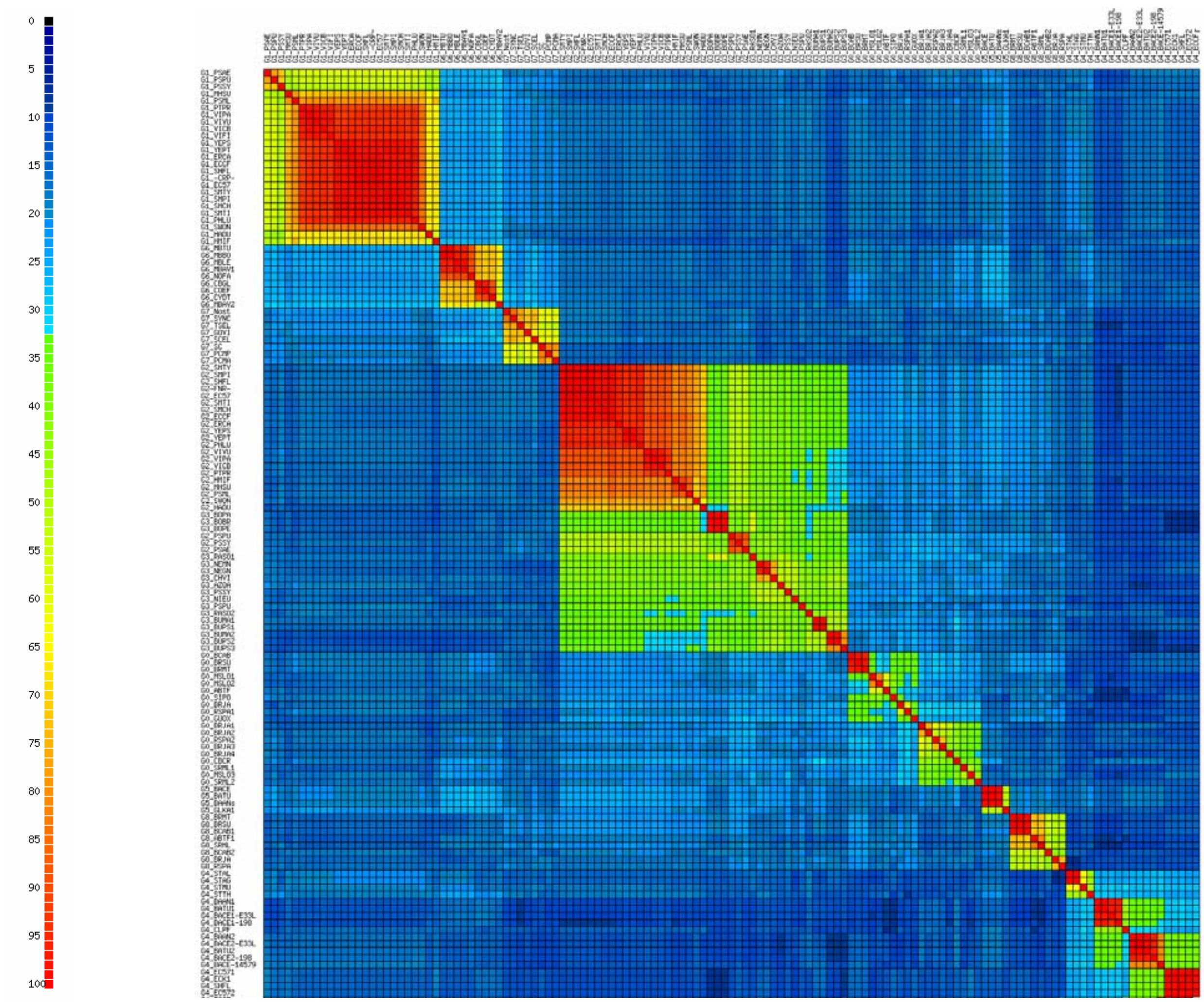


Figura 4R. **Gráfica de comparación de identidades entre pares de secuencias**. En la diagonal roja están las identidades máximas (1) y > 0.85 (rojo), de 0.75 a 0.65 (naranja-amarillo), y < 0.65 (tonos azules).

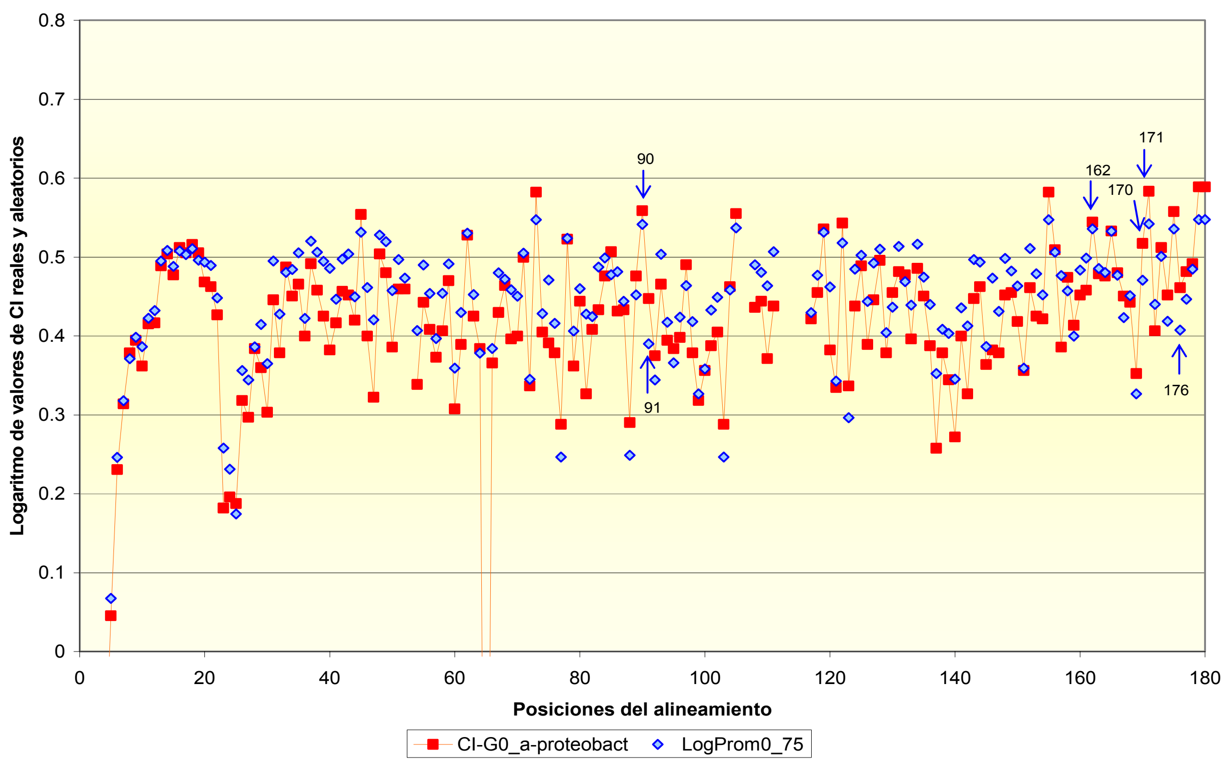
El cálculo de estas secuencias pseudos-aleatorias se realizó conforme a lo explicado en Métodos (3.5.a y 3.5.b) con la ρ = 0.75 con el programa en lenguaje de programación PERL Calc-PseudoRandom.pl (Anexo 7.6.2). Posteriormente se calcularon sus valores de *CI* con el mismo programa ContInfo-aa-alignment.pl (Anexo 7.6.1). A estos se les llamó *CI-referencia*, porque respecto a ellos se hizo la comparación estadística de los CIs de la familia real de Crp/Fnr, los *CIs-reales*.

Posteriormente, se procedió a la obtención de los valores Z (el valor con que el *CI-real* es tan similar o diferente al promedio de los *CIs-referencia*) para cada uno de los *CIs* de cada posición del alineamiento con la fórmula 2 explicada en Métodos (3.5.c), utilizando el programa MedsStats.pl (Anexo 7.6.4).

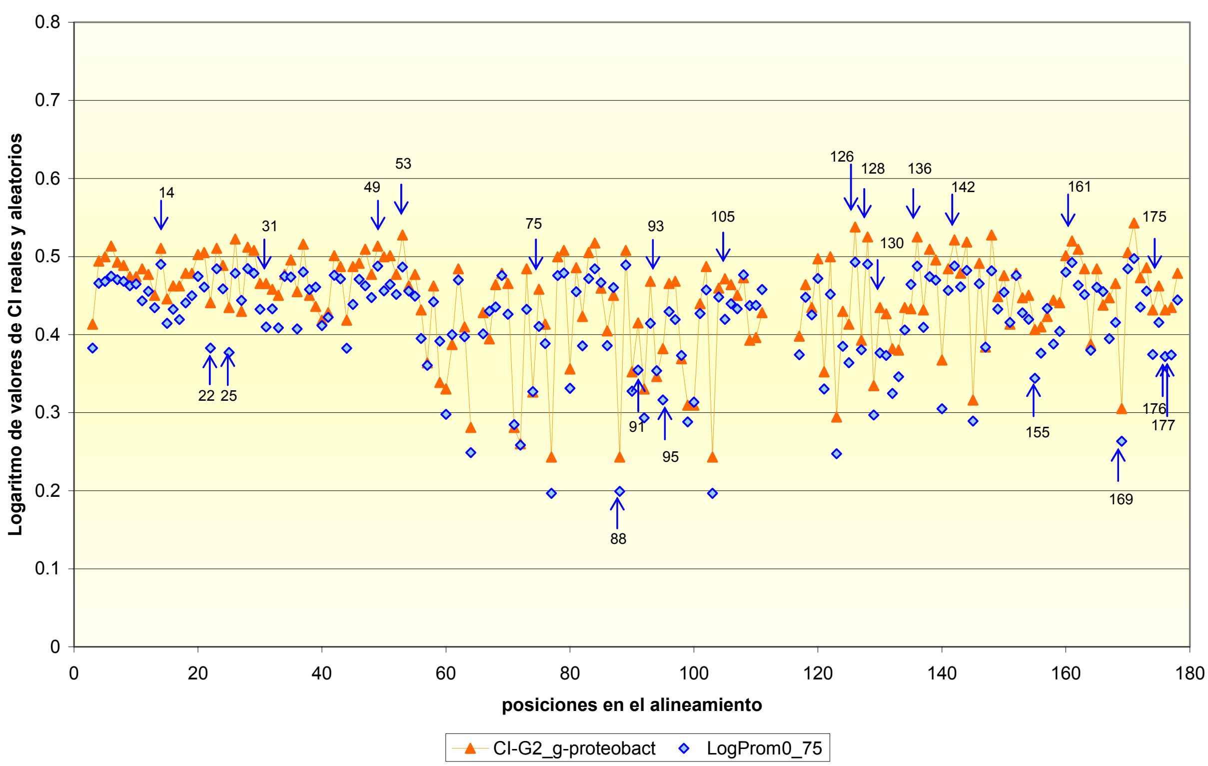
De los valores Z obtenidos para los *CI* (ver tablas en anexo 7.5), se consideraron *CI*s de residuos funcionalmente importantes aquellos que fueran ≥ 3 desviaciones estándar con un P-valor ≤ 0.05 (obtenido de tablas de valores Z, (Triola 1999)).

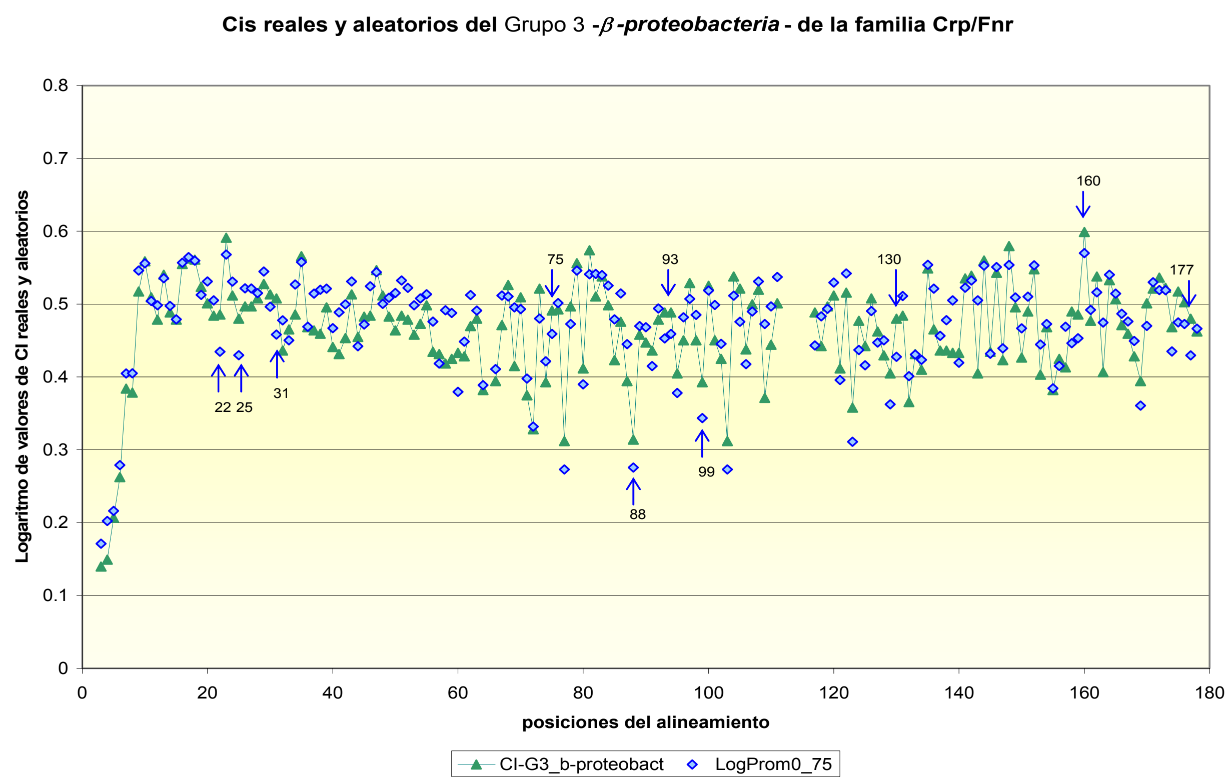
Para los grupos G1, G2, G3, G4 y G0, que son los grupos con más de 17 secuencias, sí se encontraron *CIs* con valores Z ≥ 3 con un P-valor ≤ 0.05, y se muestran en las gráficas 5R. Para los grupos G5, G6, G7 y G8, que tienen menos de 10 secuencias, se obtuvieron scores Z con valor máximo de +/-1.8, que quedan cerca del valor promedio y no se pueden diferenciar significativamente del azar (ver Anexo 7.5, tablas con valores Z). Por lo tanto, en este estudio no se puede reportar ninguna predicción de RFIs para estos grupos. Se necesitará esperar a que haya más genomas secuenciados de estos grupos filogenéticos (*Actinobacteria, Cianobacteria, α-proteobacteria* y *Firmicutes*), para volver a buscar ortólogos de Crp/Fnr y aumentar así el número de secuencias.

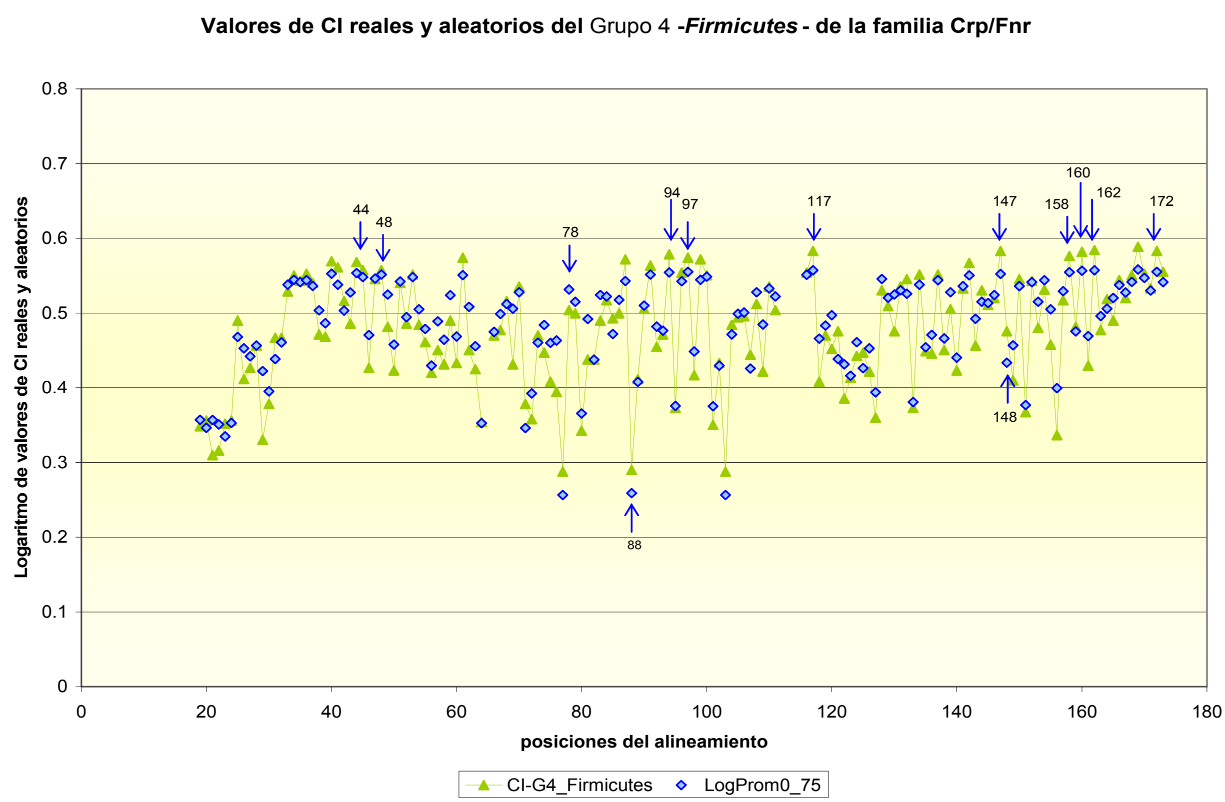
CIs reales y aleatorios del Grupo 0 **-α*-proteobacteria* - de la familia Crp/Fnr**



CIs reales y aleatorios del Grupo 2 **-γ*-proteobacteria* - de la familia Crp/Fnr**




Cis reales y aleatorios del Grupo 3 -β-proteobacteria - de la familia Crp/Fnr
Logaritmo de valores de CI reales y aleatorios
posiciones del alineamiento
CI-G3_b-proteobact
LogProm0_75
22 25
31
75
88
93
99
130
160
177


Valores de CI reales y aleatorios del Grupo 4 -Firmicutes - de la familia Crp/Fnr
Logaritmo de valores de CI reales y aleatorios
posiciones del alineamiento
CI-G4_Firmicutes
LogProm0_75
44
48
78
88
94
97
117
147
148
158
160
162
172

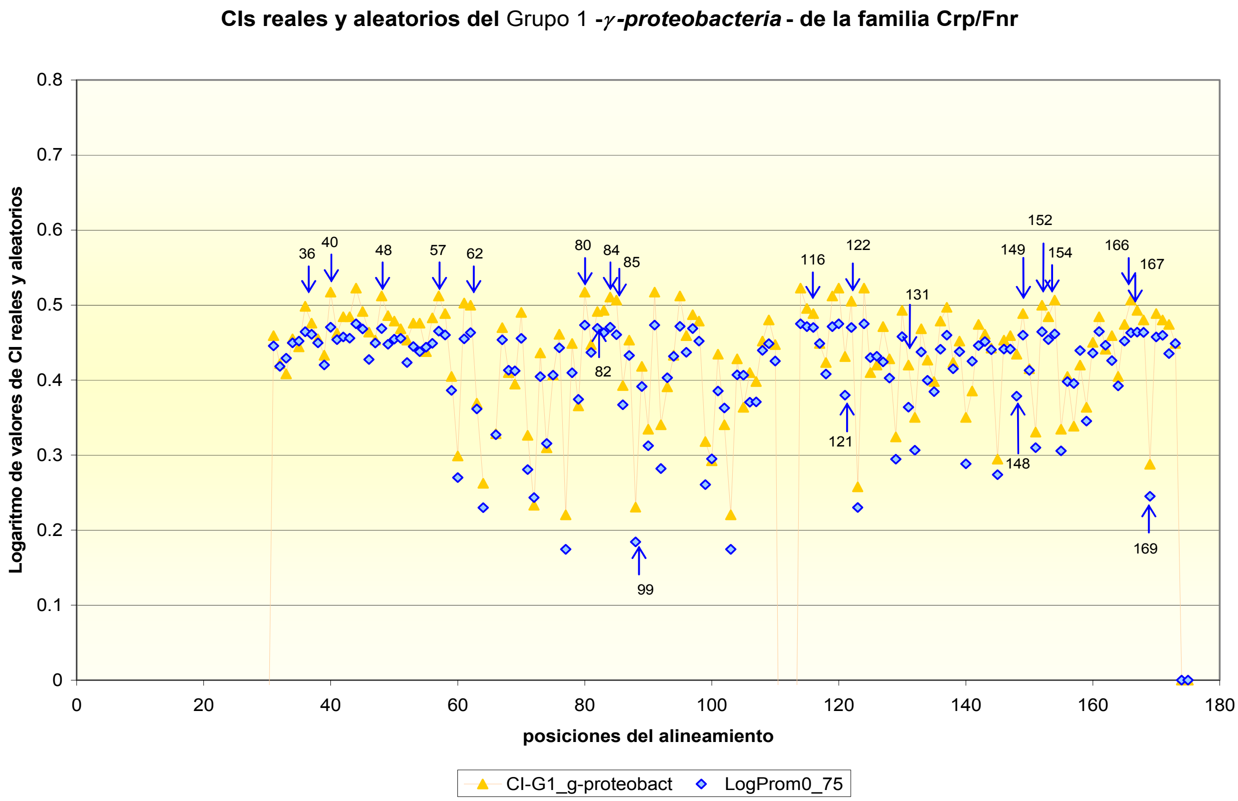


Figura 5R. Gráficas de los *CI-reales* y *CI-referencia* (con variación < 0.75) de las 180 posiciones del alineamiento para los grupos 2, 3, 0, 4 y 1. Los *CI* indicados con ↑ tienen valores Z ≥ 3 y corresponden a las posiciones donde se encuentran residuos funcionalmente importantes.

Posterior a la identificación de los CI-reales con valores Z ≥ 3 (con un valor p ≤ 0.05), se procedió a identificar a que residuos corresponden en el alineamiento para cada grupo (figura 6R).

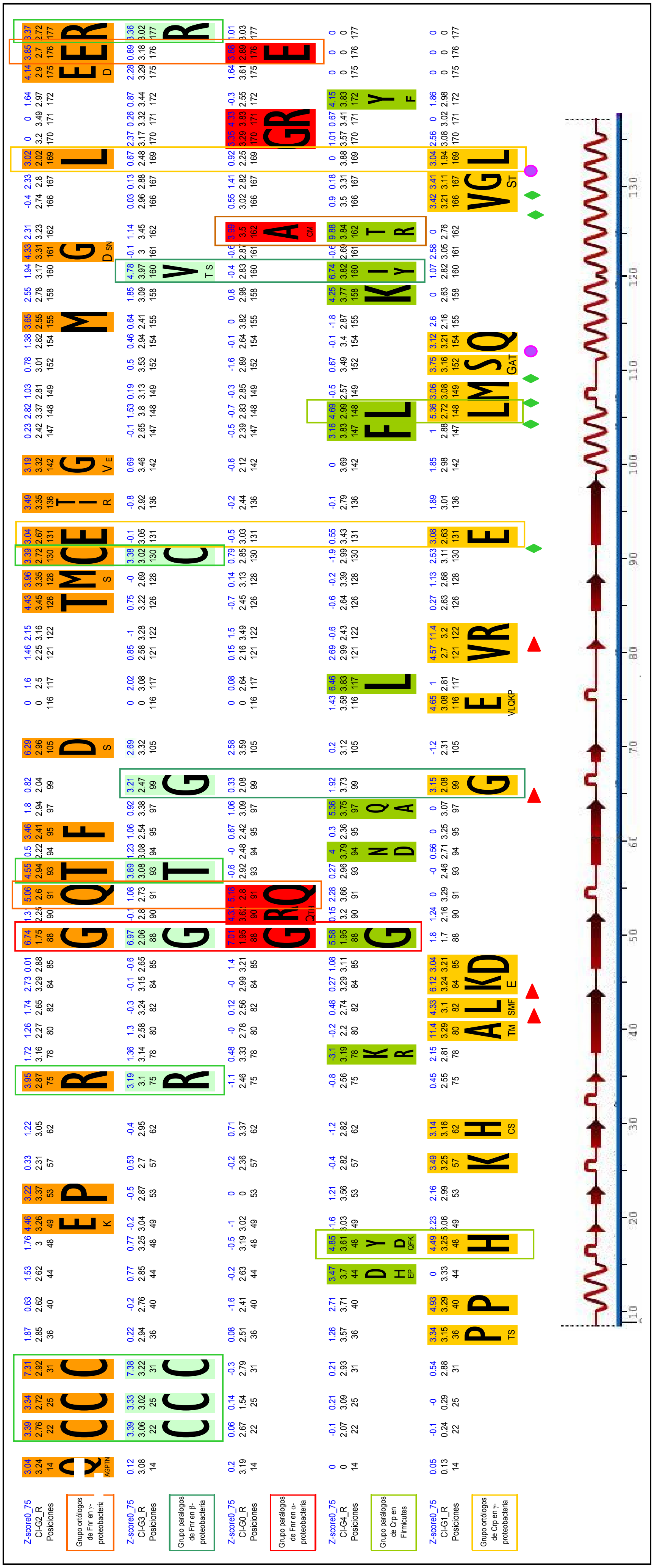


Figura 6R. Residuos Funcionalmente Importantes identificados para los grupos del alineamiento de la familia Crp/Fnr:

G2 (naranja, tipo Fnr en γ-*Proteobacteria*),
G3 (acua, tipo Fnr en β-*Proteobacteria*),
G0 (rojo, tipo Fnr en α-*Proteobacteria*),
G4 (verde, tipo Crp en *Firmicutes*) y
G1 (amarillo, tipo Crp en γ-*Proteobacteria*).
Las posiciones, los valores de CI y Z se indican arriba de cada grupo.
Los residuos comunes entre grupos están enmarcados en rectángulos entre los grupos correspondientes.
En la parte inferior está la estructura secundaria de Crp, puesta como referencia de posible ubicación de los RFIs encontrados para los otros grupos.

4.3.1. Comparando entre grupos los residuos con *CI*s significativos, que se predicen como RFIs

Se encontró que entre cuatro de los cinco grupos, los grupos 2, 3, 0 y 4, comparten la Glicina 88 con *CI* significativo (figura 6R, marcada en rectángulo rojo). En particular, este aminoácido se ve conservado en todas las secuencias del alineamiento (ver Anexo 7.2). Es interesante que a pesar de ello, esta Glicina no haya resultado un residuo con *CI* significativo para el grupo 1 de Crp (ver tabla de G1 en Anexo 7.4).

En cuanto al grupo 2 de γ-*Proteobacteria* de los 24 residuos encontrados, se encontró que comparte específicamente siete RFIs iguales con el grupo 3 de β-*Proteobacteria* (figura 6R, rectángulos verde acua): Cisteinas 22, 25, 31 y 130, Argininas 75 y 177, y Treonina 93; dos específicamente con el grupo 0 de α-*Proteobacteria* (figura 6R, rectángulo rojo-naranja): Glutamina 91 y Aspartato 176; y dos específicamente con el grupo 1 de γ-*Proteobacteria* (figura 6R, rectángulos naranjas): Aspartato 131 y 176. Los restantes 12 residuos son específicos del grupo.

En cuanto a los 10 residuos encontrados para el grupo 3 de β-*Proteobacteria*, además de los ya mencionados que comparte específicamente con el grupo 2, comparte la Glicina 99 con el grupo 1 de γ-*Proteobacteria* (figura 6R, rectángulo verde agua) y la posición 160 con el grupo 4 de *Firmicutes* (figura 6R, rectángulos verde agua; Valina para el grupo 3, Isoleucina o Tirosina para el grupo 4); sólo queda un residuo único de este grupo.

Para los 7 residuos del grupo 0 de α-*Proteobacteria* comparte, además de los ya mencionados con los grupos 2 y 3, la posición 162 con el grupo 4 *Firmicutes* (figura 6R, rectángulos verde rojo-naranja; Alanina para grupo 0 y Treonina o Arginina para el grupo 4). Los otros tres residuos son únicos de este grupo.

De los 13 residuos encontrados para el grupo 4 de *Firmicutes*, además de los ya mencionados el grupo 0, comparte específicamente con el grupo 1 de γ-*Proteobacteria* (figura 6R, rectángulos verde) la leucina 148, y la posición 48 (Tirosina o Aspartato para el grupo 4, e Histidina para el 1); los restantes 8 residuos son únicos de este grupo.

Finalmente, de los 21 residuos encontrados en el grupo 1, 16 son específicos de este grupo; el resto como ya se mencionó se comparten con los grupos 2,3 y 4 (figura 6R, rectángulos amarillos, agua y verdes).

Antes de poder hacer el análisis de estas coincidencias y diferencias mencionadas en esta sección de los RFIs encontrados en los cinco grupos de la familia, es necesario explicar lo que se conoce sobre Crp_{ECO} y Fnr_{ECO}, y mapear los RFIs encontrados para estos dos grupos en un modelo de estructura tridimensional.

4.3.2. Residuos con evidencia experimental, ligandos y conformaciones de Crp y Fnr

En general los estudios de mutagénesis de proteínas se han seguido de dos formas: con un análisis previo de la estructura para seleccionar los residuos a estudiar, o mutando residuos al azar. Con estas estrategias, los residuos que se identifican con algún efecto en la función o conformación, provienen de una selección si se parte de un estudio de estructura y dependen del tipo y condiciones del experimento realizado. Las características que se determinan a partir de estos experimentos pueden ser ambiguos funcionalmente, de función específica o sin relevancia en su función.

De los cinco grupos (2, 3, 0, 4 y 6) a los que se les pudo identificar residuos con CI con valores significativos, solo se cuenta con la referencia de residuos ya caracterizados para el DUL de Crp_{ECO} y algunos de los de Fnr_{ECO}; también se conocen detalles de su estado dentro de la célula y de cómo regulan.

Para Crp_{ECO}, la mayoría de los estudios parten de análisis de su estructura e inicialmente se habían enfocado en la identificación de los residuos que interaccionan específicamente con su ligando AMPc (Weber and Steitz 1987; Belduz, Lee et al. 1993;

Lee, Glasgow et al. 1994), luego para los que específicamente hacen la interacción entre las subunidades (Heyduk, Heyduk et al. 1992) y más recientemente para conocer como pasa la señal de interacción con el AMPc al dominio de interacción con el DNA y que modulan el cambio de conformación de la proteína (Baker, Tomlinson et al. 2001; Lin, Kovac et al. 2002; Gekko, Obu et al. 2004) (ver figura 4c en Introducción). Además, se sabe que Crp_{ECO} interacciona con AMPc sólo cuando la concentración de este metabolito aumenta, se mantiene como dímero unido o despegado del DNA y sólo cambia un poco su conformación para unirse al DNA (ver figura 4 en Introducción).

Para Fnr_{ECO}, se tienen descritas cuatro Cisteínas indispensables para la interacción con el grupo prostético 4Fe4S (Lazazzera, Beinert et al. 1996; Khoroshilova, Popescu et al. 1997); además Fnr_{ECO} siempre tiene unido a su grupo prostético 4Fe4S que cambia sólo de estado oxido-reducción, lo que provoca que Fnr_{ECO} cuando no está unido al DNA esté como monómero y que se conforme como dímero cuando se une al DNA (Kiley et. al. 1999).

4.3.3. De los residuos con evidencia experimetnal en Crp_{ECO} y los RFIs predichos para su grupo en γ-*Proteobacteria*

Para apoyar el análisis de los residuos encontrados para el grupo 1 de Crp en γ-*Proteobacterias*, estos se mapearon en una de las estructuras tridimensionales ya reportadas para Crp de *E. coli*, la 1J59 en el PDB, utilizando el software de Protein Explorer 2.8 [www.proteinexplorer.org], (tabla 2R y figura 7R).

Tabla 2R. **Residuos encontrados en este trabajo con *CI*s significativos**

Región en Estructura	Orientación	Posición en Tabla RFIs	Residuo	Posic Estructura	Descripción Función
	E	36	Prolina	--	
α-hélice	E	40	Prolina	9	
	E	48	Histidina	17	
hojas β y asas	E	57	Lisina	26	

	E	62	Histidina	31	
	O	80	Alanina	48	
	Cúmulo espacial	82	Leucina	50	
	Cúmulo espacial	84	Lisina	52	Se han determinado como residuos importantes en la transmisión de la señal de activación posterior a la unión de AMPc (Gekko 2004, Lin 2002)
	poco E	85	Aspartato	53	Se han determinado como residuos importantes en la transmisión de la señal de activación posterior a la unión de AMPc (Gekko 2004, Lin 2004)
	Cúmulo espacial	99	Glicina	67	
			Glutamato	72	Después del estudio de la estructura de Crp, se probó por mutagénesis que es importante para la interacción con el AMPc (Belduz et. al. 1993)
	E	116	Glutamato	81	
			Arginina	82	Después del estudio de la estructura de Crp, se probó por mutagénesis que es importante para la interacción con el AMPc (Belduz et. al. 1993)
			Serina	83	Después del estudio de la estructura de Crp, se probó por mutagénesis que es importante para la interacción con el AMPc (Lee et. al. 1994)
	orientada AMPc	121	Valina	86	
	Cúmulo espacial	122	Arginina	87	
	poco E	131	Glutamato	96	
			Tirosina	99	Contribuye en la estabilidad de la estructura dimerica, en la asociación entre las subunidades (Baker 2004)
α-hélice	contacto I-subU	148	Leucina	113	
	contacto I-subU	149	Metionina	114	
	contacto I-subU	152	Serina	117	es blanco de proteolisis cuando es mutada (Heyduk et. al. 1992)
	hacia adentro	154	Glutamina	119	
			Treonina	127	Después del estudio de la estructura de Crp, se probó por mutagénesis que es importante para la interacción con el AMPc (Lee et. al. 1994)
			Serina	128	Después del estudio de la estructura de Crp, se probó por mutagénesis que es importante para la interacción con el AMPc (Lee et. al. 1994)
	contacto I-subU	166	Valina	131	
	contacto I-subU	167	Glicina	132	residuo esencial para transmitir las señales de las interacciones alostéricas (Yu et. al 2004)
	hacia adentro	169	Leucina	134	
región visagra			Fenilanina	136	es blanco de proteolisis cuando es mutada (Heyduk et. al. 1992)

Residuos ordenados según su ubicación (columna 1) y orientación (columna 2; E: externo, O: oculto, I-subU: inter-subunidad) en la estructura de Crp (de la figura 7R); el número de su posición en el alineamiento (columna 3) (figura 6R); su número de posición en la estructura (columna 5); tipo de función encontrada con estudios de mutagénesis (columna 6). Las lineas en gris corresponden a los residuos no identificados con el análisis de *CI*s.

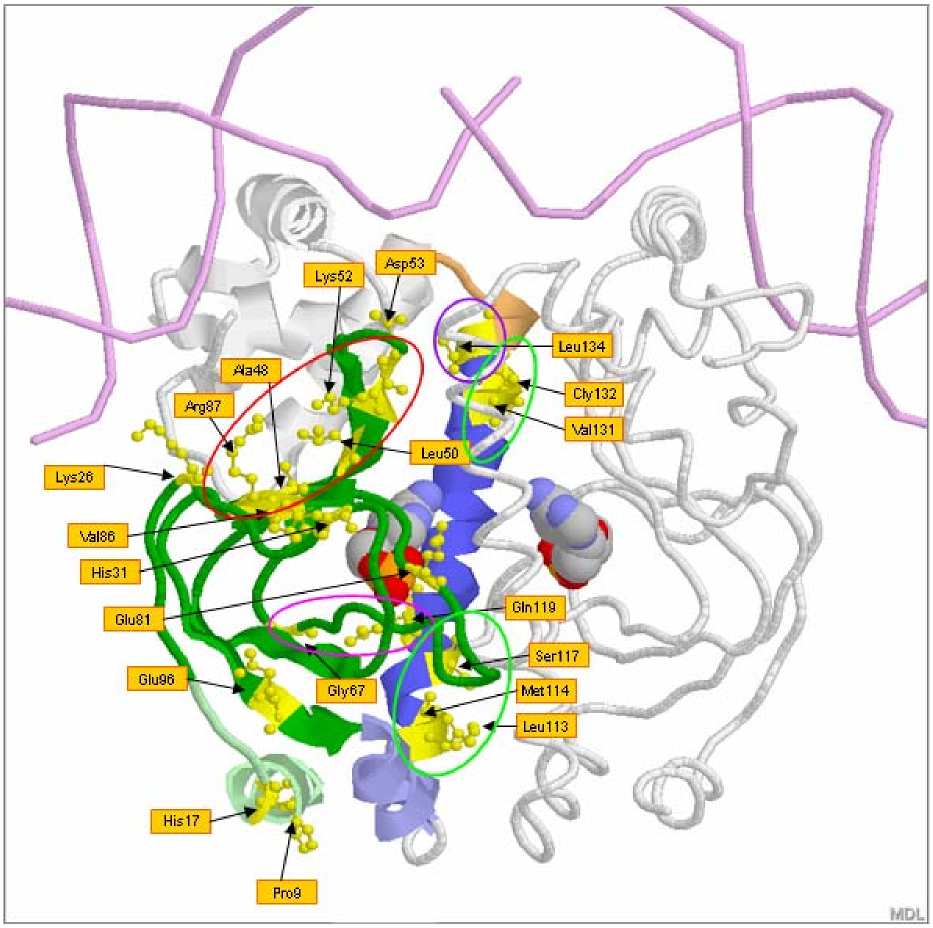


Figura 7R. Gráfica de la estructura de Crp, 1J59, con los RFIs predichos en este trabajo mapeados en color morado, en la gráfica se anotan con las posiciones en la estructura y aquí se anota la posición a la que corresponden en el alineamiento: Pro9 = P40, His17 = H48, Lys26 = K57, His31 = H62, Ala48 = A80, Leu50 = L82, Lys52 = K84, Asp53 = D85, Gly67= G99, Glu81 = E116, Val86 = V121, Arg87 = R122, Glu96 = E131, Leu113 = L148, Met114 = M149, Ser117 = S152, Gln119 = Q154, Val131 = V166, Gly132 = G167, Leu134 = L169. Ver en texto anotaciones de los círculos de colores.

Los residuos dentro de la α-hélice con que interaccionan los dominios se marcan dentro de ovalos verdes (marcados con en figura 6R) y los que están orientados hacia el propio monómero se marcan en ovalos morado y rosa (marcados con en la figura 6R), el grupo que se encuentra formando un cúmulo de residuos cercanos espacialmente se indican dentro de círculo rojo (marcados con en figura 6R). Los residuos comunes con el grupo de Fnr en γ-*Proteobacteria* se indican en los dos círculos rojos, y los comunes con el grupo 4 tipo Crp en *Firmicutes* en círculos verdes.

4.3.4. De los residuos con referencia en Fnr_{ECO} y los RFIs predichos para su grupo en *γ-Proteobacteria*

Aunque para Fnr *E. coli* se cuenta con poca información respecto a su estructura y conformación, respecto a los residuos ya reportados que interaccionan con su grupo prostético, para el grupo de Fnr en *γ-Proteobacteria* (G2) se encontraron las mismas Cisteínas con CI significativos; y no sólo en el grupo 2, sino que también se comparten con CI significativos con el grupo G3 en *β-Proteobacteria*.

Se propone que las proteínas del grupo G3 también unen un grupo prostético 4Fe4S dado que también tienen las cuatro cisteínas con CI significativos y en las mismas posiciones que el grupo 2 de Fnr.

4.4. Eficiencia del método, comparando con lo ya reportado y las características de las proteínas

De acuerdo a la comparación de los residuos ya reportados *versus* los residuos aquí encontrados como funcionalmente importantes (con valores de CI significativos), se encontraron pocas coincidencias. Por una parte, de los residuos conocidos para Crp_{ECO} hicieron falta varios que aunque en el alineamiento se conservan, no se califican como significativos. Esto puede deberse a que los RFIs detectados con este método no están involucrados con la interacción del ligando sino a que están más relacionados con la interacción inter-subunidades, el núcleo de conformación y/o la velocidad de plegamiento.

Por otra parte, para Fnr_{ECO} se encontraron las cuatro cisteínas que interaccionan con su grupo prostético y otros residuos, que por falta de caracterización estructural, no se puede decir si están relacionados con el núcleo de conformación y/o la habilidad de plegamiento rápido.

La diferencia de haber identificado a los residuos que interaccionan con el ligando para el grupo 2 de Fnr y no para el grupo 1 de Crp se puede explicar por el hecho de que Fnr en cuanto es sintetizada se une a un grupo prostético Hierro-Azufre que se oxida y reduce (2Fe-2S → 4Fe4S) en respuesta los niveles de Oxígeno, mientras que Crp sólo se une al AMPc cuando la concentración de este metabolito aumenta (ver Introducción). En el caso de Fnr, estas cuatro cisteínas permanecen ocultas por su interacción constante con el grupo prostético; además que cuando la célula detecta que Fnr no tiene a su grupo unido (Fnr-apo) lo degrada. En el caso de Crp, los residuos que interaccionan con el AMPc están expuestos al medio, si alguno de ellos se muta Crp puede perder la especificidad (reconocer otros nucleótidos monofosfatados) o quedar en la conformación activa y en algunos casos podría perder su capacidad de regulación, pero no es degrada. Está diferencia puede marcar la diferente presión selectiva en estas proteínas para conservar ciertos residuos funcionales.

5. Conclusiones generales y perspectivas

5.1. Conclusiones generales

En el presente estudio se calculó el Contenido Informacional para identificar los residuos funcionalmente importantes en los dominios de unión al ligando para los miembros de la familia de los factores de transcripción Crp/Fnr.

Esta herramienta es dependiente del número de secuencias (tamaño de la muestra de las secuencias de la familia Crp/Fnr), por lo que su análisis estadístico es indispensable. Esta dependencia hizo que la identificación de estos residuos fuera inviable para los grupos que tuvieron menos de 10 secuencias, los grupos 5, 6, 7 y 8. El hecho de tener pocas secuencias implicó que no tuvieran una significancia estadística. Por esta dependencia, sólo se predijeron diferentes residuos potencialmente funcionales y estadísticamente significativos para los grupos 2, 3, 0, 4 y 6.

La dependencia de *CIs* al número de secuencia fue observable desde su graficación (figura 3R.B) con *CIs* que se gráfican separados a la tendencia general los otros grupos.

En general los RFIs identificados corresponden a todos los aminoácidos biológicamente conocidos, excepto triptófano.

Todavía falta completar la caracterización estructural de los diferentes grupos de la familia; sin embargo, se tiene la certeza de que los valores de *CI* que han sido calificados como significativos, han sido aquellos que cumplen con criterios estadísticos estrictos (con valores Z por arriba de 3 desviaciones estándar con un p-valor menor a 0.05), lo cual nos permite hacer predicciones sobre estos residuos en los grupos aún sin conocer su ubicación espacial dentro de la estructura.

Para varios grupos no se tiene referencia de residuos caracterizados. Falta una mayor caracterización experimental de las interacciones de estos dominios; así que tener predicciones de RFIs con valores de CI estadísticamente significativos, se pueden utilizar como guía para identificar residuos que puedan ser importantes para la conformación de estas proteínas, así como aquellos que no participen en ninguna función.

Pocos de los residuos identificados han sido reportados previamente con evidencia experimental. La gran mayoría corresponde a residuos que tentativamente se encuentran fuera del sitio físico de interacción con el ligando, pero que pueden estar involucrados en el núcleo de conformación y/o en la habilidad de plegamiento rápido de estas proteínas.

En general, los residuos que se predicen como funcionalmente importantes en este trabajo pertenecen a regiones de las proteínas que aún no se han descrito por completo bioquímicamente, en un momento en que las herramientas con las que se cuenta son sólo la observación de su conservación y posición en la estructura terciaria.

La identificación de residuos funcionalmente importantes compete al área de la genómica que nos permite completar la caracterización de los dominios de unión al ligando de proteínas como factores transcripcionales. Esto nos permite conocer los detalles bioquímicos de las interacciones con los ligandos, entre subunidades y otras funciones que hacen específica la respuesta a la regulación transcripcional de estos factores en diferentes organismos.

El método propuesto y utilizado en este trabajo permitió la identificación de RFIs sin depender de una conservación de posiciones en un alineamiento múltiple entre los miembros de los diferentes grupos. Inclusive se pudieron identificar varios residuos en posiciones comunes entre los grupos de forma independiente; además, de ser un método independiente de la disponibilidad de estructuras tridimensionales de las proteínas de la familia.

5.2. Perspectivas

El análisis computacional para identificar posibles Residuos Funcionalmente Importantes (RFIs) en la familia Crp/Fnr nos ha permitido abrir diversas preguntas que pueden continuar, tales como, la caracterización estructural de diversos miembros representativos de cada grupo (obtención de su cristal, construcción y comparación de modelos tridimensionales entre los diferentes grupos de proteínas) y la comprobación experimental de los residuos aquí determinados como importantes para la función. A su vez, estos estudios podrían arrojar luz en el conocimiento del comportamiento de estas proteínas en cuanto a sus interacciones específicas con diversos compuestos, o aquellos involucrados en los cambios de conformación como consecuencia de la unión a su ligando. Por otra parte, este enfoque se puede extender a las otras familias de factores transcripcionales que se han descrito hasta la fecha.

Finalmente, el método podría ser mejorado implementando una matriz de peso para medir los cambios de aminoácidos en lugar de sólo medir los cambios de identidades de aminoácidos.

6. Referencias

Altschul, S. F., W. Gish, et al. (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-10.

Armon, A., D. Graur, et al. (2001). "ConSurf: an algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information." J Mol Biol **307**(1): 447-63.

Baker, C. H., S. R. Tomlinson, et al. (2001). "Amino acid substitution at position 99 affects the rate of CRP subunit exchange." Biochemistry **40**(41): 12329-38.

Barker, W. C., J. S. Garavelli, et al. (1999). "The PIR-International Protein Sequence Database." Nucleic Acids Res **27**(1): 39-43.

Belduz, A. O., E. J. Lee, et al. (1993). "Mutagenesis of the cyclic AMP receptor protein of Escherichia coli: targeting positions 72 and 82 of the cyclic nucleotide binding pocket." Nucleic Acids Res **21**(8): 1827-35.

Benson, D. A., I. Karsch-Mizrachi, et al. (2003). "GenBank." Nucleic Acids Res **31**(1): 23-7.

Berezin, C., F. Glaser, et al. (2004). "ConSeq: the identification of functionally and structurally important residues in protein sequences." Bioinformatics **20**(8): 1322-4.

Berman, H. M., J. Westbrook, et al. (2000). "The Protein Data Bank." Nucleic Acids Res **28**(1): 235-42.

Boeckmann, B., A. Bairoch, et al. (2003). "The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003." Nucleic Acids Res **31**(1): 365-70.

Brennan, R. G. and B. W. Matthews (1989). "The helix-turn-helix DNA binding motif." J Biol Chem **264**(4): 1903-6.

Browning, D. F. and S. J. Busby (2004). "The regulation of bacterial transcription initiation." Nat Rev Microbiol **2**(1): 57-65.

Burgess, R. R. and L. Anthony (2001). "How sigma docks to RNA polymerase and what sigma does." Curr Opin Microbiol **4**(2): 126-31.

Burgess, R. R., A. A. Travers, et al. (1969). "Factor stimulating transcription by RNA polymerase." Nature **221**(5175): 43-6.

Crick, F. (1970). "Central dogma of molecular biology." Nature **227**(5258): 561-3.

Chothia, C. and A. M. Lesk (1986). "The relation between the divergence of sequence and structure in proteins." Embo J **5**(4): 823-6.

Danot, O. (2001). "A complex signaling module governs the activity of MalT, the prototype of an emerging transactivator family." Proc Natl Acad Sci U S A **98**(2): 435-40.

Dibden, D. P. and J. Green (2005). "In vivo cycling of the Escherichia coli transcription factor FNR between active and inactive states." Microbiology **151**(Pt 12): 4063-70.

Eddy, S. R. (1998). "Profile hidden Markov models." Bioinformatics **14**(9): 755-63.

Garges, S. and S. Adhya (1988). "Cyclic AMP-induced conformational change of cyclic AMP receptor protein (CRP): intragenic suppressors of cyclic AMP-independent CRP mutations." J Bacteriol **170**(4): 1417-22.

Gekko, K., N. Obu, et al. (2004). "A linear correlation between the energetics of allosteric communication and protein flexibility in the Escherichia coli cyclic AMP receptor protein revealed by mutation-induced changes in compressibility and amide hydrogen-deuterium exchange." Biochemistry **43**(13): 3844-52.

Gough, J. (2002). "The SUPERFAMILY database in structural genomics." Acta

Crystallogr D Biol Crystallogr **58**(Pt 11): 1897-900.
Green, J., C. Scott, et al. (2001). "Functional versatility in the CRP-FNR superfamily of transcription factors: FNR and FLP." Adv Microb Physiol **44**: 1-34.
Gruber, T. M. and C. A. Gross (2003). "Multiple sigma subunits and the partitioning of bacterial transcription space." Annu Rev Microbiol **57**: 441-66.
Gunsalus, R. P. and S. J. Park (1994). "Aerobic-anaerobic gene regulation in Escherichia coli: control by the ArcAB and Fnr regulons." Res Microbiol **145**(5-6): 437-50.
Haldenwang, W. G. (1995). "The sigma factors of Bacillus subtilis." Microbiol Rev **59**(1): 1-30.
Harrison, S. C. (1991). "A structural taxonomy of DNA-binding domains." Nature **353**(6346): 715-9.
Heyduk, E., T. Heyduk, et al. (1992). "Intersubunit communications in Escherichia coli cyclic AMP receptor protein: studies of the ligand binding domain." Biochemistry **31**(14): 3682-8.
Holm, L. (1998). "Unification of protein families." Curr Opin Struct Biol **8**(3): 372-9.
Hutchison, C. A., 3rd, S. Phillips, et al. (1978). "Mutagenesis at a specific position in a DNA sequence." J Biol Chem **253**(18): 6551-60.
Innis, C. A., J. Shi, et al. (2000). "Evolutionary trace analysis of TGF-beta and related growth factors: implications for site-directed mutagenesis." Protein Eng **13**(12): 839-47.
James, L. C. and D. S. Tawfik (2003). "Conformational diversity and protein evolution--a 60-year-old hypothesis revisited." Trends Biochem Sci **28**(7): 361-8.
Johnson, J. M. and G. M. Church (2000). "Predicting ligand-binding function in families of bacterial receptors." Proc Natl Acad Sci U S A **97**(8): 3965-70.
Jones, D. T. (1999). "Protein secondary structure prediction based on position-specific scoring matrices." J Mol Biol **292**(2): 195-202.
Kawabata, T., M. Ota, et al. (1999). "The Protein Mutant Database." Nucleic Acids Res **27**(1): 355-7.
Khoroshilova, N., C. Popescu, et al. (1997). "Iron-sulfur cluster disassembly in the FNR protein of Escherichia coli by O2: [4Fe-4S] to [2Fe-2S] conversion with loss of biological activity." Proc Natl Acad Sci U S A **94**(12): 6087-92.
Kiley, P. J. and H. Beinert (1999). "Oxygen sensing by the global regulator, FNR: the role of the iron-sulfur cluster." FEMS Microbiol Rev **22**(5): 341-52.
Korner, H., H. J. Sofia, et al. (2003). "Phylogeny of the bacterial superfamily of Crp-Fnr transcription regulators: exploiting the metabolic spectrum by controlling alternative gene programs." FEMS Microbiol Rev **27**(5): 559-92.
Kumar, S., K. Tamura, et al. (2004). "MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment." Brief Bioinform **5**(2): 150-63.
Landgraf, R., D. Fischer, et al. (1999). "Analysis of heregulin symmetry by weighted evolutionary tracing." Protein Eng **12**(11): 943-51.
Landgraf, R., I. Xenarios, et al. (2001). "Three-dimensional cluster analysis identifies interfaces and functional residue clusters in proteins." J Mol Biol **307**(5): 1487-502.
Lawson, C. L., D. Swigon, et al. (2004). "Catabolite activator protein: DNA binding and transcription activation." Curr Opin Struct Biol **14**(1): 10-20.
Lazazzera, B. A., H. Beinert, et al. (1996). "DNA binding and dimerization of the Fe-S-

containing FNR protein from Escherichia coli are regulated by oxygen." J Biol Chem **271**(5): 2762-8.

Lee, E. J., J. Glasgow, et al. (1994). "Mutagenesis of the cyclic AMP receptor protein of Escherichia coli: targeting positions 83, 127 and 128 of the cyclic nucleotide binding pocket." Nucleic Acids Res **22**(15): 2894-901.

Lewin, B. (2000). Genes VII. New York, Oxford University Press.

Lewis, M. (2005). "The lac repressor." C R Biol **328**(6): 521-48.

Li, B., H. Wing, et al. (1998). "Transcription activation by Escherichia coli FNR protein: similarities to, and differences from, the CRP paradigm." Nucleic Acids Res **26**(9): 2075-81.

Li, W., L. Jaroszewski, et al. (2002). "Tolerating some redundancy significantly speeds up clustering of large protein databases." Bioinformatics **18**(1): 77-82.

Lichtarge, O., H. Yao, et al. (2003). "Accurate and scalable identification of functional sites by evolutionary tracing." J Struct Funct Genomics **4**(2-3): 159-66.

Lin, S. H., L. Kovac, et al. (2002). "Ability of E. coli cyclic AMP receptor protein to differentiate cyclic nucelotides: effects of single site mutations." Biochemistry **41**(9): 2946-55.

Madan Babu, M. and S. A. Teichmann (2003). "Evolution of transcription factors and the gene regulatory network in Escherichia coli." Nucleic Acids Res **31**(4): 1234-44.

Madan Babu, M. and S. A. Teichmann (2003). "Functional determinants of transcription factors in Escherichia coli: protein families and binding sites." Trends Genet **19**(2): 75-9.

Makarova, K. S., A. A. Mironov, et al. (2001). "Conservation of the binding site for the arginine repressor in all bacterial lineages." Genome Biol **2**(4): RESEARCH0013.

McGuffin, L. J., K. Bryson, et al. (2000). "The PSIPRED protein structure prediction server." Bioinformatics **16**(4): 404-5.

Mirny, L. A. and E. I. Shakhnovich (1999). "Universally conserved positions in protein folds: reading evolutionary signals about stability, folding kinetics and function." J Mol Biol **291**(1): 177-96.

Mironov, A. A., E. V. Koonin, et al. (1999). "Computer analysis of transcription regulatory patterns in completely sequenced bacterial genomes." Nucleic Acids Res **27**(14): 2981-9.

Moreno-Campuzano, S., S. C. Janga, et al. (2006). "Identification and analysis of DNA-binding transcription factors in Bacillus subtilis and other Firmicutes--a genomic approach." BMC Genomics **7**: 147.

Mount, D. W. (2001). Bioinformatics: Sequence and Genome Analysis., Cold Spring Harbor Laboratory.

Murzin, A. G., S. E. Brenner, et al. (1995). "SCOP: a structural classification of proteins database for the investigation of sequences and structures." J Mol Biol **247**(4): 536-40.

Pearl, F. M., C. F. Bennett, et al. (2003). "The CATH database: an extended protein family resource for structural and functional genomics." Nucleic Acids Res **31**(1): 452-5.

Pearson, W. R. (2000). "Flexible sequence similarity searching with the FASTA3 program package." Methods Mol Biol **132**: 185-219.

Perez-Rueda, E. (1999). Reconocimiento de motivos estructurales en reguladores transcripcionales. Centro de Investigación sobre Fijación de Nitrógeno.

Cuernavaca, Universidad Nacional Autónoma de México. **Doctorado:** 70.
Perez-Rueda, E. and J. Collado-Vides (2000). "The repertoire of DNA-binding transcriptional regulators in Escherichia coli K-12." <u>Nucleic Acids Res</u> **28**(8): 1838-47.
Perez-Rueda, E. and J. Collado-Vides (2001). "Common history at the origin of the position-function correlation in transcriptional regulators in archaea and bacteria." <u>J Mol Evol</u> **53**(3): 172-9.
Perez-Rueda, E., J. Collado-Vides, et al. (2004). "Phylogenetic distribution of DNA-binding transcription factors in bacteria and archaea." <u>Comput Biol Chem</u> **28**(5-6): 341-50.
Rodionov, D. A., A. G. Vitreschak, et al. (2004). "Comparative genomics of the methionine metabolism in Gram-positive bacteria: a variety of regulatory systems." <u>Nucleic Acids Res</u> **32**(11): 3340-53.
Saito, S., M. Sasai, et al. (1997). "Evolution of the folding ability of proteins through functional selection." <u>Proc Natl Acad Sci U S A</u> **94**(21): 11324-8.
Salgado, H., S. Gama-Castro, et al. (2006). "RegulonDB (version 5.0): Escherichia coli K-12 transcriptional regulatory network, operon organization, and growth conditions." <u>Nucleic Acids Res</u> **34**(Database issue): D394-7.
Shaw, D. J., D. W. Rice, et al. (1983). "Homology between CAP and Fnr, a regulator of anaerobic respiration in Escherichia coli." <u>J Mol Biol</u> **166**(2): 241-7.
Sonnhammer, E. L., S. R. Eddy, et al. (1998). "Pfam: multiple sequence alignments and HMM-profiles of protein domains." <u>Nucleic Acids Res</u> **26**(1): 320-2.
Spiro, S. and J. R. Guest (1990). "FNR and its role in oxygen-regulated gene expression in Escherichia coli." <u>FEMS Microbiol Rev</u> **6**(4): 399-428.
Suzuki, M. and N. Yagi (1996). "An in-the-groove view of DNA structures in complexes with proteins." <u>J Mol Biol</u> **255**(5): 677-87.
Tan, K., G. Moreno-Hagelsieb, et al. (2001). "A comparative genomics approach to prediction of new members of regulons." <u>Genome Res</u> **11**(4): 566-84.
Thompson, J. D., T. J. Gibson, et al. (1997). "The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools." <u>Nucleic Acids Res</u> **25**(24): 4876-82.
Triola, M. F. (1999). <u>Estadística</u>. Mexico, Pearson Educación.
Weber, I. T. and T. A. Steitz (1987). "Structure of a complex of catabolite gene activator protein and cyclic AMP refined at 2.5 A resolution." <u>J Mol Biol</u> **198**(2): 311-26.
Wilson, C. A., J. Kreychman, et al. (2000). "Assessing annotation transfer for genomics: quantifying the relations between protein sequence, structure and function through traditional and probabilistic scores." <u>J Mol Biol</u> **297**(1): 233-49.
Yasuzawa, K., N. Hayashi, et al. (1992). "Histone-like proteins are required for cell growth and constraint of supercoils in DNA." <u>Gene</u> **122**(1): 9-15.
Zubay, G., D. Schwartz, et al. (1970). "Mechanism of activation of catabolite-sensitive genes: a positive control system." <u>Proc Natl Acad Sci U S A</u> **66**(1): 104-10.
Zvelebil, M. J., G. J. Barton, et al. (1987). "Prediction of protein secondary structure and active sites using the alignment of homologous sequences." <u>J Mol Biol</u> **195**(4): 957-61.

Anexo 7.1 Tabla de secuencias de la familia Crp/Fnr

Grupo	Siglas org.	Organismo	Nombre secuencia	dentificador G	Secuencia
G6	MBTU	[Mycobacterium tuberculosis H37Rv]	probable transcriptional regula	[15610812]	MDEILARAGIFQGVEPSAIAALTKQLQPVDFPRGHTVFAEGEPGDRLYIIISGKVKIGRRAPDGRENLLTIMGPSDMFGELSIFDPGPRTSSATTITEVRAVSMDRDALRSWIADRPEISEQLLRVLARRLRRTNNNLADLIFT
G6	MBB	[Mycobacterium bovis AF2122/97]	probable transcriptional regula	[31794846]	MDEILARAGIFQGVEPSAIAALTKQLQPVDFPRGHTVFAEGEPGDRLYIIISGKVKIGRRAPDGRENLLTIMGPSDMFGELSIFDPGPRTSSATTITEVRAVSMDRDALRSWIADRPEISEQLLRVLARRLRRTNNNLADLIFT
G6	MBLE	[Mycobacterium leprae TN]	putative Crp/Fnr-family transc	[15828234]	MDEILARAGIFQGVEPSAVAALTKQLQPVDFPRGHTVFTEGEPGDRLYIIVSGKVKIGRRSPDGRENLLTIMGPSDMFGELSIFDPGPRTSSATTITEVRAVSMDRDALRAWIADRPEIAEQLLRVLARRLRRTNNNLADLIFT
G6	MBAV1	[Mycobacterium avium subsp. paratuberculosis K-10]	hypothetical protein MAP0398	[41406496]	MDEILARAGIFQGVEPSAVTALTKQLQPVDFPRGHTVFAEGEPGDRLYIIVSGKVKIGRRSPDGRENLLTIMGPSDMFGELSIFDPGPRTSSATTITEVRAVSMDRDALRAWIADRPEIAEQLLRVLARRLRRTNNNLADLIFT
G6	NOFA	[Nocardia farcinica IFM 10152]	putative transcriptional regula	[54022310]	MDEALARAGIFQGVEPTAVAALAKQLQPVDFPRGHTVIFNEGEPGDRLYIITSGKVKIGRRSPDGRENLLTIMGPSDMFGELSIFDPGPRTSSTATTVTEVRAVTMDRDALKSWIDQRPEIAEQLLRVLARRLRRTNNNLADLIFT
G6	CBGL	[Corynebacterium glutamicum ATCC 13032]	cAMP-binding domain contain	[19551541]	MEGVQEILSRAGIFQGVDPTAVNNLIQDMETVRFPRGATIFDEGEPGDRLYIITSGKVKLARHAPDGRENLLTIMGPSDMFGELSIFDPGPRTSSAVCVTEVHAATMNSDMLRNWVADHPAIAEQLLRVLARRLRRTNASLADLIFT
G6	COEF	[Corynebacterium efficiens YS-314]	putative regulatory protein	[25026843]	MEGVQDILSRAGIFQGVDPTAVNNLIQDMETVRFPRGATIFEEGEPGDRLYIITSGKVKLARHANDGRENLLTIMGPSDMFGELSIFDPGPRTSSAVCVTEVHAATMNSEMLRNWVSDHPAIAEQLLRVLARRLRRTNASLADLIFT
G6	CYDT	[Corynebacterium diphtheriae NCTC 13129]	Putative transcription regulato	[38232924]	MEGVQDILSRAGIFQGVDPIAVNNLIEDLESVRFPRGATIFEEGEPGDRLYIITSGKVKLARHSMDGRENLLTVMGPSDMFGELSIFDPGPRTSSAVCVTEVHAATMNSEMLRQWIAEHPEISEQLLRVLARRLRRTNNSLADLIFT
G6	MBAV2	[Mycobacterium avium subsp. paratuberculosis K-10]	hypothetical protein MAP0082	[41406180]	MDDILAQATIFQGIDPEAVAALSRHLQHVSFPRRRTVFVEGEPGDDLYVIVSGAVKIRHQTADGRETVFAVLGPGDVFGELALFDPGPRTSVTILTEVEALRMDRNALRTWIIERPEIAEQLLRVLARRLHRTNNTLCDLIFT
total	9				
G1	PSAE	[Pseudomonas aeruginosa PAO1]	transcriptional regulator Vfr	[15595849]	MVAITHTPKLKHLDKLLAHCHRRRYTAKSTIIYAGDRCETLFFIIKGSVTILIEDDDGREMIIGYLNSGDFFGELGLFEKEGSEQERSAWVRAKVECEVAEISYAKFRELSQQDSEILYTLGSQMADRLRKTTRKVGDLAFLD
G1	PSPU	[Pseudomonas putida KT2440]	catabolite gene activator Crp	[26987165]	MVASALPAKIKNIDKLLVHCQRRRYTAKSNIICAGDRAETLSFIIKGSVTILIEDDDGHEMIIAYLNHGDFFGELGLFEPVDGEQQRSAWVRAKTECEVAEISYEKFRELARQDPEILYALGSQMAQRLRNTTRKVGDLAFFD
G1	PSSY	[Pseudomonas syringae pv. tomato str. DC3000]	catabolite gene activator Crp	[28867824]	MVAISFTPKANGLDKLLPHAERRRCPAKTNIISAGDRSESLFFILKGTVTVLIEDDEREMIVAYLNNGDFFGELGLFEQPGHESVRSAWVRAKTECEVAEIPYETFRELTRKDPEILYAIGSQMAERLRNTTRKVGDLAFLD
G1	MHSU	[Mannheimia succiniciproducens MBEL55E]	Crp protein	[52425989]	MLEQVNAHQTNVLPQTQPVQPASPMDPTLDWFLSHCHIHKYPAKTTLIHAGERADTLYYIVKGSAAVMVKDEEGKEMILSYLSQGEFFGEVGLFEEGQVRSAWVKAKNACEIAEVSYKKFRQLLQVNPEILMYLSAQLSRRLQNTSKQVSNLAFLD
G1	PSML	[Pasteurella multocida subsp. multocida str. Pm70]	Crp	[15603022]	MQTTPSIDPTLEWFLSHCHIHKYPSKSTLIHAGEKAETLYYLIKGSVAVLVKDEDGKEMILTYLSQGDFFGEAGLFEEGQLRSAWIKAKSPCEIAEISYKKFRQLIQVNPDILMHLSAQLARRLQNTSRQVSNLAFLD
G1	VIPA	[Vibrio parahaemolyticus RIMD 2210633]	cyclic AMP receptor protein	[28899567]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHAGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEDQERTAWVRAKSPCEVAEISFKKFRQLIQVNPDILMRLSAQMARRLQVTSQKVGDLAFLD
G1	VIVU	[Vibrio vulnificus CMCP6]	cAMP-binding proteins - catab	[27364717]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHAGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEDQERTAWVRAKSPCEVAEISFKKFRQLIQVNPDILMRLSAQMASRLQVTSQKVGDLAFLD
G1	PTPR	[Photobacterium profundum SS9]	putative cyclic AMP receptor p	[54307514]	MVPGKPQTDPTLEWFLSHCHIHKYPSKSTLIHAGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEDQERTAWVRAKSPCEVAEISFKKFRQLIQVNPDILMRLSSQMATRLQVTSQKVGDLAFLD
G1	VICB	[Vibrio cholerae O1 biovar eltor str. N16961]	cyclic AMP receptor protein	[15642609]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHAGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERTAWVRAKTPCEVAEISFKKFRQLIQVNPDILMRLSAQMARRLQVTSQKVGDLAFLD
G1	VIFI	[Vibrio fischeri ES114]	catabolite gene activator Vfr	[59712887]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHAGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKSPCEVAEISFKKFRQLIQVNPDILMRLSGQMATRLQITSQKVGDLAFLD
G1	YEPS	[Yersinia pestis CO92]	cAMP-regulatory protein	[16120516]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSSQMANRLQITSEKVGNLAFLD
G1	YEPT	[Yersinia pseudotuberculosis IP 32953]	cAMP-regulatory protein	[51598019]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSSQMANRLQITSEKVGNLAFLD
G1	ERCA	[Erwinia carotovora subsp. atroseptica SCRI1043]	cyclic AMP receptor protein	[50122984]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSAQMANRLQITSEKVGNLAFLD
G1	ECCF	[Escherichia coli CFT073]	Catabolite gene activator	[26249956]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSAQMARRLQVTSEKVGNLAFLD
G1	SHFL	[Shigella flexneri 2a str. 2457T]	cyclic AMP receptor protein	[30065357]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSAQMARRLQVTSEKVGNLAFLD
G1	ECK1	[Escherichia coli K12]	DNA-binding transcriptional du	[16131236]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSAQMARRLQVTSEKVGNLAFLD
G1	EC57	[Escherichia coli O157:H7 str. Sakai]	cyclic AMP receptor protein	[15833462]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSAQMARRLQVTSEKVGNLAFLD
G1	SMTY	[Salmonella typhimurium LT2]	catabolite activator protein	[16766754]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSSQMARRLQVTSEKVGNLAFLD
G1	SMPI	[Salmonella enterica subsp. enterica serovar Paratyphi A str. ATCC 9150]	cyclic AMP receptor protein,ca	[56415384]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSSQMARRLQVTSEKVGNLAFLD
G1	SMCH	[Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67]	catabolite activator protein (C	[62181970]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSSQMARRLQVTSEKVGNLAFLD
G1	SMTI	[Salmonella enterica subsp. enterica serovar Typhi str. CT18]	cyclic AMP receptor protein,ca	[16762818]	MVLGKPQTDPALEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERSAWVRAKTACEVAEISYKKFRQLIQVNPDILMRLSSQMARRLQVTSEKVGNLAFLD
G1	PHLU	[Photorhabdus luminescens subsp. laumondii TTO1]	Catabolite gene activator (cAM	[37524404]	MVLGKPQTDPTLEWFLSHCHIHKYPSKSTLIHQGEKAETLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEGQERTAWVRAKSACEVAEISYKKFRQLIQVNPDILMRLSAQMANRLQTTSEKVGNLAFLD
G1	SWON	[Shewanella oneidensis MR-1]	catabolite gene activator	[24372215]	MALIGKPKPDPTLEWFLSHCHIHKYPSKSTLIHAGEDSDTLYYIVKGSVAVLIKDEEGKEMILSYLNQGDFIGELGLFEEQAERTAWVRAKQACEIAEISYKKFKQLIQVNPEILMKLSAQMAYRLQSTSQKVGNLAFLD
G1	HMIF	[Haemophilus influenzae Rd KW20]	catabolite gene activator	[16272895]	MSNELTEIDEVVTSSQEEATQRDPVLDWFLTHCHLHKYPAKSTLIHAGEDATTLYYVIKGSVMVSSKDDEGKEMILTYLGAGQFFGEAGLFDEGSKRSAWVKTKTTCEIAEISYKKYRQLIQANPEILMFLTAQLARRLQNTSRQVTNLAFLD
G1	HADU	[Haemophilus ducreyi 35000HP]	cyclic AMP receptor protein	[33151907]	MQMQDISISTPRLDGEVSPPPSIHDPAVEWFLTHCHIHRYPSKYTLIHAGEEANSLFYIVSGSVAVFIKDDEGKEMILSHLGQGEFVGEISLFEENTQPRTAWVKTKDACEIAEISYKKFKQLVHVNPDILMYLAAQMACRLRQTSRQVSNLAFLD
	25				
G7	Nost	[Nostoc sp. PCC 7120]	nitrogen-responsive regulatory	[17231884]	MIVTQDKALANVFRQMATGAFPPVVETFERNKTIFFPGDPAERVYFLLKGAVKLSRVYEAGEEITVALLRENSVFGVLSLLTGNKSDRFYHAVAFTPVELLSAPIEQVEQALKENPELSMLMLRGLSSRILQTEMMIETLAHRD
G7	SYNC	[Synechocystis sp. PCC 6803]	global nitrogen regulator	[16330603]	MDQSLTQDRPLAAVFRRLGSELMPPVVETFDRSKTIFFPGDPAERVYFLLKGAVKLSRVYEAGEEITVALLRENSVFGVLSLVTGQRSDRFYHAVAFTPVELLSAPIEQVEQALKEHPDLSLLMLQGLSSRILQTEMMIETLAHRD
G7	TSEL	[Thermosynechococcus elongatus BP-1]	global nitrogen regulator	[22299193]	MPTGPYAIATLSIESPWIDRRLMNKDGPLAPVFRHMASGLFPSTTETYERGKTIFFPGDPAEKVYFLLKGAVKLSRVYEAGEEITVALLRENTVFGVLSLITGTRSDRFYHAVALTEPVELLAVPIEQVEKAMHEDPDLPMFMIQGLSSRILQTEMMIETLAHRD
G7	SCEL	[Synechococcus elongatus PCC 6301]	2-oxoglutarate-dependent tran	[56751386]	MLANENSLLTMFRELGSGKLPLQIEQFERGKTIFFPGDPAERVYLLVKGAVKLSRVYESGEEITVALLRENSVFGVLSLLTGQRSDRFYHAVAFTPVQLFSVPIEFMQKALIERPELANVMLQGLSSRILQTEMMIETLAHRD
G7	GOVI	[Gloeobacter violaceus PCC 7421]	nitrogen-responsive regulatory	[37522854]	MSWDDPFASLFRQAESGSQFPPVVEHYDRGKTIFFPGDPAVRVYFLKRGAVKLSRVYEHGEEITVALLRENSLFGVLSLITGHRSDRFYHAIAFTPVDLIAVGVEQMERIIQEDPDISRLLFQGLASRILQTEMMIETLAHRD
G7	SC	[Synechococcus sp. WH 8102]	Global nitrogen regulatory pro	[33864811]	MVASSGFSRYSPQQPQVAGAAERSRTLLDVIRDLDGASTELVERNKTIFFPGDPAERVYLIRRGAVRLSRVYESGEEITVALLRENSLFGVLSLLTGHRSDRFYHAVAFTRVEMVTAPAASVRAAIEADTGVGLRLLQGLSSRILQTETMIETLTHRD
G7	PCMP	[Prochlorococcus marinus subsp. pastoris str. CCMP1986]	Global nitrogen regulatory pro	[33860804]	MSPASRGFSRLTPQPSGQNTINTIGERFPINRTLMEVIKGLEGASTEMVERSKTIFFPGDPAERVYLIRRGAVRLSRVYESGEEITVALLRENSLFGVLSLLTGHRSDRFYHAIAFTRVEMITAPANSVLRAIEADASVGLLLLQGLSSRILQTETMIETLTHRD
G7	PCMA	[Prochlorococcus marinus subsp. marinus str. CCMP1375]	Global nitrogen regulator	[33239729]	MTGSANSFSRYPSQSLGMQSQNQLNPNTQVKKTLLDVIRGLDGASNEAVERSKTIFFPGDPAERVYLIRRGAVRLTRVYESGEEITVALLRENSLFGVLSLLTCHRSDRFYHAVAFTRVEMSSAPAGSVRNAIEADSGVGLLLLQGLSSRILQTETMIETLTHRD
	8				
G2	SMTY	[Salmonella typhimurium LT2]	transcriptional regulator	[39546323]	MIPEKRIIRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQTLFKAGDELKSLYAIRSGTIKSYTITEQGDEQITGFHLAGDLVGFDAIGSGHHPSFAQALETSMVCEIPFETLDDLSGKMPNLRQQMMRLMSGEIKGDQEMILLLSKKNAEERL
G2	SMPI	[Salmonella enterica subsp. enterica serovar Paratyphi A str. ATCC 9150]	fumarate and nitrate reductior	[56413415]	MIPEKRIIRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQTLFKAGDELKSLYAIRSGTIKSYTITEQGDEQITGFHLAGDLVGFDAIGSGHHPSFAQALETSMVCEIPFETLDDLSGKMPNLRQQMMRLMSGEIKGDQEMILLLSKKNAEERL
G2	SHFL	[Shigella flexneri 2a str. 2457T]	global transcription regulator f	[30062858]	MIPEKRIIRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQTLFKAGDELKSLYAIRSGTIKSYTITEQGDEQITGFHLAGDLVGFDAIGSGHHPSFAQALETSMVCEIPFETLDDLSGKMPNLRQQMMRLMSGEIKGDQEMILLLSKKNAEERL
G2	ECK1	[Escherichia coli K12]	DNA-binding transcriptional du	[16129295]	MIPEKRIIRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQTLFKAGDELKSLYAIRSGTIKSYTITEQGDEQITGFHLAGDLVGFDAIGSGHHPSFAQALETSMVCEIPFETLDDLSGKMPNLRQQMMRLMSGEIKGDQEMILLLSKKNAEERL
G2	EC57	[Escherichia coli O157:H7 EDL933]	transcriptional regulation of ae	[15801834]	MIPEKRIIRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQTLFKAGDELKSLYAIRSGTIKSYTITEQGDEQITGFHLAGDLVGFDAIGSGHHPSFAQALETSMVCEIPFETLDDLSGKMPNLRQQMMRLMSGEIKGDQEMILLLSKKNAEERL
G2	SMTI	[Salmonella enterica subsp. enterica serovar Typhi Ty2]	fumarate and nitrate reductior	[29142004]	MIPEKRIIRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQTLFKAGDELKSLYAIRSGTIKSYTITEQGDEQITGFHLAGDLVGFDAIGSGHHPSFAQALETSMVCEIPFETLDDLSGKMPNLRQQMMRLMSGEIKGDQEMILLLSKKNAEERL
G2	SMCH	[Salmonella enterica subsp. enterica serovar Choleraesuis str. SC-B67]	transcriptional regulation of ae	[62180226]	MIPEKRIIRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQTLFKAGDELKSLYAIRSGTIKSYTITEQGDEQITGFHLAGDLVGFDAIGSGHHPSFAQALETSMVCKIPFETLDDLSGKMPNLRQQMMRLMSGEIKGDQDMILLLSKKNAEERL
G2	ECCF	[Escherichia coli CFT073]	fumarate and nitrate reductio	[26247669]	QSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQTLFKAGDELKSLYAIRSGTIKSYTITEQGDEQITGFHLAGDLVGFDAIGSGHHPSFAQALETSMVCEIPFETLDDLSGKMPNLRQQMMRLMSGEIKGDQDMILLLSKKNAEERL
G2	ERCA	[Erwinia carotovora subsp. atroseptica SCRI1043]	fumarate and nitrate reductior	[50121133]	MIPEKRIIRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQALFKAGDELRSLYAIRSGTIKSYTITEQGDEQITGFHLAGDLVGFDAIGTAQHPSFAQALETSMVCEIPFETLDDLSGKMPNLRQQMMRLMSGEIRGDQDMILLLSKKNAEERL
G2	YEPS	[Yersinia pestis CO92]	fumarate and nitrate reductior	[16122524]	MIPEKRVIRRIQSGGYAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQALFKAGDELKSLYAIRSGTIKSYTITEEGDEQITGFHLAGDLVGFDAISNLQHPSFAQALETSMVCEIPFDTLDDLSGKMPNLRQQIMRLMSGEIKGDQEMILLLSKKNAEERL
G2	YEPT	[Yersinia pseudotuberculosis IP 32953]	fumarate and nitrate reductior	[51596546]	MIPEKRVIRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQALFKAGDELKSLYAIRSGTIKSYTITEEGDEQITGFHLAGDLVGFDAISNLQHPSFAQALETSMVCEIPFDTLDDLSGKMPNLRQQIMRLMSGEIKGDQEMILLLSKKNAEERL
G2	PHLU	[Photorhabdus luminescens subsp. laumondii TTO1]	fumarate and nitrate reductior	[37526095]	MIQEKRVIRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKGQTLFKAGDELKSLYAIRSGTIKSYTITEQGDEQITGFHLAGDLVGFDAILNTQHPSFAQALETSMVCEIPFETLDDLSGKMPNLRQQMMRLMSGEIKGDQEMILLLSKKNAEERL
G2	VIPA	[Vibrio parahaemolyticus RIMD 2210633]	fumarate and nitrate reductior	[28898310]	MISEKPATKRIQSGGCAIHCQDCSISQLCIPFTLNESELDQLDQIIERKKPIQKGQELFKAGDELKSLYAIRSGTIKSYTITEQGDEQITAFHLAGDLVGFDAITGDCHPSFAQALETSMVCEIPYEILDDLSGKMPKLRQQIMRLMSNEIKGDQEMILLLSKKNAEERL
G2	VIVU	[Vibrio vulnificus YJ016]	fumarate and nitrate reductior	[37679863]	MISEKPATKRIQSGGCAIHCQDCSISQLCIPFTLNEAELDQLDQIIERKKPIQKGQELFKAGDDLKSLYAIRSGTIKSYTITEQGDEQITAFHLAGDLVGFDAITGDSHPSFAQALETSMVCEIPYEILDDLSGKMPKLRQQIMRLMSNEIKGDQEMILLLSKKNAEERL
G2	VICB	[Vibrio cholerae O1 biovar eltor str. N16961]	fumarate and nitrate reductior	[15641445]	MISEKPAAKRIQSGGCAIHCQDCSISQLCIPFTLNESELDQLDQIIERKKPIQKGQELFKAGDELRSLYAIRSGTIKSYTITEQGDEQITAFHLAGDLVGFDAITGDQHPSFAQALETSMVCEIPYEILDDLSGKMPKLRQQIMRLMSNEIKGDQEMILLLSKKNAEERL
G2	PTPR	[Photobacterium profundum SS9]	putative fumarate and nitrate r	[54309024]	MISDKLTTKRIQSGGCAIHCQDCSISQLCIPFTLNESELDQLDQIIERKKPIQKGQELFKAGDELKSLFAIRSGTIKSYTITEQGDEQITAFHLAGDLVGFDAINEMMHPSFAQSLETSMVCEIPFEILDDLSGKMPKLRQQIMRLMSHEIKGDQEMILLLSKKNAEERL
G2	HMIF	[Haemophilus influenzae Rd KW20]	fumarate/nitrate reduction reg	[16273330]	FVAGAQSGRKIQSGGCAIHCQDCSISQLCIPFTLNEQELDQLDNIIERKKPIQKSQVLFKAGDSLNSIYAIRSGTIKSYTISESGEEQITSFHLPGDLVGFDAITQMQHPSFAQALETAMVCEIPFDILDDLAGKMPKLRQQILRLMSGEIKGDQDMILLLSKKNAEERL
G2	MHSU	[Mannheimia succiniciproducens MBEL55E]	Crp protein	[52425132]	TPDAAKTGRRIQSGGCAIHCQDCSISQLCIPFTLNEHELDQLDNIIERKKPIQKSQILFKAGDELTSLYAIRSGTIKSYTISETGEEQITSFHLPGDLVGFDAITNMSHPSFAQALETAMVCEIPFDILDDLTGKMPKLRQQILRLMSGEIKGDQDMILLLSKKNAEERL
G2	PSML	[Pasteurella multocida subsp. multocida str. Pm70]	Fnr	[15602533]	LFMDTKIGRKVQSGGCAIHCQDCSISQLCIPFTLNEQELDQLDNIIERKKPIQKSQVLFKAGDPLNSLYAIRSGTIKTYTISETGEEQITSFQLPGDLVGFDAIMNMQHPSFAQALETSMVCEIPFDILDDLSGKMPKLRQQIMRLMSNEIKNDQEMILLLSKKMNAEERL
G2	SWON	[Shewanella oneidensis MR-1]	electron transport regulator a	[24373903]	MTIEQNKNRRSAASGCAIHCHDCSMGTLCIPFTLNANELDQLDDIIERKKPIQKGEKIFKSGDPLKSLFAIRSGTIKSYTITEQGDEQITGFHLAGDVIGFDGIHAQSHQSFAQALETSMVCEIPFNILDELSGTMPKLRQQIMRLMSNEIMSDQEMILLLSKKNAEERL
G2	HADU	[Haemophilus ducreyi 35000HP]	fnr-like transcriptional regulato	[33152491]	MKIISDAKNVGRTACTIHCQNCSISQLCLPFTLNEHELTQLDNIIERKKPVQKSQIIFQSGDELRSIYAIRSGTIKSYTISESGEEQITAFHLPGNLVGFDAITNMRHLGFAQALETSMICEIPFDILDDLAGKMPKIRHQIMRLMSNEIKSDQEMILLLSKMNAEERL
G2	PSPU	[Pseudomonas putida KT2440]	transcriptional regulator Anr	[26990956]	MSEPVKLRPHNQAHCKDCSLAPLCLPLSLNLEDMDALDEIVKRGRPLKKGEFLFRQGDSFGSVYAVRSGALKTFSLSDSGEEQITGFHLPSELVGLSGMDTEAYPVSAQAQETTSVCEIPFERLDELSVQLPQLRRQLMRVMSREIRDDQQMMLLLSKKTADERI
G2	PSSY	[Pseudomonas syringae pv. tomato str. DC3000]	transcriptional activator Anr	[28869195]	MSEPVKLRAHTQAHCKDCSLAPLCLPLSLNTDDMDSLDQIVKRGRPLKKGEFLFRQGDTFESVYAVRSGALKTFNISDNGEEQLTGFHLPSELVGLSGMDAEAYPVSAQAETTTSVCEIPFERLDELSVQLPQLRRQLMRVMSREIRDDQQMMLLLSKKTADERI
G2	PSAE	[Pseudomonas aeruginosa PAO1]	transcriptional regulator Anr	[15596741]	MAETIKVRALPQAHCKDCSLAPLCLPLSLTVEDMDSLDEIVKRGRPLKKGEFLFRQGDPFGSVFAVRSGALKTFSITDAGEEQITGFHLPSELVGLSGMDTETYPVSAQALETTSVCEIPFERLDELSEQLPQLRRQLMRLMSREIRDDQQMMLLLSKKTADERI
	24				
G3	BOPA	[Bordetella parapertussis 12822]	transcriptional regulatory prot	[33596440]	MQRRVPLSPDAAHCSSCMLGHVCVPVGMPANEVEKLDELVKERVRVERGKSLYELDDPLDAVYGVRFGSLKTQLEDSSGQLQITGFHLPGEIVGLDGMIESKHVSSAVALEDSEVCVIRLPEIDRVSTQLPSLQQQFRRLMSREITRSHQMLATVGAMRSEQRL
G3	BOBR	[Bordetella bronchiseptica RB50]	transcriptional regulatory prot	[33602269]	MQRRVPLSPDAAHCSSCMLGHVCVPVGMPANEVEKLDELVKERVRVERGKSLYELDDPLDAVYGVRYGSLKTQLEDSSGQLQITGFHLPGEIVGLDGMIESKHVSSAVALEDSEVCVIRLPEIDRVSTQLPSLQQQFRRLMSREITRSHQMLATVGAMRSEQRL
G3	BOPE	[Bordetella pertussis Tohama I]	transcriptional regulatory prot	[33592327]	MQRRVPLSPDAAHCSSCMLGHVCVPVGMPANEVEKLDELVKERVRVERGKTLYELDDPLDAVYGVRFGSLKTQLEDSSGQLQITGFHLPGEIVGLDGMIESKHVSSAVALEDSEVCVIRLPEIDRVSTQLPSLQQQFRRLMSREITRSHQMLATVGAMRSEQRL
G3	RASO1	[Ralstonia solanacearum GMI1000]	putative fumarate and nitrate r	[17546002]	MLTRIALTDQTSRCSTCVLGQICLPVGMNADDVRKMDDLVTERIRIKKGQVLYALGEPLEAVYGIRFGTLKTHLTLEDGRHQITGFHLPGEIVGLDGIGEMRHVSDATALEDTEVCVVRYAELQQLSRRLPSLQHQFLRIMSKEISQDLHMLLTLGSMRAEERL
G3	NEMN	[Neisseria meningitidis MC58]	transcriptional regulator, Crp/F	[15676294]	MASHNTTHQMKTLCSSCSLRELCLPVGLLPNELSQLDAVIRQSRRLKKGEYLFCVGEAFTSLFAISGFFKTTVASQDGRDQVTGFFMSGELIGMDGICSHVHSCDAVALEDSEVCELPFTHIEELGQNIPSLRTHFFRMMSREIVRDQGVMLLLGNMRAEERI
G3	NEGN	[Neisseria gonorrhoeae FA 1090]	putative anaerobic transcriptic	[59801908]	MASHNTTHQMKTLCSSCSLRELCLPVGLLPNELSQLDAVIRQSRRLKKGEYLFRAGGAFTSLFAISGFFKTTVASQDGRDQVTGFFMSGELIGMDGICSYVHSCDAVALENSEVCELPFTHIEELGQNIPSLRTHFFRMISREIVRDQGVMLLLGNMRAEERI
G3	CHVI	[Chromobacterium violaceum ATCC 12472]	probable transcription regulato	[34499102]	MSEQNHTSLHTLKISCSSCSLRELCLPVGLNREEMGQLDAVIRQSRRLKKGEYLFRSGESFKSLYAIRTGFFKTCVASQDGREQVTGFLMSGELMGLDGISSSNHSCDAVALEDSEVCELPFTRMESLCRDIPSLQHHFYRLMSREIVRDQNVMLLLGNMKAEERI
G3	AZOA	[Azoarcus sp. EbN1]	transcription factor Fnr	[56478364]	MQMKATVPITVASLKVACSQCNLVELCLPFGMSENEIDRLDELVGARRKIKRQHHLYRAGDPFEAIYAIRAGSFKTNVLLEDGREQVTGFQMTGEMLGLDGISTELHSCNAIALEDSEVCVIAYGKLEELSRVVEGLQHQFHKVMSREIVRDHGVMTLLGSMRAEERL
G3	NIEU	[Nitrosomonas europaea ATCC 19718]	putative anaerobic transcriptic	[30249681]	MHNCAIPGHPTAPQQRSSCTTCSLRELCLPVGLTDEEIEYVGNLTNHKLRVRRGEYLLHAGADFRSLYAVKNGSFKTYTLRVDGRQQVTGFYMTGELLGLDAISPEKHTCDTVALVDSEVCEVPFIKLEEIGRSIPSLMSQFHKIMSREIVQDRGVMMLLGSMRAEERL
G3	PSSY	[Pseudomonas syringae pv. tomato str. DC3000]	transcriptional regulator, Crp/F	[28868872]	MTAHLAAINGLRLKAECSNCGVRELCLVSGMSLTDTDRLSAIIPQPIKVKKGAALYHAGDPLSSLYAVRSGFFKTTISSQDGRDQLTGFQMSGELLGCDAISNDCHVCHAVALEDSEVCPFRFTHLERLSREMPSLQRHLLNNRALSKEIVRDHEMLLLLGNLNAEERM
G3	BUMA1	[Burkholderia mallei ATCC 23344]	transcriptional regulator, Crp/F	[53715989]	AAAAPVHASSTWAPRQAAHCSTCAMRHLCMPQGLAPEALARLESVICTARPVRRGETLFREGDTFDNLYAVRSGSLKTIATRHDGREQVTGLHLAGEALGLDGICDDAHPRTAVALEDSSVCVIPYSALKTLCSEAGTIQLRMHKLMSEQIVRETSQTMVLGSLNAEERV
G3	BUPS1	[Burkholderia pseudomallei K96243]	anaerobic growth regulatory p	[53721070]	AAAAPVHASSTWAPRQAAHCSTCAMRHLCMPQGLAPEALARLESVICTARPVRRGETLFREGDTFDNLYAVRSGSLKTIATRHDGREQVTGLHLAGEALGLDGICDDAHPRTAVALEDSSVCVIPYSALKTLCSEAGTMQLRMHKLMSEQIVRETSQTMVLGSLNAEERV
G3	BUMA2	[Burkholderia mallei ATCC 23344]	transcriptional regulator, Crp/F	[53716734]	VIRLHPVTADDAPKLPATRCSTCAIRNVCMPAGLTPIEFARLDAMVCTTRHVKRGDALFRTHDTFQSIYAVRTGSFKTVVMHRDGREQVTGFQIVGETLGLDGVCGGRHNSDAIALEDSTVCIIPFAQLEAVCREVKPMQHHVHQLMSGEIVRESSLMMLLGTMTAEQRV
G3	BUPS2	[Burkholderia pseudomallei K96243]	crp-family transcriptional regu	[53722938]	VIRLHPVTADDAPKLPATRCSTCAIRNVCMPAGLTPIEFARLDAMVCTTRHVKRGDALFRTHDTFQSIYAVRTGSFKTVVMHRDGREQVTGFQIVGETLGLDGVCGGRHNSDAIALEDSTVCIIPFAQLEAVCREVKPMQHHVHQLMSGEIVRESSLMMLLGTMTAEQRV
G3	BUPS3	[Burkholderia pseudomallei K96243]	fumarate and nitrate reductior	[53722191]	TVPLRAAKRDDAAKRTAARCSSCAMRSVCMPVDLNTHEFAKLDALICSTRQVKRGDTLFRTNDGFQSIYAVRTGSFKTVVMHRDGQEHVTGFQIVGETLGLDGVCAGHHNSDAVALEDSTVCIIPFAQLEAVCREVKPMQHHVHRLMSGEIVRESSQMMLLGTMTAEQRV
G3	RASO2	[Ralstonia solanacearum GMI1000]	probable transcription regulato	[17544909]	QARQTVSPALSSQCANCMMHWVCIAGAVPSTQQERVDRMVHTWRKVGKGEALFRAGDTFHALYAIRSGSFKTVVSHPNGSSHVTGFQLAGETLGLDGIAADRHTCDAIALEDSTVCSMPFHCVEDLCQAIRPLQHRLHQLLAEEIVRESGLMLLARLTAEARV
G3	PSPU	[Pseudomonas putida KT2440]	transcriptional regulator, Crp/F	[26990005]	SPNPLSKANPIHLKPLVNHCMDCRVSGLCLPTGLPVDENSCLGALIGPRMIRIRKGAALFNANDPLTMLYALRCGSFKTSLNTAEGQGVVINFWMPGDVIGLDAIATEHHVCDAIALEDSEVCPVPYRRLQTLAREFPDLQQSLNRLMSREIVREHGRVLMLCNLTAEQRL
	17				
G0	BRSU	[Brucella suis 1330]	transcriptional regulator, Crp/F	[23501541]	MPFALDQNQKLAARTPCKVCEDCCVRSMAVCSALDDGDLAALEAIMTSKKLDTNEMLVEEGEPKLRVYSLTSGMLRIYTSLPDGRRQIAGFLFPGDFLGLADDEVYSQSAEAVVPSALCAFSAKEMERLMERFPKLKERLYQMTRLALRTARDNQLVLGRLAPVEKLAS
G0	BCAB	[Brucella abortus biovar 1 str. 9-941]	transcriptional regulator, Crp/F	[62289614]	MPFALDQNQKLAARTPCNVCEDCCVRSMAVCSALDDGDLAALEAIMTSKKLDTNEMLVEEGEPKLRVYSLTSGMLRIYTSLPDGRRQIAGFLFPGDFLGLADDEVYSQSAEAVVPSALCAFSPKEMERLMERFPKLKERLYQMTRLALRTARDNQLVLGRLAPVEKLAN
G0	BRMT	[Brucella melitensis 16M]	nitrogen fixation regulation pro	[17987577]	VVPFALDQNQKLAARTPCNVCEDCCVRSMAVCSALDDGDLAALEAIMTSKKLDTNEMLVEEGEPKLRVYSLTSGMLRIYTSLPDGRRQIAGFLFPGDFLGLADDEVYSLSAEAVVPSALCAFSAKEIERLMERFPKLKERLYQMTRLALRTARDNQLVLGRLAPVEKLAS
G0	BRJA	[Bradyrhizobium japonicum USDA 110]	transcriptional regulatory prot	[27381172]	MKPSVVMIEPNGHFCSDCAIRTSAVCSSLDAAELREFEHLGRRVHFSSGETVFSEEDITTSFYNVLEGVMRLYKLLPDGRRQIVGFALPGDFLGMNLSGRHNFSADAIGAVTVCQFAKAPFGRFIEERPQLLRRINELAIRELSQARDHMVLLGRRSADEKVAA
G0	RSPA1	[Rhodopseudomonas palustris CGA009]	anaerobic aromatic degradati	[39937294]	MPHLAYPTTTCEGFRCETHCAVRGLAICGELGPADHEFERLAQHVRYGPKEALFSEDEVADSVYSLIEGIARLYKLLPDGRRQIIGFALPGDFLGMAPGNRYSFSADSIGGVTVCKFFRGPFLRFIENRPQMLLRMNDFATRELSLAQDQMLLLGRRSAEEKVAA
G0	GUOX	[Gluconobacter oxydans 621H]	Transcriptional regulator Aadf	[58039438]	MSASHSPEDVHDRCAHCVGRRLSICNSIDDCDLAVLANESSRMTVADGRSFIEEGAPAHDFFVVTGGRVKLFTLLPDGRRQITGFAEGGDFLGAASTSYAFGAEALGSATLCRFPHAGMQRLTERFPSLERHLREEASRELALMQARMTLLGRKTARERVAT
G0	MSLO1	[Mesorhizobium loti MAFF303099]	Fnr-type transcriptional regula	[13475537]	MTVRQDVHSAGIPVLCQSCEARHRGVCGALDPDQLVGLARISSKHTIEPGVELIGDAETVDSYSNVLSGVVKLTKSLSDGRQQIVGLQFAPDFLGRPFKVESAINAEAATAVSLCSFPRASIERMMRESPELEHRLLKQTLNELDEARDWMVTLGRKTAAEKVAS
G0	MSLO2	[Mesorhizobium loti MAFF303099]	transcriptional regulator	[13475362]	MTVRQDIHSSGIPVLCLSCAARLRGVCAALDPNHLVGLAKTSSRHKEPVGEVLIADAEAVDSYSNLLSGVVKLTKCLSDGRQQIVGLQFAPDFLGRPFKVKSTINVEAATSVSLCSFPKAAVERIMKESPGLEHRLLKQTLDELDEALEWMVTLGRKTAPEKVAS
G0	ABTF	[Agrobacterium tumefaciens str. C58]	hypothetical protein AGR_C_2	[15888914]	MDTLQKTIHNSEYPVVCRSCEARHGGLCSTLTPQQLCDLNRHSSRKKLEAGNELLGQGELVTSYGNILNGVVKLSKMMSDGRQQIVGLQFAPDFLGRPFMAESKMTAEAATDVEICLFPRRIVDRMVTEVPDMERKLHSQSLKELDEARDWMLTLGRKSAQEKVAS
G0	SIPO	[Silicibacter pomeroyi DSS-3]	transcriptional activator protei	[56698353]	MSTQPQLCHTSCDDCPIRHRAVCAHCDADELEKLEEIKYYRSFEAGQTVIWSGDQMSFVGSVVSGIATLTQTMEDGRTQMVGLLLPSDFVGRPGRDGAPYDVVATTDLVMCCFRKKPFEELMARTPHIAHRLLEMTLDELDAAREWMVLVGRKTAREKIAS
G0	BRJA1	[Bradyrhizobium japonicum USDA 110]	transcriptional regulator FixK	[27378577]	MFVRITTDASPRPNSLKDLGMTSDSNPIVSLNEFTYKKGTEIYGEKEPADYVYQVKSGAVRSYKLLSDGRRQIGAFHLAGDIFGLENGSEHRFTAEAIVDTTVRLIKRQSLELVAESDAMVARNLLNMTTSNLQHAEDHMLLLGRKTSLERVAA
G0	BRJA2	[Bradyrhizobium japonicum USDA 110]	transcriptional regulator FixK	[27377220]	LVQECRTMFVSITTDHRSLINSLSELGISSGSNPVSLSEFTYKKGTKIYIESEPAEYVYQVRKGAVRTCKLLSDGRRQIGAFHLPGDIFGLENGGVHRFRAEAIQTAVRLIRRQHLETVAENDAVVWRTLLSLTTNNLQHAEDHMLLLGRKTALERVAA
G0	RSPA2	[Rhodopseudomonas palustris CGA009]	transcriptional regulator FixK	[39937310]	MLNQSLRTPATTHAPIDPSSCSPAEQFFATTGHAGLIASEFTYKKEEEIYGEDEPSEYVYQVVSGAVRSYKLLSDGRRQIGAFYLPGDVFGLESGATHRLTAEAIVDTTVRLIKRRSLEQAAATDVNIARGLWSMTAGELRHAEDHMLLLGRKNAMERVAN
G0	BRJA3	[Bradyrhizobium japonicum USDA 110]	transcriptional regulator FixK	[27377868]	MAQMLTQTLKTQVINTQIGGKIAPPHQVSDQFGAITGHVGLVATEFSYRKDEEIYGEDEPAEYVYQVVTGAVRSYKLLSDGRRQIGAFHLPGDVFGLESGPSHRLAAEAIIDTSVRLVKRSSLEKAAGIDVQVARKLWAMTAGELRHAEDHMLLLGRKTAMERVAT
G0	BRJA4	[Bradyrhizobium japonicum USDA 110]	transcriptional regulator FixK	[27382807]	PKAAGFPKGWRLPFVRIFPASQEKHPPPVDYVGPPLADLAGERVFCSEFKYRKGIEIFGEGGHADEYVYQITSGAVRTYKMLPDGRHQINAFHLQGDMFGFENGVAHRFSADALVETHVRIIKRGSILDAMQNRVVSTKNLICFVTQNLQHAEDHMLLLGRKTSLERVAA
G0	CBCR	[Caulobacter crescentus CB15]	transcriptional regulator FixK	[16125005]	MLSPIRAKTFENDYSLDAVLPAPGACARFHRDEEIFGEGEAADYVYQVISGAVRTYRILRDGRRQIDEFHFAGDYFGLEMTETHRINAEAMSDATVRMIRRGALSDLAAQRGDAARALFRLTAEGLQRCQDHVLMLGRRSAQERVVG
G0	SRML1	[Sinorhizobium meliloti 1021]	transcriptional regulator FixK	[16262861]	MYAAAQAKPQSIEVEHLGPAPMSGPHLVATYKPGREIYAQGDLNDKCYQVSTGAVRVYRLLSDGRRQVVSFHLPGEMFGFEAGSNHSFFAEAITETTLAIFGRRNMQERSRELLALALTGMARAQQHLLVIGRQCAVERIAAFLVDLCE

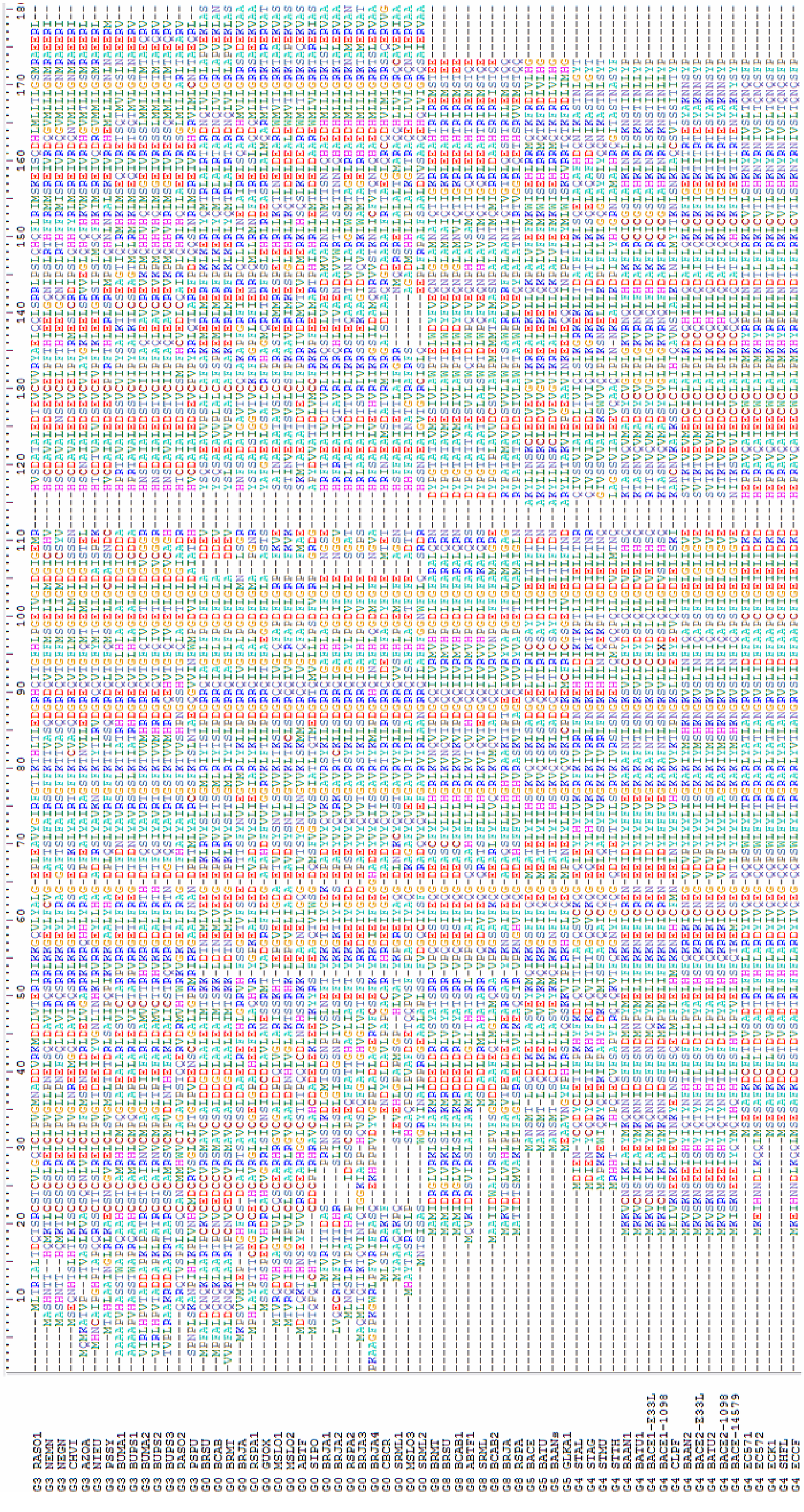
2

Anexo 7.1 Tabla de secuencias de la familia Crp/Fnr

G0	MSLO3	[Mesorhizobium loti MAFF303099]	transcriptional regulator FixK	[13475493]	MHAYTASRISLPSHSAQPSLPVAFDSETQQPVSFFPAGSEIYAQGEKAGAFYQVEFGAVRIYRLLADGRRQISAFHLAGETFGFEADTTHHFFAEAINATGVRAFRFSAGADMSHHLLPLALKGLTRAQEHLLVLGRQNAIERVAAFLADMAE
G0	SRML2	[Sinorhizobium meliloti 1021]	transcriptional regulator FixK	[16262801]	MNTSISASWGRYINAIPPRSGEAALAPLSSFVDGQCIYSFGERADQVYQVEFGAVRVYRLAANGRRQILAFHFGGNWFGLQSRDRHSSNAEAIGVTGVRCISLQEEPLFRPALFSAALDNVSAAQEHQLVIGRQSAIERVAAFLLEMSE
	19				
G8	BRMT	[Brucella melitensis 16M]	transcription regulator, crp fan	[17989331]	MAMIDRGLVRKLSLFAKMRDDELDRLVSYATSRRVPQGESIFEQGDEASCFYLLLHGRLKVNQVTPDGQQIIVRMVHPGDLFGFARALQRNDYPGTATAVMESIVLAWPTELWDYFVEQNPGLAMNAIQTIGKRLEEAHTRIREMSTEE
G8	BRSU	[Brucella suis 1330]	transcriptional regulator, Crp/I	[23500019]	MAMIDRGLVRKLSLFAKMRDDELDRLVSYATSRRVPQGESIFEQGDEASCFYLLLHGRLKVNQVTPDGQQIIVRMVHPGDLFGFARALQRNDYPGTATAVMESIVLAWPTELWDYFVEQNPGLAMNAIQTIGKRLEEAHTRIREMSTEE
G8	BCAB1	[Brucella abortus biovar 1 str. 9-941]	transcriptional regulator, Crp/I	[62317811]	MAMIDRGLVRKLSLFAKMRDDELDRLVSYATSRRVPQGESIFEQGDEASCFYLLLHGRLKVNKVTPDGQQIIVRMVHPGDLFGFARALQRNDYPGTATAVMESIVLAWPTELWDYFVEQNPGLAMNVIQTIGKRLEEAHTRIREMSTEE
G8	ABTF1	[Agrobacterium tumefaciens str. C58]	hypothetical protein AGR_L_9	[15890603]	MQVKIDRSVIRSLALFAKMADDELDGLLTHATSRLVPPGEAIFEQGQSAAHFFLLIHGRLKVTQVTHDGQQIIVRMVHPGDLFGFARALQRSDYPGTATTAAESVILSWQTDLWPQFVEQNPHLAVSAMQTIGQRLEEAHTRIREMSTQE
G8	SRML	[Sinorhizobium meliloti 1021]	putative crp/fnr-like transcripti	[16263132]	MSDADLDRMLAHATARRVPQGDAVFEQGQRATSFFLLLHGRLKVTQVTEDGQQIIVRVVHPGDLFGFAKALQRSDYPGTATAATESLALSWPTDLWPQFVEQNPHLAVSTMQTIGQRLEEAHTRIREMSTQE
G8	BCAB2	[Brucella abortus biovar 1 str. 9-941]	NnrR, transcriptional regulator	[62317794]	MAAIIDWALVRSLPVFSGMDDAAFEQLMGLAMQRAVPKGSAVFKQGEEARLFYLLVQGRLKVMQVTADGQQLIVRVVNPGEFFGFAMALGREDYPGTPLAVVDCSVIAWPMEMMTAFMASNPALAVNTMQMIGRRLDAAHSRIREISTQE
G8	BRJA	[Bradyrhizobium japonicum USDA 110]	transcriptional regulatory prote	[27382195]	MAKVDTSLVAHLPLFAGVTPEALDEILREARSARYPKNSAIFEQGADAQSFFLLLHGHVRAAKTTPTGEQIVVRYVAPGETFGLAMAIGAGRYPATAMAVDDSVVLIWPTSAWPRLVERFPALAANTLQTVGTRLQESHTRILEMSTQQ
G8	RSPA	[Rhodopseudomonas palustris CGA009]	putative nitric oxide responsiv	[39937203]	MATIDTSMVAKLPLFAGLSRAELDAVLKEAQSIRVPKNGHVFEQGEDAHSFFVLLHGHVRASKVTPAGEQIVVRYVAPGETLGVAMAIGLNRYPATATAVDDSIVLAWPTAAWPRLVEQYPALATNTLRTVGGRLQETHSRVVEMSTQQ
	8				
G5	BACE	[Bacillus cereus ATCC 14579]	Transcription regulator, Crp fa	[30020257]	MANSMTLSQDLKELLASVEYKMQIKKGSFIFQEGMEATELYIIHSGKVQISKLSADGQELTLRICSAYDIIGELTLFTDNAKYLLNSKCLEDVEVGVIKREALEKELLQKPALVFEFMKWISEHLRRMQTKFRDLVLHG
G5	BATU	[Bacillus thuringiensis serovar konkukian str. 97-27]	transcriptional regulator, Crp/I	[49477526]	MANSMTLSQDLKELLASVEYKMQIKKGSFIFQEGMEATELYIIHSGKVQISKLSADGQELTLRICSAYDIIGELTLFTDNAKYLLNSKCLEDVEVGVIKREALEKELLQKPALVFEFMKWISEHLRRMQTKFRDLVLHG
G5	BACE	[Bacillus cereus E33L]	transcriptional regulator, Crp/I	[52143301]	MANSMTLSQDLKELLASVEYKMQIKKGSFIFQEGMEATELYIIHSGKVQISKLSADGQELTLRICSAYDIIGELTLFTDNAKYLLNSKCLEDVEVGVIKREALEKELLQKPALVFEFMKWISEHLRRMQTKFRDLVLHG
G5	BAANs	[Bacillus anthracis str. Sterne]	transcriptional regulator, Crp/I	[49184995]	MANSMTLSQDLKELLASVEYKMQIKKGNFIFQEGMEATELYIIHSGKVQISKLSADGQELTLRICSAYDIIGELTLFTDNAKYLLNSKCLEDVEVGVIKREALEKELLQKPALVFEFMKWISEHLRRMQTKFRDLVLHG
G5	GLKA	[Geobacillus kaustophilus HTA426]	transcriptional regulator of ana	[56419303]	MEAAVNGYFPHLRSLLQSSKKVIPLRKHSFLFQEGMPANELYLILSGKVQVSKICPDGKEMCFRICGSGEIVGELTLFTNDARYLLTAKVIEPGEAAVINKEELEQQLLINPPLTLEYMLWISKHVRRTQTKFRDLIMHG
	5				
G4	STAL	[Streptococcus agalactiae 2603V/R]	cyclic nucleotide-binding doma	[22538032]	MDIENYCQDYCLETIFPKHYFKDLQVITLTAGQSVCKQGEQLEYLHYIVKGRFKIVRRLFNGKEHILDIKTKPTLIGDIELLTNRQIVSSVIALEDLTVIQLSLKGRKEKLLTDATFLLKLSQELAQAFHDQNIKASTNLGYT
G4	STAG	[Streptococcus agalactiae NEM316]	hypothetical protein gbs1882	[25011921]	MDIENYCQDYCLETIFPKHYFKDLQVITLTAGQSVCKQGEQLEYLHYIVKGRFKIVRRLFNGKEHILDIKTKPTLIGDIELLTNRQIVSSVIALEDLTVIQLSLKGRKEKLLTDATFLLKLSQELAQAFHDQNIKASTNLGYT
G4	STMU	[Streptococcus mutans UA159]	putative transcriptional regula	[24378661]	MAPKKEWIQKYQLEKIFPAHYFDKLQMISFSAGQAICQQGEELQALSYFVKGKIKIVRRLFNGKEHILNIQEKPTIIGDIELLTNQGIVSSVVALEKSWVVQLPLKSNKELLLKDPLFLLKLGQGLAYSLYQQNIKASTNLSYT
G4	STTH	[Streptococcus thermophilus LMG 18311]	transcriptional regulator, putat	[55820250]	MRHHTIHPQLEQVFSAPYLRQLQKVTFQKGNYIGTQGQDITELTYILSGRVKIVRSLSNGKEHILQMPQQPQLLGDVELMTNQLAGSSVIALEEVQAAQLPLNNKEALLKDPIFLYHIGRNLAMALHRQGITASTNASYF
G4	BAAN1	[Bacillus anthracis str. Ames]	cyclic nucleotide-binding doma	[30263413]	MKKVCNSIKLAEYMKQNNIDSFFSNDMNPYMELIFFKKNEFICRENEEIDYLYFFVEGKAKAFNTLSNGKSVLLCFYDSLQLLGDVELIHSQKTASNVQVMADSYCVGLPLGKVRNQLFHDAKFLRCICGSLAHKLNRLSKNSTINLLYP
G4	BATU1	[Bacillus thuringiensis serovar konkukian str. 97-27]	possible transcriptional regula	[49478992]	MKKVCNSIKLAEYMKQNNIDSFFSNDMNPYMELIFFKKNEFICRENEEIDYLYFFVEGKAKAFNTLSNGKSVLLCFYDSLQLLGDVELIHSQKTASNVQVMADSYCVGLPLGKVRNQLFHDAKFLRCICGSLAHKLNRLSKNSTINLLYP
G4	BACE1-E33	[Bacillus cereus E33L]	possible transcriptional regula	[52142078]	MKKVCNSIKLAEYMKQNNIDSFFSNDMQPYMELIFFKKNEFICRENEEIDYLYFFVEGKAKAFNTLSNGKSVLLCFYDSLQLLGDVELIHSQRITSNVQVMADSYCVGLPLGKVRNQLFHDAKFLRCICGSLAHKLNRLSKNSTINLLYP
G4	BACE1-109	[Bacillus cereus ATCC 10987]	cyclic nucleotide-binding doma	[42782514]	MKKICNSIKLAEYMKQNNIDSFFSNDMKPYMELIFFKKNEFICRENEEIDYLFFFVEGKAKAFNTLSNGKSVLLCFXDSLQLLGDVSLIHSQKTASNVQVMADSYCVGLPLGKVRNQLFHDAKFLRCICGSLAHKLNRLSKNSTINLLYP
G4	CLPF	[Clostridium perfringens str. 13]	probable transcriptional regula	[18309329]	MLLVKNEEKLLEVINKYELNSIFSQEIMPFLELHMFKKNEHICFMNEELNYLYFFVHGKAKVYKTLPNGKSLLLEFYEPVQLIGDLELISKIKANCNVKALKSSLCIAIHIQAVNSIALKDLNFLMYITKNLSNKLAQCSVTSSITITYP
G4	BAAN2	[Bacillus anthracis str. Ames]	cyclic nucleotide-binding doma	[30265430]	MKVSKNSEEISHYIQTYNFHTLFSFDILPYAELHSFQKREKICNEGVDVPYLYYLISGKAKIYMSHKNGKVSLINFIQAPSFIGELGLIGVESVTKTVEVMEDCICLALPLKDCQHILLQDAIFLQKLCKFIGEKTITRTESYAKNNSYP
G4	BACE2-E33	[Bacillus cereus E33L]	cyclic nucleotide-binding doma	[52145218]	MKVSKNSEEISHYIQTYNFHTLFSFDILPYAELHSFQKREKICNEGVDVPYLYYLISGKAKIYMSHKNGKVSLINFIQAPSFIGELGLIGVESVTKTVEVMEDCICLALPLKDCQHILLQDAIFLQKLCKFIGEKTITRTESYAKNNSYP
G4	BATU2	[Bacillus thuringiensis serovar konkukian str. 97-27]	cyclic nucleotide-binding doma	[49480446]	MKVSKNSEEISHYIQTYNFHTLFSYDILPYAELHSFQKREKICNEGVDVPYLYYLISGKAKIYMSHKNGKVSLINFIQAPSFIGELGLIGVESVTKTVEVMEDCICLALPLKDCQHILLQDATFLQKLCKFIGEKTITRTESYAKNNSYP
G4	BACE2-109	[Bacillus cereus ATCC 10987]	cyclic nucleotide-binding doma	[42784583]	MKVSRNSEEISHYIQTYNFHTLFSFDILPYAELHSFQKREKICNEGVDVPYLYYLIAGKAKIYMSHKNGKVSLINFIQAPSFIGELGLIGVESITKTVEVIEDCICLALPLKDCQHLLLQDATFLQKLCKFIGEKTITRTESYAKNNSYP
G4	BACE-1457	[Bacillus cereus ATCC 14579]	Catabolite gene activator	[30023439]	MKISKNEEEIYQYMQIHQFHTLFSFHVLPYVELHSFQTKEMICSEGNTLPYLYYLISGKAKIYMSHKNGKVSLINFIQAPSFIGELGLIGVENITKSVEVLEDCVCLALPLKDCQQLLLQDATFLQHLCKFIGEKTITRTENYAKNYSYP
G4	EC571	[Escherichia coli O157:H7 EDL933]	putative transcriptional regula	[15802719]	MSESAFKDCFLTDVSADTRLFHFLARDYIVQEGQQPSWLFYLTRGRARLYATLANGRVSLIDFFAAPCFIGEIELIDKDHEPRAVQAIEECWCLALPMKHYRPLLLNDTLFLRKLCVTLSHKNYRNIVSLTQNQSFP
G4	ECK1	[Escherichia coli K12]	DNA-binding transcriptional ac	[49176193]	MSESAFKDCFLTDVSADTRLFHFLARDYIVQEGQQPSWLFYLTRGRARLYATLANGRVSLIDFFAAPCFIGEIELIDKDHEPRAVQAIEECWCLALPMKHYRPLLLNDTLFLRKLCVTLSHKNYRNIVSLTQNQSFP
G4	EC572	[Escherichia coli O157:H7 str. Sakai]	putative transcriptional regula	[38704059]	MKEIHNNDLKQQLMSESAFKDCFLTDVSADTRLFHFLARDYIVQEGQQPSWLFYLTRGRARLYATLANGRVSLIDFFAAPCFIGEIELIDKDHEPRAVQAIEECWCLALPMKHYRPLLLNDTLFLRKLCVTLSHKNYRNIVSLTQNQSFP
G4	SHFL	[Shigella flexneri 2a str. 2457T]	putative transcriptional regula	[30063604]	MSESAFKDCFSTDVSADTRLFHFLARDYIVQEGQQPSWLFYLTRGRARLYATLANGRVSLIDFFAAPCFIGEIELIDKDHEPRAVQAIEECWCLALPMKHYRPLLLNDTLFLRKLCVTLSHKNYRNIVSLTQNQSFP
G4	ECCF	[Escherichia coli CFT073]	Regulatory protein nsr	[26248545]	MKEIHNNDLKQQLMSESAFKDCFSTDVSADTRLFHFLARDYIVQEGQQPSWLFYLTRGRARLYVTLANGRVSLIDFFAAPCFIGEIELIDKDHEPRAVQAIEECWCLALPMKHYRPLLLNDTLFLRKLCVTLSHKNYRNIVSLTQNQSFP
	19				

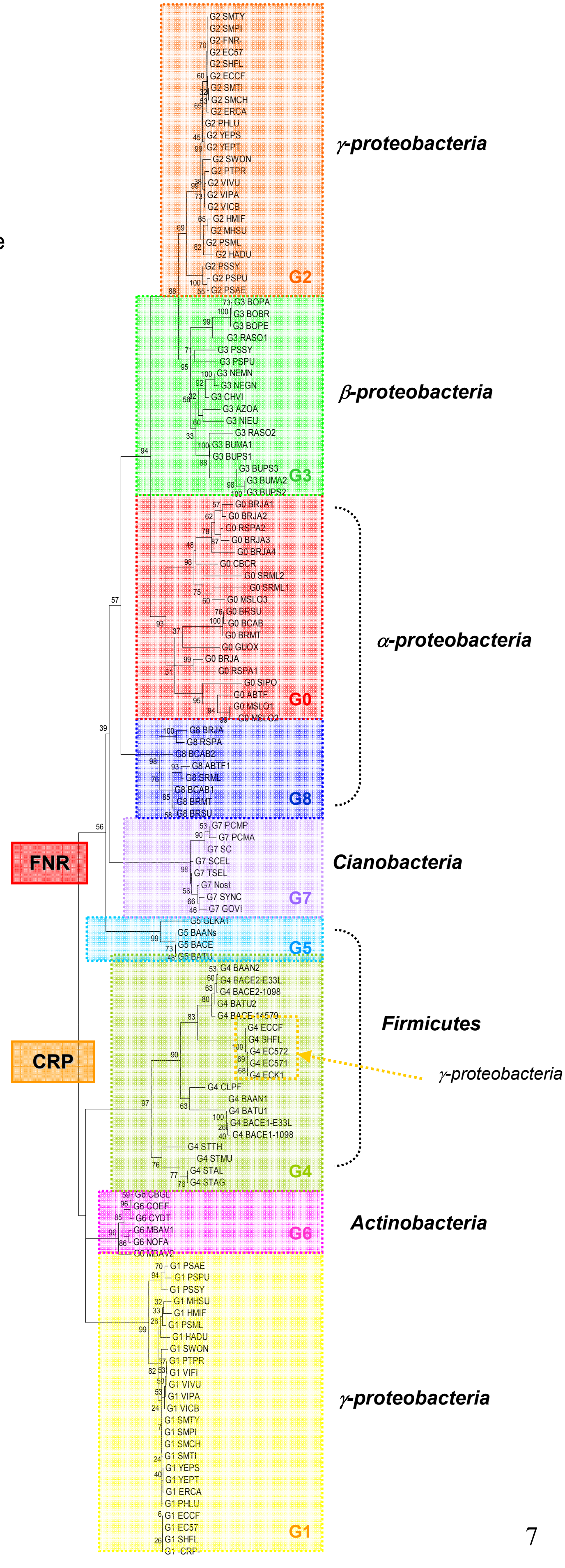
Anexo 7.2 Alineamiento de las secuencias de la familia Crp/Fnr

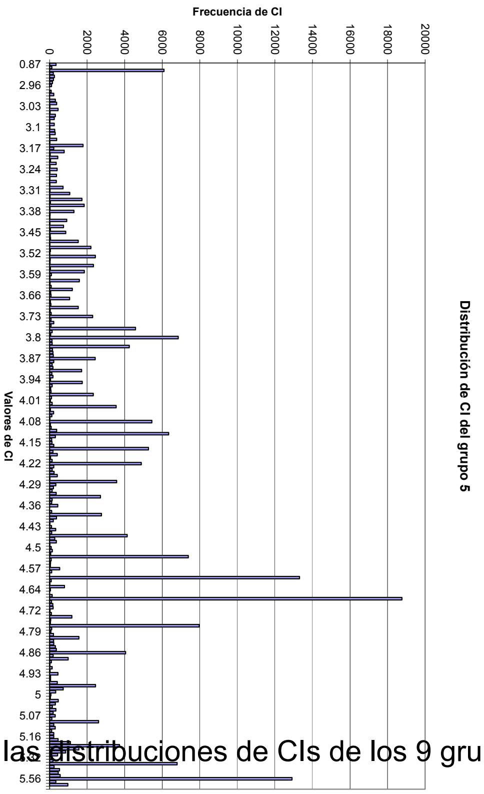
Anexo 7.2 Alineamiento de las secuencias de la familia Crp/Fnr

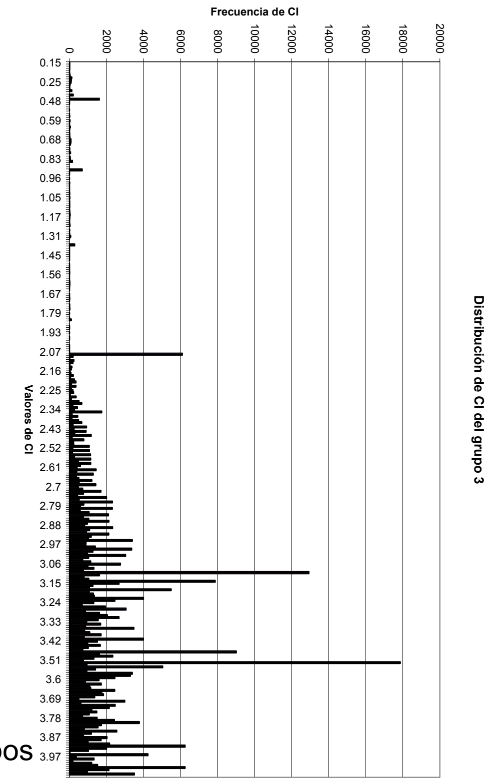


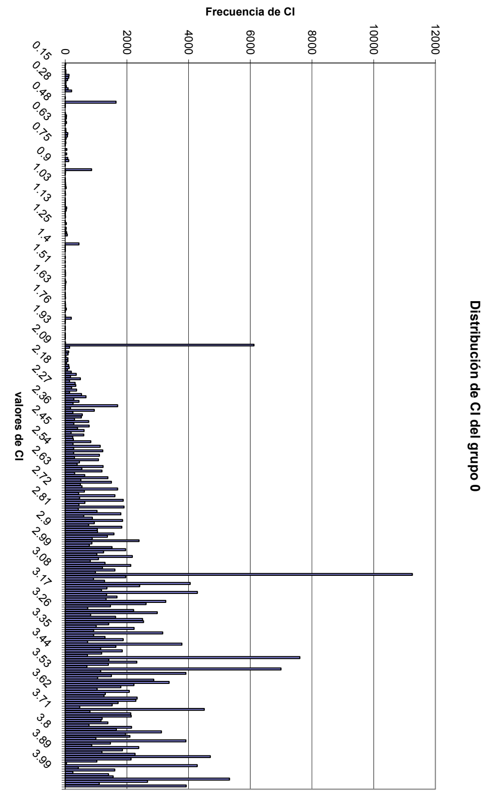
Anexo 7.3 Representación del árbol filogenético de la familia Crp/Fnr

Árbol de 134 secuencias de la familia Crp/Fnr, con los valores de 'bootstrap' que ayudaron a determinar los nueve grupos. Generado por Neighbor Joining siguiendo el modelo JTT (gama = 2.0) con el software Mega3.1

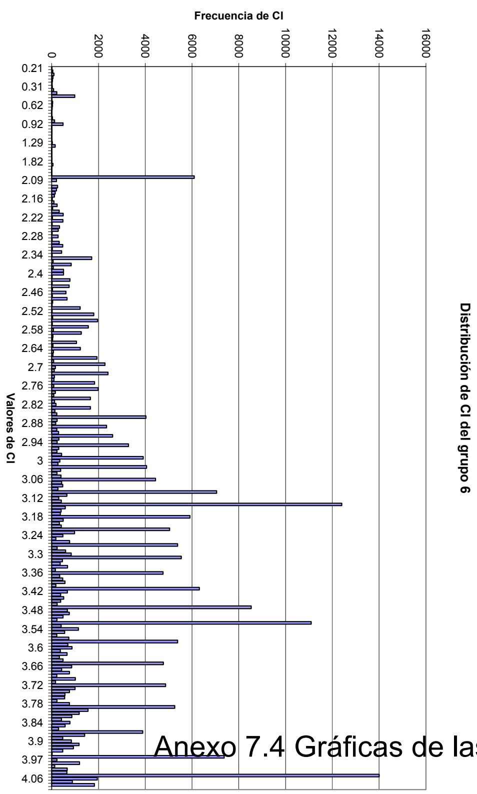


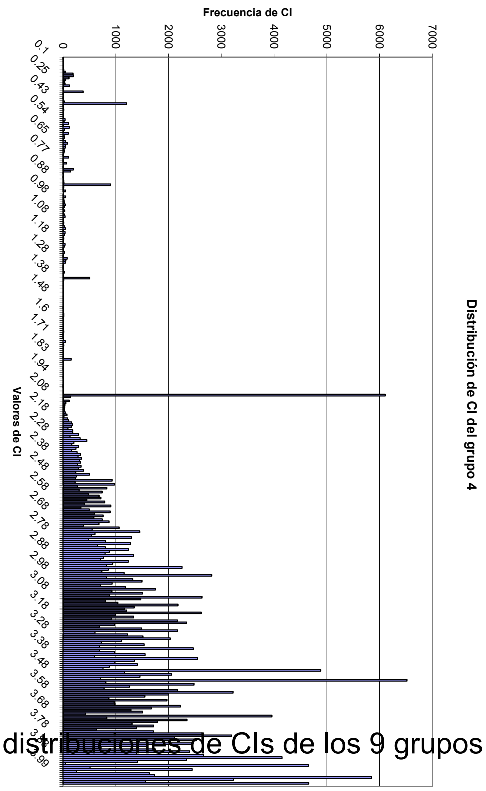


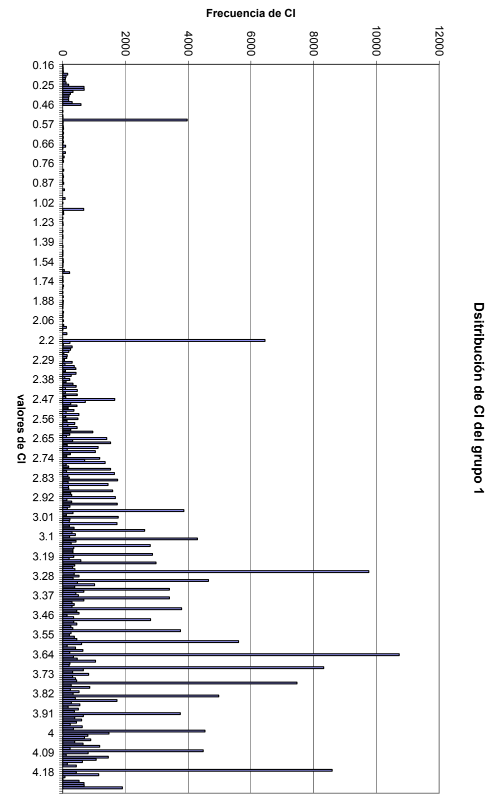




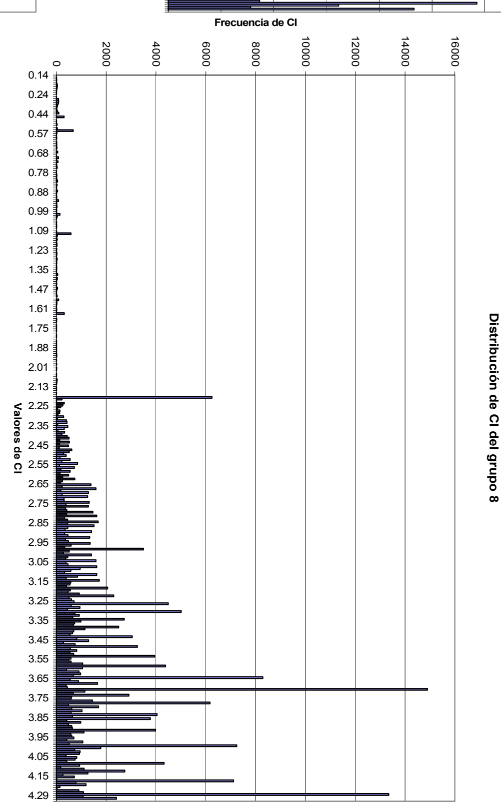
as de las distribuciones de CIs de los 9 grupos

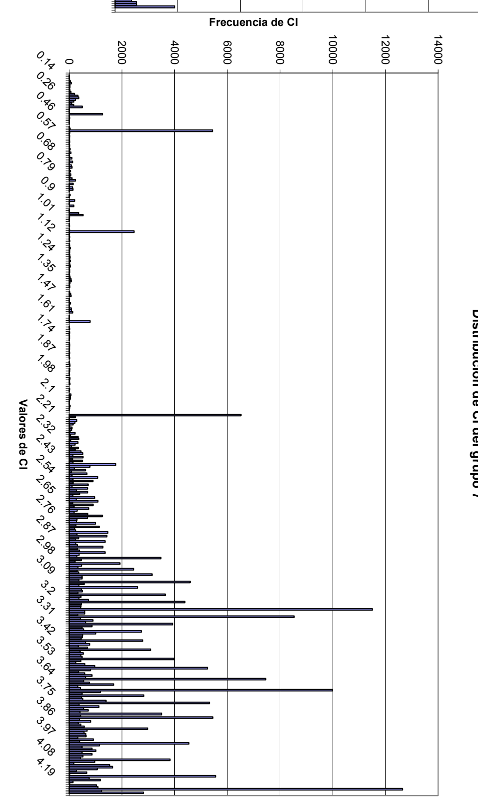






Anexo 7.4 Gráficas de las distribuciones de CIs de los 9 grupos





Anexo 7.6 Códigos de programas en lenguaje PERL

7.6.1 ContInfo-aa-alignment-PRO.pl

```
=========================================================================
=========================================================================
NAME
    ContInfo-aa-alignment-PRO.pl

DESCRIPTION
    Calculates and assigns a score for the amino acids content of each
    column in an alignment

AUTHORS
    heladia@ccg.unam.mx; mipspin@ccg.unam.mx

CATEGORY
    Multiple Sequence Alignment analysis; Informational Content

USAGE
    CalculateLogAligment [-i AlignmentFileName] [-help] [-h] [-f
    InputFormat] [-o OutputFile]

    perl ContInfo-aa-alignment-PRO.pl -i alignment.txt -f
    pipe/id-tab/tab3/tab2 -o OuptputFilename.txt

    Formula Used
    ScoreCol = PropAAGroupCol * (log (PropAAGroupCol/ (PropAACol *
    totalElementByGroup)))

    Where

    PropAAGroupCol = ratio of the amino acid in each column by the group it
    belongs to

    PropAACol = ratio of the amino acid in each column

    ScoreCol = score of amino acid

    SumScoreCol = sumatory of the scores of the aminoacid by column
```

```
    -help
    Same as -h

    -i AlignmentFileName
    name of the file [don't use spaces in the name]

    -f InputFormat
    input Formats: "pipe" group|identifier|name|organism sequence "id-tab"
    group_identifier sequence "tab3" group identifier sequence "tab2" group
    sequence

    -o output
    The output file for the CI results

=========================================================================
#!/usr/local/bin/perl

=pod

=head1 NAME

ContInfo-aa-alignment-PRO.pl

=head1 DESCRIPTION

Calculates and assigns a score for the amino acids content of each column
in an alignment

=head1 AUTHORS

heladia@ccg.unam.mx; mipspin@ccg.unam.mx

=head1 CATEGORY

Multiple Sequence Alignment analysis; Informational Content

=head1 USAGE

CalculateLogAligment [-i AlignmentFileName] [-help] [-h] [-f InputFormat]
[-o OutputFile]
```

Anexo 7.6 Códigos de programas en lenguaje PERL

7.6.1 ContInfo-aa-alignment-PRO.pl

```
perl ContInfo-aa-alignment-PRO.pl -i alignment.txt -f pipe/id-tab/tab3/tab2 -o
OuptputFilename.txt

=cut


################################################################
###        MAIN
################################################################
&ReadArguments;
&OpenOutPut;
&ReadData;
&CalculateLog;


################################################################
#### Calcula el score para cada columna

=pod

=item B<Formula Used>

ScoreCol = PropAAGroupCol * (log (PropAAGroupCol/
                                              (PropAACol * totalElementByGroup)))

Where

PropAAGroupCol = ratio of the amino acid in each column by the group it
                 belongs to

PropAACol = ratio of the amino acid in each column

ScoreCol = score of amino acid

SumScoreCol = sumatory of the scores of the aminoacid by column

=cut

sub CalculateLog {
```

```
$totalElementByGroup = 0;
@SumScore_tot = ();

print OUT "Posiciones\t";
foreach $col ( sort { $a <=> $b } keys %{ $FreqAAByGroupCol{$group} } ){
        $col_imprimir = $col + 1;
        print OUT "$col_imprimir\t";
}
print OUT "\n";

# recorremos los grupos
foreach $group ( sort keys %GroupGICol ) {

        # total de elementos por grupo
        $totalElementByGroup =  keys %{$GroupGICol{$group}};

        print OUT "$group\t";

        ## recorrer las columnas de manera numerica 0,1,2,3..
        foreach $col ( sort { $a <=> $b } keys %{ $FreqAAByGroupCol{$group} } ){

                $SumlogAACol = 0;
                $PropAACol = 0;
                $PropAAGroupCol = 0;
                $SumScoreCol = 0;

                foreach $aa ( sort keys %{ $FreqAAByGroupCol{$group}{$col} } ){

                        $FreAA_Gn_ColN = $FreqAAByGroupCol{$group}{$col}{$aa};

                        # proporcion del aminoacido por grupo y en columna
                        $PropAA_GroupCol = $FreAA_Gn_ColN / $totalElementByGroup;

                        $FrecAA_All_ColN = $FreqAAByColAllGroup{$col}{$aa};

                        # proporcion del aminoacido por columna
                        $PropAA_AllCol = $FrecAA_All_ColN / $TotalElementsAllGroup;

                        #proporcion de secuencias de cada grupo respecto al total de secs
                #de alin
                        $propElemEnGrup = $totalElementByGroup / $TotalElementsAllGroup;

                        $FE = ($PropAA_AllCol * $propElemEnGrup);

                        $FO_FE = ($PropAA_GroupCol / $FE );

                        $log = log ($FO_FE);

                        #score del aminoacido
                        $ScoreCol = ($PropAA_GroupCol * $log);

                        if ($aa eq '-') {$ScoreCol = 0;}

                        #sumatoria de  todos los scores de los aminoacidos en la columna n
                        $SumScoreCol += $ScoreCol;

                        }

                $SumScore_tot[$col] += $SumScoreCol;

                $SumScoreCol_dec = sprintf("%3.2f",$SumScoreCol);
                print OUT "$SumScoreCol_dec\t";

                }

        print OUT "\n";

        }

print OUT "GT\t";

$num_cols = scalar (@SumScore_tot);
$col = 0;

while ($col < $num_cols) {

        $SumScore_tot_dec = sprintf ("%3.2f",$SumScore_tot[$col]);

        print OUT "$SumScore_tot_dec\t";
        $col++;
        }
```

```
        print OUT "\n";
        }



############################
### Read the info form file name
sub ReadData {

    $TotalElementsAllGroup = 0;

    open(IN,"sort $filename |") || die "I can't open the file\n";

    while(<IN>){
        chomp;

      if ($informat =~ /pipe/i) {
      #ARVHIVO EN FORMATO +|+|+|+s+*

        next if $_ !~ /^([^\|]+)\|([^\|]+)\|([^\|]+)\|([^\s]+)\s+(.*)$/;
        #print "$1 .. $2 .. $3 .. $4 ..$5\n";     #CHECKPOINT
        $group = $1;
        $gi = $2;
        $alignment = $5;
      }

      if ($informat =~ /id-tab/i) {
      #ARCHIVO EN FORMATO + +s+*

        next if $_ !~ /^([^\ ]+)\ ([^\s]+)\s+(.*)$/;
        $group = $1;
        $gi = $2;
        $alignment = $3;
        #print "$group .. $gi .. $alignment\n";   #CHECKPOINT

      }

      if ($informat =~ /tab3/i) {
      #OCUPAR ESTAS LINEAS EN CASO DE ARCHIVO EN FORMATO 'identif de grupo \t secuencia'

        ($group, $gi, $alignment) = split (/\t/, $_);
        #print "$group .. $gi .. $alignment\n";   #CHECKPOINT

      }
      if ($informat =~ /tab2/i) {
      #OCUPAR ESTAS LINEAS EN CASO DE ARCHIVO EN FORMATO 'identif de grupo \t secuencia'

        ($group, $alignment) = split (/\t/, $_);
        #print "$group  .. $alignment\n"; #CHECKPOINT
     }


        $TotalElementsAllGroup++;

        # separamos los aminoacidos en columnas
        $i = 0;
        foreach $aa (split(//,$alignment)) {

                #print "$group $i = $aa\n";

                # lleva los aminoacidos por columna
                $GroupGICol{$group}{$gi}[$i] = $aa;

                # Frecuencia de AA en cada grupo x columna
                $FreqAAByGroupCol{$group}{$i}{$aa}++;

                # Frecuencia de AA en todo el alineamiento x columna
                $FreqAAByColAllGroup{$i}{$aa}++;

                $i++;
        }

    }


}

################################################################
#### Open Output File
```

```
sub OpenOutPut {
        open (OUT, ">$outFile") || die "No pude abrir el archivo de salida";
}



################################################################
#### display full help message
sub PrintHelp {
    system "pod2text -c $0";
    exit()
}


################################################################
#### Read arguments
sub ReadArguments {

    my $arg = "";
    while ($arg = shift (@ARGV)) {
        if (($arg eq "-h") || ($arg eq "-help")) {
            &PrintHelp();

            ## List of options
=pod

=item B<-help>

Same as -h

=cut
        } elsif ($arg eq "-i") {
            $fileName = shift (@ARGV);


=pod

=item B<-i AlignmentFileName>

name of the file [don't use spaces in the name]

=cut
        } elsif ($arg eq "-f") {
            $informat = lc (shift (@ARGV));

=pod

=item B<-f InputFormat>

input Formats: "pipe" group|identifier|name|organism       sequence "id-tab"
group_identifier sequence "tab3" group    identifier       sequence "tab2" group    sequence

=cut

        } elsif ($arg eq "-o") {
            $outFile = shift (@ARGV);

=item B<-o output>

The output file for the CI results

=cut



        } else {
            &PrintHelp();

        }
    }

}
=========================================================================
=========================================================================
NAME
    Calc-seqs-PseudoRandom.pl

DESCRIPTION
    Generates pseudo-aleatory sequences starting from a multiple sequence
    alignment, in accordance to a probability given
```

```
AUTHORS
    mipspin@ccg.unam.mx;heladia@ccg.unam.mx

CATEGORY
    Multiple Sequence Alignment analysis; Generate Pseudo-random sequences

USAGE
    Calc-seqs-PseudoRandom.pl [-help] [-h] [-i AlignmentFileName] [-f
    InputFormat] [-p probability] [-n NumberRepetitions] [-o OutputFile]

    perl Calc-seqs-PseudoRandom.pl [-help] [-h] [-i AlignmentFileName] [-f
    InputFormat] [-p probability] [-n NumberRepetitions] [-o OutputFile]

    -help
    Same as -h

    -i AlignmentFileName
    name of the file [don't use spaces in the name]

    -f inputFormat
    inputFormat:"pipe" group|identifier|name|organism sequence "id-tab"
    group_identifier sequence "tab3" group identifier sequence "tab2" group
    sequence

    -p probability
    probability: real number between 0 and 1

    -n NumberRepetitions
    the times that the generation of sequences will be done

    -o OutputFilename
    name of the output file [don't use spaces in the name]

=========================================================================
#!/usr/local/bin/perl

=pod

=head1 NAME

Calc-seqs-PseudoRandom.pl

=head1 DESCRIPTION

Generates pseudo-aleatory sequences starting from a multiple sequence
alignment, in accordance to a probability given

=head1 AUTHORS

mipspin@ccg.unam.mx;heladia@ccg.unam.mx

=head1 CATEGORY

Multiple Sequence Alignment analysis; Generate Pseudo-random sequences

=head1 USAGE

Calc-seqs-PseudoRandom.pl [-help] [-h] [-i AlignmentFileName]
[-f InputFormat] [-p probability] [-n NumberRepetitions] [-o OutputFile]

perl Calc-seqs-PseudoRandom.pl [-help] [-h] [-i AlignmentFileName]
[-f InputFormat] [-p probability] [-n NumberRepetitions] [-o OutputFile]


=cut


################################################################
###        MAIN
################################################################
&ReadArguments;
&ReadData;
&CalculateProb_AA;

for ($r = 0; $r <= $numberRep; $r++) {

        $out_file = "$r".".out";
        open (OUT_S, ">$out_file") || die "No fue posible abrir archivo salida num $r\n";

        &Genera_Seq_Parental;
```

```
        &Genera_Seqs_hijas;

        close (OUT_S);
}
################################################################
####

#############################
###
sub Genera_Seqs_hijas {

 @seq_hija = ();

 foreach $group ( sort keys %GroupGICol ) {

        $totalSecsByGroup{$group} =  keys %{$GroupGICol{$group}};

        $g = 0;
           while ($g < ($totalSecsByGroup{$group} - 1)) {

                $col = 0;
                $prob_a_busc = 0;

                $largo_sec = (keys %{ $AcumAAGroupCol{$group} });

                while ($col < $largo_sec) {

                   $prob_a_busc = rand ();

                   if ($prob_a_busc <= $prob) {
                      $aa_ahora = $seq_parental{$group}[$col];
                      $seq_hija{$group}[$col] = $aa_ahora;

                      print OUT_S "$seq_hija{$group}[$col]";

                      }

                   else {

                      $prob_a_busc_2nd = rand ();

                      $i = 0;

                      $aa_ahora = $aa_array{$group}{$col}[$i];
                      $prob_a_probar = $AcumAAGroupCol{$group}{$col}[$i];

                      while ($prob_a_busc_2nd >= $prob_a_probar && $prob_a_busc_2nd != 1) {

                         $i++;
                         $aa_ahora = $aa_array{$group}{$col}[$i];
                         $prob_a_probar = $AcumAAGroupCol{$group}{$col}[$i];

                         if ($prob_a_probar == "") {      #Este ciclo lo pongo
                                                          #por si llega a pasar
                                                          #que la probab de algun
                                                          #aa se pierda por ahi y
                                                          #asi saber al menos
                                                          #donde buscar y reparar

                                print STDERR "Error de P_aa en G:$group C:$col i:$i\n";
                         }

                         if ($prob_a_probar == 1 || $prob_a_busc_2nd == 1 ) { last }

                         }

                      $seq_hija{$group}[$col] = $aa_ahora;

                      print OUT_S "$seq_hija{$group}[$col]";
                        }
                   $col++;
                   }
                print OUT_S "\n";
                $g++;
                }
        }
 }


################################################################
```

```
####

sub Genera_Seq_Parental {

  @seq_parental = ();

  foreach $group ( sort keys %GroupGICol ) {

        print OUT_S "$group\t00\t";

        $col = 0;
        $prob_a_busc = 0;

        $largo_sec = (keys %{ $AcumAAGroupCol{$group} });

        while ($col < $largo_sec) {
                $i = 0;

                $prob_a_busc = rand ();

                $aa_ahora = $aa_array{$group}{$col}[$i];
                $prob_a_probar = $AcumAAGroupCol{$group}{$col}[$i];

                while ($prob_a_busc >= $prob_a_probar && $prob_a_busc != 1) {
                        $i++;

                        $aa_ahora = $aa_array{$group}{$col}[$i];
                        $prob_a_probar = $AcumAAGroupCol{$group}{$col}[$i];


                        if ($prob_a_probar == "") {       #Este ciclo lo pongo
                                                          #por si llega a pasar
                                                          #que la probab de algun
                                                          #aa se pierda por ahi y
                                                          #asi saber al menos
                                                          #donde buscar y reparar

                         print STDERR "Error de P_aa en G:$group C:$col i:$i\n";
                         }

                        if ($prob_a_probar == 1 || $prob_a_busc == 1 ) { last }

                        }

                        $seq_parental{$group}[$col] = $aa_ahora;

                        print OUT_S "$seq_parental{$group}[$col]";

                        $col++;
                        }

        print OUT_S "\n";
        }
   }

################################################################
#### Calcula probabilidad de aa en cada columna


sub CalculateProb_AA {
        @totalSecsByGroup = ();
        @AcumAAGroupCol = ();
        @aa_array = ();

        $g = 0;
        # recorremos los grupos
        foreach $group ( sort keys %GroupGICol ) {

                # total de elementos por grupo
                $totalElementByGroup{$group} =  keys %{$GroupGICol{$group}};

                $totSecs_x_group{$group}[$g] = $totalElementByGroup{$group};

                ## recorrer las columnas de manera numerica 0,1,2,3..
                foreach $col ( sort { $a <=> $b } keys %{ $FreqAAByGroupCol{$group} } ){

                        $PropAAGroupCol_1x1 = 0;

                        $acumulando_propAA = 0;
```

```
                    $i = 0;
                    foreach $aa ( sort keys %{ $FreqAAByGroupCol{$group}{$col} } ) {

                            $aa_array{$group}{$col}[$i] = $aa;

                            $FreAA_Gn_ColN = $FreqAAByGroupCol{$group}{$col}{$aa};

                            # proporcion del aminoacido por grupo y en columna
                            $PropAAGroupCol_1x1 = $FreAA_Gn_ColN / $totalElementByGroup{$group};

                            $acumulando_propAA += $PropAAGroupCol_1x1;

                            $AcumAAGroupCol{$group}{$col}[$i] = $acumulando_propAA;

                            $i++;
                            }
                    }

        $g++;
        }

  }

############################
### Read the info form file name
sub ReadData {

    $TotalElementsAllGroup = 0;

    open(IN,"$filename") || die "I can't open the file\n";
    while(<IN>){
        chomp;

      if ($informat =~ /pipe/i) {
      #ARVHIVO EN FORMATO +|+|+|+s+*

        next if $_ !~ /^([^\|]+)\|([^\|]+)\|([^\|]+)\|([^\s]+)\s+(.*)$/;
        #print "$1 .. $2 .. $3 .. $4 ..$5\n";     #CHECKPOINT
        $group = $1;
        $gi = $2;
        $alignment = $5;
      }

      if ($informat =~ /id-tab/i) {
      #ARCHIVO EN FORMATO +_+s+*

        next if $_ !~ /^([^\_]+)\_([^\s]+)\s+(.*)$/;
        $group = $1;
        $gi = $2;
        $alignment = $3;
        #print "$group .. $gi .. $alignment\n";   #CHECKPOINT

      }

      if ($informat =~ /tab3/i) {
      #ARCHIVO EN FORMATO 'grupo \t identif de grupo \t secuencia'

        ($group, $gi, $alignment) = split (/\t/, $_);
        #print "$group .. $gi .. $alignment\n";   #CHECKPOINT

      }

      if ($informat =~ /tab2/i) {
      #ARCHIVO EN FORMATO 'identif de grupo \t secuencia'

        ($group, $alignment) = split (/\t/, $_);
        #print "$group  .. $alignment\n"; #CHECKPOINT
     }


        $TotalElementsAllGroup++;

        # separamos los aminoacidos en columnas
        $i = 0;
        foreach $aa (split(//,$alignment)) {

                #print "$group $gi $i = $aa\n";

                # lleva los aminoacidos por columna
                $GroupGICol{$group}{$gi}[$i] = $aa;
```

```
                # lleva la frecuencia por columna
                $FreqAAByGroupCol{$group}{$i}{$aa}++;

                # Frecuencia de AA en todos los grupos
                $FreqAAByColAllGroup{$i}{$aa}++;

                $i++;
        }

    }
```

Anexo 7.6 Códigos de programas en lenguaje PERL

7.6.2 Calc-seqs-PseudoRandom.pl

```
################################################################
#### display full help message
sub PrintHelp {
    system "pod2text -c $0";
    exit()
}


################################################################
#### Read arguments
sub ReadArguments {

    my $arg = "";
    while ($arg = shift (@ARGV)) {
        if (($arg eq "-h") || ($arg eq "-help")) {
            &PrintHelp();

            ## List of options
=pod

=item B<-help>

Same as -h

=cut
        }  elsif ($arg eq "-i") {
            $filename = shift(@ARGV);

=pod

=item B<-i AlignmentFileName>

name of the file [don't use spaces in the name]

=cut

        }  elsif ($arg eq "-f") {
            $informat = shift(@ARGV);

=pod

=item B<-f inputFormat>

inputFormat:"pipe" group|identifier|name|organism  sequence
               "id-tab" group_identifier  sequence
               "tab3" group      identifier       sequence
               "tab2" group      sequence


=cut

        } elsif ($arg eq "-p") {
            $prob = shift(@ARGV);
             $probname = ($prob * 100);

=pod

=item B<-p probability>

probability: real number between 0 and 1

=cut

        } elsif ($arg eq "-n") {
            $numberRep = shift(@ARGV);
```

```
=pod

=item B<-n NumberRepetitions>

the times that the generation of sequences will be done

=cut

        } elsif ($arg eq "-o") {
            $outFile = shift(@ARGV);

=pod

=item B<-o OutputFilename>

name of the output file [don't use spaces in the name]

=cut

        } else {
            &PrintHelp();

        }
    }

}
==========================================================================

==========================================================================
NAME
    Shuffling_alin_PRO.pl

DESCRIPTION
    TOMAR UN ALINEAMIENTO Y MEZCLAR SUS RESIDUOS ALEATORIAMENTE PARA SER
    COLOCADOS ALEATORIAMENTE; Y ASI GENERAR SECUENCIAS 'SHUFFLED'

AUTHORS
    mipspin@ccg.unam.mx

CATEGORY
    Multiple Sequence Alignment analysis; To shuffle an alignment

USAGE
    perl Shuffling_alin_PRO.pl -h -i [alignment-file] -f [inputFormat]

    -help
    Same as -help

    -i AlignmentFileName
    name of the file [don't use spaces in the name]

    -f inputFormat
    inputFormat:"pipe" group|identifier|name|organism sequence "id-tab"
    group_identifier sequence "tab3" group identifier sequence "tab2" group
    sequence
==========================================================================

#!/usr/bin/perl
=pod

=head1 NAME

Shuffling_alin_PRO.pl

=head1 DESCRIPTION

TOMAR UN ALINEAMIENTO Y MEZCLAR SUS RESIDUOS ALEATORIAMENTE
PARA SER COLOCADOS ALEATORIAMENTE; Y ASI GENERAR SECUENCIAS
'SHUFFLED'


=head1 AUTHORS

mipspin@ccg.unam.mx

=head1 CATEGORY

Multiple Sequence Alignment analysis; To shuffle an alignment

=head1 USAGE
```

```
perl Shuffling_alin_PRO.pl -i [alignment-file] -f [inputFormat]

=cut


#############################################################
###        MAIN
#############################################################
&ReadArguments;
&ReadData;
&RandomizeData;
&DoRandAlin;
```

Anexo 7.6 Códigos de programas en lenguaje PERL

7.6.3 Shuffling_alin_PRO.pl

```
############################
### Read the info form file name and split alignment
sub ReadData {

@aa_x_posic = ();
@ids = ();

open (ALIN, "$filename") || die "No se pudo abrir el alineamiento:( \n";

$j = 0;
while (<ALIN>) {
        chomp;
      if ($informat =~ /pipe/i) {
      #ARVHIVO EN FORMATO +|+|+|+s+*

        next if $_ !~ /^([^\|]+)\|([^\|]+)\|([^\|]+)\|([^\s]+)\s+(.*)$/;
        #print "$1 .. $2 .. $3 .. $4 ..$5\n";     #CHECKPOINT
        $group = $1;
        $gi = $2;
        $alignment = $5;
      }

      if ($informat =~ /id-tab/i) {
      #ARCHIVO EN FORMATO +_+s+*

        next if $_ !~ /^([^\_]+)\_([^\s]+)\s+(.*)$/;
        $group = $1;
        $gi = $2;
        $alignment = $3;
        #print "$group .. $gi .. $alignment\n";   #CHECKPOINT

      }

      if ($informat =~ /tab3/i) {
      #ARCHIVO EN FORMATO 'grupo \t identif de grupo \t secuencia'

        ($group, $gi, $alignment) = split (/\t/, $_);
        #print "$group .. $gi .. $alignment\n";   #CHECKPOINT

      }

      if ($informat =~ /tab2/i) {
      #ARCHIVO EN FORMATO 'identif de grupo \t secuencia'

        ($group, $alignment) = split (/\t/, $_);
        #print "$group  .. $alignment\n"; #CHECKPOINT
     }

        $ids[$j] = $id;


        # se separan las columnas de la secuencia
        $i = 0;
        foreach $aa (split(//,$alignment)) {

                $aa_x_posic{$id}[$i] = $aa;
                #print "aas: $aa_x_posic{$id}[$i]\t";     #CHECK_POINT

                $i++;
        }
        #print "\n";     #parte de CHECKPOINT ultimo-anterior

    #print "$ids[$j]\n"; #CHECK_POINT
```

```
    $j++;
    }

close ALIN;

} ############FIN_ReadData


#Randomizing an array as Perl says it can be done
sub RandomizeData {

$l =0;
%bolsa_nums_movida = ();
while ($l < $j){
                #rango de numeros asignados segun el numero de columnas del alineamiento
  @bolsa_nums = (1..180);
  @moviendo_nums = ();


  while (@bolsa_nums) {

        push (@moviendo_nums, splice (@bolsa_nums, rand (@bolsa_nums), 1))

        }
        #print "nums-mov: @moviendo_nums \n";     #CHECKPOINT

   $h = 0;
   foreach $moviendo_nums (@moviendo_nums) {
        $bolsa_nums_movida{$l}[$h] = $moviendo_nums[$h];

        #print "$bolsa_nums_movida{$l}[$h] \t";   #CHECKPOINT

        $h++;
        }
   #print "\n";  #salto de linea del CHECKPOINT de la linea 90

  $l++;
  }

#
} ############FIN_RandomizeData



####SUB-MODULO 3#####
#Assigning the random order to the secuence of each @aa_x_posic{$identificador}
#
sub DoRandAlin {
$n = 0;
open (OUT, ">alin_shuf.aln" ) || die "no se pudo hacer un archivo de salida \n";

while ($n < $j){

   $id = $ids[$n];
   print OUT "$id\t";    #CHECKPOINT

   $col = 0;
   while ($col < 180) {

        $num = $bolsa_nums_movida{$n}[$col];
        #print "numero: $num\t"; #CHECKPOINT

        $aa_x_mov{$id}[$col] = $aa_x_posic{$id}[$num];
        #print "aa: $aa_x_posic{$id}[$num] de:$aa_x_mov{$id}[$col]\t";     #CHECKPOINT

        print OUT "$aa_x_posic{$id}[$num]";
        $col++;
        }
   print OUT "\n";
   $n++;
   }

close (OUT);

} ########Fin DoRandAlin

################################################################
#### display full help message
sub PrintHelp {
    system "pod2text -c $0";
```

```
    exit()
}


################################################################
#### Read arguments
sub ReadArguments {

    my $arg = "";
    while ($arg = shift (@ARGV)) {
        if (($arg eq "-h") || ($arg eq "-help")) {
            &PrintHelp();
```

Anexo 7.6 Códigos de programas en lenguaje PERL

7.6.3 Shuffling_alin_PRO.pl

7.6.4 Meds-Stats.pl

```
            ## list of options
=pod

=item B<-help>

Same as -h

=cut
        }  elsif ($arg eq "-i") {
            $filename = shift(@ARGV);

=pod

=item B<-i AlignmentFileName>

name of the file [don't use spaces in the name]

=cut

        }  elsif ($arg eq "-f") {
            $informat = shift(@ARGV);

=pod

=item B<-f inputFormat>

inputFormat:"pipe" group|identifier|name|organism  sequence
               "id-tab" group_identifier  sequence
               "tab3" group      identifier       sequence
               "tab2" group      sequence


=cut

        }  else {
            &PrintHelp();

        }
    }

}

========================================================================
========================================================================
NAME
    MedsStats.pl

DESCRIPTION
    It obtains the average, standard deviation and Z-score of the
    Informational Content of the real alignment compared with the
    Informational Content of the speudo-random sequences generated with
    Calc-PseudoRandom.pl

AUTHORS
    mipspin@ccg.unam.mx

CATEGORY
    Multiple Sequence Alignment analysis; Statistic Measures

USAGE
    MedsStats.pl[-help] [-h] [-ir InfContFileReal] [-ip
    InfContFilePseudoRandSeqs] [-g GroupIdentifier] [-n NumberRepetitions]

    perl Calc-seqs-PseudoRandom.pl [-help] [-h] [-ir InfContFileReal] [-ip
    InfContFilePseudoRandSeqs] [-g GroupIdentifier] [-n NumberRepetitions]

    Output is automatically saved in a file named Meds-Stats0_[percentage of
    variablity of the pseudorandom seqs]_G[group identifier].out
```

```
    -help
    Same as -h

    -ir InfContFileReal
    name of the file with CI of real data set [don't use spaces in the name]

    -ip InfContFilePseudoRandSeqs
    name of the file with CI of pseudo random seqs [don't use spaces in the
    name]

    -g GroupIdentifier
    The group identifier, usually a single number or letter

    -p probability
    probability: real number between 0 and 1 The probability at which the
    pseudo-random sequences were done, just to have the reference of the
    corresponding population and calculus

    -n NumberRepetitions
    the times that the calculation must be done
=========================================================================

#!/usr/bin/perl

=pod

=head1 NAME

MedsStats.pl

=head1 DESCRIPTION

It obtains the average, standard deviation and Z-score of the Informational
Content of the real alignment compared with the Informational Content of the
speudo-random sequences generated with Calc-PseudoRandom.pl

=head1 AUTHORS

mipspin@ccg.unam.mx

=head1 CATEGORY

Multiple Sequence Alignment analysis; Statistic Measures

=head1 USAGE

MedsStats.pl[-help] [-h] [-ir InfContFileReal]
[-ip InfContFilePseudoRandSeqs] [-g GroupIdentifier] [-n NumberRepetitions]


perl Calc-seqs-PseudoRandom.pl [-help] [-h] [-ir InfContFileReal]
[-ip InfContFilePseudoRandSeqs] [-g GroupIdentifier] [-n NumberRepetitions]

Output is automatically saved in a file named Meds-Stats0_[percentage of
variablity of the pseudorandom seqs]_G[group identifier].out

=cut

###############
##
&ReadArguments;
&OpenOutPut;
&ArrayContInfReal;
&OpenCIPseudoRand;
&MedsStats;

##############
#
sub OpenOutPut {
   open (SALIDA, ">Meds-Stats0_${probname}_G${group}.out") || die "No pude abrir el archivo";

   print SALIDA "CI-G${group}_R\t";

}

#########################
### M O D U L O 1 ####### Abre y pone en arreglo los Cont.Inf. real del grupo
#########################

sub ArrayContInfReal {
```

```
@ci_Gs_lines =();
open (GS, "$filenameCI_real") || die "Can't open the file :( \n"; # prueba_sta

   while(<GS>){
        chomp($_);
        push(@ci_Gs_lines,$_);
        #print "$_\n";   #CHECKPOINT
       }

close GS;

@part_elem_real_col = ();

foreach $curr_line (@ci_Gs_lines){

        next if $curr_line !~ /^G${group}\t(.*)$/; #busca que se cumpla el
                                                      # patron 'GX' donde X= numero o          #letra,
                                          seguido de un tabulador,                #seguido de cualquier
                                          caracter               #0 o mas veces y el resto de la
                                                #linea

        $valores = $1;                  #asigna como $valores lo siguiente a la
                                 # primer columna, en este caso
                                 #los aminoacidos de la secuencia

        print SALIDA "$valores\n";        #CHECKPOINT

        @part_elem_real_col = split (/\t/,$valores);

        }

}
###########################
##FiN M O D U L O 1 #######
###########################



####MINI-M O D U L O #### Abre y pone en arreglo el archivo con los "Random'
#de ContInf

sub OpenCIPseudoRand {

@out_lines =();
open (IN,"$filenameCI_pseudoR") || die "No pude abrirlo :( \n";

   while(<IN>){
        chomp($_);
        push(@out_lines,$_);
       }

close IN;
}
#### FIN MINI-M O D U L O ####


#########################
### M O D U L O 2 ####### Pone en arreglo y hace todas las operaciones con
#########################'Random' de ContInf

sub MedsStats {
   chomp @out_lines;
   $tot_secs_alin = scalar (@out_lines);#asigno variable con numero de Cont.
                                         #Infs de secuencias pseudo-aleatorias
                                         # que tengo
   $veces_repetir = $numberRep;

   @Random_en_arreglo = ();
   @part_elem_col = ();

   $i = 0;
   while ($i < $veces_repetir) {

        next if $out_lines[$i] !~ /^G${group}\s+(.*)$/; #busca el patron 'G

                                         #seguido de un número y
                                         # uno o mas espacios,
                                         #seguido de cualquier
                                         #caracter 0 o mas veces
```

```
                                        #y el resto de la linea.
     $valores = $1;                             #asigna como $valores lo
                                        #siguiente a la primer
                                        #columna, en este
                                        #caso los numeritos.

     @part_elem_col = split (/\t/,$valores);     #parte $valores por
                                  #columnas y los pone
                                  #en arreglo

     $h = 0;
     foreach $part_elem_col (@part_elem_col) {  #aqui se pondra cada
                                   #elemento partido
                                   #del arreglo
                                   #part_elem_col dentro
                                   # de un arreglo
                                   #bidimensional
             $Random_en_arreglo[$i][$h] = $part_elem_col[$h];

             $h++;
             }
     $i++;
     }

$tot_cols = scalar (@part_elem_col);   #CALCULO DE SUMA por COLUMNA
$tot_lines = $veces_repetir;

print SALIDA "Posiciones\t";
$l = 0;
while ($l < $tot_cols) {
     $col_imprimir = $l + 1;
     print SALIDA "$col_imprimir\t";
     $l++;
     }


print SALIDA "\nSum_Rand0_${probname}\t";

@suma_rand_arreglo = ();

$l = 0;
while ($l < $tot_cols) {


     $Suma_random_col = 0;
     $i = 0;
     while ($i < $tot_lines) {

             $Suma_random_col += $Random_en_arreglo[$i][$l];

             $i++;
             }

     $suma_rand_arreglo[$l] = $Suma_random_col;


     printf SALIDA "%.3f\t","$suma_rand_arreglo[$l]\t";  #CHECKPOINT

     $l ++;
     }

print SALIDA "\n";

print SALIDA "Prom_Rand0_${probname}\t";

@promedio_arreglo = ();

$l = 0;
while ($l < $tot_cols) {               #CALCULO DE PROMEDIO por COLUMNA

     $promedio = ($suma_rand_arreglo[$l] / $veces_repetir);
     $promedio_arreglo[$l] = $promedio;

     printf SALIDA "%.3f\t","$promedio_arreglo[$l]\t";    #CHECKPOINT
     $l ++;
     }

print SALIDA "\n";

print SALIDA "\Desv_Est_0${probname}\t";
```

```
   @desv_est_arreglo = ();

   $l = 0;
   while ($l < $tot_cols) {

         $cuadrado_resta = 0;
         $Resta_x_promedio = 0;
         $sumatoria_cuadrados = 0;

         $i = 0;
         while ($i < $tot_lines) {

                 $Resta_x_promedio = ($Random_en_arreglo[$i][$l] - $promedio_arreglo[$l]);
                 $cuadrado_resta = (($Resta_x_promedio)**2);
                 $sumatoria_cuadrados += $cuadrado_resta;

                 $i++;
                 }

         $div_suma_cuad_N = $sumatoria_cuadrados / $veces_repetir;

         $desv_est[$l] = sqrt($div_suma_cuad_N);

         #printf "%.6f\t","$desv_est[$l]\n";  #CHECKPOINT
         $l ++;
         }

   print SALIDA "\n";

   print SALIDA "Z-score0_${probname}\t"; #imprime en archivo  la linea deseada

   @z_score = ();

   $l = 0;
   while ($l < $tot_cols) {       #CALCULO DE Z-SCORE por COLUMNA

         if ($desv_est[$l] > 0) {

             $z_score[$l] = (($part_elem_real_col[$l] - $promedio_arreglo[$l]) / $desv_est[$l]);


         } else {$Z_score[$l] = 0;}

         printf SALIDA "%.3f\t","$z_score[$l]";

         $l++;
         }

print SALIDA "\n\n";

}
###########################
##FiN M O D U L O 2 #######
###########################

##############################################################
#### display full help message

sub PrintHelp {
    system "pod2text -c $0";
    exit()
}


##############################################################
#### Read arguments
sub ReadArguments {

    my $arg = "";
    while ($arg = shift (@ARGV)) {
        if (($arg eq "-h") || ($arg eq "-help")) {
            &PrintHelp();

            ## List of options
=pod

=item B<-help>

Same as -h

=cut
```

```
        } elsif ($arg eq "-ir") {
            $filenameCI_real = shift(@ARGV);

=pod

=item B<-ir InfContFileReal>

name of the file with CI of real data set [don't use spaces in the name]

=cut

        } elsif ($arg eq "-ip") {
            $filenameCI_pseudoR = shift(@ARGV);

=pod

=item B<-ip InfContFilePseudoRandSeqs>

name of the file with CI of pseudo random seqs [don't use spaces in the name]

=cut

        } elsif ($arg eq "-g") {
            $group = shift(@ARGV);

=pod

=item B<-g GroupIdentifier>

The group identifier, usually a single number or letter

=cut

        } elsif ($arg eq "-p") {
            $prob = shift(@ARGV);
             $probname = ($prob * 100);

=pod

=item B<-p probability>

probability: real number between 0 and 1
The probability at which the pseudo-random sequences were done, just
to have the reference of the corresponding population and calculus

=cut

        } elsif ($arg eq "-n") {
            $numberRep = shift(@ARGV);

=pod

=item B<-n NumberRepetitions>

the times that the calculation must be done

=cut

        } else {
            &PrintHelp();

        }
    }

}

========================================================================
========================================================================
Calculo_rho.pl
------------------------------------------------------------------------

#! /usr/bin/perl


### PROGRAMA QUE CALCULA LAS RELACIONES DE IDENTIDAD DE UN ALINEAMIENTO DE SECUENCIAS DE FORMA PAREADA.

### Programa escrito en lenguaje PERL
### Irma Lozada Chavez: ilozada@ccg.unam.mx

### Ejecucion:
### .\Calculo_rho.pl
```

```
# /home/ilozada/GENETIC_REGULATION/TRN_ORTHOLOGY/TFs_alignments/TFs_ECO_alignments/

#  crp     son_SO0624      DVAGRIAQTLLHLAKQPDAMTHPDGMQIKITRQEIGQIVGCSRETVGRILKMLEEQNLIQ
#  crp     hin_HI0957      DVAGRIAQTLMNLAKQPEAMTHPDGMQIKITRQEIGQMVGCSRETVGRIIKMLEDQNLIH
#  lexA    sty_STY4433     MKALTARQQEVFDLIRDHISQTGMPPTRAEIAQRLGFRSPNAAEEHLKALARKGVLEIVS
#  lexA    stm_STM4237     MKALTARQQEVFDLIRDHISQTGMPPTRAEIAQRLGFRSPNAAEEHLKALARKGVLEIVS
#  lexA    eco_b4043|1-72| MKALTARQQEVFDLIRDHISQTGMPPTRAEIAQRLGFRSPNAAEEHLKALARKGVIEIVS


system "rm Gs-G_ORG-1a_SL.rho.txt";

open ENTRA, "Gs-G_ORG-1a_SL.txt" || die "No pude abrir Gs-G_ORG-1a_SL.txt\n";

open SALE, ">>Gs-G_ORG-1a_SL.rho.txt" || die "No pude abrir Gs-G_ORG-1a_SL.rho.txt\n";

@aln= ();
@aln= <ENTRA>;
chomp @aln;

print SALE "\t";
print "\t";

@horizontal= ();
$z= 0;
foreach $all (@aln) {
    chomp $all;
    $all=~ s/\s+/\t/gis;
    ($idh0, $sec0)= split "\t", $all;
    chomp $idh0, $sec0;

    #print "\t$idh0";
    $horizontal[$z++]= $idh0;
}

foreach $h (@horizontal){
    print SALE "$h\t";
    print "$h\t";
}
print SALE "\n";
print "\n";

@aln2= ();
@aln2= @aln;

$j= 0;
$k= 0;
@vertical= ();
foreach $query (@aln) {
    chomp $query;
    $query=~ s/\s+/\t/gis;
    ($idh, $sec)= split "\t", $query;
    chomp $idh, $sec;
    #print "$sec\n";
    @aas= ();
    @aas= split "", $sec;         # A V K G K Y I T I E N S D A L A A L A G H T R
    chomp @aas;
    $aas= @aas;

    print SALE "$idh";
    print "$idh";
    #@aln2= ();
    #@aln2= @aln;

    %hashline= ();
    foreach $subject (@aln2) {
        chomp $subject;
        $subject=~ s/\s+/\t/gis;
        ($idv, $secs)= split "\t", $subject;
        chomp $idv, $secs;
        #print "$secs\n";
        @aas2= ();
        @aas2= split "", $secs;           # A V K G K Y I T I E N S D A L A A L A G H T R
        chomp @aas2;

        $aln= 0;
        $total= 0;
        $iden= 0;
        $i= 0;
        AA:
```

```
        for ($i>= 0; $i<= $aas; $i++) {
            #print "*$i-$aas*\t";
            #print "$aas[$i]-$aas2[$i]\t";

            if ($aas[$i] eq "-" || $aas2[$i] eq "-" ||  $aas[$i] eq "-" && $aas2[$i] eq "-") {
                #print "$aas[$i]-$aas2[$i]\t";
                next AA;
            }
            if ($aas[$i] eq $aas2[$i]) {
                #print "$aas[$i]-$aas2[$i]\t";
               $iden+= 1;
               $total++;
            }
            else {
                $total++;
            }
        }
        #print "$idh=$iden/$total\n";
        $identity= $iden/$total;
        $identity= sprintf ("%.2f",$identity);
        print SALE "\t$identity";
        print "\t$identity";
        #<STDIN>;
    }
    print SALE "\n";
    print "\n";
    #shift @aln;
}

close ENTRA;
close SALE;


========================================================================
========================================================================
NAME
    color_SeqRelation.pl

DESCRIPTION
    It colors a pairwise-identity matrix table within the format: . line1:
    OrganismsIdentifiers[tab separated] Further lines:
    SequenceIdentifier[tab ]identity1[tab]identitity2[tab]identity...
    Separation Line between groups:------[tab]0[tab]0[tab] Further lines:
    SequenceIdentifier[tab ]identity1[tab]identitity2[tab]identity...

AUTHORS
    Irma Lozada Chavez: ilozada@ccg.unam.mx

CATEGORY
    To color sequence relationships

USAGE
    color_SeqRelation.pl -h -i [InputMatrix]

    -help
    Same as -h

    -i InputMatrix
    name of the file [don't use spaces in the name]

========================================================================

#!/usr/bin/perl

=pod

=head1 NAME

color_SeqRelation.pl

=head1 DESCRIPTION


It colors a pairwise-identity matrix table within the format:                  .
line1:  OrganismsIdentifiers[tab separated]
Further lines: SequenceIdentifier[tab ]identity1[tab]identitity2[tab]identity...
Separation Line between groups:------[tab]0[tab]0[tab]
Further lines: SequenceIdentifier[tab ]identity1[tab]identitity2[tab]identity...


=head1 AUTHORS
```

```
Irma Lozada Chavez:
ilozada@ccg.unam.mx


=head1 CATEGORY

To color sequence relationships

=head1 USAGE

color_SeqRelation.pl -h -i [InputMatrix]

=cut
################################################
&ReadArguments;

################################################
# GD library
use GD;

open (IN, "$filename") or die "Can't open ENTRADA!!\n";
@arr1 = (<IN>);
$arr2 = join ("", @arr1);
$arr2 =~ s/^\t+/\t/;
@arr = split ("\n", $arr2);


#print "Mandar salida a:\nPantalla\t(1)\nArchivo\t(2)\n";
#$opt = <STDIN>;
#chomp $opt;

##### Sacando dimensiones
$filaMasGrande = 1;
$columnaMasGrande = 1;
foreach (@arr) {
$ac = 0;
        if ($_ =~ /^(\S+)/) {
        @a = split ("", $1);
                foreach (@a) {
                $ac++;
                }

                if ($ac > $filaMasGrande) {
                $filaMasGrande = $ac;            ### Tamanio del encabezado mas grande de las filas
                }

        $filas++;                                        ### Numero de filas
        }
}


$an = join ("", @arr[0]);
@anc = split ("\t", $an);
foreach (@anc) {
$bc = 0;
        if ($_ =~ /(\S+)/) {
        @b = split ("", $1);
                foreach (@b) {
                $bc++;
                }

                if ($bc > $columnaMasGrande) {
                $columnaMasGrande = $bc; ### Tamanio del encabezado mas grande de las columnas
                }
        $columnas++;                                     ### Numero de columnas
        }
}

$alto = ((($filas * 10) + ($columnaMasGrande * 7)) - 10);  ### ANCHO DE LA FIGURA
$ancho = (20 + ($columnas * 10) + ($filaMasGrande * 7));   ### ALTO DE LA FIGURA

#print "Ancho: $ancho\nAlto: $alto\nFilas: $filas\nColumnas: $columnas\nfilaMasGrande =
$filaMasGrande\ncolumnaMasGrande = $columnaMasGrande\n";

########## Definiendo las variables
$box = 10;                                                                                  #TAMANO DE CUADRO

#if ($opt == 2) {
open(OUT,"$filename.png") || die "No display... arghh!\n";
```

```
#}else{
#open(OUT,"| display") || die "No display... arghh!\n";
#}

binmode OUT;

$image = new GD::Image($ancho,$alto);

$white = $image -> colorAllocate(255,255,255);                                     #EL PRIMER COLOR
DEFINIDO ES EL DEL FONDO
$gray =  $image -> colorAllocate(225,225,225);
$graydark =  $image -> colorAllocate(120,120,120);
$black = $image -> colorAllocate(0,0,0);

$text_color = $black;                                                              #COLOR DEL TEXTO

#POSITIVE
$c[0] = $image -> colorAllocate(0,0,0);
$c[1] = $image -> colorAllocate(0,0,150);
$c[2] = $image -> colorAllocate(0,10,150);
$c[3] = $image -> colorAllocate(0,17,150);
$c[4] = $image -> colorAllocate(0,24,150);
$c[5] = $image -> colorAllocate(0,31,150);
$c[6] = $image -> colorAllocate(0,38,150);
$c[7] = $image -> colorAllocate(0,45,150);
$c[8] = $image -> colorAllocate(0,52,150);
$c[9] = $image -> colorAllocate(0,59,150);
$c[10] = $image -> colorAllocate(0,66,200);
$c[11] = $image -> colorAllocate(0,73,200);
$c[12] = $image -> colorAllocate(0,80,200);
$c[13] = $image -> colorAllocate(0,87,200);
$c[14] = $image -> colorAllocate(0,94,200);
$c[15] = $image -> colorAllocate(0,101,200);
$c[16] = $image -> colorAllocate(0,108,200);
$c[17] = $image -> colorAllocate(0,115,200);
$c[18] = $image -> colorAllocate(0,122,200);
$c[19] = $image -> colorAllocate(0,129,200);
$c[20] = $image -> colorAllocate(0,136,200);
$c[21] = $image -> colorAllocate(0,143,255);
$c[22] = $image -> colorAllocate(0,150,255);
$c[23] = $image -> colorAllocate(0,157,255);
$c[24] = $image -> colorAllocate(0,164,255);
$c[25] = $image -> colorAllocate(0,171,255);
$c[26] = $image -> colorAllocate(0,178,255);
$c[27] = $image -> colorAllocate(0,185,255);
$c[28] = $image -> colorAllocate(0,192,255);
$c[29] = $image -> colorAllocate(0,199,255);
$c[30] = $image -> colorAllocate(0,206,255);
$c[31] = $image -> colorAllocate(0,213,255);
$c[32] = $image -> colorAllocate(0,220,255);
$c[33] = $image -> colorAllocate(95,255,0);
$c[34] = $image -> colorAllocate(100,255,0);
$c[35] = $image -> colorAllocate(105,255,0);
$c[36] = $image -> colorAllocate(110,255,0);
$c[37] = $image -> colorAllocate(115,255,0);
$c[38] = $image -> colorAllocate(120,255,0);
$c[39] = $image -> colorAllocate(125,255,0);
$c[40] = $image -> colorAllocate(130,255,0);
$c[41] = $image -> colorAllocate(135,255,0);
$c[42] = $image -> colorAllocate(140,255,0);
$c[43] = $image -> colorAllocate(145,255,0);
$c[44] = $image -> colorAllocate(150,255,0);
$c[45] = $image -> colorAllocate(155,255,0);
$c[46] = $image -> colorAllocate(160,255,0);
$c[47] = $image -> colorAllocate(165,255,0);
$c[48] = $image -> colorAllocate(170,255,0);
$c[49] = $image -> colorAllocate(175,255,0);
$c[50] = $image -> colorAllocate(180,255,0);
$c[51] = $image -> colorAllocate(185,255,0);
$c[52] = $image -> colorAllocate(190,255,0);
$c[53] = $image -> colorAllocate(195,255,0);
$c[54] = $image -> colorAllocate(200,255,0);
$c[55] = $image -> colorAllocate(205,255,0);
$c[56] = $image -> colorAllocate(210,255,0);
$c[57] = $image -> colorAllocate(215,255,0);
$c[58] = $image -> colorAllocate(220,255,0);
$c[59] = $image -> colorAllocate(225,255,0);
$c[60] = $image -> colorAllocate(230,255,0);
$c[61] = $image -> colorAllocate(235,255,0);
$c[62] = $image -> colorAllocate(240,255,0);
$c[63] = $image -> colorAllocate(245,255,0);
```

```
$c[64] = $image -> colorAllocate(250,255,0);
$c[65] = $image -> colorAllocate(255,248,0);
$c[66] = $image -> colorAllocate(255,241,0);
$c[67] = $image -> colorAllocate(255,234,0);
$c[68] = $image -> colorAllocate(255,227,0);
$c[69] = $image -> colorAllocate(255,220,0);
$c[70] = $image -> colorAllocate(255,213,0);
$c[71] = $image -> colorAllocate(255,206,0);
$c[72] = $image -> colorAllocate(255,199,0);
$c[73] = $image -> colorAllocate(255,192,0);
$c[74] = $image -> colorAllocate(255,185,0);
$c[75] = $image -> colorAllocate(255,178,0);
$c[76] = $image -> colorAllocate(255,171,0);
$c[77] = $image -> colorAllocate(255,164,0);
$c[78] = $image -> colorAllocate(255,157,0);
$c[79] = $image -> colorAllocate(255,150,0);
$c[80] = $image -> colorAllocate(255,143,0);
$c[81] = $image -> colorAllocate(255,136,0);
$c[82] = $image -> colorAllocate(255,129,0);
$c[83] = $image -> colorAllocate(255,122,0);
$c[84] = $image -> colorAllocate(255,115,0);
$c[85] = $image -> colorAllocate(255,108,0);
$c[86] = $image -> colorAllocate(255,101,0);
$c[87] = $image -> colorAllocate(255,94,0);
$c[88] = $image -> colorAllocate(255,87,0);
$c[89] = $image -> colorAllocate(255,80,0);
$c[90] = $image -> colorAllocate(255,73,0);
$c[91] = $image -> colorAllocate(255,66,0);
$c[92] = $image -> colorAllocate(255,59,0);
$c[93] = $image -> colorAllocate(255,52,0);
$c[94] = $image -> colorAllocate(255,45,0);
$c[95] = $image -> colorAllocate(255,38,0);
$c[96] = $image -> colorAllocate(255,31,0);
$c[97] = $image -> colorAllocate(255,24,0);
$c[98] = $image -> colorAllocate(255,17,0);
$c[99] = $image -> colorAllocate(255,10,0);
$c[100] = $image -> colorAllocate(255,3,0);


foreach (@arr) {
        chomp;
        unless (/^\S/) {
                my ($id,@data) = (split/\t/);
                foreach $i (1..$#data+1) {
                        if ($i =~ /\S/) {
#COORDENADAS DEL TEXTO EN X (ESPECIES)
                        $inicio_texto_X_horizontal = (8 + $filaMasGrande * 6);
                        $inicio_texto_X_vertical = ($columnaMasGrande * 6); #igual que
$inicio_vertical_cuadro_1
                        $image ->
stringUp(gdSmallFont,$i*$box+$inicio_texto_X_horizontal,$inicio_texto_X_vertical,$data[$i-1],$text_color);
                        }
                }
                next;
        }
        $line++;
        my ($id,@data) = (split/\t/);

#COORDENADAS DEL TEXTO EN Y (DOMINIOS)
        $inicio_texto_Y_vertical = (($columnaMasGrande * 6) - 2);
        $inicio_texto_Y_horizontal = ($filaMasGrande * 6);

        $image -> string(gdSmallFont,$box,$line*$box+$inicio_texto_Y_vertical,$id,$text_color);

        foreach $i (1..$#data+1) {

#COORDENADAS DE LOS CUADROS DEL PERFIL

                $inicio_horizontal_cuadro_1 = (10 + $filaMasGrande * 6);
                $inicio_vertical_cuadro_1 = ($columnaMasGrande * 6);

                $image ->
filledRectangle($i*$box+$inicio_horizontal_cuadro_1,$line*$box+$inicio_vertical_cuadro_1,$i*$box+$box+$inicio_h
orizontal_cuadro_1,$line*$box+$box+$inicio_vertical_cuadro_1,$c[$data[$i-1]]);
                $image ->
rectangle($i*$box+$inicio_horizontal_cuadro_1,$line*$box+$inicio_vertical_cuadro_1,$i*$box+$box+$inicio_horizon
tal_cuadro_1,$line*$box+$box+$inicio_vertical_cuadro_1,$black);
        }
}

#print "\n";
```

```
######## ESCALA DE POSITIVOS #############

foreach $i (0..100) {
#foreach $i (0..35) {

#COORDENADAS DE LOS CUADROS DE LA ESCALA
        if ($i =~ /^\d+$/) {
        $inicio_horizontal_cuadro_escala_1 = 10;
        $inicio_vertical_cuadro_escala_1 = 20;
        $fin_vertical_cuadro_escala_1 = 30;
        $image ->
filledRectangle($inicio_horizontal_cuadro_escala_1+$i*$box,$inicio_vertical_cuadro_escala_1,$inicio_horizontal_
cuadro_escala_1+$i*$box+$box,$fin_vertical_cuadro_escala_1,$c[$i]);
        $image ->
rectangle($inicio_horizontal_cuadro_escala_1+$i*$box,$inicio_vertical_cuadro_escala_1,$inicio_horizontal_cuadro
_escala_1+$i*$box+$box,$fin_vertical_cuadro_escala_1,$white);

        $text = $i;

        }


#COORDENADAS DEL TEXTO DE LA ESCALA
        if ($i =~ /^\d+$/) {
                if ($i =~ /^0$/ or $i =~ /^5$/ or $i =~ /^10$/ or $i =~ /^15$/ or $i =~ /^20$/ or $i =~ /^25$/
or $i =~ /^30$/ or $i =~ /^35$/ or $i =~ /^40$/ or $i =~ /^45$/ or $i =~ /^50$/ or $i =~ /^55$/ or $i =~ /^60$/
or $i =~ /^65$/ or $i =~ /^70$/ or $i =~ /^75$/ or $i =~ /^80$/ or $i =~ /^85$/ or $i =~ /^90$/ or $i =~ /^95$/
or $i =~ /^100$/) { #Inicio y final de los valore de la escala
                $text = "$i";
                $inicio_texto_escala_horizontal = 8;
                $inicio_texto_escala_vertical = 45;

                $image ->
stringUp(gdSmallFont,$inicio_texto_escala_horizontal+$i*$box,$inicio_texto_escala_vertical,$text,$text_color);
                }
        }

}


# Send to output
print OUT $image->png;


#############################################################
#### display full help message
sub PrintHelp {
    system "pod2text -c $0";
    exit()
}


#############################################################
#### Read arguments
sub ReadArguments {

    my $arg = "";
    while ($arg = shift (@ARGV)) {
        if (($arg eq "-h") || ($arg eq "-help")) {
            &PrintHelp();

            ## List of options
=pod

=item B<-help>

Same as -h

=cut
        }  elsif ($arg eq "-i") {
            $filename = shift(@ARGV);


=pod

=item B<-i InputMatrix>

name of the file [don't use spaces in the name]
```

```
=cut
        }  else {
            &PrintHelp();

        }

    }
}

========================================================================
```

Anexo 7.4 Tablas de CIs con valores Z de los 9 grupos de la familia Crp/Fnr

Posics	CI-r_G1	Sum-0.75	Prom0.75	Des.E0.75	valor-Z0.75	Area acum
1	0	0	0	0	0	
...	...	...	...	...	...	
13	0	0	0	0	0	
14	0.13	110.5	0.11	0.42	0.05	
15	0.16	215.8	0.22	0.59	-0.10	
16	0.21	225.5	0.23	0.59	-0.03	
17	0.34	312.7	0.31	0.69	0.04	
18	0.26	348.6	0.35	0.72	-0.12	
19	0.37	352.4	0.35	0.73	0.02	
20	0.16	263.1	0.26	0.64	-0.16	
21	0.13	235.9	0.24	0.60	-0.18	
22	0.24	317.6	0.32	0.69	-0.11	
23	0.19	264.1	0.26	0.63	-0.12	
24	0.20	304.9	0.31	0.68	-0.16	
25	0.29	320.2	0.32	0.69	-0.04	
26	0.16	286.7	0.29	0.65	-0.19	
27	0.16	266.3	0.27	0.63	-0.17	
28	0.36	333.0	0.33	0.70	0.04	
29	0.24	301.5	0.30	0.67	-0.09	
30	0.56	673.3	0.67	0.91	-0.12	
31	2.88	2638.9	2.64	0.45	0.54	
32	2.64	2450.8	2.45	0.23	0.81	
33	2.56	2556.2	2.56	0.27	0.01	
34	2.85	2642.0	2.64	0.19	1.11	
35	2.78	2712.5	2.71	0.15	0.46	
36	3.15	2758.8	2.76	0.12	3.34	1.00
37	2.99	2721.8	2.72	0.12	2.31	0.99
38	2.85	2648.1	2.65	0.18	1.14	
39	2.71	2494.6	2.50	0.25	0.86	
40	3.29	2776.2	2.78	0.10	4.93	1.00
41	2.90	2691.1	2.69	0.16	1.29	
42	3.05	2708.2	2.71	0.16	2.08	0.98
43	3.05	2721.6	2.72	0.14	2.28	0.99
44	3.33	2810.0	2.81	0.00	0.00	
45	3.10	2810.0	2.81	0.00	0.00	
46	2.91	2558.9	2.56	0.23	1.54	
47	2.84	2672.5	2.67	0.17	0.99	
48	3.25	2775.7	2.78	0.11	4.49	1.00
49	3.06	2632.4	2.63	0.19	2.23	0.99
50	3.01	2686.8	2.69	0.17	1.90	
51	2.94	2720.1	2.72	0.13	1.68	
52	2.84	2501.5	2.50	0.24	1.39	
53	2.99	2626.4	2.63	0.17	2.16	0.98
54	2.99	2596.6	2.60	0.21	1.86	
55	2.74	2769.7	2.77	0.11	-0.28	
56	3.04	2648.0	2.65	0.18	2.22	0.99
57	3.25	2740.3	2.74	0.15	3.49	1.00
58	3.08	2766.6	2.77	0.11	2.81	1.00
59	2.54	2385.3	2.39	0.28	0.55	
60	1.99	1708.7	1.71	0.14	2.01	0.98
61	3.18	2682.1	2.68	0.18	2.71	1.00
62	3.16	2730.3	2.73	0.14	3.14	1.00
63	2.34	2173.1	2.17	0.25	0.66	
64	1.83	1640.1	1.64	0.15	1.29	
65	0.00	0.0	0.00	0.00	0.00	
66	2.13	2025.9	2.03	0.26	0.40	
67	2.95	2671.3	2.67	0.18	1.51	
68	2.57	2572.1	2.57	0.21	-0.01	
69	2.48	2481.0	2.48	0.25	0.00	
70	3.09	2683.6	2.68	0.23	1.80	
71	2.12	1788.7	1.79	0.15	2.24	0.99
72	1.71	1655.0	1.66	0.23	0.24	
73	2.73	2384.7	2.39	0.16	2.22	0.99
74	2.04	1988.8	1.99	0.21	0.24	
75	2.55	2448.4	2.45	0.23	0.45	
76	2.89	2639.2	2.64	0.25	1.02	
77	1.66	1410.0	1.41	0.00	0.00	
78	2.81	2395.5	2.40	0.19	2.15	0.98
79	2.32	2359.8	2.36	0.22	-0.19	
80	3.29	2800.0	2.80	0.04	11.43	1.00
81	2.80	2647.7	2.65	0.20	0.76	
82	3.10	2781.4	2.78	0.07	4.33	1.00
83	3.11	2719.4	2.72	0.14	2.84	1.00
84	3.24	2778.9	2.78	0.08	6.12	1.00
85	3.21	2702.1	2.70	0.17	3.04	1.00
86	2.47	2202.3	2.20	0.18	1.50	
87	2.84	2549.4	2.55	0.27	1.07	
88	1.70	1441.0	1.44	0.14	1.80	
89	2.62	2323.3	2.32	0.25	1.21	
90	2.16	1986.6	1.99	0.14	1.24	
91	3.29	2810.0	2.81	0.00	0.00	
92	2.19	1783.7	1.78	0.14	2.91	1.00
93	2.46	2460.8	2.46	0.22	0.00	
94	2.71	2560.4	2.56	0.27	0.56	
95	3.25	2810.0	2.81	0.00	0.00	
96	2.88	2655.0	2.66	0.20	1.12	
97	3.07	2810.0	2.81	0.00	0.00	
98	3.01	2665.9	2.67	0.22	1.59	
99	2.08	1695.0	1.70	0.12	3.15	1.00
100	1.96	1887.8	1.89	0.19	0.38	
101	2.72	2252.6	2.25	0.24	1.93	
102	2.19	2209.9	2.21	0.24	-0.08	
103	1.66	1410.0	1.41	0.00	0.00	
104	2.68	2489.9	2.49	0.21	0.91	
105	2.31	2554.1	2.55	0.20	-1.20	
106	2.57	2181.5	2.18	0.22	1.74	
107	2.50	2270.0	2.27	0.00	0.00	
108	2.83	2704.4	2.70	0.17	0.72	
109	3.02	2622.4	2.62	0.21	1.91	

Posics	CI-r-G4	Sum0.75	Prom0.75	Des.E0.75	valor-Z-0.75	area acum
1	0	0	0	0	0	
...	...	...	...	...	...	
13	0	0	0	0	0	
14	0	0	0	0	0	
15	0	0	0	0	0	
16	0	0	0	0	0	
17	0	0	0	0	0	
18	0	0	0	0	0	
19	2.23	2314.7	2.32	1.26	-0.07	
20	2.27	2173.8	2.17	1.27	0.08	
21	2.04	2201.1	2.20	1.25	-0.13	
22	2.07	2252.7	2.25	1.24	-0.15	
23	2.25	2250.1	2.25	1.27	0.00	
24	2.27	2151.4	2.15	1.27	0.09	
25	3.09	2879.7	2.88	1.00	0.21	
26	2.58	2789.4	2.79	0.96	-0.22	
27	2.67	2680.6	2.68	1.03	-0.01	
28	2.86	2861.7	2.86	1.03	0.00	
29	2.14	2577.2	2.58	1.01	-0.44	
30	2.39	2544.9	2.55	1.17	-0.13	
31	2.93	2696.3	2.70	1.10	0.21	
32	2.93	2707.8	2.71	0.65	0.34	
33	3.38	3358.3	3.36	0.17	0.13	
34	3.55	3361.1	3.36	0.16	1.20	
35	3.49	3356.6	3.36	0.18	0.74	
36	3.57	3378.4	3.38	0.15	1.26	
37	3.47	3330.4	3.33	0.19	0.72	
38	2.96	3121.6	3.12	0.30	-0.54	
39	2.94	2929.8	2.93	0.36	0.03	
40	3.71	3429.5	3.43	0.10	2.71	
41	3.64	3291.8	3.29	0.23	1.51	
42	3.28	3051.2	3.05	0.34	0.67	
43	3.06	3292.4	3.29	0.23	-1.02	
44	3.70	3447.6	3.45	0.07	3.47	
45	3.61	3429.8	3.43	0.10	1.85	
46	2.67	2837.0	2.84	0.35	-0.48	
47	3.51	3389.3	3.39	0.16	0.77	
48	3.61	3462.4	3.46	0.03	4.85	
49	3.03	3322.5	3.32	0.19	-1.58	
50	2.65	2759.9	2.76	0.35	-0.32	
51	3.47	3337.3	3.34	0.23	0.57	
52	3.06	2997.5	3.00	0.32	0.19	
53	3.56	3424.3	3.42	0.11	1.21	
54	3.05	3224.1	3.22	0.27	-0.65	1.00
55	2.89	2853.6	2.85	0.32	0.11	
56	2.63	2553.7	2.55	0.35	0.22	
57	2.82	2994.4	2.99	0.45	-0.39	
58	2.70	2807.3	2.81	0.33	-0.32	1.00
59	3.09	3245.2	3.25	0.22	-0.70	
60	2.71	2963.3	2.96	0.34	-0.75	
61	3.75	3407.9	3.41	0.13	2.72	
62	2.82	3154.8	3.16	0.27	-1.23	
63	2.66	2832.4	2.83	0.31	-0.56	
64	2.26	2187.8	2.19	0.45	0.16	
65	0.00	0.0	0.00	0.00	0.00	
66	2.95	2909.0	2.91	0.43	0.10	
67	3.00	3021.6	3.02	0.30	-0.07	
68	3.28	3099.5	3.10	0.34	0.53	
69	2.70	3113.1	3.11	0.31	-1.32	
70	3.43	3247.6	3.25	0.24	0.77	
71	2.39	2117.8	2.12	0.15	1.84	
72	2.28	2341.6	2.34	0.52	-0.12	
73	2.95	2769.6	2.77	0.23	0.77	
74	2.80	3053.8	3.05	0.39	-0.65	
75	2.56	2813.0	2.81	0.32	-0.80	
76	2.48	2858.0	2.86	0.43	-0.88	
77	1.94	1730.0	1.73	0.00	0.00	
78	3.19	3445.4	3.45	0.08	-3.05	
79	3.16	3145.8	3.15	0.35	0.04	
80	2.20	2246.1	2.25	0.23	-0.20	
81	2.74	3092.8	3.09	0.29	-1.21	
82	2.74	2582.1	2.58	0.33	0.48	
83	3.09	3262.6	3.26	0.21	-0.81	
84	3.29	3225.4	3.23	0.24	0.27	
85	3.11	2834.8	2.84	0.26	1.08	
86	3.16	3183.2	3.18	0.24	-0.10	
87	3.73	3338.3	3.34	0.23	1.67	
88	1.95	1739.3	1.74	0.04	5.58	0.00
89	2.58	2428.1	2.43	0.24	0.63	
90	3.20	3136.3	3.14	0.43	0.15	
91	3.66	3420.7	3.42	0.11	2.28	
92	2.85	3030.4	3.03	0.51	-0.36	
93	2.96	2876.3	2.88	0.31	0.27	
94	3.79	3440.8	3.44	0.09	4.00	
95	2.36	2264.4	2.26	0.33	0.30	
96	3.58	3333.2	3.33	0.19	1.30	
97	3.75	3459.1	3.46	0.05	5.36	
98	2.61	2644.3	2.64	0.36	-0.10	1.00
99	3.73	3352.3	3.35	0.20	1.92	
100	3.55	3443.5	3.44	0.10	1.10	
101	2.24	2252.1	2.25	0.22	-0.05	
102	2.71	2567.8	2.57	0.30	0.48	
103	1.94	1730.0	1.73	0.00	0.00	
104	3.05	2864.4	2.86	0.29	0.64	1.00
105	3.12	3059.2	3.06	0.30	0.20	
106	3.13	3017.5	3.02	0.43	0.26	
107	2.78	2600.0	2.60	0.00	0.00	1.00
108	3.25	3331.6	3.33	0.20	-0.41	
109	2.64	3010.1	3.01	0.30	-1.23	

110	2.80	2500.6	2.50	0.26	1.16	
111	0.27	450.8	0.45	0.80	-0.23	
112	0.40	380.2	0.38	0.75	0.03	
113	0.40	332.4	0.33	0.70	0.10	
114	3.33	2810.0	2.81	0.00	0.00	
115	3.13	2810.0	2.81	0.00	0.00	
116	3.08	2786.2	2.79	0.06	4.65	1.00
117	2.81	2810.0	2.81	0.00	1.00	
118	2.65	2416.4	2.42	0.25	0.92	
119	3.25	2810.0	2.81	0.00	0.00	
120	3.33	2810.0	2.81	0.00	0.00	
121	2.70	2242.4	2.24	0.10	4.57	1.00
122	3.20	2799.6	2.80	0.04	11.43	1.00
123	1.81	1595.9	1.60	0.22	0.99	
124	3.33	2810.0	2.81	0.00	0.00	
125	2.57	2630.4	2.63	0.18	-0.34	
126	2.63	2573.7	2.57	0.21	0.27	
127	2.96	2491.6	2.49	0.21	2.28	0.99
128	2.68	2451.1	2.45	0.20	1.13	
129	2.11	1827.7	1.83	0.27	1.06	
130	3.11	2739.0	2.74	0.15	2.53	0.99
131	2.63	2130.1	2.13	0.16	3.08	1.00
132	2.24	1871.9	1.87	0.17	2.18	0.99
133	2.94	2583.2	2.58	0.28	1.27	
134	2.67	2330.6	2.33	0.30	1.14	
135	2.50	2277.5	2.28	0.24	0.92	
136	3.01	2585.0	2.59	0.23	1.89	
137	3.14	2695.9	2.70	0.17	2.62	1.00
138	2.65	2517.3	2.52	0.21	0.64	
139	2.83	2630.9	2.63	0.20	0.97	
140	2.24	1776.9	1.78	0.17	2.71	1.00
141	2.43	2622.5	2.62	0.21	-0.91	
142	2.98	2636.7	2.64	0.19	1.85	
143	2.89	2654.5	2.65	0.20	1.15	
144	2.77	2674.5	2.67	0.19	0.52	
145	1.97	1797.7	1.80	0.18	0.97	
146	2.84	2672.3	2.67	0.17	0.97	
147	2.88	2690.7	2.69	0.19	1.00	
148	2.72	2229.2	2.23	0.09	5.36	1.00
149	3.08	2767.1	2.77	0.10	3.06	1.00
150	2.60	2422.4	2.42	0.23	0.78	
151	2.14	1927.1	1.93	0.20	1.07	
152	3.16	2756.1	2.76	0.11	3.75	1.00
153	3.05	2669.8	2.67	0.16	2.40	0.99
154	3.21	2719.5	2.72	0.16	3.12	1.00
155	2.16	1858.2	1.86	0.12	2.60	1.00
156	2.54	2420.0	2.42	0.29	0.42	
157	2.18	2370.7	2.37	0.25	-0.77	
158	2.63	2810.0	2.81	0.00	0.00	
159	2.31	2159.6	2.16	0.17	0.87	
160	2.82	2604.6	2.61	0.20	1.07	
161	3.05	2745.4	2.75	0.12	2.58	1.00
162	2.76	2810.0	2.81	0.00	0.00	
163	2.88	2481.6	2.48	0.23	1.73	
164	2.54	2285.2	2.29	0.28	0.92	
165	2.98	2703.5	2.70	0.15	1.81	
166	3.21	2749.0	2.75	0.14	3.42	1.00
167	3.11	2745.9	2.75	0.11	3.41	1.00
168	3.02	2810.0	2.81	0.00	0.00	
169	1.94	1661.9	1.66	0.09	3.04	1.00
170	3.08	2749.2	2.75	0.13	2.56	0.99
171	3.02	2810.0	2.81	0.00	0.00	
172	2.98	2544.9	2.55	0.23	1.86	
173	2.81	2810.0	2.81	0.00	1.00	
…	…	…	…	…	…	
180	0	0	0	0	0	

110	3.43	3268.6	3.27	0.22	0.75	
111	3.19	3202.4	3.20	0.26	-0.05	
112	0.00	0.0	0.00	0.00	0.00	
113	0.00	0.0	0.00	0.00	0.00	
114	0.00	0.0	0.00	0.00	0.00	
115	0.00	0.0	0.00	0.00	0.00	
116	3.58	3437.3	3.44	0.10	1.43	
117	3.83	3457.2	3.46	0.06	6.46	
118	2.56	2819.3	2.82	0.35	-0.75	
119	2.95	2925.5	2.93	0.35	0.07	
120	2.83	3104.2	3.10	0.27	-1.04	
121	2.99	2590.3	2.59	0.15	2.69	
122	2.43	2560.1	2.56	0.23	-0.57	
123	2.59	2499.2	2.50	0.62	0.15	
124	2.77	2793.6	2.79	0.41	-0.06	
125	2.80	2558.1	2.56	0.34	0.71	
126	2.64	2806.0	2.81	0.30	-0.56	
127	2.29	2354.1	2.35	0.26	-0.24	1.00
128	3.39	3415.0	3.42	0.13	-0.20	
129	3.23	3234.4	3.23	0.42	-0.01	
130	2.99	3380.7	3.38	0.21	-1.89	
131	3.43	3319.8	3.32	0.20	0.55	
132	3.51	3200.8	3.20	0.37	0.83	
133	2.36	2327.5	2.33	0.23	0.14	
134	3.56	3328.3	3.33	0.20	1.16	
135	2.81	2723.5	2.72	0.31	0.28	
136	2.79	2868.7	2.87	0.64	-0.12	
137	3.56	3393.1	3.39	0.14	1.23	
138	2.82	2850.5	2.85	0.30	-0.10	
139	3.20	3260.8	3.26	0.26	-0.23	
140	2.65	2639.5	2.64	0.55	0.02	
141	3.41	3397.0	3.40	0.16	0.08	
142	3.69	3470.0	3.47	0.00	0.00	
143	2.86	2990.2	2.99	0.32	-0.41	
144	3.39	3196.1	3.20	0.30	0.64	
145	3.24	3136.1	3.14	0.44	0.23	
146	3.31	3199.2	3.20	0.26	0.42	
147	3.83	3430.7	3.43	0.13	3.16	1.00
148	2.99	2559.1	2.56	0.09	4.69	1.00
149	2.57	2743.9	2.74	0.35	-0.51	
150	3.51	3293.0	3.29	0.22	0.97	
151	2.33	2262.9	2.26	0.32	0.21	
152	3.49	3364.6	3.37	0.19	0.67	
153	3.02	3167.3	3.17	0.31	-0.48	
154	3.40	3408.5	3.41	0.13	-0.06	
155	2.87	3292.1	3.29	0.24	-1.77	
156	2.17	2469.8	2.47	0.39	-0.78	
157	3.29	3281.3	3.28	0.23	0.04	
158	3.77	3442.5	3.44	0.08	4.25	1.00
159	3.03	2891.6	2.89	0.43	0.32	
160	3.82	3457.2	3.46	0.05	6.74	1.00
161	2.69	2856.9	2.86	0.27	-0.62	
162	3.84	3464.1	3.46	0.04	9.88	1.00
163	3.00	3010.4	3.01	0.34	-0.03	
164	3.30	3063.0	3.06	0.38	0.62	
165	3.09	3267.9	3.27	0.25	-0.70	
166	3.50	3335.5	3.34	0.18	0.90	
167	3.31	3271.5	3.27	0.21	0.18	1.00
168	3.55	3381.8	3.38	0.15	1.13	
169	3.88	3470.0	3.47	0.00	0.00	
170	3.57	3393.3	3.39	0.18	1.01	
171	3.41	3264.2	3.26	0.22	0.67	
172	3.83	3445.9	3.45	0.09	4.15	
173	3.59	3359.7	3.36	0.20	1.17	
…	…	…	…	…	…	
180	0	0	0	0	0	

Posics	CI-r_G2	Sum-0.75	Prom0.75	Des.E0.75	valor-Z0.75	Area acum
1	0	0	0	0	0	
2	0	0	0	0	0	
3	2.59	2244.5	2.24	0.96	0.36	
4	3.12	2738.7	2.74	0.52	0.73	
5	3.16	2815.2	2.82	0.45	0.77	
6	3.26	2814.5	2.81	0.50	0.89	
7	3.11	2767.2	2.77	0.53	0.65	
8	3.08	2776.3	2.78	0.50	0.61	
9	2.98	2770.9	2.77	0.40	0.53	
10	2.98	2758.7	2.76	0.45	0.49	
11	3.05	2604.2	2.60	0.56	0.80	
12	3	2703.3	2.70	0.48	0.62	
13	2.82	2617.5	2.62	0.51	0.39	
14	3.24	2922.2	2.92	0.11	3.04	1.00
15	2.79	2418.8	2.42	0.81	0.46	
16	2.9	2601.0	2.60	0.74	0.41	
17	2.9	2528.2	2.53	0.77	0.49	
18	3.01	2578.8	2.58	0.79	0.54	
19	3.01	2659.3	2.66	0.23	1.55	
20	3.18	2830.9	2.83	0.19	1.80	
21	3.2	2710.8	2.71	0.25	1.98	
22	2.76	2249.4	2.25	0.15	3.39	0.98
23	3.24	2888.5	2.89	0.15	2.39	1.00
24	3.08	2757.3	2.76	0.24	1.36	
25	2.72	2211.0	2.21	0.15	3.34	0.91
26	3.33	2838.9	2.84	0.19	2.62	1.00
27	2.69	2705.1	2.71	0.24	-0.06	
28	3.25	2897.1	2.90	0.15	2.30	0.47
29	3.22	2839.6	2.84	0.20	1.95	
30	2.92	2617.4	2.62	0.24	1.26	
31	2.92	2406.5	2.41	0.07	7.31	0.90
32	2.87	2570.5	2.57	0.26	1.15	
33	2.82	2422.7	2.42	0.18	2.16	0.87
34	3	2944.3	2.94	0.11	0.53	
35	3.13	2847.4	2.85	0.19	1.51	
36	2.85	2385.9	2.39	0.25	1.87	
37	3.28	2858.6	2.86	0.17	2.42	0.97
38	2.82	2774.1	2.77	0.22	0.21	
39	2.73	2726.6	2.73	0.25	0.01	
40	2.62	2482.5	2.48	0.22	0.63	
41	2.68	2579.8	2.58	0.25	0.40	
42	3.17	2910.1	2.91	0.14	1.80	
43	3.07	2824.8	2.83	0.19	1.28	
44	2.62	2311.1	2.31	0.20	1.53	
45	3.07	2580.6	2.58	0.19	2.55	0.94
46	3.1	2813.2	2.81	0.19	1.53	
47	3.23	2741.4	2.74	0.24	2.06	0.94
48	3	2652.7	2.65	0.20	1.76	
49	3.26	2951.5	2.95	0.07	4.46	0.96
50	3.16	2679.8	2.68	0.26	1.86	
51	3.17	2785.6	2.79	0.22	1.74	
52	3	2674.8	2.68	0.25	1.29	
53	3.37	2911.5	2.91	0.14	3.22	0.90
54	2.9	2725.6	2.73	0.23	0.76	
55	3	2663.1	2.66	0.25	1.34	
56	2.7	2367.6	2.37	0.28	1.18	
57	2.31	2232.2	2.23	0.24	0.33	
58	2.9	2611.9	2.61	0.29	0.98	
59	2.18	2331.8	2.33	0.30	-0.50	
60	2.14	1830.1	1.83	0.16	1.91	
61	2.44	2538.3	2.54	0.25	-0.40	
62	3.05	2820.5	2.82	0.19	1.22	
63	2.57	2341.9	2.34	0.28	0.81	
64	1.91	1722.6	1.72	0.15	1.27	
65	0	0.0	0.00	0.00	0.00	
66	2.68	2443.4	2.44	0.25	0.94	
67	2.48	2609.7	2.61	0.26	-0.49	
68	2.91	2563.7	2.56	0.24	1.45	
69	3.01	2838.6	2.84	0.19	0.91	
70	2.92	2510.5	2.51	0.28	1.45	
71	1.91	1818.6	1.82	0.22	0.41	
72	1.82	1733.6	1.73	0.22	0.39	
73	3.05	2543.3	2.54	0.17	2.90	0.65
74	2.12	2026.2	2.03	0.20	0.48	
75	2.87	2433.4	2.43	0.11	3.95	0.68
76	2.59	2369.2	2.37	0.28	0.78	
77	1.75	1490.0	1.49	0.00	0.00	
78	3.16	2827.8	2.83	0.19	1.72	
79	3.22	2867.1	2.87	0.17	2.13	0.96
80	2.27	2077.5	2.08	0.15	1.26	
81	3.06	2695.3	2.70	0.26	1.43	
82	2.65	2260.7	2.26	0.22	1.74	
83	3.2	2805.6	2.81	0.21	1.91	
84	3.29	2908.4	2.91	0.14	2.73	0.97
85	2.88	2879.1	2.88	0.14	0.01	
86	2.54	2310.4	2.31	0.21	1.08	
87	2.82	2743.7	2.74	0.25	0.31	
88	1.75	1499.2	1.50	0.04	6.74	0.62
89	3.22	2980.0	2.98	0.00	0.00	
90	2.25	2068.7	2.07	0.14	1.31	
91	2.6	2094.5	2.10	0.10	5.06	0.90
92	2.14	1840.1	1.84	0.14	2.14	1.00
93	2.94	2432.6	2.43	0.11	4.55	0.98
94	2.22	2111.9	2.11	0.22	0.50	
95	2.41	1904.8	1.91	0.15	3.46	0.69
96	2.92	2536.8	2.54	0.27	1.42	
97	2.94	2469.2	2.47	0.26	1.80	
98	2.34	2256.6	2.26	0.25	0.34	
99	2.04	1837.0	1.84	0.25	0.82	

Posics	CI-r_G3	Sum-0.75	Prom0.75	Des.E0.75	valor-Z0.75	Area acum
1	0	0	0	0	0	
2	0	0	0	0	0	
3	1.38	1378.9	1.38	1.34	0.00	
4	1.41	1544.0	1.54	1.32	-0.10	
5	1.61	1534.1	1.53	1.38	0.06	
6	1.83	1899.7	1.90	1.37	-0.05	
7	2.42	2441.6	2.44	1.23	-0.02	
8	2.39	2447.9	2.45	1.25	-0.05	
9	3.29	3363.5	3.36	0.23	-0.32	
10	3.62	3440.3	3.44	0.18	0.98	
11	3.23	3017.7	3.02	0.73	0.29	
12	3.01	2992.7	2.99	0.72	0.02	
13	3.47	3270.2	3.27	0.29	0.70	
14	3.08	2972.8	2.97	0.93	0.12	
15	3.01	2939.0	2.94	0.91	0.08	
16	3.59	3449.0	3.45	0.18	0.80	
17	3.65	3519.9	3.52	0.12	1.10	
18	3.64	3493.8	3.49	0.15	0.97	
19	3.34	3077.3	3.08	0.31	0.85	
20	3.17	3293.7	3.29	0.28	-0.44	
21	3.05	3028.4	3.03	0.34	0.06	
22	3.06	2551.2	2.55	0.15	3.39	1.00
23	3.9	3528.7	3.53	0.14	2.64	
24	3.25	3283.6	3.28	0.30	-0.11	
25	3.02	2513.3	2.51	0.15	3.33	1.00
26	3.14	3183.4	3.18	0.32	-0.14	
27	3.14	3190.3	3.19	0.33	-0.16	
28	3.22	3155.7	3.16	0.30	0.21	
29	3.37	3369.6	3.37	0.26	0.00	
30	3.26	2976.3	2.98	0.31	0.92	
31	3.22	2706.7	2.71	0.07	7.38	1.00
32	2.73	2872.5	2.87	0.39	-0.36	
33	2.92	2697.2	2.70	0.19	1.18	
34	3.06	3237.6	3.24	0.31	-0.57	
35	3.68	3476.9	3.48	0.19	1.09	
36	2.94	2857.6	2.86	0.37	0.22	
37	2.91	3133.6	3.13	0.32	-0.69	
38	2.88	3186.7	3.19	0.32	-0.97	
39	3.13	3183.9	3.18	0.35	-0.15	
40	2.76	2821.2	2.82	0.27	-0.23	
41	2.7	2928.6	2.93	0.39	-0.59	
42	2.84	3068.9	3.07	0.34	-0.67	
43	3.26	3315.0	3.32	0.29	-0.19	
44	2.85	2650.1	2.65	0.26	0.77	
45	3.04	2807.4	2.81	0.23	1.03	
46	3.05	3243.4	3.24	0.28	-0.70	
47	3.52	3352.1	3.35	0.29	0.59	
48	3.25	3035.4	3.04	0.28	0.77	
49	3.04	3090.4	3.09	0.33	-0.15	
50	2.91	3192.7	3.19	0.31	-0.92	
51	3.05	3346.5	3.35	0.28	-1.07	
52	3.01	3279.0	3.28	0.29	-0.93	
53	2.87	3032.8	3.03	0.36	-0.45	
54	2.97	3179.3	3.18	0.33	-0.64	
55	3.15	3160.5	3.16	0.30	-0.04	
56	2.72	2994.7	3.00	0.45	-0.61	
57	2.7	2553.4	2.55	0.28	0.53	
58	2.62	2989.3	2.99	0.35	-1.05	
59	2.66	2953.2	2.95	0.41	-0.72	
60	2.71	2228.8	2.23	0.18	2.64	
61	2.68	2739.5	2.74	0.29	-0.21	
62	2.95	3096.3	3.10	0.35	-0.42	
63	3.02	2954.6	2.96	0.43	0.15	
64	2.41	2382.5	2.38	0.53	0.05	
65	0	0.0	0.00	0.00	0.00	
66	2.48	2475.1	2.48	0.33	0.02	
67	2.96	3173.2	3.17	0.30	-0.72	
68	3.36	3054.5	3.06	0.34	0.90	
69	2.6	3031.6	3.03	0.36	-1.20	
70	3.23	2938.2	2.94	0.37	0.79	
71	2.37	2380.6	2.38	0.47	-0.02	
72	2.13	2066.8	2.07	0.26	0.25	
73	3.32	2850.5	2.85	0.21	2.25	
74	2.47	2477.1	2.48	0.40	-0.02	
75	3.1	2728.7	2.73	0.12	3.19	1.00
76	3.11	3008.0	3.01	0.44	0.23	
77	2.05	1790.0	1.79	0.00	0.00	
78	3.14	2796.1	2.80	0.25	1.36	
79	3.6	3418.2	3.42	0.22	0.82	
80	2.58	2380.7	2.38	0.15	1.30	
81	3.75	3284.0	3.28	0.32	1.45	
82	3.24	3332.4	3.33	0.34	-0.28	
83	3.45	3293.9	3.29	0.30	0.51	
84	3.15	3186.0	3.19	0.29	-0.12	
85	2.65	2888.1	2.89	0.39	-0.61	
86	2.99	3131.1	3.13	0.39	-0.36	
87	2.48	2727.3	2.73	0.26	-0.95	
88	2.06	1799.3	1.80	0.04	6.97	1.00
89	2.87	2903.7	2.90	0.31	-0.11	
90	2.8	2831.4	2.83	0.50	-0.06	
91	2.73	2438.8	2.44	0.27	1.08	
92	3.01	2992.5	2.99	0.55	0.03	
93	3.08	2681.3	2.68	0.10	3.89	1.00
94	3.08	2692.0	2.69	0.32	1.23	
95	2.54	2253.7	2.25	0.27	1.06	
96	2.82	2848.0	2.85	0.39	-0.07	
97	3.38	3019.2	3.02	0.39	0.92	
98	2.82	2928.9	2.93	0.39	-0.28	
99	2.47	2082.2	2.08	0.12	3.21	1.00

100	2.04	1967.4	1.97	0.20	0.36	
101	2.75	2606.5	2.61	0.26	0.56	
102	3.07	2718.1	2.72	0.28	1.26	
103	1.75	1490.0	1.49	0.00	0.00	
104	2.88	2704.6	2.71	0.27	0.65	
105	2.96	2458.3	2.46	0.08	6.29	0.74
106	2.91	2592.4	2.59	0.36	0.89	
107	2.82	2599.4	2.60	0.18	1.20	
108	2.97	2823.7	2.82	0.17	0.88	
109	2.47	2618.5	2.62	0.26	-0.58	
110	2.49	2615.0	2.62	0.28	-0.45	
111	2.68	2738.1	2.74	0.20	-0.29	
112	0	0.0	0.00	0.00	0.00	
113	0	0.0	0.00	0.00	0.00	
114	0	0.0	0.00	0.00	0.00	
115	0	0.0	0.00	0.00	0.00	
116	0	0.0	0.00	0.00	0.00	
117	2.5	2225.6	2.23	0.17	1.60	
118	2.91	2723.9	2.72	0.24	0.77	
119	2.72	2550.7	2.55	0.26	0.65	
120	3.14	2817.7	2.82	0.20	1.62	
121	2.25	2091.8	2.09	0.11	1.46	
122	3.16	2674.6	2.68	0.23	2.15	0.93
123	1.97	1661.5	1.66	0.20	1.56	
124	2.69	2296.2	2.30	0.16	2.48	0.94
125	2.59	2176.2	2.18	0.15	2.79	0.99
126	3.45	2949.0	2.95	0.11	4.43	1.00
127	2.47	2284.4	2.28	0.21	0.88	
128	3.35	2926.9	2.93	0.11	3.96	0.81
129	2.16	1861.1	1.86	0.19	1.55	
130	2.72	2211.2	2.21	0.15	3.39	0.94
131	2.67	2200.2	2.20	0.15	3.04	1.00
132	2.41	1964.6	1.97	0.17	2.60	1.00
133	2.4	2124.5	2.13	0.20	1.42	
134	2.72	2381.3	2.38	0.29	1.17	
135	2.71	2914.1	2.91	0.15	-1.35	
136	3.35	2910.2	2.91	0.13	3.49	0.09
137	2.7	2509.5	2.51	0.25	0.76	
138	3.23	2851.9	2.85	0.19	2.04	0.78
139	3.13	2780.9	2.78	0.22	1.62	
140	2.33	1859.9	1.86	0.17	2.76	0.95
141	3.05	2708.8	2.71	0.23	1.48	
142	3.32	2921.5	2.92	0.13	3.19	0.93
143	3.01	2737.7	2.74	0.24	1.13	
144	3.3	2873.3	2.87	0.16	2.68	0.87
145	2.07	1874.2	1.87	0.17	1.17	
146	3.1	2752.1	2.75	0.20	1.71	
147	2.42	2361.4	2.36	0.26	0.23	
148	3.37	2878.2	2.88	0.17	2.82	0.59
149	2.81	2531.3	2.53	0.27	1.03	
150	2.99	2766.4	2.77	0.21	1.06	
151	2.59	2462.0	2.46	0.38	0.33	
152	3.01	2928.1	2.93	0.11	0.78	
153	2.8	2580.0	2.58	0.23	0.94	
154	2.82	2495.2	2.50	0.24	1.38	
155	2.55	2034.0	2.03	0.14	3.65	0.92
156	2.57	2287.9	2.29	0.18	1.59	
157	2.65	2570.1	2.57	0.29	0.28	
158	2.78	2286.5	2.29	0.19	2.55	0.61
159	2.76	2430.0	2.43	0.00	0.00	
160	3.17	2881.8	2.88	0.15	1.94	
161	3.31	2946.3	2.95	0.08	4.33	0.97
162	3.23	2737.5	2.74	0.21	2.31	1.00
163	3.05	2679.9	2.68	0.25	1.46	
164	2.44	2232.6	2.23	0.27	0.76	
165	3.05	2795.8	2.80	0.21	1.23	
166	2.74	2830.9	2.83	0.21	-0.44	
167	2.8	2316.6	2.32	0.21	2.33	0.33
168	2.92	2434.6	2.44	0.25	1.93	
169	2.02	1747.2	1.75	0.09	3.02	0.97
170	3.2	2980.0	2.98	0.00	0.00	
171	3.49	2980.0	2.98	0.00	0.00	
172	2.97	2575.5	2.58	0.24	1.64	
173	3.06	2700.3	2.70	0.24	1.49	
174	2.7	2204.0	2.20	0.19	2.66	0.93
175	2.9	2435.3	2.44	0.11	4.14	1.00
176	2.7	2189.8	2.19	0.13	3.85	1.00
177	2.72	2217.7	2.22	0.15	3.37	1.00
178	3.01	2619.6	2.62	0.24	1.63	
179	0	0	0	0	0	
180	0	0	0	0	0	

100	3.35	3176.0	3.18	0.42	0.42	
101	2.82	3082.4	3.08	0.37	-0.71	
102	2.66	2661.9	2.66	0.39	-0.01	
103	2.05	1790.0	1.79	0.00	0.00	
104	3.45	3086.1	3.09	0.31	1.17	
105	3.32	2810.7	2.81	0.19	2.69	
106	2.74	2440.9	2.44	0.21	1.40	
107	3.16	2955.4	2.96	0.25	0.81	
108	3.31	3243.7	3.24	0.28	0.24	
109	2.35	2832.0	2.83	0.36	-1.35	
110	2.78	3011.5	3.01	0.37	-0.63	
111	3.17	3298.7	3.30	0.25	-0.51	
112	0	0.0	0.00	0.00	0.00	
113	0	0.0	0.00	0.00	0.00	
114	0	0.0	0.00	0.00	0.00	
115	0	0.0	0.00	0.00	0.00	
116	0	0.0	0.00	0.00	0.00	
117	3.08	2601.3	2.60	0.24	2.02	
118	2.77	2934.5	2.93	0.40	-0.42	
119	3.14	3000.2	3.00	0.36	0.39	
120	3.25	3277.4	3.28	0.28	-0.10	
121	2.58	2422.2	2.42	0.19	0.85	
122	3.28	3463.2	3.46	0.19	-0.99	
123	2.28	1935.4	1.94	0.12	2.90	
124	3	2606.5	2.61	0.18	2.21	
125	2.77	2466.7	2.47	0.18	1.71	
126	3.22	2983.9	2.98	0.31	0.75	
127	2.9	2689.3	2.69	0.29	0.74	
128	2.69	2699.6	2.70	0.26	-0.04	
129	2.54	2172.1	2.17	0.20	1.88	
130	3.02	2513.7	2.51	0.15	3.38	1.00
131	3.05	3115.1	3.12	0.46	-0.14	
132	2.32	2375.4	2.38	0.41	-0.14	
133	2.7	2561.6	2.56	0.40	0.35	
134	2.57	2488.4	2.49	0.27	0.30	
135	3.54	3410.7	3.41	0.21	0.63	
136	2.92	3196.7	3.20	0.33	-0.84	
137	2.73	2793.0	2.79	0.32	-0.20	
138	2.73	2920.3	2.92	0.33	-0.57	
139	2.71	3091.0	3.09	0.35	-1.10	
140	2.71	2466.6	2.47	0.50	0.48	
141	3.43	3159.3	3.16	0.32	0.85	
142	3.46	3244.5	3.24	0.31	0.69	
143	2.54	3117.0	3.12	0.32	-1.81	
144	3.63	3409.5	3.41	0.23	0.97	
145	2.72	2549.2	2.55	0.52	0.33	
146	3.49	3425.7	3.43	0.23	0.28	
147	2.65	2676.0	2.68	0.31	-0.08	
148	3.8	3415.1	3.42	0.25	1.53	
149	3.13	3064.6	3.07	0.35	0.19	
150	2.67	2816.5	2.82	0.36	-0.40	
151	3.09	3124.8	3.13	0.49	-0.07	
152	3.53	3434.8	3.44	0.19	0.50	
153	2.53	2729.9	2.73	0.27	-0.73	
154	2.94	2806.4	2.81	0.29	0.46	
155	2.41	2283.6	2.28	0.20	0.64	
156	2.66	2513.5	2.51	0.20	0.75	
157	2.59	2871.1	2.87	0.36	-0.79	
158	3.09	2626.9	2.63	0.25	1.85	
159	3.06	2730.0	2.73	0.00	0.00	
160	3.97	3549.9	3.55	0.09	4.78	1.00
161	3	3038.9	3.04	0.30	-0.13	
162	3.45	3107.0	3.11	0.30	1.14	
163	2.55	2879.3	2.88	0.42	-0.79	
164	3.41	3355.0	3.36	0.32	0.17	
165	3.21	3137.0	3.14	0.33	0.22	
166	2.96	2950.4	2.95	0.33	0.03	
167	2.88	2824.0	2.82	0.42	0.13	
168	2.68	2696.7	2.70	0.32	-0.05	
169	2.48	2218.8	2.22	0.39	0.67	
170	3.17	2792.7	2.79	0.16	2.37	
171	3.32	3249.5	3.25	0.27	0.26	
172	3.44	3112.5	3.11	0.38	0.87	
173	3.33	3122.3	3.12	0.34	0.62	
174	2.94	2550.0	2.55	0.28	1.42	
175	3.29	2807.0	2.81	0.21	2.28	
176	3.18	2791.4	2.79	0.44	0.89	
177	3.02	2520.2	2.52	0.15	3.36	1.00
178	2.9	2755.6	2.76	0.28	0.51	
179	0	0	0	0	0	
180	0	0	0	0	0	

Posics	CI-r_G0	Sum-0.75	Prom0.75	Des.E0.75	valor-Z0.75	Area acum
1	0.2	179.7	0.18	0.59	0.03	
2	0.41	324.0	0.32	0.76	0.11	
3	0.47	588.9	0.59	0.98	-0.12	
4	0.68	653.5	0.65	1.01	0.03	
5	1.11	1029.9	1.03	1.22	0.07	
6	1.7	1716.9	1.72	1.29	-0.01	
7	2.06	1949.3	1.95	1.25	0.09	
8	2.39	2301.4	2.30	1.18	0.08	
9	2.48	2346.0	2.35	1.14	0.12	
10	2.3	2367.1	2.37	1.15	-0.06	
11	2.6	2508.0	2.51	1.03	0.09	
12	2.61	2525.5	2.53	1.06	0.08	
13	3.08	2953.4	2.95	0.66	0.19	
14	3.19	3081.2	3.08	0.55	0.20	
15	3	2908.9	2.91	0.62	0.15	
16	3.25	3037.2	3.04	0.63	0.34	
17	3.2	3016.4	3.02	0.59	0.31	
18	3.28	3120.6	3.12	0.56	0.29	
19	3.2	2974.3	2.97	0.66	0.34	
20	2.94	2955.3	2.96	0.59	-0.03	
21	2.9	2950.3	2.95	0.57	-0.09	
22	2.67	2647.0	2.65	0.40	0.06	
23	1.52	1671.3	1.67	1.26	-0.12	
24	1.57	1523.1	1.52	1.16	0.04	
25	1.54	1402.2	1.40	0.97	0.14	
26	2.08	2088.4	2.09	1.20	-0.01	
27	1.98	2143.9	2.14	1.16	-0.14	
28	2.42	2281.7	2.28	1.02	0.14	
29	2.29	2423.2	2.42	1.05	-0.13	
30	2.01	2244.4	2.24	1.12	-0.21	
31	2.79	3006.5	3.01	0.64	-0.34	
32	2.39	2540.8	2.54	0.60	-0.25	
33	3.07	2893.6	2.89	0.70	0.25	
34	2.82	2926.1	2.93	0.57	-0.19	
35	2.92	3081.2	3.08	0.27	-0.60	
36	2.51	2483.5	2.48	0.33	0.08	
37	3.1	3162.3	3.16	0.20	-0.31	
38	2.87	3105.7	3.11	0.22	-1.08	
39	2.66	3013.0	3.01	0.31	-1.14	
40	2.41	2961.1	2.96	0.35	-1.58	
41	2.61	2700.2	2.70	0.31	-0.29	
42	2.86	3037.5	3.04	0.25	-0.71	
43	2.83	3073.8	3.07	0.26	-0.94	
44	2.63	2710.4	2.71	0.37	-0.22	
45	3.58	3251.3	3.25	0.15	2.17	0.99
46	2.51	2764.6	2.77	0.63	-0.40	
47	2.1	2514.1	2.51	0.58	-0.72	
48	3.19	3272.8	3.27	0.18	-0.45	
49	3.02	3200.6	3.20	0.18	-1.03	
50	2.43	2766.1	2.77	0.36	-0.94	
51	2.88	3017.9	3.02	0.29	-0.47	
52	2.88	2823.0	2.82	0.33	0.17	
53	0	0.0	0.00	0.00	0.00	
54	2.18	2451.9	2.45	0.28	-0.97	
55	2.77	2999.8	3.00	0.29	-0.80	
56	2.56	2695.2	2.70	0.41	-0.33	
57	2.36	2418.8	2.42	0.34	-0.17	
58	2.55	2724.7	2.73	0.34	-0.51	
59	2.95	2928.0	2.93	0.37	0.06	
60	2.03	2218.4	2.22	0.41	-0.46	
61	2.45	2557.8	2.56	0.36	-0.30	
62	3.37	3272.4	3.27	0.14	0.71	
63	2.66	2739.9	2.74	0.28	-0.28	
64	2.42	2312.9	2.31	0.53	0.20	
65	0.2	189.7	0.19	0.60	0.02	
66	2.32	2305.2	2.31	0.31	0.05	
67	2.69	2962.4	2.96	0.28	-0.98	
68	2.91	2881.9	2.88	0.29	0.10	
69	2.49	2775.9	2.78	0.33	-0.87	
70	2.51	2736.1	2.74	0.36	-0.63	
71	3.16	3144.6	3.15	0.24	0.07	
72	2.17	2123.3	2.12	0.42	0.11	
73	3.82	3360.0	3.36	0.00	0.00	
74	2.54	2550.7	2.55	0.38	-0.03	
75	2.46	2855.2	2.86	0.35	-1.13	
76	2.39	2511.1	2.51	0.35	-0.35	
77	1.94	1680.0	1.68	0.00	0.00	
78	3.33	3241.3	3.24	0.18	0.48	
79	2.3	2480.1	2.48	0.22	-0.81	
80	2.78	2790.1	2.79	0.46	-0.02	
81	2.12	2600.6	2.60	0.33	-1.44	
82	2.56	2523.4	2.52	0.30	0.12	
83	2.71	3043.2	3.04	0.27	-1.22	
84	2.99	3001.5	3.00	0.29	-0.04	
85	3.21	2861.6	2.86	0.25	1.40	
86	2.7	2949.7	2.95	0.33	-0.75	
87	2.71	2764.7	2.77	0.26	-0.21	
88	1.95	1689.2	1.69	0.04	7.01	1.00
89	2.99	2733.8	2.73	0.28	0.91	
90	3.62	3341.1	3.34	0.06	4.33	1.00
91	2.8	2287.8	2.29	0.10	5.18	1.00
92	2.37	2083.7	2.08	0.26	1.12	
93	2.92	3116.0	3.12	0.32	-0.62	
94	2.48	2494.6	2.50	0.34	-0.04	
95	2.42	2197.9	2.20	0.33	0.67	
96	2.5	2501.1	2.50	0.33	0.00	
97	3.09	2747.0	2.75	0.32	1.06	
98	2.39	2571.4	2.57	0.36	-0.51	
99	2.08	1986.0	1.99	0.28	0.33	

Posics	CI-r_G5	Sum-0.75	Prom0.75	Des.E0.75	valor-Z0.75	Area acum
1	0	0.0	0.00	0.00	0.00	
2	0	0.0	0.00	0.00	0.00	
3	0	0.0	0.00	0.00	0.00	
4	0	0.0	0.00	0.00	0.00	
5	0	0.0	0.00	0.00	0.00	
6	0	0.0	0.00	0.00	0.00	
7	0	0.0	0.00	0.00	0.00	
8	0	0.0	0.00	0.00	0.00	
9	0	0.0	0.00	0.00	0.00	
10	0	0.0	0.00	0.00	0.00	
11	0	0.0	0.00	0.00	0.00	
12	0	0.0	0.00	0.00	0.00	
13	0	0.0	0.00	0.00	0.00	
14	0	0.0	0.00	0.00	0.00	
15	0	0.0	0.00	0.00	0.00	
16	0	0.0	0.00	0.00	0.00	
17	0	0.0	0.00	0.00	0.00	
18	0	0.0	0.00	0.00	0.00	
19	0	0.0	0.00	0.00	0.00	
20	0	0.0	0.00	0.00	0.00	
21	0	0.0	0.00	0.00	0.00	
22	0	0.0	0.00	0.00	0.00	
23	0	0.0	0.00	0.00	0.00	
24	0	0.0	0.00	0.00	0.00	
25	0	0.0	0.00	0.00	0.00	
26	0	0.0	0.00	0.00	0.00	
27	0	0.0	0.00	0.00	0.00	
28	0	0.0	0.00	0.00	0.00	
29	0	0.0	0.00	0.00	0.00	
30	5.98	6250.0	6.25	0.69	-0.39	
31	5.4	5663.1	5.66	0.59	-0.45	
32	6.43	6725.3	6.73	0.41	-0.72	
33	5.68	6160.2	6.16	0.64	-0.75	
34	6.33	6614.2	6.61	0.59	-0.48	
35	4.95	5398.4	5.40	0.62	-0.73	
36	1.17	1444.3	1.44	2.20	-0.13	
37	5.95	6285.0	6.29	0.64	-0.52	
38	6.29	6648.9	6.65	0.46	-0.79	
39	5.72	6252.7	6.25	0.67	-0.80	
40	5.42	5854.7	5.86	0.68	-0.64	
41	4.55	4578.0	4.58	0.29	-0.10	
42	6.72	6917.9	6.92	0.23	-0.86	
43	4.85	5315.4	5.32	0.60	-0.78	
44	4.37	4391.5	4.39	0.18	-0.12	
45	5.38	5439.6	5.44	0.30	-0.20	
46	5.09	5612.3	5.61	0.64	-0.82	
47	5.12	5219.6	5.22	0.48	-0.21	
48	5.44	5906.8	5.91	0.65	-0.72	
49	4.98	5414.1	5.41	0.55	-0.78	
50	6.6	6839.9	6.84	0.36	-0.67	
51	4.77	5207.5	5.21	0.58	-0.75	
52	6.2	6599.1	6.60	0.56	-0.72	
53	5.8	6218.3	6.22	0.79	-0.53	
54	4.71	5228.6	5.23	0.54	-0.97	
55	4.95	5474.1	5.47	0.67	-0.78	
56	4.45	4472.5	4.47	0.25	-0.09	
57	4.62	4811.1	4.81	1.04	-0.19	
58	5.04	5462.1	5.46	0.73	-0.58	
59	5.98	6268.4	6.27	0.67	-0.43	
60	3.81	4245.6	4.25	0.38	-1.15	0.12561469
61	4.19	4202.9	4.20	0.13	-0.10	
62	5.98	6239.1	6.24	0.65	-0.40	
63	4.67	4703.9	4.70	0.28	-0.12	
64	3.66	3658.1	3.66	0.11	0.02	
65	0	0.0	0.00	0.00	0.00	
66	6.99	6990.0	6.99	0.00	0.00	
67	4.37	4840.2	4.84	0.50	-0.95	
68	4.47	4482.8	4.48	0.21	-0.06	
69	5.55	6124.6	6.13	0.66	-0.87	
70	6.77	6861.4	6.86	0.34	-0.27	
71	3.95	3958.4	3.96	0.13	-0.07	
72	3.73	3745.6	3.75	0.15	-0.11	
73	5.4	5777.0	5.78	0.35	-1.09	0.13725099
74	4.13	4148.7	4.15	0.19	-0.10	
75	6.3	6565.8	6.57	0.61	-0.44	
76	4.34	4351.9	4.35	0.22	-0.06	
77	3.5	3500.0	3.50	0.00	0.00	
78	5.24	5327.4	5.33	0.39	-0.22	
79	4.27	4287.6	4.29	0.16	-0.11	
80	6.99	6990.0	6.99	0.00	0.00	
81	4.56	5074.7	5.08	0.57	-0.90	
82	5.67	5752.8	5.75	0.40	-0.21	
83	5.49	5628.5	5.63	0.57	-0.24	
84	4.94	5464.9	5.47	0.63	-0.83	
85	5.69	5941.0	5.94	0.71	-0.35	
86	5.13	5750.5	5.75	0.71	-0.87	
87	4.43	4445.9	4.45	0.19	-0.08	
88	3.5	3506.3	3.51	0.03	-0.21	
89	4.93	5373.1	5.37	0.66	-0.67	
90	4	3998.9	4.00	0.11	0.01	
91	5.8	6203.4	6.20	0.79	-0.51	
92	6.02	6111.7	6.11	0.38	-0.24	
93	4.82	5052.8	5.05	0.81	-0.29	
94	5.89	5890.0	5.89	0.00	0.00	
95	5.74	5850.6	5.85	0.46	-0.24	
96	6.99	6990.0	6.99	0.00	0.00	
97	5.99	6391.1	6.39	0.62	-0.65	
98	4.49	5036.2	5.04	0.68	-0.80	
99	5.88	6086.5	6.09	1.24	-0.17	

100	2.27	2196.4	2.20	0.23	0.32	
101	2.44	2601.1	2.60	0.35	-0.46	
102	2.54	2808.4	2.81	0.33	-0.82	
103	1.94	1680.0	1.68	0.00	0.00	
104	2.9	2727.5	2.73	0.28	0.62	
105	3.59	3323.8	3.32	0.10	2.58	1.00
106	0	0.0	0.00	0.00	0.00	
107	0	0.0	0.00	0.00	0.00	
108	2.73	2929.1	2.93	0.31	-0.63	
109	2.78	2938.7	2.94	0.28	-0.57	
110	2.35	2788.9	2.79	0.34	-1.28	
111	2.74	3069.2	3.07	0.24	-1.40	
112	0	0.0	0.00	0.00	0.00	
113	0	0.0	0.00	0.00	0.00	
114	0	0.0	0.00	0.00	0.00	
115	0	0.0	0.00	0.00	0.00	
116	0	0.0	0.00	0.00	0.00	
117	2.64	2610.0	2.61	0.36	0.08	
118	2.85	2865.6	2.87	0.35	-0.05	
119	3.43	3266.9	3.27	0.23	0.72	
120	2.41	2816.4	2.82	0.29	-1.38	
121	2.16	2135.6	2.14	0.16	0.15	
122	3.49	3122.2	3.12	0.25	1.50	
123	2.17	1859.6	1.86	0.22	1.44	
124	2.74	2971.8	2.97	0.35	-0.66	
125	3.08	3055.4	3.06	0.26	0.09	
126	2.45	2677.7	2.68	0.34	-0.68	
127	2.79	3042.9	3.04	0.33	-0.77	
128	3.13	3091.9	3.09	0.28	0.14	
129	2.39	2437.3	2.44	0.50	-0.09	
130	2.85	2545.5	2.55	0.38	0.79	
131	3.03	3137.6	3.14	0.21	-0.52	
132	3	2780.2	2.78	0.50	0.44	
133	2.49	2642.7	2.64	0.45	-0.34	
134	3.06	3221.4	3.22	0.24	-0.66	
135	2.82	2855.8	2.86	0.36	-0.10	
136	2.44	2653.2	2.65	0.95	-0.23	
137	1.81	2120.9	2.12	0.86	-0.36	
138	2.39	2425.6	2.43	0.91	-0.04	
139	2.21	2524.3	2.52	0.87	-0.36	
140	1.87	2108.5	2.11	0.95	-0.25	
141	2.51	2595.0	2.60	0.96	-0.09	
142	2.12	2634.6	2.64	0.91	-0.57	
143	2.8	3038.0	3.04	0.24	-1.01	
144	2.9	3021.8	3.02	0.29	-0.42	
145	2.31	2286.1	2.29	0.45	0.05	
146	2.41	2852.1	2.85	0.31	-1.41	
147	2.39	2596.8	2.60	0.39	-0.53	
148	2.83	3095.2	3.10	0.36	-0.74	
149	2.85	2928.3	2.93	0.27	-0.29	
150	2.62	2769.9	2.77	0.31	-0.49	
151	2.27	2182.7	2.18	0.21	0.42	
152	2.89	3187.0	3.19	0.19	-1.60	
153	2.66	2899.5	2.90	0.32	-0.75	
154	2.64	2674.5	2.67	0.38	-0.09	
155	3.82	3360.0	3.36	0.00	0.00	
156	3.23	3025.1	3.03	0.38	0.54	
157	2.43	2879.3	2.88	0.34	-1.31	
158	2.98	2681.5	2.68	0.38	0.80	
159	2.59	2451.4	2.45	0.20	0.70	
160	2.83	2951.7	2.95	0.30	-0.41	
161	2.87	3043.7	3.04	0.29	-0.59	
162	3.5	3354.2	3.35	0.04	3.99	1.00
163	3.01	2946.3	2.95	0.31	0.21	
164	2.99	2844.4	2.84	0.38	0.39	
165	3.41	3266.2	3.27	0.14	1.01	
166	3.02	2856.5	2.86	0.30	0.55	
167	2.82	2491.2	2.49	0.23	1.41	
168	2.77	2672.5	2.67	0.35	0.28	
169	2.25	2007.3	2.01	0.26	0.92	
170	3.29	2787.1	2.79	0.15	3.35	1.00
171	3.83	3321.0	3.32	0.12	4.33	1.00
172	2.55	2631.8	2.63	0.30	-0.27	
173	3.25	2994.4	2.99	0.29	0.87	
174	2.83	2462.3	2.46	0.29	1.26	
175	3.61	3263.0	3.26	0.21	1.64	
176	2.89	2381.5	2.38	0.13	3.88	1.00
177	3.03	2644.3	2.64	0.38	1.01	
178	3.1	2880.3	2.88	0.33	0.68	
179	3.88	3360.0	3.36	0.00	0.00	
180	3.88	3360.0	3.36	0.00	0.00	

100	3.75	4153.5	4.15	0.41	-1.00	0.15925345
101	5.98	6243.6	6.24	0.64	-0.41	
102	4.23	4672.5	4.67	0.42	-1.04	0.14805509
103	3.5	3500.0	3.50	0.00	0.00	
104	4.51	4527.1	4.53	0.16	-0.11	
105	4.51	4538.8	4.54	0.23	-0.13	
106	6.99	6990.0	6.99	0.00	0.00	
107	4.34	4334.8	4.34	0.04	0.14	
108	4.67	4674.0	4.67	0.12	-0.03	
109	5.29	5372.1	5.37	0.43	-0.19	
110	5.51	6041.8	6.04	0.69	-0.77	
111	5.57	6049.4	6.05	0.58	-0.83	
112	0	0.0	0.00	0.00	0.00	
113	0	0.0	0.00	0.00	0.00	
114	0	0.0	0.00	0.00	0.00	
115	6.77	6863.8	6.86	0.36	-0.26	
116	5.91	6390.6	6.39	0.60	-0.80	
117	5.38	5463.6	5.46	0.40	-0.21	
118	6.77	6863.8	6.86	0.36	-0.26	
119	6.43	6599.9	6.60	0.55	-0.31	
120	5.14	5697.2	5.70	0.70	-0.80	
121	5.9	6106.9	6.11	1.20	-0.17	
122	6.18	6431.4	6.43	0.64	-0.39	
123	5.9	6328.5	6.33	0.56	-0.77	
124	4.12	4572.7	4.57	0.47	-0.96	
125	4.34	4339.6	4.34	0.13	0.00	
126	4.41	4884.0	4.88	0.54	-0.87	
127	5.37	5537.7	5.54	0.71	-0.24	
128	4.51	4521.0	4.52	0.17	-0.07	
129	3.9	4282.2	4.28	0.61	-0.63	
130	6.13	6363.4	6.36	0.85	-0.28	
131	5.98	6153.0	6.15	0.63	-0.28	
132	4.16	4168.4	4.17	0.15	-0.05	
133	6.01	6450.7	6.45	0.56	-0.78	
134	5.26	5707.5	5.71	0.63	-0.72	
135	4.88	4968.2	4.97	0.41	-0.22	
136	5.22	5704.4	5.70	0.58	-0.83	
137	4.45	4455.8	4.46	0.19	-0.03	
138	5.05	5138.2	5.14	0.46	-0.19	
139	5.61	6067.5	6.07	0.82	-0.56	
140	6.59	6845.7	6.85	0.35	-0.74	
141	5.12	5146.1	5.15	0.25	-0.10	
142	5.43	5486.5	5.49	0.44	-0.13	
143	5.26	5607.4	5.61	0.71	-0.49	
144	6.14	6430.7	6.43	0.81	-0.36	
145	3.81	3823.0	3.82	0.13	-0.10	
146	5.86	6299.0	6.30	0.54	-0.82	
147	4.34	4363.4	4.36	0.20	-0.12	
148	6.82	6950.1	6.95	0.17	-0.75	
149	6.39	6687.7	6.69	0.51	-0.59	
150	6.77	6863.3	6.86	0.35	-0.27	
151	5.95	6192.1	6.19	0.62	-0.39	
152	5.05	5084.4	5.08	0.30	-0.12	
153	5.79	6090.1	6.09	0.70	-0.43	
154	6.99	6990.0	6.99	0.00	0.00	
155	5.61	5757.0	5.76	0.51	-0.29	
156	4.32	4330.4	4.33	0.15	-0.07	
157	5.37	5932.6	5.93	0.67	-0.84	
158	6.99	6990.0	6.99	0.00	0.00	
159	4.47	4709.3	4.71	0.98	-0.24	
160	5.12	5218.1	5.22	0.43	-0.23	
161	4.69	4733.4	4.73	0.28	-0.15	
162	5.83	6207.2	6.21	0.96	-0.39	
163	4.8	4834.7	4.84	0.22	-0.16	
164	5.81	5881.9	5.88	0.39	-0.18	
165	5.05	5096.5	5.10	0.39	-0.12	
166	6.99	6990.0	6.99	0.00	0.00	
167	6.08	6128.0	6.13	0.36	-0.13	
168	5.33	5432.7	5.43	0.39	-0.27	
169	3.77	3773.4	3.77	0.07	-0.05	
170	6.39	6610.3	6.61	0.51	-0.44	
171	6.36	6597.6	6.60	0.54	-0.44	
172	6.99	6990.0	6.99	0.00	0.00	
173	6.99	6990.0	6.99	0.00	0.00	
174	0	0	0	0	0	
175	0	0	0	0	0	
176	0	0	0	0	0	
177	0	0	0	0	0	
178	0	0	0	0	0	
179	0	0	0	0	0	
180	0	0	0	0	0	

Posics	CI-r_G6	Sum-0.75	Prom0.75	Des.E0.75	valor-Z0.75	Area acum
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	
5	0	0	0	0	0	
6	0	0	0	0	0	
7	0	0	0	0	0	
8	0	0	0	0	0	
9	0	0	0	0	0	
10	0	0	0	0	0	
11	0	0	0	0	0	
12	0	0	0	0	0	
13	0	0	0	0	0	
14	0	0	0	0	0	
15	0	0	0	0	0	
16	0	0	0	0	0	
17	0	0	0	0	0	
18	0	0	0	0	0	
19	0	0	0	0	0	
20	0	0	0	0	0	
21	1.56	1600.6	1.60	1.93	-0.02	
22	1.69	1770.9	1.77	2.00	-0.04	
23	1.69	1900.8	1.90	2.07	-0.10	
24	5.37	5370.0	5.37	0.00	1.00	
25	5.17	5277.0	5.28	0.22	-0.50	
26	4.48	4809.9	4.81	0.43	-0.77	
27	3.88	4100.0	4.10	0.38	-0.58	
28	4.74	4810.0	4.81	0.40	-0.18	
29	4.37	4710.3	4.71	0.44	-0.77	
30	4.87	5105.1	5.11	0.36	-0.66	
31	4.57	4640.8	4.64	0.39	-0.18	
32	5.1	5207.6	5.21	0.25	-0.43	
33	5	5106.9	5.11	0.40	-0.27	
34	4.13	4180.5	4.18	0.30	-0.17	
35	5.27	5298.1	5.30	0.21	-0.14	
36	5.27	5272.3	5.27	0.14	-0.02	
37	4.89	5062.8	5.06	0.29	-0.60	
38	3.7	4067.0	4.07	0.40	-0.93	
39	5	5062.0	5.06	0.37	-0.17	
40	4.48	4946.5	4.95	0.52	-0.90	
41	5	5107.5	5.11	0.35	-0.30	
42	4.85	5096.3	5.10	0.34	-0.73	
43	4.3	4809.9	4.81	0.43	-1.20	0.12
44	5.27	5304.1	5.30	0.20	-0.17	
45	4.57	4629.3	4.63	0.30	-0.20	
46	4.6	5016.9	5.02	0.41	-1.03	0.15
47	4.84	5170.6	5.17	0.29	-1.14	0.13
48	4.42	4864.8	4.87	0.48	-0.93	
49	4.55	4843.8	4.84	0.41	-0.71	
50	4	4391.3	4.39	0.48	-0.82	
51	4.64	5069.9	5.07	0.33	-1.31	0.09
52	5	5098.8	5.10	0.37	-0.26	
53	4.6	4948.5	4.95	0.42	-0.83	
54	4.04	4079.0	4.08	0.28	-0.14	
55	3.85	3853.7	3.85	0.16	-0.02	
56	4.2	4226.5	4.23	0.26	-0.10	
57	3.17	3403.2	3.40	0.38	-0.61	
58	4.98	5230.0	5.23	0.23	-1.09	0.14
59	3.52	3693.5	3.69	0.42	-0.42	
60	3.75	4102.6	4.10	0.64	-0.55	
61	3.38	3391.6	3.39	0.13	-0.09	
62	4.35	4870.0	4.87	0.42	-1.23	0.11
63	3.85	3892.9	3.89	0.28	-0.15	
64	2.85	2847.1	2.85	0.11	0.03	
65	0	0.0	0.00	0.00	0.00	
66	3.51	3552.2	3.55	0.26	-0.17	
67	3.98	4036.3	4.04	0.33	-0.17	
68	5.37	5370.0	5.37	0.00	1.00	
69	4.27	4393.6	4.39	0.47	-0.26	
70	4.57	4701.7	4.70	0.29	-0.46	
71	3.14	3147.0	3.15	0.13	-0.06	
72	2.92	2934.2	2.93	0.15	-0.10	
73	4.97	5069.4	5.07	0.17	-0.59	
74	3.32	3337.6	3.34	0.19	-0.09	
75	3.55	4192.4	4.19	0.54	-1.19	0.12
76	3.53	3540.4	3.54	0.21	-0.05	
77	2.69	2690.0	2.69	0.00	0.00	
78	4.09	4336.8	4.34	0.38	-0.65	
79	3.46	3477.3	3.48	0.16	-0.11	
80	3.21	3221.5	3.22	0.12	-0.09	
81	3.77	4144.8	4.15	0.41	-0.92	
82	5.28	5324.7	5.33	0.12	-0.38	
83	4.39	4467.5	4.47	0.36	-0.22	
84	5.11	5223.7	5.22	0.26	-0.44	
85	4.26	4638.4	4.64	0.51	-0.74	
86	4.25	4551.5	4.55	0.43	-0.70	
87	3.62	3634.8	3.64	0.19	-0.08	
88	2.69	2696.2	2.70	0.03	-0.21	
89	3.74	3772.5	3.77	0.23	-0.14	
90	3.19	3188.3	3.19	0.11	0.02	
91	5.37	5370.0	5.37	0.00	1.00	
92	4.05	4304.1	4.30	0.36	-0.70	
93	3.8	3908.2	3.91	0.38	-0.28	
94	4.82	5062.5	5.06	0.45	-0.55	
95	4.68	4849.3	4.85	0.28	-0.61	
96	4.96	5163.2	5.16	0.42	-0.48	
97	5.08	5132.9	5.13	0.32	-0.16	
98	3.88	3927.9	3.93	0.34	-0.14	
99	4.51	4761.2	4.76	0.72	-0.35	

Posics	CI-r_G7	Sum-0.75	Prom0.75	Des.E0.75	valor-Z0.75	Area acum
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	
5	1.13	1271.2	1.27	1.83	-0.08	
6	2.53	2824.7	2.83	2.24	-0.13	
7	2.34	2645.1	2.65	2.19	-0.14	
8	2.07	2461.7	2.46	2.11	-0.19	
9	2.29	2629.7	2.63	2.24	-0.15	
10	2.18	2686.0	2.69	2.19	-0.23	
11	2.48	2917.6	2.92	2.23	-0.20	
12	2.63	2756.8	2.76	2.24	-0.06	
13	2.3	2492.5	2.49	2.18	-0.09	
14	2.42	2652.8	2.65	2.20	-0.11	
15	2.32	2643.8	2.64	2.20	-0.15	
16	2.06	2657.9	2.66	2.16	-0.28	
17	2.36	2732.6	2.73	2.20	-0.17	
18	2.22	2540.4	2.54	2.10	-0.15	
19	1.81	2182.3	2.18	2.19	-0.17	
20	1.7	2023.5	2.02	2.14	-0.15	
21	2.02	2172.6	2.17	2.20	-0.07	
22	2.57	2815.3	2.82	2.23	-0.11	
23	3.04	3299.6	3.30	2.12	-0.12	
24	2.15	2835.4	2.84	1.88	-0.37	
25	2.54	2679.9	2.68	2.24	-0.06	
26	3.42	3902.9	3.90	1.92	-0.25	
27	3.8	4432.1	4.43	1.48	-0.43	
28	4.37	4954.6	4.96	0.44	-1.32	
29	4.06	4921.9	4.92	0.47	-1.84	0.03
30	4.78	5256.3	5.26	0.40	-1.20	
31	5.38	5566.3	5.57	0.12	-1.54	0.06
32	4.27	4714.1	4.71	1.42	-0.31	
33	2.95	3527.7	3.53	1.29	-0.45	
34	3.78	4322.1	4.32	1.46	-0.37	
35	4.06	4781.0	4.78	0.54	-1.35	
36	4.38	5115.0	5.12	0.49	-1.50	0.07
37	4.4	4927.6	4.93	0.39	-1.36	
38	5.08	5332.3	5.33	0.32	-0.79	
39	5.51	5583.7	5.58	0.11	-0.65	
40	3.94	4669.4	4.67	0.67	-1.09	
41	3.86	4579.0	4.58	0.67	-1.08	
42	4.38	5040.8	5.04	0.44	-1.51	0.07
43	4.27	4964.0	4.96	0.46	-1.52	0.06
44	4.99	5349.6	5.35	0.38	-0.95	
45	0.61	746.2	0.75	1.55	-0.09	
46	3.54	4151.9	4.15	0.49	-1.25	
47	4.7	5272.9	5.27	0.39	-1.46	
48	5.42	5535.1	5.54	0.19	-0.62	
49	4.34	5022.7	5.02	0.46	-1.48	
50	4.29	4900.5	4.90	0.48	-1.28	
51	4.48	4988.7	4.99	0.50	-1.02	
52	5.61	5610.0	5.61	0.00	1.00	
53	4.8	5306.1	5.31	0.35	-1.44	
54	3.22	3851.5	3.85	0.37	-1.71	0.04
55	4.8	5125.3	5.13	0.43	-0.75	
56	4.32	4346.0	4.35	0.26	-0.10	
57	3.65	4266.5	4.27	0.81	-0.76	
58	5.2	5300.7	5.30	0.40	-0.25	
59	3.7	3721.2	3.72	0.25	-0.08	
60	3.75	3783.8	3.78	0.26	-0.13	
61	3.5	3510.5	3.51	0.13	-0.08	
62	5.49	5525.3	5.53	0.24	-0.15	
63	5.61	5610.0	5.61	0.00	1.00	
64	2.97	2965.8	2.97	0.11	0.04	
65	0	0.0	0.00	0.00	0.00	
66	3.73	3746.6	3.75	0.21	-0.08	
67	4.1	4154.9	4.16	0.33	-0.16	
68	3.77	3790.4	3.79	0.21	-0.10	
69	4.15	4248.3	4.25	0.47	-0.21	
70	4.69	4835.8	4.84	0.31	-0.47	
71	4.39	4488.9	4.49	0.47	-0.21	
72	3.04	3052.3	3.05	0.15	-0.08	
73	4.51	4854.8	4.86	0.40	-0.86	
74	3.31	3711.9	3.71	0.51	-0.79	
75	3.4	4094.6	4.10	0.59	-1.17	
76	4.16	4506.2	4.51	0.56	-0.62	
77	2.8	2800.0	2.80	0.00	-1.00	
78	4.64	4757.5	4.76	0.48	-0.24	
79	3.58	3593.9	3.59	0.16	-0.09	
80	3.11	3494.7	3.50	0.45	-0.86	
81	4.55	4657.7	4.66	0.48	-0.22	
82	4.67	4916.7	4.92	0.35	-0.71	
83	4.51	4567.7	4.57	0.37	-0.16	
84	4.98	4992.9	4.99	0.18	-0.07	
85	5.61	5610.0	5.61	0.00	1.00	
86	3.73	3763.5	3.76	0.26	-0.13	
87	4.96	5246.4	5.25	0.40	-0.71	
88	2.81	2808.5	2.81	0.03	0.05	
89	4.85	4983.8	4.98	0.49	-0.27	
90	3.3	3306.1	3.31	0.11	-0.06	
91	4.98	4995.9	5.00	0.19	-0.09	
92	5.29	5297.6	5.30	0.15	-0.05	
93	4.55	4603.8	4.60	0.30	-0.18	
94	4.55	4687.7	4.69	0.51	-0.27	
95	5.05	5158.9	5.16	0.45	-0.24	
96	4.02	4043.0	4.04	0.17	-0.14	
97	5.61	5610.0	5.61	0.00	1.00	
98	5.61	5610.0	5.61	0.00	1.00	
99	5.61	5610.0	5.61	0.00	1.00	

100	3.25	3263.9	3.26	0.17	-0.08	
101	4.83	5045.1	5.05	0.40	-0.54	
102	3.9	3953.6	3.95	0.30	-0.18	
103	2.69	2690.0	2.69	0.00	0.00	
104	3.7	3714.9	3.72	0.15	-0.10	
105	3.7	3728.1	3.73	0.23	-0.12	
106	4.3	4526.5	4.53	0.32	-0.71	
107	3.51	3757.4	3.76	0.36	-0.68	
108	3.85	3860.8	3.86	0.12	-0.09	
109	4.52	4667.2	4.67	0.48	-0.31	
110	5.17	5226.7	5.23	0.31	-0.18	
111	5.17	5205.8	5.21	0.28	-0.13	
112	0	0.0	0.00	0.00	0.00	
113	0	0.0	0.00	0.00	0.00	
114	0	0.0	0.00	0.00	0.00	
115	0	0.0	0.00	0.00	0.00	
116	5.27	5301.0	5.30	0.23	-0.14	
117	3.83	3830.0	3.83	0.00	1.00	
118	4.35	4482.5	4.48	0.50	-0.27	
119	3.78	3841.7	3.84	0.32	-0.19	
120	3.9	4253.0	4.25	0.47	-0.75	
121	2.9	3117.6	3.12	0.15	-1.41	0.08
122	3.95	4429.5	4.43	0.40	-1.21	0.11
123	4.87	5057.7	5.06	0.29	-0.65	
124	3.7	4222.5	4.22	0.52	-1.01	0.16
125	3.93	3978.6	3.98	0.34	-0.14	
126	4.35	4450.8	4.45	0.43	-0.23	
127	4.31	4343.6	4.34	0.27	-0.13	
128	4.85	5094.1	5.09	0.49	-0.50	
129	5.08	5119.3	5.12	0.29	-0.14	
130	3.82	4350.0	4.35	0.54	-0.98	
131	4.59	4908.3	4.91	0.35	-0.91	
132	5.27	5278.2	5.28	0.24	-0.03	
133	5.28	5320.3	5.32	0.10	-0.42	
134	4.6	4748.5	4.75	0.41	-0.37	
135	4.03	4504.3	4.50	0.55	-0.87	
136	4.62	4798.1	4.80	0.37	-0.49	
137	3.64	3644.3	3.64	0.19	-0.02	
138	3.65	3868.5	3.87	0.33	-0.66	
139	4.08	4797.9	4.80	0.49	-1.46	0.07
140	5.37	5370.0	5.37	0.00	1.00	
141	3.79	4062.0	4.06	0.36	-0.77	
142	4.83	5069.7	5.07	0.35	-0.69	
143	4.41	4810.1	4.81	0.51	-0.78	
144	4.87	5047.9	5.05	0.35	-0.51	
145	3	3012.0	3.01	0.13	-0.09	
146	4.37	4674.4	4.67	0.46	-0.67	
147	3.9	3912.6	3.91	0.16	-0.08	
148	4.24	4510.4	4.51	0.34	-0.79	
149	5.08	5156.0	5.16	0.34	-0.22	
150	3.93	3942.1	3.94	0.16	-0.08	
151	3.28	3296.7	3.30	0.18	-0.09	
152	4.27	4331.8	4.33	0.37	-0.17	
153	3.74	3754.8	3.76	0.18	-0.08	
154	5.08	5153.2	5.15	0.36	-0.20	
155	3.98	4012.9	4.01	0.27	-0.12	
156	3.74	3764.9	3.77	0.23	-0.11	
157	3.93	4040.8	4.04	0.44	-0.25	
158	3.66	3660.0	3.66	0.00	1.00	
159	3.33	3339.1	3.34	0.13	-0.07	
160	4.31	4407.3	4.41	0.43	-0.23	
161	3.74	3964.3	3.96	0.37	-0.60	
162	3.78	3786.6	3.79	0.06	-0.12	
163	5.08	5155.9	5.16	0.34	-0.22	
164	5.18	5268.0	5.27	0.23	-0.38	
165	4.26	4723.7	4.72	0.42	-1.10	0.13
166	4.52	4638.4	4.64	0.47	-0.25	
167	4.76	4862.6	4.86	0.44	-0.23	
168	4.52	4622.1	4.62	0.39	-0.26	
169	2.96	2961.0	2.96	0.07	-0.02	
170	5.17	5199.6	5.20	0.24	-0.12	
171	4.04	4040.0	4.04	0.00	1.00	
172	5.37	5370.0	5.37	0.00	1.00	
173	3.83	3830.0	3.83	0.00	1.00	
174	0	0	0	0	0	
175	0	0	0	0	0	
176	0	0	0	0	0	
177	0	0	0	0	0	
178	0	0	0	0	0	
179	0	0	0	0	0	
180	0	0	0	0	0	

100	4.96	5195.1	5.20	0.42	-0.56	
101	4.18	4559.9	4.56	0.68	-0.56	
102	4.02	4071.0	4.07	0.30	-0.17	
103	2.8	2800.0	2.80	0.00	-1.00	
104	5.49	5516.2	5.52	0.19	-0.14	
105	3.81	3846.1	3.85	0.23	-0.16	
106	4.8	4824.7	4.83	0.26	-0.10	
107	3.64	3643.0	3.64	0.04	-0.08	
108	4.28	4659.0	4.66	0.46	-0.82	
109	4.6	4679.4	4.68	0.43	-0.19	
110	4.73	4890.9	4.89	0.52	-0.31	
111	4.01	4776.1	4.78	0.51	-1.51	0.07
112	0	0.0	0.00	0.00	0.00	
113	0	0.0	0.00	0.00	0.00	
114	5.61	5610.0	5.61	0.00	1.00	
115	5.38	5464.1	5.46	0.33	-0.26	
116	5.05	5061.4	5.06	0.23	-0.05	
117	3.95	3950.0	3.95	0.00	1.00	
118	5.61	5610.0	5.61	0.00	1.00	
119	5.49	5534.3	5.53	0.22	-0.20	
120	5.61	5610.0	5.61	0.00	1.00	
121	3.3	3304.0	3.30	0.08	-0.05	
122	4.3	4584.2	4.58	0.40	-0.71	
123	3.03	3034.8	3.04	0.11	-0.05	
124	5.61	5610.0	5.61	0.00	1.00	
125	4.05	4097.3	4.10	0.34	-0.14	
126	4.89	5184.1	5.18	0.44	-0.68	
127	4.43	4463.2	4.46	0.27	-0.12	
128	3.96	4105.1	4.11	0.54	-0.27	
129	4.95	5199.7	5.20	0.40	-0.62	
130	4.3	4817.3	4.82	0.49	-1.07	
131	4.14	4664.6	4.67	0.45	-1.16	
132	5.24	5437.9	5.44	0.29	-0.68	
133	3.52	3648.9	3.65	0.50	-0.26	
134	5.3	5476.9	5.48	0.26	-0.68	
135	3.68	4335.8	4.34	0.50	-1.30	
136	4.91	5257.1	5.26	0.38	-0.92	
137	4.66	5000.2	5.00	0.48	-0.71	
138	3.46	4165.4	4.17	0.56	-1.27	
139	3.63	4488.0	4.49	0.55	-1.55	0.06
140	4.97	5263.3	5.26	0.42	-0.71	
141	3.47	4025.9	4.03	0.39	-1.42	
142	4.23	4837.1	4.84	0.54	-1.12	
143	4.51	4838.4	4.84	0.48	-0.68	
144	3.57	4162.1	4.16	0.42	-1.42	
145	3.11	3651.7	3.65	0.78	-0.70	
146	3.65	4430.5	4.43	0.52	-1.50	0.07
147	3.56	4075.8	4.08	0.73	-0.71	
148	4.86	5187.4	5.19	0.45	-0.73	
149	4.06	4673.9	4.67	0.58	-1.07	
150	5.09	5363.7	5.36	0.46	-0.60	
151	3.46	3878.5	3.88	0.72	-0.58	
152	4.08	4472.0	4.47	0.41	-0.97	
153	4.09	4359.9	4.36	0.50	-0.54	
154	5.49	5542.3	5.54	0.21	-0.25	
155	4.1	4132.1	4.13	0.28	-0.12	
156	3.28	3519.0	3.52	0.22	-1.07	
157	5.38	5451.5	5.45	0.34	-0.21	
158	3.77	3770.0	3.77	0.00	0.00	
159	3.81	3810.0	3.81	0.00	1.00	
160	5.61	5610.0	5.61	0.00	1.00	
161	5.05	5190.9	5.19	0.46	-0.31	
162	3.9	3903.7	3.90	0.06	-0.06	
163	5.12	5203.8	5.20	0.44	-0.19	
164	5.06	5223.8	5.22	0.34	-0.48	
165	4.1	4118.1	4.12	0.16	-0.12	
166	4.13	4151.7	4.15	0.23	-0.09	
167	5.61	5610.0	5.61	0.00	1.00	
168	4.74	4898.4	4.90	0.50	-0.32	
169	3.08	3080.4	3.08	0.07	-0.01	
170	4.45	4609.5	4.61	0.54	-0.30	
171	5.61	5610.0	5.61	0.00	1.00	
172	5.29	5373.9	5.37	0.37	-0.23	
173	3.95	3950.0	3.95	0	1	
174	0	0	0	0	0	
175	0	0	0	0	0	
176	0	0	0	0	0	
177	0	0	0	0	0	
178	0	0	0	0	0	
179	0	0	0	0	0	
180	0	0	0	0	0	

Anexo 7.4 Tablas de CIs con valores Z de los 9 grupos de la familia Crp/Fnr

Posics	CI-r_G8	Sum-0.75	Prom0.75	Des.E0.75	valor-Z0.75	Area acum
1	0	0	0	0	0	
2	0	0	0	0	0	
3	0	0	0	0	0	
4	0	0	0	0	0	
5	0	0	0	0	0	
6	0	0	0	0	0	
7	0	0	0	0	0	
8	0	0	0	0	0	
9	0	0	0	0	0	
10	0	0	0	0	0	
11	0	0	0	0	0	
12	0	0	0	0	0	
13	0	0	0	0	0	
14	0	0	0	0	0	
15	0	0	0	0	0	
16	0	0	0	0	0	
17	0	0	0	0	0	
18	1.23	1377.09	1.38	1.93	-0.08	
19	3.54	4062.23	4.06	1.37	-0.38	
20	3.88	4236.3	4.24	1.42	-0.25	
21	4.2	4472.62	4.47	1.66	-0.16	
22	4.25	4560.61	4.56	1.47	-0.21	
23	4.79	4822.84	4.82	1.51	-0.02	
24	4.5	4737.85	4.74	1.48	-0.16	
25	4.42	4626.46	4.63	1.49	-0.14	
26	4.43	4705.07	4.71	1.54	-0.18	
27	3.88	4261.74	4.26	1.45	-0.26	
28	3.63	3951.44	3.95	1.41	-0.23	
29	4.02	4468.6	4.47	1.48	-0.30	
30	3.36	3439.84	3.44	1.12	-0.07	
31	4.11	4526.42	4.53	1.45	-0.29	
32	3.58	3946.37	3.95	1.35	-0.27	
33	4.91	4903.83	4.90	1.50	0.00	
34	4.05	4384.66	4.39	1.45	-0.23	
35	3.34	3702.54	3.70	1.31	-0.28	
36	4.53	4884.69	4.89	0.61	-0.58	
37	4.57	5201.6	5.20	0.41	-1.54	0.06
38	4.29	4736.37	4.74	0.49	-0.91	
39	3.81	4376.74	4.38	0.58	-0.99	
40	3.69	4181.03	4.18	0.56	-0.88	
41	3.56	3800.93	3.80	0.29	-0.84	
42	4.03	4232.9	4.23	0.27	-0.76	
43	3.7	4560.59	4.56	0.55	-1.57	0.06
44	3.85	4165.84	4.17	0.69	-0.46	
45	4.19	4678.18	4.68	0.44	-1.11	
46	4.15	4942.13	4.94	0.47	-1.67	0.05
47	4.96	5325.87	5.33	0.36	-1.01	
48	5.2	5304.94	5.31	0.40	-0.26	
49	4.39	4949.32	4.95	0.47	-1.20	
50	4.59	4967.98	4.97	0.47	-0.80	
51	4.28	4797.53	4.80	0.53	-0.97	
52	4.48	4712.81	4.71	0.52	-0.45	
53	0	0	0	0	0	
54	4.13	4399.06	4.40	0.46	-0.59	
55	3.97	3972.48	3.97	0.16	-0.02	
56	4.34	4662.42	4.66	0.77	-0.42	
57	3.31	3658.16	3.66	0.55	-0.63	
58	3.62	4242.79	4.24	0.56	-1.11	
59	4.48	4981.16	4.98	0.48	-1.04	
60	3.64	3997.16	4.00	0.54	-0.66	
61	3.5	3510.45	3.51	0.13	-0.08	
62	4.5	4828.24	4.83	0.52	-0.63	
63	4.28	4410.97	4.41	0.51	-0.26	
64	2.97	2965.78	2.97	0.11	0.04	
65	0	0	0.00	0.00	0.00	
66	3.21	4013.84	4.01	0.76	-1.06	
67	3.78	4511.65	4.51	0.54	-1.35	
68	3.77	3790.36	3.79	0.21	-0.10	
69	4.32	5079.41	5.08	0.45	-1.69	0.05
70	4.63	4892.95	4.89	0.67	-0.39	
71	5.29	5362.18	5.36	0.38	-0.19	
72	3.2	3598.52	3.60	0.78	-0.51	
73	5	5168.08	5.17	0.30	-0.55	
74	4.28	4386.92	4.39	0.42	-0.25	
75	4.09	4641.5	4.64	0.56	-0.99	
76	5.49	5544.24	5.54	0.20	-0.27	
77	2.8	2800	2.8	0	0	
78	4.92	5086	5.09	0.45	-0.37	
79	3.93	4316.81	4.32	0.70	-0.56	
80	3.06	3388.24	3.39	0.38	-0.87	
81	4.05	4329.1	4.33	0.45	-0.62	
82	4.52	5142.52	5.14	0.44	-1.42	
83	4.82	5055.11	5.06	0.47	-0.50	
84	4.6	4800.97	4.80	0.24	-0.84	
85	4.39	4449.92	4.45	0.40	-0.15	
86	3.76	4367.99	4.37	0.56	-1.09	
87	3.65	3961.85	3.96	0.52	-0.61	
88	2.81	2808.46	2.81	0.03	0.05	
89	4.44	4754.74	4.76	0.42	-0.75	
90	5.2	5309.81	5.31	0.40	-0.27	
91	4.74	4887.11	4.89	0.12	-1.26	
92	3.16	3469.94	3.47	0.45	-0.69	
93	4.55	4603.83	4.60	0.30	-0.18	
94	5.2	5200	5.20	0.00	0.00	
95	4.68	5036.73	5.04	0.53	-0.67	
96	5.61	5610	5.61	0.00	0.00	
97	4.91	5256.2	5.26	0.46	-0.76	
98	4	4046.35	4.05	0.34	-0.14	
99	3.22	3224.4	3.22	0.11	-0.04	

100	3.09	3456.89	3.46	0.41	-0.89	
101	3.47	3992.62	3.99	0.55	-0.95	
102	3.71	3968.73	3.97	0.30	-0.86	
103	2.8	2800	2.80	0.00	0.00	
104	3.73	4180.72	4.18	0.37	-1.24	
105	4.98	5121.49	5.12	0.47	-0.30	
106	5.61	5610	5.61	0.00	0.00	
107	5.61	5610	5.61	0.00	0.00	
108	4.27	4615.62	4.62	0.45	-0.76	
109	4.78	5066.62	5.07	0.51	-0.56	
110	5.3	5536.91	5.54	0.18	-1.29	
111	4.2	4862.02	4.86	0.42	-1.56	0.06
112	0	0	0.00	0.00	0.00	
113	0	0	0.00	0.00	0.00	
114	0	0	0.00	0.00	0.00	
115	0	0	0.00	0.00	0.00	
116	4.8	4948.23	4.95	0.19	-0.78	
117	4.69	4771.23	4.77	0.40	-0.20	
118	4.08	4104	4.10	0.29	-0.08	
119	4.84	5092.53	5.09	0.56	-0.45	
120	4.69	4798.6	4.80	0.48	-0.23	
121	3.48	3517.57	3.52	0.55	-0.07	
122	5.09	5171.16	5.17	0.32	-0.26	
123	2.96	3180.11	3.18	0.47	-0.47	
124	4.8	5086.74	5.09	0.39	-0.73	
125	4.57	5115.93	5.12	0.57	-0.96	
126	3.64	4017.7	4.02	0.42	-0.91	
127	3.46	3707.31	3.71	0.30	-0.84	
128	4.73	5255.31	5.26	0.39	-1.35	
129	3.08	3568.07	3.57	0.59	-0.82	
130	4.42	4693.03	4.69	0.42	-0.65	
131	3.99	4519.83	4.52	0.45	-1.18	
132	5.61	5610	5.61	0.00	1.00	
133	3.6	3647.28	3.65	0.49	-0.10	
134	5.38	5535.62	5.54	0.21	-0.75	
135	3.56	4296.53	4.30	0.49	-1.50	0.07
136	4.91	5252.37	5.25	0.40	-0.86	
137	5.36	5515.54	5.52	0.22	-0.70	
138	4.74	4934.26	4.93	0.59	-0.33	
139	4.14	4742.66	4.74	0.62	-0.98	
140	4.49	4800.27	4.80	0.85	-0.37	
141	4.94	5224.68	5.23	0.40	-0.72	
142	4.45	4750.8	4.75	0.46	-0.66	
143	4.19	4621.46	4.62	0.49	-0.88	
144	4.18	4472.34	4.47	0.47	-0.62	
145	3.12	3129.41	3.13	0.13	-0.07	
146	4.82	5193.84	5.19	0.32	-1.18	
147	3.64	3670.15	3.67	0.20	-0.15	
148	4.69	4767.28	4.77	0.39	-0.20	
149	4.63	4945.62	4.95	0.58	-0.55	
150	5.2	5392.15	5.39	0.34	-0.57	
151	5.43	5573.89	5.57	0.12	-1.18	
152	3.93	4468.32	4.47	0.57	-0.94	
153	4.09	4373.54	4.37	0.49	-0.59	
154	4.85	5141.5	5.14	0.45	-0.64	
155	4.77	4975.95	4.98	0.37	-0.55	
156	5.12	5240.18	5.24	0.42	-0.29	
157	3.97	4803.87	4.80	0.61	-1.36	
158	3.77	3770	3.77	0.00	0.00	
159	3.45	3457.59	3.46	0.13	-0.06	
160	4.75	5090.63	5.09	0.55	-0.62	
161	5.16	5383.88	5.38	0.34	-0.66	
162	3.95	4357.88	4.36	0.47	-0.87	
163	4.98	5084.53	5.09	0.43	-0.24	
164	4.71	4958.17	4.96	0.28	-0.88	
165	5.38	5466.05	5.47	0.30	-0.29	
166	3.78	4015.96	4.02	0.26	-0.91	
167	4.57	4943.28	4.94	0.50	-0.75	
168	5.61	5610	5.61	0.00	1.00	
169	5.47	5580.81	5.58	0.10	-1.07	
170	4.25	4250	4.25	0.00	0.00	
171	5.2	5319.13	5.32	0.39	-0.31	
172	5.31	5443.77	5.44	0.30	-0.45	
173	5.61	5610	5.61	0.00	0.00	
174	0	0	0	0	0	
175	0	0	0	0	0	
176	0	0	0	0	0	
177	0	0	0	0	0	
178	0	0	0	0	0	
179	0	0	0	0	0	
180	0	0	0	0	0	