



**CENTRO DE INVESTIGACION Y DE ESTUDIOS AVANZADOS DEL  
INSTITUTO POLITECNICO NACIONAL**

**UNIDAD ZACATENCO**

**DEPARTAMENTO DE GENÉTICA Y BIOLOGÍA MOLECULAR**

**“Análisis funcional de la proteína antiCRISPR G2 que inhibe el sistema  
CRISPR-cas I-F de *Pseudomonas aeruginosa*”**

**T E S I S**

Que presenta

**IBT. Wendy Figueroa Chávez**

para obtener el Grado de

**Maestra en Ciencias**

en la Especialidad de

**Genética y Biología Molecular**

Director de Tesis:

**Dr. Luis Yoshio Kameyama Kawabe**

Ciudad de México

AGOSTO, 2017

## **ASESORES**

Dr. Gabriel Guarneros Peña

Departamento de Genética y Biología Molecular

CINVESTAV

Dra. Rosa Ma. del Refugio Bermúdez Cruz

Departamento de Genética y Biología Molecular

CINVESTAV

Dr. Rodolfo García Contreras

Departamento de microbiología y parasitología, facultad de medicina

UNAM

Este trabajo se realizó bajo la dirección del Dr. Luis Yoshio Kameyama Kawabe en el laboratorio 8 del Departamento de Genética y Biología Molecular del Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, el cual está a cargo del Dr. Gabriel Guarneros Peña. El autor de este trabajo agradece que durante la realización del mismo fue becario del Consejo Nacional de Ciencia y Tecnología (no. de registro: 591274).

## AGRADECIMIENTOS

Al Dr. Gabriel Guarneros por recibirme en su grupo de investigación aún antes de iniciar la maestría, y por permitirme seguir en el durante la realización de la presente tesis. Por ser un crítico duro, pero justo y objetivo hacia mi trabajo. Por inculcarme la importancia del uso de controles y el rigor del método científico.

Al doctor Kameyama por dirigir la presente tesis, y permitirme desarrollar el proyecto que yo quería. Por su tiempo y sus recomendaciones. Por su disposición para apoyarme en los siguientes pasos de mi carrera. Por sus consejos personales.

Al Dr. Rodolfo García por su tiempo, por abrirme las puertas de su laboratorio. Por escribirme las mejores cartas de recomendación, por su disposición para ayudarme en todo momento, y su amistad.

A la doctora Rosa M. Bermúdez por los comentarios y sugerencias que contribuyeron a mejorar el presente trabajo, y por la disposición y tiempo dedicado al asesoramiento de esta tesis.

A el doctor Pablo Vinuesa por recibirme en su laboratorio a pesar de no ser oficialmente mi asesor. Por proporcionarme conocimientos invaluable de su área de investigación.

A la Dra. Eva Jacinto, y al M. en C. Marco Magos por compartirme su experiencia y conocimiento, y por su amistad.

A Guadalupe Aguilar González, encargada de la Unidad de Ácidos Nucleicos del DGBM, y a la doctora Dulce Delgadillo Álvarez del UGPM, por toda su ayuda en la obtención de las secuencias analizadas en el presente trabajo.

Al biólogo Miguel Ángel Moreno Galeana por todos sus consejos para la optimización de la secuenciación de las mutantes obtenidas durante el proyecto.

A Gabriela Mora, por su auxilio constante durante toda la maestría.

A la señora Flor por su arduo trabajo que mantiene a flote el laboratorio.

## **DEDICATORIAS**

A mi padre Dios, toda la Gloria es tuya, tú me llevas de triunfo en triunfo, y de victoria en victoria. Gracias por permitirme ser instrumento de tu obra.

A mis padres. Por siempre impulsarme a seguir mis sueños, y darme todas las herramientas para lograrlos. Por todos sus sacrificios, por darlo todo por mí y por mis hermanos. Por la confianza que han tenido en mí siempre. Por ser un ejemplo excepcional de seres humanos. Por ser los mejores padres.

A mis hermanos. Por todo su apoyo incondicional en mi proyecto de vida. Por impulsarme a ser mejor cada día. Por mostrarme que la felicidad y la plenitud se alcanzan en formas muy diversas, pero que el ingrediente principal es la cercanía a Dios, y una familia unida.

A mi novio, por ser mi coach, mi colaborador, mi compañero de vida. Por soñar conmigo, y por trabajar a mi lado para cumplir nuestros anhelos. Por impulsar mi competitividad, y mostrarme que el trabajo y la disciplina son la única manera de alcanzar nuestras metas. Por tu apoyo infinito en cada aspecto de mi vida, por ser parte de ella.

A Dani, por ser el mejor cuñado que pudiera tener. Por permitirme ser parte de tu familia y de tus proyectos. Por tus valiosos consejos y tu amistad.

A la señora Rosalba, por abrirme las puertas de su casa y permitirme ser parte de su familia. Por ser un ejemplo de persona.

## INDICE

INDICE.....	6
INDICE DE FIGURAS .....	8
INDICE DE TABLAS .....	10
INDICE DE ANEXOS .....	11
LISTA DE ABREVIATURAS .....	12
RESUMEN.....	13
ABSTRACT .....	14
1. INTRODUCCIÓN .....	14
1.1 Clasificación .....	18
1.2 Mecanismo de Acción.....	18
2. ANTECEDENTES.....	21
3. JUSTIFICACIÓN .....	24
4. HIPÓTESIS .....	24
5.1 OBJETIVO GENERAL .....	24
5.2 OBJETIVOS ESPECÍFICOS .....	24
6. ESTRATEGIA EXPERIMENTAL .....	25
7. MATERIALES Y MÉTODOS .....	28
7.1 Cepas .....	28
7.2 Bacteriófagos .....	28
7.3 Plásmidos.....	28
7.4 Oligonucleótidos.....	28
7.5 Propagación de fagos por placa confluyente .....	29
7.6 Concentración de fagos con Polietilenglicol.....	30
7.7 Mutagénesis al Azar.....	30
7.8 Purificación de fragmentos amplificados por PCR .....	32
7.9 Restricción .....	32
7.10 Ligación .....	33
7.11 Preparación de células quimiocompetentes de <i>E. coli</i> .....	33
7.12 Transformación química en <i>E. coli</i> .....	34
7.13 PCR de colonia .....	34
7.14 Geles de agarosa y electroforesis.....	35

7.15 Extracción de plásmido.....	35
7.16 Preparación de células electrocompetentes de <i>P. aeruginosa</i> .....	36
7.17 Electroporación en <i>P. aeruginosa</i> .....	36
7.18 Preparación de células quimiocompetentes de <i>P. aeruginosa</i> .....	36
7.19 Transformación química en <i>P. aeruginosa</i> .....	37
7.20 Colony Blot en <i>P. aeruginosa</i> .....	37
7.21 Ensayo de funcionalidad de G2.....	39
7.22 Secuenciación de variantes de G2.....	40
7.23 Mutagénesis sitio dirigida.....	40
8. RESULTADOS.....	42
8.1 Búsqueda de aminoácidos conservados mediante análisis de alineamiento múltiple de secuencias aminoacídicas.....	42
8.2 Obtención de biblioteca de mutantes.....	44
8.3 Secuenciación de candidatas.....	55
8.4 Mutagénesis sitio dirigida.....	62
9. DISCUSION.....	70
10. CONCLUSIONES.....	70
11. PERSPECTIVAS.....	79
REFERENCIAS.....	80
ANEXOS.....	84

## INDICE DE FIGURAS

<b>Fig. 1.</b> Organización característica del locus CRISPR.	16
<b>Fig. 2.</b> Mecanismos de acción de las proteínas antiCRISPR de las familias AcrF1, AcrF2 y AcrF3.	22
<b>Fig. 3.</b> AntiCRISPR locus del bacteriófago H70.	23
<b>Fig. 4.</b> Metodología general empleada durante la realización del presente trabajo.	26
<b>Fig. 5.</b> Protocolo empleado en el ensayo de Colony Blot.	27
<b>Fig. 6.</b> Alineamiento Múltiple de Secuencias aminoacídicas de familias de antiCRISPR.	44
<b>Fig. 7.</b> Productos de PCR mutagénica.	45
<b>Fig. 8.</b> Candidatas de <i>E. coli</i> con g2 proveniente de las diferentes condiciones de mutagénesis.	45
<b>Fig. 9.</b> PCR del sitio múltiple de clonación de los distintos pools de plásmido.	46
<b>Fig. 10.</b> Candidatas de <i>P. aeruginosa</i> .	47
<b>Fig. 11.</b> Membranas analizadas por Colony Blot.	48
<b>Fig. 12.</b> Ensayo de fenotipificación del antiCRISPR G2.	49
<b>Fig. 13.</b> Ensayo de fenotipificación.	50
<b>Fig. 14.</b> Perfiles de infección de las candidatas analizadas.	52
<b>Fig. 15.</b> PCR de candidatas secuenciadas.	56
<b>Fig. 16.</b> Alineamiento nucleotídico de las variantes de g2 secuenciadas.	57
<b>Fig. 17.</b> Alineamiento aminoacídico de las candidatas secuenciadas.	58

<b>Fig. 18.</b> Mapeo de candidatas secuenciadas sobre la secuencia de G2 silvestre.	59
<b>Fig. 19.</b> Estructura tridimensional de G2.	62
<b>Fig. 20.</b> Diseño de oligos para el kit <i>Q5 Site-Directed Mutagenesis</i> .	63
<b>Fig. 21.</b> DH5 $\alpha$ transformadas con las diferentes mutantes sitio dirigidas	66
<b>Fig. 22.</b> PA14 transformadas con los diferentes pools de cada mutante sitio dirigida.	67
<b>Fig. 23.</b> PCR de candidatas provenientes de mutagénesis Y32.	68
<b>Fig. 24.</b> Fenotipificación de mutantes Y32.	68
<b>Fig. 25.</b> Alineamiento nucleotídico de las candidatas secuenciadas pertenecientes a la mutagénesis sitio dirigida Y32.	69
<b>Fig. 26.</b> Fenotipificación de mutantes V40/45D.	70
<b>Fig. 27.</b> Estructura tridimensional de Mut-A7.	74
<b>Fig. 28.</b> Estructura tridimensional de AcrF3: Acercamiento del sitio de interacción con RecA2 de Cas3.	75
<b>Fig. 29.</b> Estructura tridimensional de AcrF1.	76
<b>Fig. 30.</b> Superficie de interacción propuesta para G2.	76

## INDICE DE TABLAS

<b>Tabla 1.</b> Oligonucleótidos empleados durante el proyecto.	29
<b>Tabla 2.</b> Condiciones de PCR empleadas en la mutagénesis aleatoria.	31
<b>Tabla 3.</b> Doble digestión de fragmentos para la clonación.	32
<b>Tabla 4.</b> Ligación de variantes de g2 en pUCP24.	33
<b>Tabla 5.</b> PCR de colonias candidatas.	34
<b>Tabla 6.</b> Marcaje de oligos con <sup>32</sup> P.	38
<b>Tabla 7.</b> Reacciones de Secuenciación de candidatas.	40
<b>Tabla 8.</b> Mutagénesis Sitio-Dirigida: PCR.	41
<b>Tabla 9.</b> Mutagénesis Sitio-Dirigida: Reacción KLD.	42
<b>Tabla 10.</b> Recopilación de mutantes obtenidas por la condición A.	53
<b>Tabla 11.</b> Recopilación de mutantes obtenidas por la condición B.	54
<b>Tabla 12.</b> Recopilación de mutantes obtenidas por la condición C.	54

## INDICE DE ANEXOS

### **ANEXO 1.** Medios y soluciones.

1.1 Medio LB (Medio Luria-Bertani)	84
1.2 Medio SOC (Super Optimal broth with Catabolite repression)	84
1.3 Buffer TMG	85
1.4 Solución Tfb I	85
1.5 Solución Tfb II	86
1.6 Buffer SB 20x	86
1.7 TG Salt	86
1.8 Solución I (Colony Blot)	87
1.9 Solución II (Colony Blot)	87
1.10 Solución III (Colony Blot)	87
1.11 Solución de pre-hibridación (Colony Blot)	87
1.12 Solución de lavado I (Colony Blot)	87
1.13 Solución de lavado II (Colony Blot)	87
1.14 Solución de lavado III (Colony Blot)	88
1.15 Solución de lavado IV (Colony Blot)	88

### **ANEXO 2.** Electroferogramas.

2.1 Mut-A1	89
2.2 Mut-A2	90
2.3 Mut-A3	91
2.4 Mut-A4	92
2.5 Mut-A5	93
2.6 Mut-A6	94

2.7 Mut-A7	95
2.8 Mut-A8	96
2.9 Mut-A9	97
2.10 Mut-A10	98
2.11 Mut-B3	99
2.12 Mut-B4	100
2.13 Mut-B12	101
2.14 Mut-C7	102
2.15 Mut-C12	103
2.16 Y32-B1	104
2.17 Y32-C2	105
2.18 Y32-D5	106
2.19 Y32-D1	107

#### LISTA DE ABREVIATURAS

AcrF	Familia anti-CRISPR
BLAST	Herramienta básica de búsqueda de alineamientos locales
Cas	Asociado a CRISPR
Cascade	Complejo asociado a CRISPR para la defensa antiviral
CRISPR	Repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas
crRNA	RNA derivado de la región CRISPR
crRNP	Complejo ribonucleoproteico CRISPR
Csy	CRISPR subtipo <i>Yersinia</i>
dNTPs	Desoxinucleótidos trifosfatados
D.O.	Densidad óptica
<i>E. coli</i>	<i>Escherichia coli</i>
LB	Luria-Bertani
MCS	Sitio múltiple de clonación
PAM	Motivo adyacente al protoespaciador
PCR	Reacción en cadena de la polimerasa
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
pb	Pares de bases
UFP	Unidades formadoras de placa
Δ	Delección

## RESUMEN

En procariontes se han caracterizado sistemas específicos que tienen como objetivo evitar la invasión por elementos genéticos móviles a los que están constantemente expuestos, como son plásmidos y bacteriófagos. Uno de éstos es el sistema CRISPR-cas, el cual se ha definido como un sistema de defensa constitutiva que provee inmunidad adaptativa contra elementos genéticos foráneos, mediante la degradación del DNA invasor guiada por secuencias complementarias a éste. Sin embargo, la constante interacción entre bacterias y sus virus, ha conducido a la generación de diversas estrategias que les permiten a los fagos evadir la defensa bacteriana; una de ellas es la expresión de proteínas denominadas antiCRISPR, las cuales se unen a diversos componentes del sistema CRISPR-cas, e impiden el reconocimiento o degradación del DNA fágico. Hasta el momento se han descrito 10 familias de genes antiCRISPR en bacteriófagos pertenecientes al grupo D3112virus, los cuales inhiben el sistema CRISPR-cas I-F de *Pseudomonas aeruginosa*. Sin embargo, a pesar de que se conocen distintos tipos de estas familias, se sabe poco de ellas. A la fecha, se ha caracterizado el mecanismo de acción y la estructura tridimensional de 3 de ellas (AcrF1, AcrF2 y AcrF3). En nuestro grupo de investigación se identificó por sintenia un gen antiCRISPR en la región de plasticidad genómica G (antiCRISPR locus) del fago H70, al cual se le denominó g2. Durante el presente trabajo se caracterizó la proteína G2 mediante un análisis funcional con el fin de dilucidar los aminoácidos implicados en la interacción de G2 con el sistema CRISPR-cas. La estrategia se basó en la generación de una colección de mutantes aleatorias que nos permitió identificar las posiciones aminoacídicas importantes para la función de la proteína de estudio y la posible superficie de interacción. Posteriormente, ésta fue comprobada por mutagénesis sitio dirigida. Con base en los resultados obtenidos, se identificaron aminoácidos de dicha región, cuya modificación resulta en la pérdida parcial o total de la actividad antiCRISPR. Asimismo, con base en el análisis de la estructura putativa de G2, inferimos que el extremo carboxilo terminal podría estar relacionado con el correcto plegamiento de la proteína.

Palabras clave: fago, CRISPR-Cas, antiCRISPR, mutagénesis, interacción proteica.

## ABSTRACT

Prokaryotes are constantly exposed to mobile genetic elements, such as plasmids and bacteriophages, in nature. To avoid the invasion of such elements, bacteria have developed specific defense systems. One of these is the CRISPR-cas system, which has been defined as a constitutive defense system that provides adaptive immunity against foreign genetic elements through the degradation of invasive DNA, guided by sequences complementary to it. Yet, the constant interaction between bacteria and the viruses infecting them, has led to the development of strategies that allow phages to evade this bacterial defense throughout the expression of proteins called antiCRISPR, which bind to components of CRISPR-cas system to prevent the recognition or degradation of the phage DNA. To date, 10 families of antiCRISPR genes that inhibit the I-F CRISPR-cas system of *Pseudomonas aeruginosa* have been described in bacteriophages of the D3112virus group. Although some representative types of these families have been studied, little is known about these proteins. Currently, both the mechanism of action and the three-dimensional structure have been characterized only for three of the recognized families (AcrF1, AcrF2 and AcrF3). In our research group, an antiCRISPR gene was identified in the region of genomic plasticity G (antiCRISPR locus) of phage H70 and was named g2. In this work, we characterize the G2 protein through a functional analysis in order to elucidate which amino acids are involved in the interaction of G2 with the CRISPR-cas system. The strategy was based on the generation of a collection of random mutants that allowed us to identify the amino acid positions which are important for the function maintenance of the anti-CRISPR protein and also the putative interaction surface of the protein. Finally, this was verified by site-directed mutagenesis. Based on the obtained results, we identified amino acids that are key for the G2 function. Modification of such residues results in the partial or total loss of antiCRISPR activity. Additionally, based on the analysis of the putative structure of G2, we could infer that the terminal carboxyl end of the protein could be involved into its correct folding.

Keywords: Phage, CRISPR-cas, antiCRISPR, mutagenesis, protein interaction.

## 1. INTRODUCCIÓN

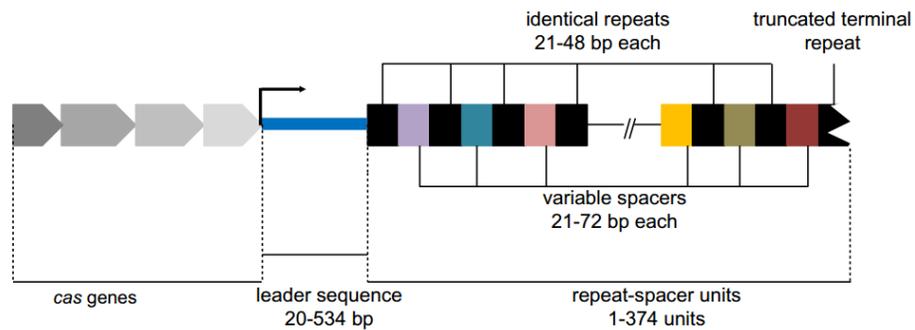
En procariotas, se han caracterizado sistemas específicos, los cuales están especialmente dirigidos a limitar la invasión de DNA proveniente de bacteriófagos, algunos de ellos son la inhibición de la adsorción, el bloqueo de la inyección, la restricción/modificación, y la infección abortiva.

En 1987, Ishino y colaboradores fueron los primeros en identificar en el genoma de *Escherichia coli*, arreglos de secuencias en clúster regularmente interespaciadas por regiones cortas palindrómicas, a las cuales se les denominó “secuencias CRISPR”. Posteriormente, Mojica et al. (2005) reportaron que estos elementos no codificantes, junto con las proteínas Cas, actúan como un sistema de defensa constitutiva contra DNA invasivo.

Se ha establecido que las regiones CRISPR-cas sirven como un sistema de inmunidad adaptativa contra elementos genéticos foráneos, en el cual, las secuencias espaciadoras juegan un rol importante como responsables del reconocimiento específico del DNA invasor (Barrangou et al., 2007; Brouns et al., 2008; Marraffini and Sontheimer, 2010). La incorporación de nuevas secuencias espaciadoras en el genoma bacteriano puede ser comparada con un tipo de memoria genética, la cual protege a la célula de infecciones posteriores. Asimismo, el sistema CRISPR-cas es heredable, y estos espaciadores están presentes en la progenie (Van der Oost et al., 2009). Además, estudios en *Thermus thermophilus* han mostrado que el sistema censa infecciones por fagos, ya que durante el proceso de infección se sobreexpresan componentes del sistema CRISPR-cas (Agari et al., 2010).

Actualmente, se sabe que cerca del 40-70% de todas las bacterias, y casi todas las arqueas, contienen CRISPRs en sus genomas (Godde and Bickerton, 2006; Grissa et al., 2007a; Mojica et al., 2005; Rousseau et al., 2009). Sin embargo, se sospecha que no todos los sistemas CRISPR-cas presentes en el genoma son activos, debido a ello, se han establecido algunas características generales que permiten distinguir casetes CRISPR funcionales y no funcionales.

Se ha establecido que un sistema funcional, debe estar constituido por uno o más loci no codificante CRISPR, acompañados por dos determinantes: los genes cas y una secuencia líder (Horvath et al., 2008; Makarova et al., 2006; Sorek et al., 2008) (Fig. 1).



**Fig. 1. Organización característica del locus CRISPR.** Szczepankowska A., 2012. En la figura se muestra una representación del sistema CRISPR-cas. En el extremo derecho se encuentra el locus CRISPR, conformado por repetidos directos cortos (en negro) separados por secuencias variables (en colores). En el extremo izquierdo se muestran los genes asociados a CRISPR (*cas*), y en el medio se muestra una secuencia líder.

El locus CRISPR está constituido por dos tipos de elementos no codificantes: (i) repetidos directos (21-48 pb) separados por (ii) secuencias espaciadoras no repetidas (21-72 pb). Dentro de cada región CRISPR, la longitud de los espaciadores y los repetidos directos es conservada, excepto el último repetido del extremo 3', el cual aparece truncado en el 30% de los casos (Horvath et al., 2008; Jansen et al., 2002). Estos repetidos directos están fuertemente conservados dentro del locus y parecen ser especie específicos (Deveau et al., 2010; Goode and Bickerton, 2006).

Por el contrario, las secuencias espaciadoras son generalmente muy diferentes entre sí, y altamente variables. De hecho, a excepción de algunos casos, los espaciadores dentro de un genoma suelen ser únicos.

Como se mencionó, otro determinante estrictamente conectado con loci CRISPR activos, es la secuencia líder. Esta región no codificante rica en A-T, está localizada en el extremo 5' del arreglo CRISPR y tiene una longitud (dependiendo de la especie), de 20 a 534 pb (Jansen et al., 2002; Lillestøl et al., 2006). Se ha propuesto que esta región está involucrada en la adquisición de nuevos espaciadores, incluso se ha sugerido que puede contener sitios de unión a proteínas específicas (por ejemplo, las proteínas Cas), por lo cual, jugaría un papel importante en la duplicación de repetidos y/o la integración de espaciadores (Barrangou et al., 2007; Pourcel et al., 2005). Adicionalmente, se ha propuesto la hipótesis de que la secuencia líder actúa como promotor de la región CRISPR (Brouns et al., 2008; Hale et al., 2008; Lillestøl et al., 2006, 2009; Marraffini and Sontheimer, 2010), la cual está soportada por el hallazgo de que loci CRISPR que no contienen secuencia líder, no incorporan nuevos espaciadores y permanecen como remanentes inactivos (Lillestøl et al., 2006).

Otro determinante importante de los sistemas CRISPR-cas son las proteínas codificadas por los genes cas, cuyo número dentro de un genoma puede variar de 4 a 20 (Barrangou et al., 2007; Haft et al., 2005; Sorek et al., 2008). Adicionalmente, se ha demostrado que estas proteínas solo están presentes en genomas que contienen CRISPR, lo cual implica su estrecha relación.

Generalmente, los genomas que portan el locus CRISPR no contienen todos los genes cas, sino combinaciones de estos, siendo los más frecuentes *cas1* y *cas2* (Haft et al., 2005; Makarova et al., 2006). De hecho, se ha empleado a *cas1* como marcador para detectar sistemas CRISPR-cas (Makarova et al., 2006; Sorek et al., 2008). La proteína Cas1 de *Pseudomonas aeruginosa* ha sido caracterizada bioquímicamente, estableciéndose como una proteína metal-dependiente, que además posee actividad endonucleasa/integrasa de secuencia no específica de DNA (Wiedenheft et al., 2009; Zegans et al., 2009). Estos hallazgos permitieron proponer que la Cas1 está implicada en el reconocimiento y/o corte del DNA foráneo y la integración de espaciadores en el locus CRISPR (Deveau et al., 2010).

## 1.1 Clasificación

La clasificación vigente de los sistemas CRISPR-cas ha sido propuesta por Makarova et al. (2011), la cual está basada en la información de secuencia y estructura de los componentes proteicos (proteínas Cas), distinguiéndose 3 tipos principales de sistemas (I, II y III), en los cuales los genes core son *cas1* y *cas2*. Además, cada tipo de sistema está caracterizado por ciertas firmas genómicas, siendo la del tipo I *cas3*, el tipo II *cas9*, y el tipo III *cas10* (Makarova et al., 2011a). A su vez, se han establecido 8 subtipos basados en un análisis de filogenia de la Cas1 y clustering de los genes *cas*: Ecoli, Ypest, Apern, Nmeni, Mtube, Tneap, Hmari, y Dvulg (Haft et al., 2005; Makarova et al., 2006).

*P. aeruginosa* porta sistemas CRISPR-cas tipo I, siendo los subtipos I-F, I-E y I-C los que se han identificado en cepas clínicas (Cady et al., 2011; Eshoh et al., 2013; van Belkum et al., 2015), sin embargo, el subtipo I-F es el prevalente (Cady et al., 2011; van Belkum et al., 2015). Los genes *cas* que lo conforman son seis: *cas1*, *cas3*, *csy1*, *csy2*, *csy3* y *csy4*, siendo los últimos cuatro, únicos de este subtipo (Haft et al., 2005). Este sistema consta de dos loci CRISPR (CRISPR1 y CRISPR2) que flanquean a los genes *cas* (Grissa et al., 2007). En 2012, Cady y colaboradores mostraron la primera evidencia de que el sistema CRISPR-cas de la cepa PA14 de *P. aeruginosa* es funcional y media la resistencia a la infección por fagos.

## 1.2 Mecanismo de Acción

Van der Oost y colaboradores describieron en 2009 las principales etapas del mecanismo de acción del Sistema CRISPR-cas, las cuales se dividen en:

**Etapa 1.** Fase de inmunización (o adaptación), durante la cual nuevos espaciadores provenientes de DNA foráneo son incorporados al locus CRISPR.

Análisis genómicos detallados han determinado que la integración de nuevos espaciadores se realiza casi siempre en el extremo proximal de la región líder, específicamente entre el líder y el primer espaciador (Barrangou et al., 2007; Deveau et al., 2008; Pourcel et al., 2005).

Se ha propuesto que las proteínas Cas1, Cas2 y Cas4 se unen dentro de la región líder y juegan un papel principal en la etapa de adaptación, donde la Cas1 corta el DNA invasor en fragmentos que sirven como precursores de espaciadores (van der Oost et al., 2009; Wiedenheft et al., 2009).

La selección de los fragmentos de DNA foráneos, los cuales serán integrados como espaciadores en el locus CRISPR, parece no ser aleatoria. Un determinante importante en los tipos I y II parece ser la presencia de las secuencias PAM (proto-spacer adjacent) (Deveau et al., 2008; Mojica et al., 2009; Semenova et al., 2009), las cuales son empleadas para el reconocimiento de proto-espaciadores, y eliminadas en los espaciadores, lo cual también parece ser un sistema de la bacteria para distinguir su genoma del exógeno.

Durante esta etapa, también es necesario duplicar las secuencias repetidas que separan cada espaciador. Se cree que esto se lleva a cabo mediante un evento de recombinación, de manera similar al proceso de integración de transposones (Steiniger-White et al., 2004).

**Etapa 2.** Expresión de CRISPR, la cual involucra la transcripción de las secuencias CRISPR y el subsecuente procesamiento para dar lugar a pequeños RNAs guías. Asimismo, implica la expresión de las proteínas Cas.

Durante esta fase, el locus CRISPR se transcribe como un mRNA largo (pre-crRNA). Este proceso se lleva a cabo de manera constitutiva e inicia en la secuencia líder. Posteriormente, el transcrito es procesado en fragmentos más pequeños, y finalmente en pequeñas moléculas de RNA maduras (crRNA).

En el sistema CRISPR-cas I-F se ha observado que los crRNA interactúan con las proteínas Cas denominadas Cascade (Cas-complex for anti-virus defense). Estas proteínas reconocen estructuras secundarias específicas dentro de las secuencias de repetidos del pre-crRNA (en las regiones palindrómicas) y las digieren, en consecuencia, los crRNA cortos son generados (35-46 nt). En *Pseudomonas aeruginosa* se determinó que la proteína responsable de llevar a cabo dicho corte es la endonucleasa Csy4 (Haurwitz et al., 2010).

El producto final de la expresión CRISPR son crRNA mono-espaciadores con repetidos parciales en el extremo 5', los cuales actúan como etiquetas que atraen a las proteínas Cas (particularmente nucleasas), y las direccionan a elementos genéticos foráneos.

**Etapas 3.** Interferencia, efectuada por los pequeños RNAs guías (generados en la etapa anterior) acoplados a las proteínas Cas, las cuales se unen y degradan el blanco invasivo.

El papel principal en esta fase se lleva a cabo por los crRNAs maduros, que sirven como guías en el reconocimiento y destrucción de elementos invasores con secuencias coincidentes.

En el tipo I, el crRNA maduro no actúa solo, sino que forma un complejo de ribonucleoproteína (crRNP) con las proteínas de la cascada y las guía al DNA blanco, degradado posteriormente por la Cas3 (Sontheimer y Marraffini, 2010). La detección de blancos por los crRNPs se basa en un proceso de apareamiento de bases, que exige complementación exacta entre espaciador y proto-espaciador, incluso una mutación puntual dentro de una de las secuencias podría suprimir la interferencia de CRISPR (Deveau et al, 2008; Marraffini y Sontheimer, 2010).

Se ha propuesto que la Cas3 tiene un papel en la interferencia al estar involucrada en el alineamiento de bases del crRNA con un proto-espaciador, y su posterior degradación (Brouns et al, 2008; Karginov y Hannon, 2010). En ensayos computacionales, la Cas3 fue identificada como una helicasa, lo cual sugirió estar implicada en el desenrollamiento del DNA blanco, permitiendo la hibridación del proto-espaciador con el crRNA. Por otra parte, la mayoría de las proteínas Cas3 se fusionan con un dominio de nucleasa (Cas2), implicado en la degradación del blanco.

## 2. ANTECEDENTES

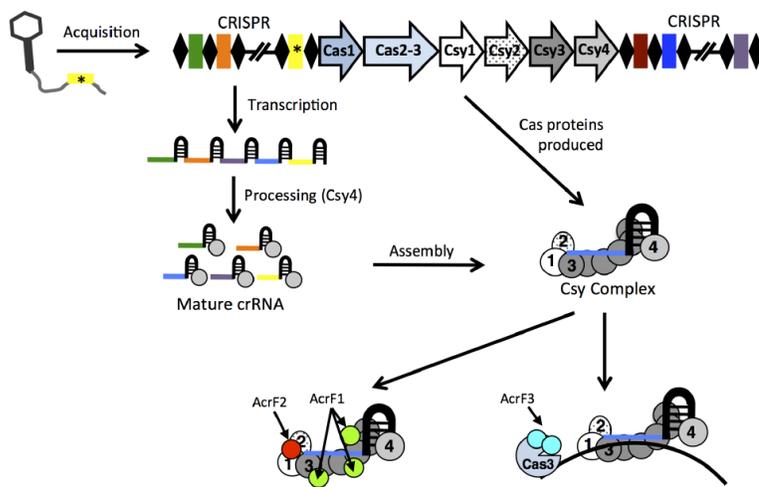
La constante interacción entre las bacterias y sus bacteriófagos eventualmente conlleva a la selección de los fagos que son aptos para evadir la resistencia mediada por CRISPR-cas. Así como los procariontes han desarrollado múltiples estrategias para evitar la infección por virus, los virus han desarrollado contramedidas para estas estrategias. Se han descrito diversos mecanismos que han demostrado evadir el sistema CRISPR-cas, siendo la forma más básica de escapar a este sistema la mutagénesis aleatoria del proto-espaciador que afecta bases clave en la interacción con el crRNA, o el reconocimiento de las regiones PAM (Barrangou et al., 2007), de hecho, se ha demostrado que una simple mutación puntual en estas regiones puede inhibir este mecanismo de defensa (Deveau et al., 2008; Lillestøl et al., 2009; Mojica et al., 2009).

Esta continua batalla evolutiva entre bacterias y los virus que las infectan dio pie a los investigadores para proponer la hipótesis sobre la existencia de sistemas antiCRISPR codificados por fagos. En 2003, Peng y colaboradores sugirieron que el profago residente en *Sulfolobus solfataricus* P2 codificaba una proteína con putativa actividad antiCRISPR, demostrando que esta se une preferencialmente a los repetidos CRISPR e induce cambios conformacionales en el DNA, inhibiendo así el sistema (Peng et al., 2003; Sorek et al., 2008). Asimismo, se descubrieron genes homólogos en otros genomas bacterianos, siempre dentro de secuencias de profagos, indicando que otros sistemas similares al putativo antiCRISPR pueden funcionar en otras especies microbianas (Sorek et al., 2008).

Posteriormente, Bondy-Denomy y colaboradores encontraron en diversos bacteriófagos del grupo D3112virus que infectan a *Pseudomonas aeruginosa*, 5 proteínas que eran capaces de inhibir el sistema CRISPR-cas I-F (AcrF1, AcrF2, AcrF3, AcrF4 y AcrF5 (Bondy-Denomy, 2013), a las cuales denominaron *antiCRISPR*. Un año más tarde, reportaron nuevas proteínas que inhiben al subtipo I-E de *P. aeruginosa*, en el mismo grupo de fagos (Pawluk et al., 2014).

Estas proteínas se encuentran codificadas en una región accesoria del genoma de dichos fagos, esta región está flanqueada río arriba por el homólogo del gen que codifica la proteína G del fago Mu, y río abajo por el gen que codifica la proteína proteasa/scaffold (Fig. 3). En el extremo 5' de la región se ha identificado un promotor putativo (Cazares et al., 2014) y en el extremo 3' se encuentra un gen que codifica una proteína con un dominio de unión a DNA hélice-vuelta-hélice, la cual se ha propuesto como un putativo regulador transcripcional, que podría regular al operón anti-CRISPR (Pawluk *et al.*, 2014), y a la que se ha designado como una firma del locus antiCRISPR.

En 2015, el grupo del Dr. Alan R. Davidson caracterizó el mecanismo molecular de acción de 3 (AcrF1, AcrF2 y AcrF3) de las 5 familias de genes antiCRISPR que inhiben el sistema CRISPR-cas I-F (Bondy-Denomy et al., 2015) (Fig. 2). Ellos encontraron que todas las familias analizadas inhiben el sistema de defensa bacteriana mediante la unión a distintas proteínas Cas; la familia AcrF1 se une a las proteínas Csy3, creando un impedimento estérico para la unión del complejo con el DNA; la familia AcrF2 se une al heterodímero Csy1-Csy2 localizado en el extremo 5' del crRNA, también bloqueando la unión con el DNA foráneo; y por último la familia AcrF3 se une a la proteína Cas3, inhibiendo su actividad de nucleasa, e impidiendo la degradación del DNA blanco.

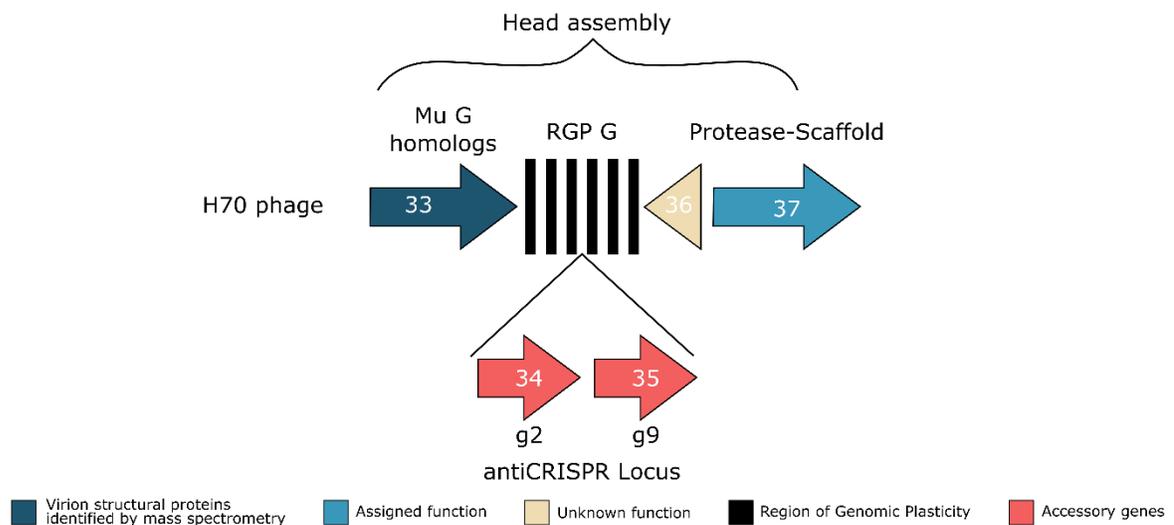


**Fig. 2. Mecanismos de acción de las proteínas antiCRISPR de las familias AcrF1, AcrF2 y AcrF3.** Bondy-Denomy et al., 2015. El mecanismo CRISPR-cas cuenta con diferentes etapas (adquisición, transcripción del locus CRISPR, maduración del crRNA, traducción de las proteínas cas, formación del complejo Csy, e interferencia), en las cuales puede darse la inhibición por las proteínas antiCRISPR.

Las proteínas antiCRISPR AcrF1 y AcrF2 interactúan directamente con el complejo previniendo la unión con el DNA. AcrF3 se une con la nucleasa Cas3, bloqueando su reclutamiento.

Por otro lado, un análisis comparativo del genoma de bacteriófagos D3112virus realizado por Cazares et al, 2014, evidenció la presencia de genes antiCRISPR localizados en la región de plasticidad genómica RGP G en los fagos pertenecientes a la colección del laboratorio, siendo g2 del fago H70, un putativo antiCRISPR no antes reportado (Fig. 3).

Posteriormente, De la Cruz (De la Cruz, C. A., 2016) demostró que dicho gen confiere resistencia ante el sistema CRISPR-cas I-F de la cepa PA14 de *Pseudomonas aeruginosa* (datos no publicados), y recientemente Pawluk y colaboradores lo han reportado como una nueva familia de proteínas antiCRISPR, AcrF7 (Pawluk et al, 2016).



**Fig. 3. AntiCRISPR locus del bacteriófago H70.** Análisis del pangenoma de bacteriófagos D3112virus reveló la presencia de un antiCRISPR nuevo (g2) en el genoma fago H70 (Cazares et al, 2014) en la región de plasticidad genómica G (RGP G). En esta misma región se encontró un ORF con putativa función de factor de transcripción (g9). Los números dentro de las flechas, corresponde al número de ORF en el genoma del fago H70. El código de colores de cada flecha (ORF) fue muestra en la parte inferior de la figura.

### **3. JUSTIFICACIÓN**

La batalla por la supervivencia entre bacterias y los virus que las infectan (*arm race*), ha conducido a la evolución de diversos sistemas de defensa bacteriana y su contraparte en los fagos. Hasta ahora han sido reportadas 10 familias de proteínas que inhiben el sistema CRISPR cas I-F de *Pseudomonas aeruginosa*, y se ha determinado el mecanismo de acción de tres de ellas. Sin embargo, debido a que G2 no ha sido caracterizado, el análisis funcional de éste permitirá ampliar el conocimiento del funcionamiento e interacción de los antiCRISPR con el sistema CRISPR-cas.

### **4. HIPÓTESIS**

Existe un conjunto de residuos aminoacídicos en el antiCRISPR G2 implicados en la interacción con el sistema CRISPR-cas. Una mutación en dichos aminoácidos afectará el reconocimiento de este, impidiendo la protección de bacteriófagos.

#### **5.1 OBJETIVO GENERAL**

Realizar un análisis funcional de la proteína anti-CRISPR G2 perteneciente al bacteriófago H70, con el fin de identificar los residuos aminoacídicos implicados en la funcionalidad de la proteína.

#### **5.2 OBJETIVOS ESPECÍFICOS**

- Realizar un alineamiento de las diferentes familias de antiCRISPR con el fin de buscar aminoácidos conservados.
- Obtener una biblioteca de mutantes de g2.
- Determinar las variantes de G2 no funcionales mediante un ensayo de fenotipificación.
- Secuenciar las variantes no funcionales de G2.
- Determinar los residuos aminoacídicos importantes para la funcionalidad de G2.

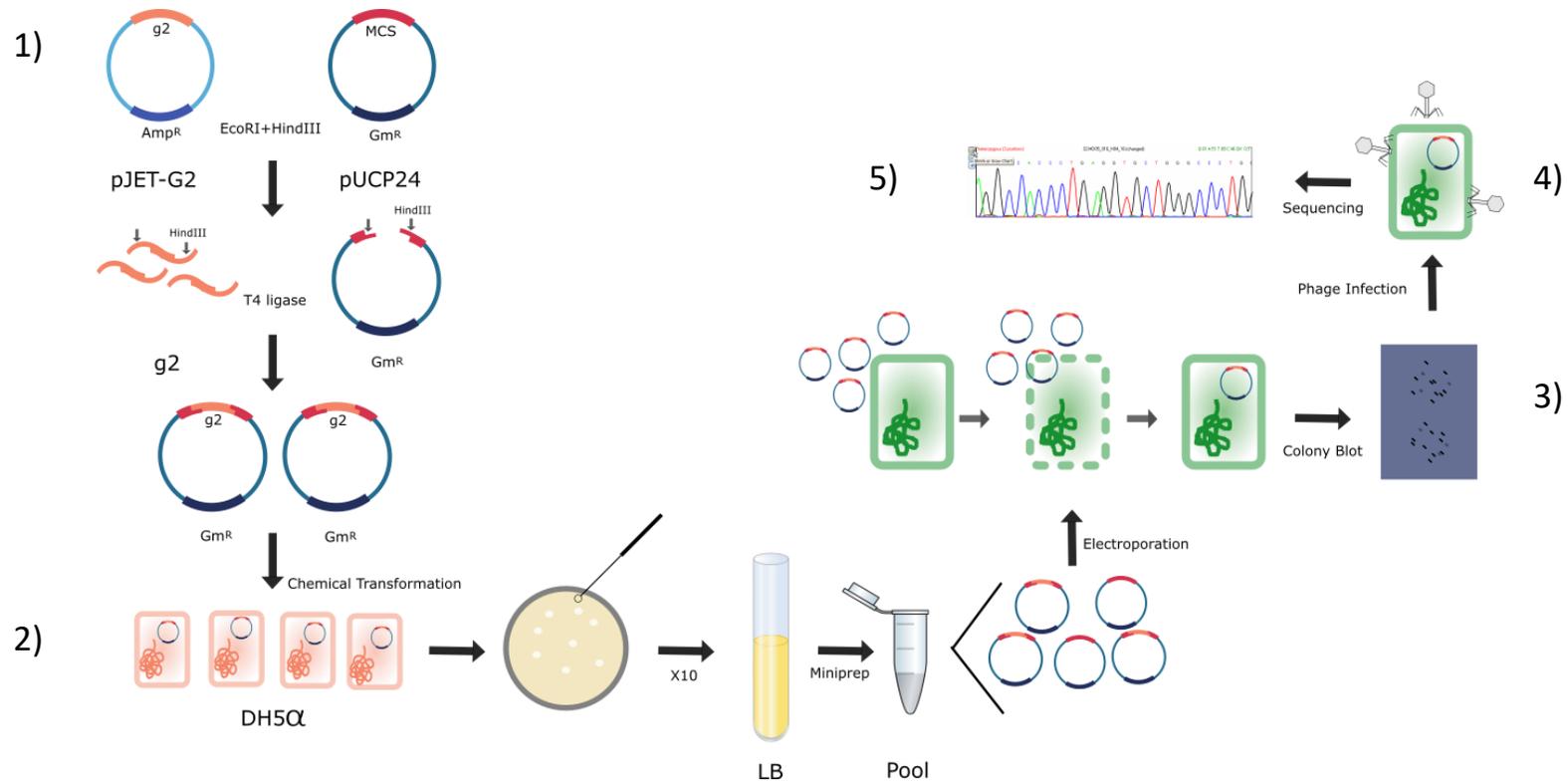
## 6. ESTRATEGIA EXPERIMENTAL

La estrategia general empleada se basó en la generación de mutantes a partir del protocolo propuesto por Cadwell y colaboradores, así como en modificaciones de éste. A través de la mutagénesis por PCR se generaron fragmentos mutantes que posteriormente fueron clonados en el vector shuttle pUCP24, el cual tiene la capacidad de replicarse tanto en *E. coli*, como en *P. aeruginosa*. Dicha clonación se transformó en la cepa DH5 $\alpha$  de *E. coli*, posteriormente se tomaron 10 colonias candidatas y se mezclaron en 10 mL de medio LB para formar un cultivo mixto, esto fue repetido al menos 30 veces para cada una de las condiciones de PCR empleadas. A continuación, se extrajo plásmido de cada cultivo mixto para dar lugar a diversos pools de plásmidos.

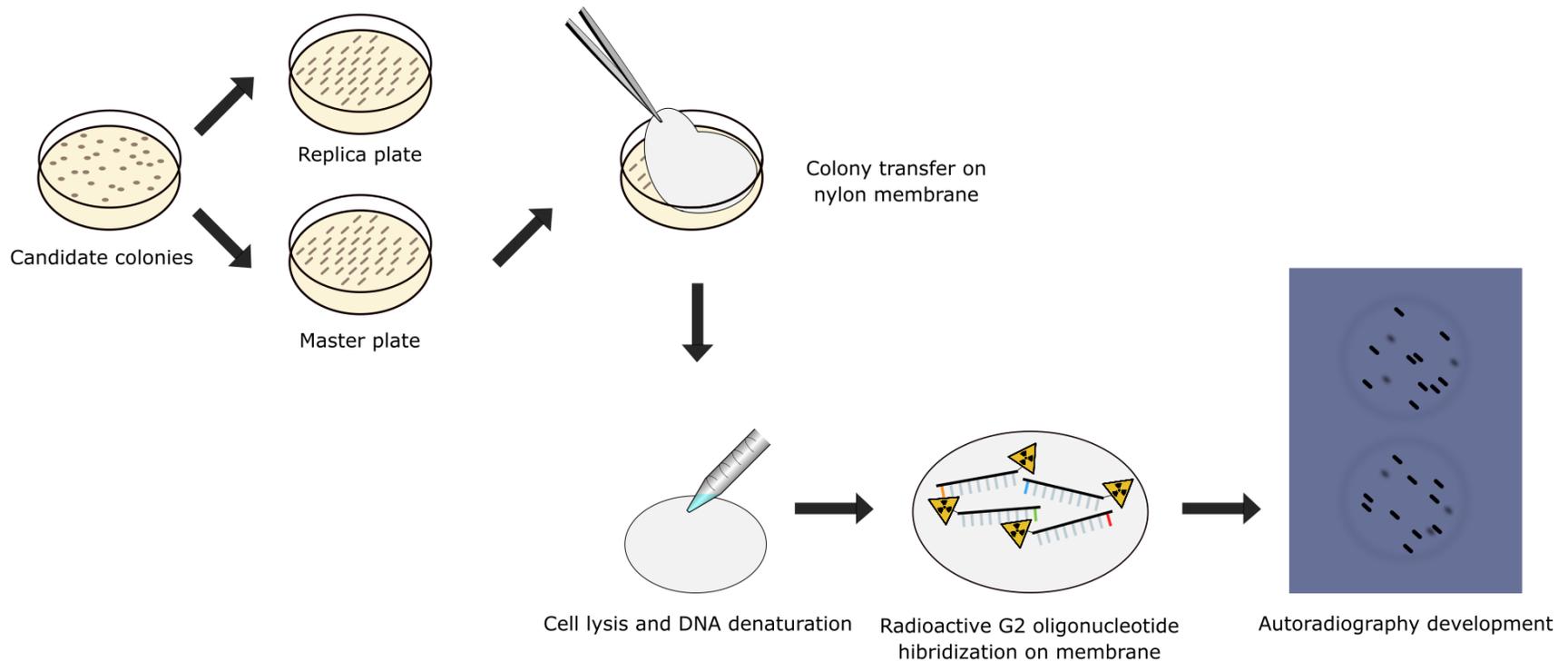
Estos pools fueron probados por PCR, empleando oligos flanqueantes al sitio de clonación múltiple del vector pUCP24 (MCS pUCP24 forward y reverse), para determinar la presencia de vectores con el inserto deseado (g2). Todos aquellos pools que contenían el antiCRISPR fueron transformados en la cepa PA14 de *P. aeruginosa*. Después, se sembraron al menos 15 candidatas provenientes de la transformación de cada pool, y en seguida se identificaron las candidatas de PA14 que portaban el gen antiCRISPR mediante la técnica colony blot (Fig. 5). Todas aquellas candidatas positivas fueron cultivadas en LB líquido para realizar tapices de cada una de ellas, y probar la eficiencia de infección de un fago sensible (JBD18) al sistema CRISPR-cas I-F, y con esto tipificar cada una de las variantes de g2 clonadas. Posteriormente, se secuenciaron las variantes no funcionales del antiCRISPR a partir del producto de PCR de colonia de cada candidata, empleando los oligos de pUCP24 antes mencionados. Adicionalmente, se secuenciaron variantes con fenotipo silvestre, para conocer la tasa de mutagénesis de cada una de las condiciones empleadas de PCR mutagénica.

Tras analizar las secuencias obtenidas, se propuso una putativa superficie de interacción. Para probar dicha hipótesis, se diseñaron oligonucleótidos para realizar mutagénesis sitio dirigida de los aminoácidos propuestos. Por último, cada una de las mutantes diseñadas fueron probadas fenotípicamente mediante un ensayo de infección del fago JBD18.

En la figura 4 se muestra un esquema de la metodología general empleada en la realización de este trabajo (Fig. 4), así como la metodología particular del ensayo de colony blot (Fig. 5).



**Fig. 4. Metodología general empleada durante la realización del presente trabajo.** En la figura se representan los pasos seguidos durante la obtención de la biblioteca de mutantes al azar y su secuenciación. La estrategia general consiste en 1) la clonación de un producto de diversas PCRs mutagénicas en el vector pUCP24, 2) su transformación en *E. coli* para la amplificación de los plásmidos, seguido de 3) la transformación en *P. aeruginosa* y su 4) fenotipificación, siendo el último paso 5) la secuenciación de candidatas.



**Fig. 5. Protocolo empleado en el ensayo de Colony Blot.** Este ensayo consiste en la resiembra de colonias candidatas en 2 cajas (replica y master), para posteriormente transferir las colonias de una de ellas a una membrana de nylon. La membrana es tratada con soluciones para promover la lisis celular y la desnaturalización del DNA, el cual hibrida con una sonda marcada con  $^{32}\text{P}$ . Las colonias que portan g2 son identificadas como manchas negras mediante la exposición de una placa autorradiográfica a las membranas.

## **7. MATERIALES Y MÉTODOS**

### **7.1 Cepas**

Las cepas de *P. aeruginosa* PA14 y PA14  $\Delta$ CRISPR-Cas (loci CRISPRs y genes *cas* deletados) provienen del laboratorio del Dr. George O'toole de Geisel School of Medicine, Dartmouth College, E.U.A. Estas fueron proporcionadas por el Dr. Alan R. Davidson de la Universidad de Toronto, Canadá, en el 2014. Adicionalmente se empleó la cepa de *E. coli* DH5 $\alpha$  para la obtención de la biblioteca de mutantes, tanto mutantes aleatorias como sitio dirigidas.

### **7.2 Bacteriófagos**

El fago portador de g2, H70, pertenece a la colección del laboratorio y proviene de una cepa clínica de *Pseudomonas aeruginosa* (Cazares *et al.*, 2014). Para la fenotipificación de la proteína G2 se empleó el fago sensible al sistema CRISPR-cas JB18, el cual pertenece al grupo B3virus y porta proto-espaciadores complementarios a espaciadores presentes en el locus CRISPR de la cepa PA14 y no codifica ningún antiCRISPR, por lo cual no puede evadir el sistema CRISPR-cas. Dicho bacteriófago fue proporcionado por el Dr. Alan R. Davidson.

### **7.3 Plásmidos**

El plásmido pUCP24 inicialmente fue proporcionado por la Dra. Herminia Loza Tavera de la Facultad de Química de la UNAM. Posteriormente fue modificado por la M.C. Ana de la Cruz Cruz (De la Cruz, C. A., 2016), para inhibir la producción de proteínas de fusión (inserto+lacZ alfa).

### **7.4 Oligonucleótidos**

En la Tabla 1 se muestra el nombre de cada oligo empleado en el proyecto, y su respectiva secuencia.

**Tabla 1. Oligonucleótidos empleados durante el proyecto.**

<b>Nombre</b>	<b>Secuencia (5' - 3')</b>
pJET1.2 forward	CGACTCACTATAGGGAGAGCGGC
pJET1.2 reverse	AAGAACATCGATTTTCCATGGCAG
MCS pUCP24 forward (M13 F (-20))	GTAAAACGACGGCCAGT
MCS pUCP24 reverse	GTCATAGCTGTTTCCTG
G2 exp forward	CGTTCACCAGCATCGTTACC
V45D forward	ACCGCCGTGGwCGAGGAAATG
V40D reverse	AGGAwCGCTAACACCGTATG
D29 forward	CAGTTCGACGbTTCAGCATATGAAG
D29 reverse	CTGTCCCGCTTCAACGTA
Y32 forward	CGATTCAGCAgnTGAAGAGGCATACGGTGTTAG
Y32 reverse	TCGAACTGCTGTCCCGCT
E33 forward	GATTCAGCATATGbAGAGGCATACG

Donde b= C, G, T; n= A, C, G, T; y w= A, T.

### **7.5 Propagación de fagos por placa confluyente**

Se tomaron 70 µL de un stock de fagos de bajo/mediano título ( $10^4$ - $10^6$  UFP/mL) y se mezclaron con 150 µL de un cultivo overnight bacteriano. Posteriormente, se añadieron 3.5 mL de medio TΦ semisólido y se vertió la mezcla sobre una caja de TΦ para formar un tapiz (x5). Las cajas se incubaron a 37°C toda la noche. Al día siguiente se rasparon los tapices con una espátula, y se colocaron en un tubo falcón.

Se añadieron 5 ml de buffer TMG a cada una de las 5 cajas, se mantuvieron en agitación a 4°C por 2 hs, y se decantaron en el tubo falcon que contenía los tapices. Se colocó en agitación el tubo a 4°C durante 2.5-4 horas. Pasado este tiempo, se centrifugó a 509 g por 10 min, se recuperó el sobrenadante y se almacenó a 4°C.

### **7.6 Concentración de fagos con Polietilenglicol.**

A 25 mL de lisado fágico obtenido por propagación por placa confluyente se le agregó DNAsa y RNAsa a una concentración final de 1µg/ml y se incubó a 37°C por 30 min. Posteriormente, se adicionó polietilenglicol 8000 (PEG-8000) y cloruro de sodio (NaCl) a una concentración final del 16% y 1M respectivamente. Se incubó en hielo durante toda la noche y se centrifugó a 10,509 g por 30 min. Se decantó el sobrenadante y la pastilla de fago se resuspendió en 1ml de buffer TMG. Se agregó un volumen de cloroformo, se agitó en vortex y se centrifugó a 8,944 g por 10 min para eliminar el PEG-8000. Se tomó la fase acuosa y se almacenó a 4°C.

### **7.7 Mutagénesis al Azar**

Se optó por realizar la mutagénesis al azar empleado el método de promoción de errores durante la reacción de PCR. Para esto nos basamos en el trabajo de Cadwell y colaboradores (Cadwell et al, 1992).

Se emplearon 3 condiciones mutagénicas (A, B, y C) a lo largo del proyecto, las cuales varían en la concentración de MgCl<sub>2</sub>, MnCl<sub>2</sub>, y/o en el número de ciclos de PCR (tabla 2).

**Tabla 2. Condiciones de PCR empleadas en la mutagénesis aleatoria.**

Reactivo	Condición A	Condición B	Condición C
dGTP	0.2mM	0.2mM	0.2mM
dATP	0.2mM	0.2mM	0.2mM
dCTP	1mM	1mM	1mM
dTTP	1mM	1mM	1mM
MgCl <sub>2</sub>	5.5 mM	6 mM	5.5 mM
MnCl <sub>2</sub>	0.5 mM	0.3 mM	0.5 mM
Taq polimerasa	5 unidades	5 unidades	5 unidades
KCl	500mM	500mM	500mM
Tris-HCl, PH 8.3	100mM	100mM	100mM
Oligo Forward/Reverse	30 pmoles	30 pmoles	30 pmoles
DNA	50 ng	50 ng	50 ng
Volumen final	10 uL	10 uL	10 uL
Ciclos de PCR	35	35	25

Las condiciones de PCR empleadas fueron las siguientes, donde n es el número de ciclos empleado en cada una de las condiciones de mutagénesis y que se muestra en la tabla anterior:

Desnaturalización	98°C	30 s	} n ciclos
Alineamiento	60°C	30 s	
Extensión	72°C	30 s	
4°C	∞		

### 7.8 Purificación de fragmentos amplificados por PCR

Para la purificación de los amplicones de PCR previamente obtenidos, se empleó el kit SAP-exo de Jena Bioscience, el cual permite la remoción de oligonucleótidos y dNTPS no incorporados. Se tomaron 5  $\mu$ L de cada uno de los productos de PCR deseados (previamente examinados en un gel de agarosa al 1% para corroborar la ausencia de productos inespecíficos) y se les añadió 1  $\mu$ L de SAP-exo. Dicha mezcla se incubó a 37°C durante 5 minutos, y posteriormente se incubó a 80°C durante 10 minutos con el fin de inactivar las enzimas presentes en la reacción. Al terminar, el DNA estaba listo para usos posteriores como secuenciación o clonación.

### 7.9 Restricción

El producto de PCR de cada una de las condiciones de mutagénesis fue purificado mediante el protocolo anterior, y dicho producto fue posteriormente restringido. Para ello, 1.5  $\mu$ g de vector (pUCP24) y 750 ng de inserto (g2-mutante) fueron doblemente restringidos con las enzimas *EcoRI* y *HindIII* de New England Biolabs. La reacción fue preparada de la siguiente manera:

**Tabla 3. Doble digestión de fragmentos para la clonación.**

Reactivo	Cantidad	Cantidad
NEBuffer 2.1 10x	2 $\mu$ L	1 $\mu$ L
<i>EcoRI</i>	1.5 $\mu$ L	1 $\mu$ L
<i>HindIII</i>	1.5 $\mu$ L	1 $\mu$ L
DNA	1.5 $\mu$ g vector	750 ng inserto
H <sub>2</sub> O	para 20 $\mu$ L	para 10 $\mu$ L

Dicha reacción se incubó a 37°C durante toda la noche, para promover una restricción total; cuidando que en el buffer empleado las enzimas no tuvieran actividad estrella. Posteriormente, la reacción se incubó a 80°C durante 20 minutos, para inactivar las enzimas de restricción, y proceder a la ligación de los fragmentos.

### 7.10 Ligación

Se ligó el producto de doble digestión del plásmido y el inserto en una relación molar 1:3, para lo cual se preparó la reacción de la siguiente manera:

**Tabla 4. Ligación de variantes de g2 en pUCP24.**

Reactivo	Cantidad
Buffer Ligasa T4 10x	1 µL
Vector (pUCP24)	300 ng
Inserto (g2-mutante)	45 ng
Ligasa T4	1 µL
H <sub>2</sub> O	para 10 µL

Posteriormente, se incubó a 16°C durante 16 horas.

### 7.11 Preparación de células quimiocompetentes de *E. coli*.

Las células súper competentes de DH5α se obtuvieron por el método de Green & Rogers (2013). A partir de un cultivo líquido overnight de *E. coli* se realizó una dilución 1:100 en 100 mL de medio SOC (véase anexo 1.2). Se incubó a 37°C hasta alcanzar una D.O.<sub>600</sub> de 0.5-0.6. Se incubó durante 5 minutos en hielo.

Después, se centrifugó a 8,944 G durante 10 min, se decantó el sobrenadante y la pastilla se resuspendió suavemente en 33 ml de solución Tfb I (véase anexo 1.4). Se incubó 10 min en hielo y se centrifugó a 2,236 G durante 10 min. Se decantó el sobrenadante y la pastilla se resuspendió en 4 ml de solución Tfb II (véase anexo 1.5). Se hicieron alícuotas de 200 µL en microtubos EPPENDORF y se congelaron con hielo seco. Finalmente, las alícuotas se guardaron a -70°C hasta su transformación.

#### **7.12 Transformación química en *E. coli*.**

Se descongeló en hielo una alícuota de 100 µL de células quimiocompetentes y se mezcló con 5 µL de reacción de ligación o 20 ng de plásmido. Se incubó en hielo durante 30 min. Después se dio un choque térmico de 2 minutos a 42°C. Inmediatamente se incubó en hielo durante 2 minutos y se agregó 1 mL de medio LB precalentado a 37°C. Las células se dejaron recuperando en agitación durante 1 h a 37°C. Se centrifugó a 2,236 G durante 5 min, por último, se plaqueó la reacción en cajas de LB con el antibiótico de selección, realizando las diluciones correspondientes en LB.

#### **7.13 PCR de colonia**

Se realizaron PCR de colonia tanto de *E. coli*, como de *P. aeruginosa*. Las PCR se ajustaron a una mezcla de reacción con volumen final de 10 µL que incluyó lo siguiente:

**Tabla 5. PCR de colonias candidatas.**

<b>Reactivo</b>	<b>Cantidad</b>
Buffer 10x	1 µL
dNTPs	0.2 µL
Taq polimerasa	0.1 µL
Oligo pUCP2410 F/R 10 µM	0.25 µL
DNA	20 ng
H <sub>2</sub> O	para 10 uL

En este caso, el DNA molde provino de una suspensión que contenía una colonia bacteriana aislada, la cual fue resuspendida en 20  $\mu$ L de agua inyectable estéril y calentada a 95°C durante 5 minutos.

#### **7.14 Geles de agarosa y electroforesis.**

Para la visualización de los productos de PCR, se realizaron geles de agarosa al 1%. Para ello se pesó 1 g de agarosa de BioRad, y se le agregaron 100 mL de buffer SB 1X (véase anexo 1.6). La mezcla se calentó en un horno de microondas por 4 minutos, para su homogenización. Se dejó enfriar un poco y se vertió en una base de electroforesis sellada en ambos costados con cinta adhesiva, colocando un peine de tamaño de pozos mediano. Una vez solidificado, se retiró la cinta adhesiva, se colocó en una cámara de electroforesis. Se llenó la cámara con buffer SB 1X hasta cubrir el gel y se retiró el peine. Se cargaron las muestras y se corrió el gel a 200 V durante 35 minutos.

#### **7.15 Extracción de plásmido.**

Se empleó el kit de extracción Wizard® Plus SV Minipreps DNA Purification System de Promega. Se centrifugaron de 5 a 10 mL de un cultivo O.N. a 8,944 G durante 5 min y se descartó el sobrenadante. La pastilla se resuspendió en 250  $\mu$ L de Cell Resuspension Solution. Después se le agregaron 250  $\mu$ L de Cell Lysis Solution y se mezcló por inversión. Se le adicionaron 10  $\mu$ L de Alkaline Protease Solution, se mezcló por inversión y se dejó incubando a temperatura ambiente durante 5 minutos. Después se le añadieron 350  $\mu$ L de Neutralization Solution y se mezcló por inversión. Se centrifugó a 16,000 G durante 10 minutos, y se recuperó el sobrenadante, transfiriéndolo a una columna adaptada a un tubo colector. Se centrifugó a 16,000 G durante 1 min y se descartó el filtrado.

Se realizó un primer lavado con 750  $\mu$ L de Wash Solution centrifugando a 16,000 G durante 1 min y después un segundo lavado con 250  $\mu$ L de Wash Solution centrifugando a 16,000 G durante 2 min. Después, se transfirió la columna a un tubo EPPENDORF de 1.5 mL y se adicionaron 100  $\mu$ L de agua libre de nucleasas. Se centrifugó a 16,000 G durante 1 min y el filtrado se guardó a -20°C.

### **7.16 Preparación de células electrocompetentes de *P. aeruginosa*.**

Para obtener células electrocompetentes de PA14 se siguió el método de Choi et al. (2006). Se inocularon 6 mL de medio LB con una asada de PA14 proveniente de un glicerol, y se incubó toda la noche a 37°C. Se centrifugó a 8,944 G durante 1 min y se decantó el sobrenadante. Se realizaron tres lavados de la pastilla, cada uno con 1 mL de sacarosa 300 mM y se centrifugó a 8,944 G durante 8 min. Por último, se resuspendió en 100 µL de sacarosa 300 mM. La alícuota se guardó a -70°C, hasta su electroporación.

### **7.17 Electroporación en *P. aeruginosa*.**

Se dejó descongelar en hielo 100 µL de células electrocompetentes y se mezclaron con 100 ng de plásmido. Se transfirieron a una celda de 1 mm de Bio-rad y se dejaron incubando en hielo durante 30 min. Se dio un pulso a 2.2kV, 25µF y 200Ω (electroporador Gene Pulser™ de Biorad). Posteriormente se dejaron recuperar las células en 1 mL de medio LB precalentado a 37°C con agitación durante 1 hora. Por último, se plaqueó la reacción en cajas de LB con el antibiótico correspondiente, y se dejó incubar a 37°C durante toda la noche.

### **7.18 Preparación de células quimiocompetentes de *P. aeruginosa*.**

Se empleó el método de Chuanchuen et al (2002). Este consistió en incubar 4 mL de cultivo líquido de la cepa PA14 en LB toda la noche a 37°C. Se dividió el cultivo en alícuotas de 1 mL en microtubos previamente enfriados a 4°C, y se centrifugó a 13,000 G durante 30 segundos a temperatura ambiente. La pastilla fue resuspendida en 1 mL de MgCl<sub>2</sub> 0.1 M frío. Posteriormente, se volvió a centrifugar a 13,000 G por 30 segundos a temperatura ambiente. Se resuspendió la pastilla con 1mL de la solución TG salts fría (véase anexo 1.7), y se mantuvo en hielo por 10 minutos. En seguida se centrifugó a 13,000 G por 30 segundos a temperatura ambiente, y la pastilla se resuspendió en 200 µL de la solución TG salts fría. La alícuota se mantuvo en hielo para su posterior transformación.

### **7.19 Transformación química en *P. aeruginosa*.**

100  $\mu$ L de las células quimiocompetentes obtenidas por la metodología anterior se mezclaron con 100-1000 ng de plásmido. Se mantuvieron en hielo por al menos 15 minutos, posteriormente, se incubaron a 37°C por 2 minutos y se les añadió 500  $\mu$ L de medio LB. Se dejaron recuperar las células a 37°C con agitación, durante 1 hora. Por último, se plaqueó la reacción en medio LB sólido con el antibiótico correspondiente.

### **7.20 Colony Blot en *P. aeruginosa*.**

Posterior a la transformación de *P. aeruginosa* con el vector de interés, se sembraron 97 candidatas por duplicado en una caja maestra y una caja replica, adicionalmente se sembraron 2 controles positivos y uno negativo (células con el vector y el inserto, y el vector vacío, respectivamente), los cuales permitieron la identificación de las candidatas positivas.

Ambas cajas se incubaron toda la noche a 37°C. A continuación, se transfirieron las colonias de la caja replica a una membrana de nylon de 8 cm de diámetro, colocándola sobre las colonias durante 2 minutos.

Transcurrido este tiempo, se colocó la membrana con las colonias hacia arriba, sobre un papel filtro humedecido con la solución I (véase anexo 1.8) por 10 minutos. Después, se colocó la membrana sobre papel filtro impregnado con la solución II (véase anexo 1.9) por 2 minutos, y finalmente se colocó sobre papel filtro con la solución III (véase anexo 1.10) por 5 minutos. Posteriormente, la membrana se expuso a luz U.V. durante 5 minutos en el crosslinker para fijar el DNA a la membrana.

Una vez fijo el DNA, se introdujo la membrana en un tubo de hibridación y se adicionó la solución de pre-hibridación (véase anexo 1.11). Se incubó a 42°C durante 2 horas en el horno de hibridación. Durante este tiempo se marcó el oligonucleótido deseado con radioactividad. Para esto, se realizó la siguiente reacción:

**Tabla 6. Marcaje de oligos con <sup>32</sup>P.**

<b>Reactivo</b>	<b>Cantidad</b>
Oligo G2 exp F 4pmol/uL	1 µL
Buffer PNK 10x	2 µL
PNK	1 µL
<sup>32</sup> P*ATP	0.5 µL
H <sub>2</sub> O	para 20 µL

Esta reacción se incubó a 37°C durante 1 hora, y posteriormente se centrifugó a 16,000 G por 20 minutos. Se decantó el sobrenadante y se resuspendió en 100 µL de agua.

Una vez pasado el tiempo de pre-hibridación de la membrana, se adicionó la sonda marcada radioactivamente, y se dejó hibridar a 50°C durante toda la noche en el horno de hibridación.

Al día siguiente, se realizaron lavados a la membrana. Para ello se añadió la solución de lavado I (véase anexo 1.12) y se incubó a temperatura ambiente por 3 minutos. Los siguientes dos lavados se realizaron con la solución de lavado II (véase anexo 1.13) durante 5 minutos cada uno. Posteriormente, se incubó la membrana con la solución de lavado III (véase anexo 1.14) durante 30 minutos. Finalmente, se realizó un último lavado con la solución de lavado IV (véase anexo 1.15) durante 15 minutos. Se dejó secar la membrana y se introdujo en un casete para uso de membranas con radioactividad, en el cual se colocó una placa auto radiográfica en condiciones de oscuridad sobre la membrana. Se dejó exponer la placa (el tiempo varió según la actividad de la marca radiactiva), y finalmente se reveló.

### 7.21 Ensayo de funcionalidad de G2

Se inocularon 3 mL de medio LB líquido con cada una de las colonias positivas en el colony blot (es decir, colonias que portan g2, incluyendo un control positivo de g2 wt), y se incubaron en agitación continua por toda la noche a 37 °C. A partir del cultivo O.N. se tomaron 100 µL y se añadieron 1.5 mL de medio semisólido TΦ, se vertió la mezcla sobre una caja Petri pequeña con medio TΦ sólido y gentamicina a una concentración de 50 µg/mL. Se dejó solidificar a temperatura ambiente.

Se realizaron 4 diluciones seriadas del fago sensible al sistema CRISPR-cas (JBD18) para formar las diluciones 10<sup>-2</sup>, 10<sup>-4</sup>, 10<sup>-6</sup> y 10<sup>-8</sup>. Se tomaron 5 µL de cada dilución y se gotearon sobre el tapiz bacteriano. Se dejaron secar las gotas a temperatura ambiente y cada caja fue incubada a 37 °C toda la noche. Los títulos finales de los lisados se determinaron aplicando la siguiente fórmula matemática (Douglas, 1975):

$$\frac{UFP}{mL} = \left( \frac{\text{No. de placas}}{\text{Vol. goteado en ml}} \right) \times (\text{Inverso de la última dilución con placas})$$

Donde:

UFP= unidades formadoras de placas

La eficiencia de inhibición del sistema CRISPR-cas de cada variante se calculó tomando en cuenta el título fágico obtenido en la cepa portando un g2 silvestre, aplicando la siguiente fórmula:

$$\text{eficiencia de inhibición de variante } i = \frac{\text{título fágico en cepa con variante de } g2 \text{ } i}{\text{título fágico con } g2 \text{ wt}}$$

### 7.22 Secuenciación de variantes de G2

Para la secuenciación de las variantes de g2 se empleó el kit BigDye™ Terminator v1.1 Cycle Sequencing Kit de ThermoFisher Scientific. Se siguió el protocolo proporcionado por el Biol. Miguel Ángel Moreno Galeana, preparando la reacción de la siguiente manera:

**Tabla 7. Reacciones de Secuenciación de candidatas.**

Reactivo	Cantidad
Buffer de Secuenciación 5x	2 $\mu$ L
Producto de PCR purificado	5 $\mu$ L
Oligo MCS pUCP24 forward 10 $\mu$ M	0.5 $\mu$ L
Mix de secuenciación	0.4 $\mu$ L
Betaína 5M	2 $\mu$ L
H <sub>2</sub> O	para 10 $\mu$ L

Las condiciones de PCR empleadas fueron las siguientes:

Desnaturalización	98°C	30 s	} 27 ciclos
Alineamiento	55°C	30 s	
Extensión	60°C	30 s	
4°C	$\infty$		

### 7.23 Mutagénesis sitio dirigida

Los oligonucleótidos empleados para la mutagénesis sitio dirigida fueron diseñados con ayuda de la herramienta NEBaseChanger (disponible en: [NEBaseChanger.neb.com](http://NEBaseChanger.neb.com)). La mutagénesis sitio dirigida fue realizada empleando el kit Q5® Site-Directed Mutagenesis Kit de New England Biolabs.

Ésta constó de tres pasos, en el primero se realizó una amplificación exponencial del vector con la mutación deseada por PCR, para ello se realizó la siguiente reacción:

**Tabla 8. Mutagénesis Sitio-Dirigida: PCR.**

Reactivo	Cantidad
Q5 Hot Start High-Fidelity 2X Master Mix	6.25 $\mu$ L
Oligo F 10 $\mu$ M	0.65 $\mu$ L
Oligo R 10 $\mu$ M	0.65 $\mu$ L
DNA (pUCP24-g2 WT)	12.5 ng
H <sub>2</sub> O	para 12.5 $\mu$ L

Las condiciones de PCR fueron las siguientes:

Desnaturalización inicial	98°C	30s	
Desnaturalización	98°C	10 s	} 25 ciclos
Alineamiento	50-72°C	10-30 s	
Extensión	72°C	20-30 s/kb	
Extensión final	72°C	2 min	
4°C	$\infty$		

Posteriormente, el producto de PCR se trató con la ligasa ligasa y la enzima *DpnI*, con el fin de eliminar el DNA molde metilado. La reacción se preparó de la siguiente forma:

**Tabla 9. Mutagénesis Sitio-Dirigida: Reacción KLD.**

<b>Reactivo</b>	<b>Cantidad</b>
Producto de PCR	0.5 µL
Buffer de reacción KLD 2x	2.5 µL
Mix de enzimas KLD 10x	0.5 µL
H <sub>2</sub> O	1.5 µL

La reacción se incubó a temperatura ambiente por 5 minutos y se prosiguió a la transformación por choque térmico del total de la reacción en células quimiocompetentes de DH5α. Se extrajo plásmido de las candidatas de *E. coli* y posteriormente se introdujeron las variantes de G2 en *P. aeruginosa* mediante transformación química en células quimiocompetentes de Pa14.

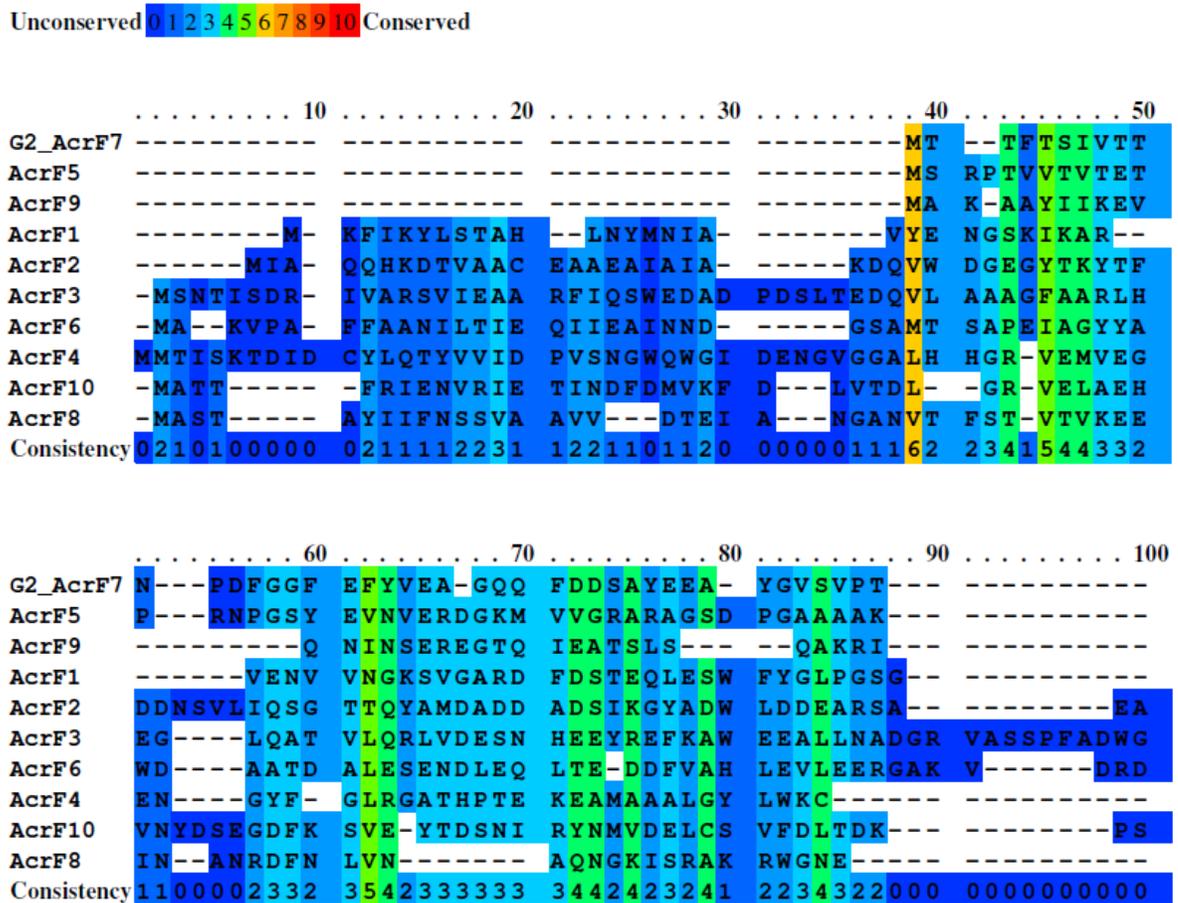
## **8. RESULTADOS**

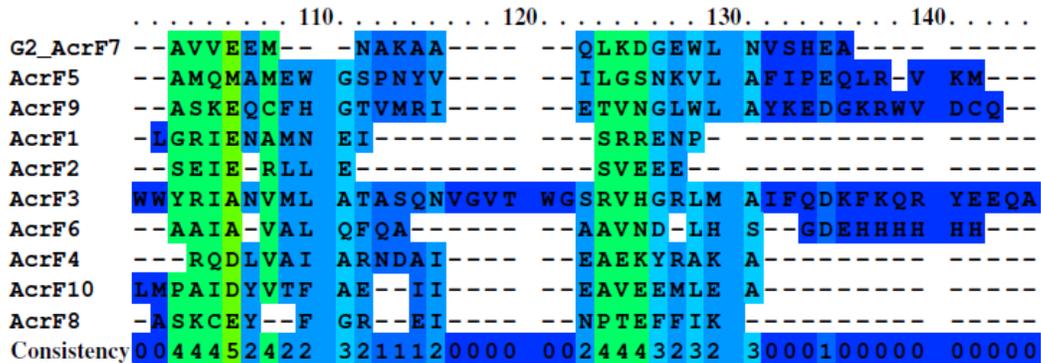
### **8.1 Búsqueda de aminoácidos conservados mediante análisis de alineamiento múltiple de secuencias aminoacídicas.**

El primer objetivo del presente proyecto fue la búsqueda de aminoácidos conservados en las distintas familias de proteínas antiCRISPR mediante un análisis de alineamiento múltiple de secuencias. Las secuencias de las primeras 5 familias de antiCRISPR (AcrF1-5) fueron tomadas de los datos suplementarios de Bondy-Denomy et al., 2013, y las secuencias de las familias AcrF6-10 fueron obtenidas de los datos suplementarios de Pawluk et al., 2016. La familia reportada como AcrF7 es la correspondiente a los homólogos de G2, de dicha familia se tomó la secuencia reportada por nuestro grupo de investigación para g2 del fago H70, ya que la secuencia reportada por Pawluk y colaboradores presenta inconsistencias entre la secuencia que reportan y la publicada bajo ese mismo número de acceso.

El alineamiento fue realizado empleando la herramienta Praline (disponible en <http://www.ibi.vu.nl/programs/pralinewww/>), mediante el cual se realizó el alineamiento múltiple de secuencias empleando PSI-BLAST con 3 iteraciones, bajo una matriz BLOSUM62.

En la figura 6, se muestra el alineamiento múltiple de las 10 familias de antiCRISPR que inhiben el sistema CRISPR-cas I-F, siendo G2 la secuencia query:





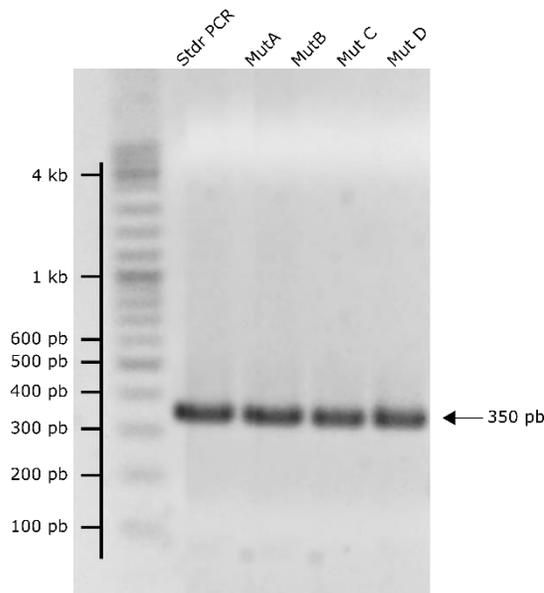
**Fig. 6. Alineamiento Múltiple de Secuencias aminoacídicas de familias de antiCRISPR.** Las secuencias del alineamiento representan cada una de las 10 familias de antiCRISPR reportadas que inhiben el sistema I-F de *P. aeruginosa*, siendo el query g2 (familia AcrF7). El código de colores se muestra en el extremo superior izquierdo. Las líneas punteadas representan gaps introducidos en el alineamiento. Los números superiores denotan la longitud del alineamiento.

Los colores en el alineamiento representan el grado de conservación de cada residuo, siendo los rojos los más conservados, y los azules los menos conservados. Como se puede observar en la figura 6, no existe ningún residuo que se encuentre conservado entre las diferentes familias, por lo cual no es posible asignar posibles sitios de interacción.

**8.2 Obtención de biblioteca de mutantes**

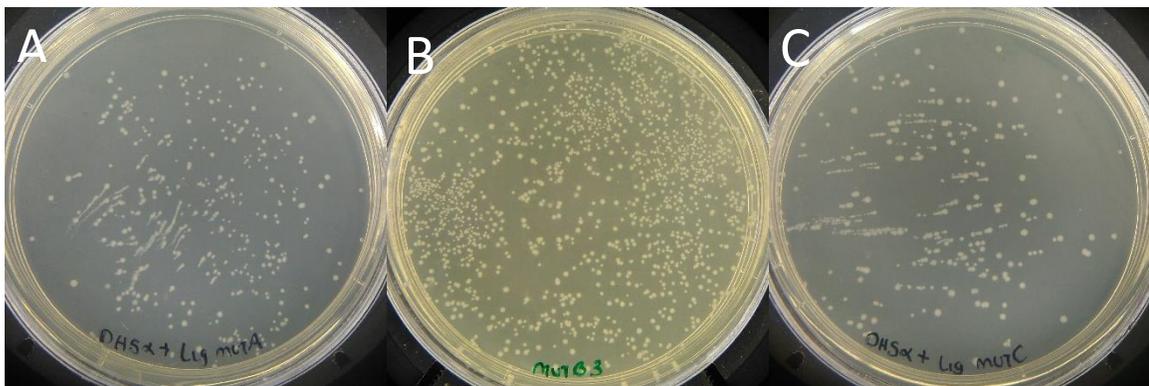
Debido a que no se observaron residuos conservados en los alineamientos de las diferentes familias antiCRISPR, se optó por realizar una mutagénesis al azar. Ésta se llevó a cabo por la técnica de promoción de errores durante la PCR (Error-Prone PCR) propuesta por Cadwell y colaboradores en 1992 (Cadwell et al, 1992).

Se emplearon 3 condiciones de PCR mutagénica (véase Métodos 7.7), las cuales se corrieron en un gel de agarosa al 1% (figura 7), para corroborar la amplificación del fragmento (debido a que se ha reportado la precipitación de polimerasas en presencia de MnCl<sub>2</sub>), así como comprobar la ausencia de productos inespecíficos. Posteriormente, se purificaron estos productos de PCR y fueron restringidos con las enzimas *EcoRI* y *HindIII*. Se desactivaron las enzimas de restricción calentando la reacción, y después fueron ligadas independientemente con el vector pUCP24L3.



Cada reacción de ligación fue transformada en células competentes de DH5 $\alpha$  previamente preparadas (véase Métodos 7.11), mediante transformación química. Cada una de las transformaciones fue dividida en 3 antes de la recuperación celular, y plaqueada de manera independiente. En la figura 8 se muestra una caja representativa de cada una de las condiciones de mutagénesis.

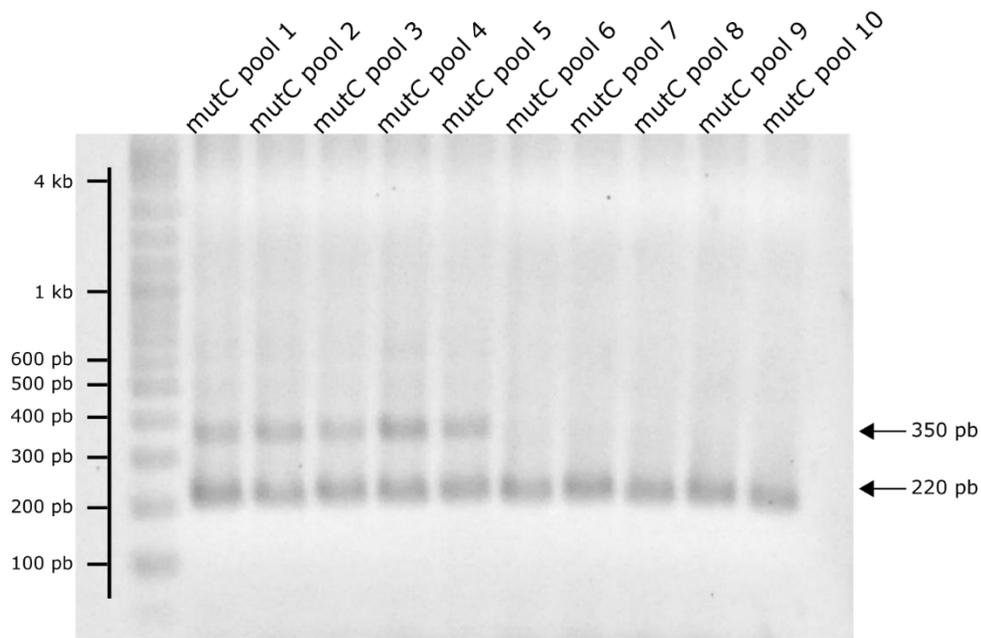
**Fig. 7. Productos de PCR mutagénica.** En la imagen se muestra un gel de agarosa al 1% con los amplicones obtenidos (350 pb) en las distintas condiciones de mutagénesis por PCR, siendo *Stdr PCR* el control de condiciones de PCR estándar.



**Fig. 8. Candidatas de *E. coli* con g2 proveniente de las diferentes condiciones de mutagénesis.** En la sección A se muestran candidatas obtenidas a partir de la condición A (tabla 2) de mutagénesis, en la B las obtenidas de la condición B (tabla 2), y en la sección C las obtenidas por la condición C (tabla 2).

Después de diversos ensayos de estandarización y optimización del proceso de identificación de candidatas que portan g2, se optó por realizar cultivos mixtos a partir de 10 colonias distintas (para cada condición de mutagénesis), para su posterior extracción plasmídica.

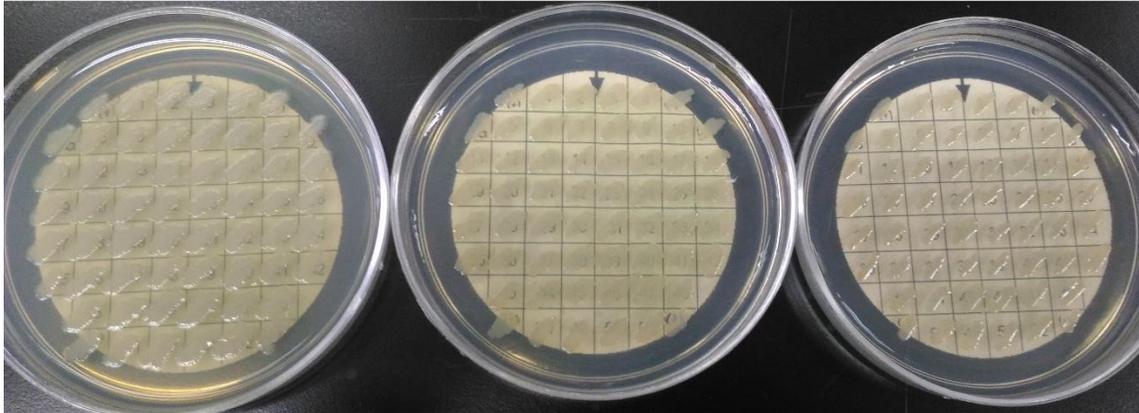
Una vez extraídos los pools de plásmidos, éstos fueron probados por PCR para determinar si alguno de los plásmidos que conformaban el pool contenía nuestro inserto de interés. Las PCR fueron llevadas a cabo con los oligos flanqueantes del sitio múltiple de clonación (MCS pUCP24 F/R), esperando una banda de 350 pb en los vectores con inserto, y de 220 en los vectores vacíos (Fig. 9). Esto se realizó debido a la baja eficiencia que se observó durante la clonación del gen de interés, analizando de esta manera, 10 candidatas en una sola PCR. En la figura 9, se muestra un gel representativo de todos los que se llevaron a cabo.



**Fig. 9. PCR del sitio múltiple de clonación de los distintos pools de plásmido.** En la figura se muestra un gel de agarosa al 1% de las PCR de 10 pools distintos de las candidatas obtenidas mediante la condición de mutagénesis C (representando en total 100 candidatas). En los primeros 5 carriles existe una mezcla de vector vacío y con inserto (dobles bandas), y en los últimos 5 solo existen vectores vacíos.

Los pools que no contuvieron clones de vector con inserto (solo una banda que migra a 220 pb) fueron desechados; aquellos que portaron una mezcla de vectores vacíos y con inserto fueron posteriormente transformados en *Pseudomonas aeruginosa*.

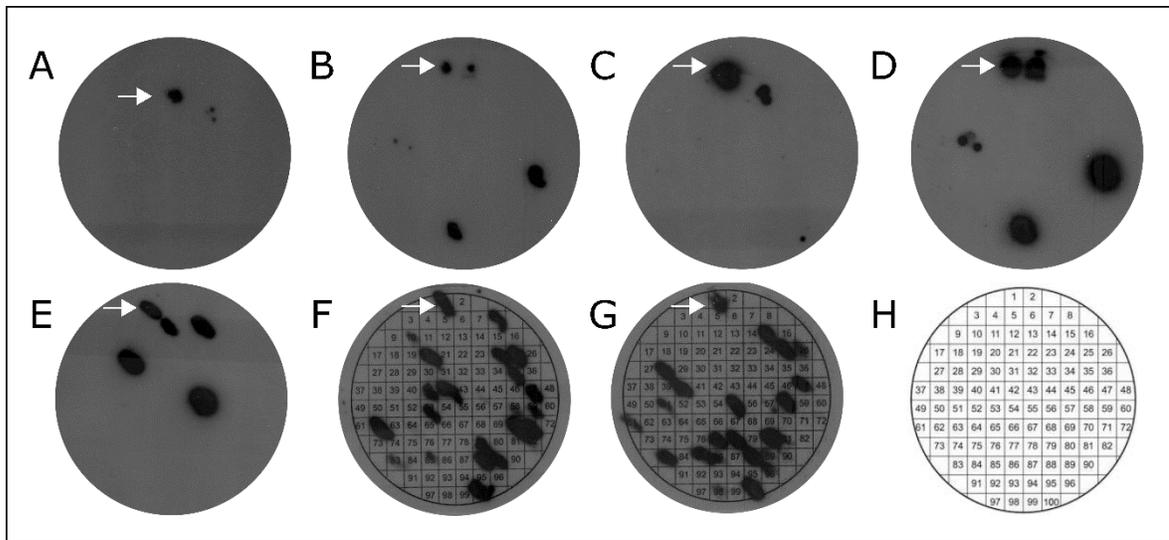
A partir de las candidatas obtenidas en la cepa PA14, se tomaron sets de 100 colonias, las cuales fueron resembradas en forma de líneas diagonales (como se muestra en la figura 10) con el fin de evitar falsos positivos (puntos que aparecen espontáneamente en el revelado de las membranas), para posteriormente identificar las candidatas que portaban plásmido con inserto mediante la técnica de *Colony Blot* (véase Métodos 7.20).



**Fig. 10. Candidatas de *P. aeruginosa*.** En la imagen se muestran cajas con 100 colonias distintas cada una, resembradas a partir de la transformación de PA14 con los pools de plásmido probados por PCR.

Se analizaron en total 7 membranas por colony blot, es decir, 700 colonias candidatas de *P. aeruginosa*. Para ello, se empleó el oligonucleótido G2 exp forward marcado radioactivamente con  $^{32}\text{P}$ , el cual está dirigido contra el extremo 5' del gen g2, por lo cual, una señal positiva de radiactividad denota la presencia del gen de interés en la colonia analizada.

En cada una de las membranas se incluyó un control positivo (una colonia que porta g2), denotado como flechas blancas en la figura 11.



**Fig. 11. Membranas analizadas por Colony Blot.** Cada círculo representa una membrana distinta, donde las manchas oscuras denotan las candidatas con las cuales hibrido el oligo contra *g2*. Las flechas blancas señalan los controles positivos. En la sección H se muestra la plantilla empleada para resembrar las colonias, y en las secciones F y G se puede observar cómo se empleó dicha plantilla para identificar las candidatas positivas.

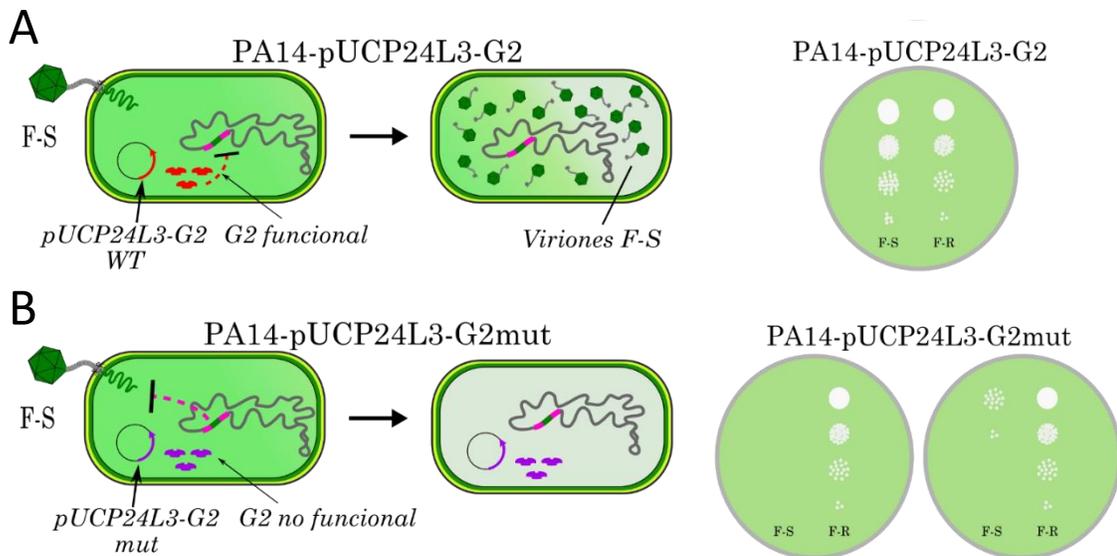
Una vez identificadas las candidatas positivas para *g2*, se realizaron PCRs de cada una de ellas (empleando los oligos del MCS), con el fin de descartar falsos positivos, y evidenciar si alguna de estas candidatas contuviese más de un inserto (concatámeros) clonado en el plásmido.

En total se analizaron (mediante pools en *E. coli* y colony blot en *P. aeruginosa*) 1000 candidatas, de las cuales sólo se encontraron 44 con un inserto único. De estos 44 candidatos se realizaron ensayos de tipificación, los cuales consisten en la elaboración de tapices bacterianos de las distintas candidatas y el goteo del fago sensible al sistema CRISPR-cas JBD18. Dicho ensayo tiene el siguiente racional:

-El bacteriófago JBD18 es un fago que no porta genes antiCRISPR, y que es blanco del sistema CRISPR-cas, es decir, PA14 porta espaciadores contra su genoma. Por lo tanto, este fago no puede infectar a la cepa, a menos que esta última posea un antiCRISPR, en este caso codificado en el plásmido pUCP24.

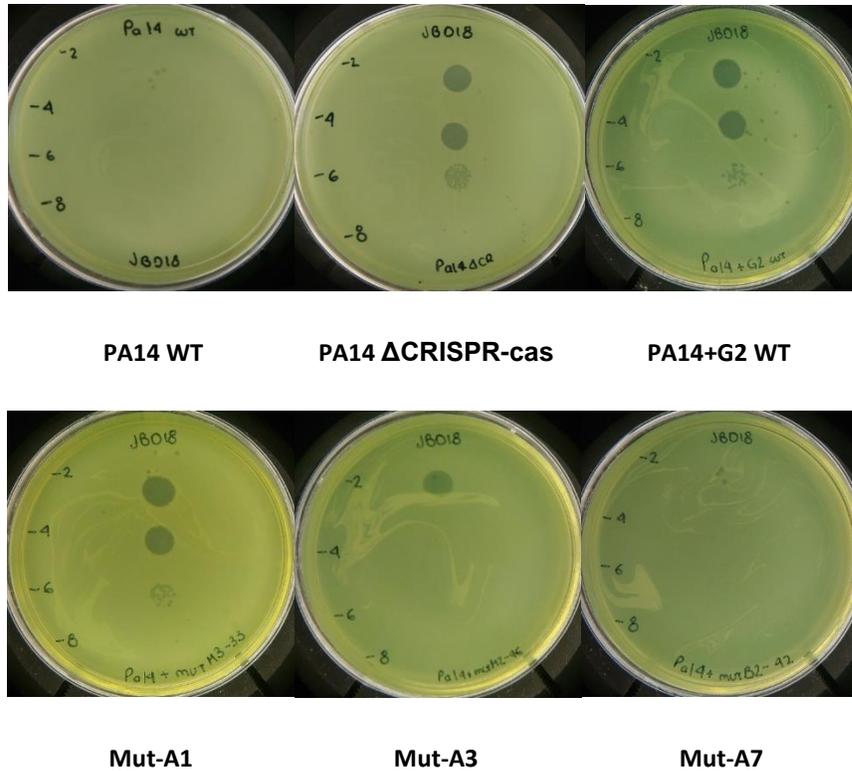
-Cuando la cepa PA14 está transformada con un plásmido que porta un antiCRISPR silvestre, éste estará inhibiendo la defensa bacteriana, y por lo tanto el fago JBD18 podrá infectarla, lo cual se ve reflejado como placas en el tapiz bacteriano (figura 12A).

-Cuando la cepa PA14 porta un antiCRISPR no funcional, o con una función disminuida, éste no podrá inhibir, o no por completo, el sistema CRISPR-cas, por lo cual JBD18 no podrá infectar, o no eficientemente. En un tapiz esto se ve reflejado como ausencia de placas o un menor título fágico (figura 12B).



**Fig. 12. Ensayo de fenotipificación del antiCRISPR G2.** En la sección A se muestra la inhibición del sistema CRISPR-cas por un antiCRISPR funcional, permitiendo la infección por un fago sensible (F-S), y su consecuente formación de placas en un tapiz. En la sección B se muestra que un antiCRISPR no funcional no puede inhibir el sistema de defensa, por lo cual sistema CRISPR-cas inhibe la infección por el fago sensible, resultando en la ausencia de placas.

En la figura 13, se muestran algunos de los perfiles de infección observados en los ensayos de fenotipificación:



**Fig. 13. Ensayo de fenotipificación.** En las imágenes se muestran los distintos fenotipos obtenidos en las candidatas. En la parte superior se muestran las cepas control (PA14 WT, PA14  $\Delta$ CRISPR-cas y PA14+G2 WT). En la parte inferior se muestran 3 de los ensayos de fenotipificación realizados, donde cada una muestra un fenotipo distinto: Mut-A1 muestra un fenotipo silvestre, Mut-A3 mutante parcial y Mut-A7 mutante total.

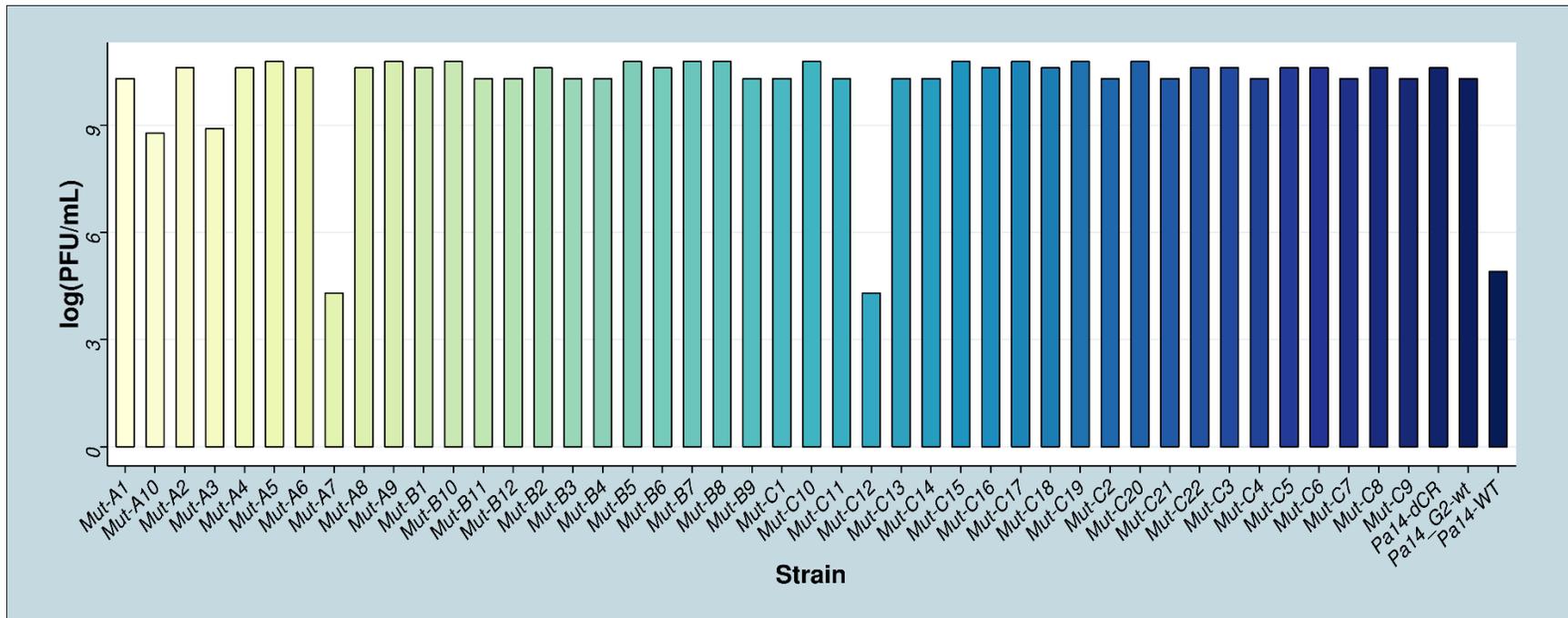
En la figura 13, en la cepa PA14 silvestre se pueden observar algunas placas en la dilución  $10^{-2}$ , lo cual nos indica que el fago JBD18 tiene una tasa basal de evasión del sistema CRISPR-cas, el cual ha sido reportado anteriormente (Bondy-Denomy et al., 2013), adicionalmente, para comprobar que este fuera nuestro caso, se picaron dichas placas con un palillo y se pasaron sobre un tapiz de PA14 silvestre y uno de la cepa  $\Delta$ CRISPR-cas, observándose amplificación del fago en ambas con la misma eficiencia, lo que nos indicó que efectivamente este fago ya no era blanco del sistema de defensa de la cepa, por lo cual debe contener mutaciones en su protoespaciador, que le permiten ya no ser reconocido por la bacteria.

En la cepa PA14  $\Delta$ CRISPR-cas, se observó que el fago es capaz de infectar con gran eficiencia a la cepa que no tiene sistema de defensa, observándose casi la misma eficiencia cuando la cepa contenía un antiCRISPR silvestre (PA14+G2 WT) (Fig. 13).

En el caso de las candidatas a mutante, en la imagen superior se muestran 3 de los perfiles observados, donde en el primer caso (Mut-A1), se presentó un perfil de infección similar al de la cepa con el antiCRISPR silvestre, por lo cual, la versión de G2 que porta la cepa es tan funcional como el silvestre.

En la candidata Mut-A3 se observó una eficiencia aproximadamente 100 veces menor que la de G2 silvestre, a lo cual se le denominó fenotipo de *mutante parcial*, es decir, que pierde parcialmente la función. En el último caso, Mut-A7, se observó la pérdida total de la función, cuyo perfil de infección es análogo al de la cepa PA14 WT. A este fenotipo se le denominó *mutante total*.

En la figura 14, se muestra una gráfica con todos los perfiles de infección de las distintas candidatas:



**Fig. 14. Perfiles de infección de las candidatas analizadas.** En la gráfica se muestran los perfiles de infección obtenidos para cada una de las 44 candidatas analizadas a través de ensayos de fenotipificación. En la abscisa se muestran los nombres de las candidatas analizadas, así como de las cepas controles (extremo derecho del eje). En la ordenada se muestra el logaritmo del título del fago JBD18 sobre cada una de las candidatas.

En la figura 14 se puede apreciar que prácticamente todas las candidatas presentaron un fenotipo silvestre, es decir, un fenotipo similar al del control de la cepa PA14\_G2-wt. Las candidatas Mut-A10 y Mut-A3 presentaron un fenotipo mutante parcial, observándose una reducción en el título del fago sensible de aproximadamente 2 logaritmos. Por otro lado, Mut-A7 y Mut-C12 presentaron un fenotipo mutante total, equiparable al de la cepa PA14-WT, lo que implica que las variantes de G2 que portan estas mutantes, ya no son funcionales.

A continuación, se muestra una recopilación de la biblioteca de mutantes, separadas por la condición de mutagénesis por la que se obtuvieron, donde en la primera columna se muestra el número de candidata, seguido del nombre, y con el símbolo de “correcto” se especifican los pasos por los cuales ha pasado cada candidata, es decir, si se encuentra clonada en *E. coli*, si se encuentra en *P. aeruginosa*, si se conoce su fenotipo y/o secuencia, y en la última columna se especifica la descripción del fenotipo.

**Tabla 10. Recopilación de mutantes obtenidas por la condición A.**

	Candidata	<i>E. coli</i>	<i>P. aeruginosa</i>	Fenotipo	Secuencia	Descripción
1	Mut-A1	✓	✓	✓	✓	Fenotipo wt
2	Mut-A2	✓	✓	✓	✓	Fenotipo wt
3	Mut-A3	✓	✓	✓	✓	<b>Fenotipo mutante</b>
4	Mut-A4	✓	✓	✓	✓	Fenotipo wt
5	Mut-A5	✓	✓	✓	✓	Fenotipo wt
6	Mut-A6	✓	✓	✓	✓	Fenotipo wt
7	Mut-A7	✓	✓	✓	✓	<b>Fenotipo mutante</b>
8	Mut-A8	✓	✓	✓	✓	Secuencia silvestre
9	Mut-A9	✓	✓	✓	✓	Fenotipo wt
10	Mut-A10	✓	✓	✓	✓	<b>Fenotipo mutante</b>

**Tabla 11. Recopilación de mutantes obtenidas por la condición B.**

	<b>Candidata</b>	<b><i>E. coli</i></b>	<b><i>P. aeruginosa</i></b>	<b>Fenotipo</b>	<b>Secuencia</b>	<b>Descripción</b>
11	Mut-B1	✓	✓	✓	-	Fenotipo wt
12	Mut-B2	✓	✓	✓	-	Fenotipo wt
13	Mut-B3	✓	✓	✓	✓	Secuencia silvestre
14	Mut-B4	✓	✓	✓	✓	Secuencia silvestre
15	Mut-B5	✓	✓	✓	-	Fenotipo wt
16	Mut-B6	✓	✓	✓	-	Fenotipo wt
17	Mut-B7	✓	✓	✓	-	Fenotipo wt
18	Mut-B8	✓	✓	✓	-	Fenotipo wt
19	Mut-B9	✓	✓	✓	-	Fenotipo wt
20	Mut-B10	✓	✓	✓	-	Fenotipo wt
21	Mut-B11	✓	✓	✓	-	Fenotipo wt
22	Mut-B12	✓	✓	✓	✓	Fenotipo wt

**Tabla 12. Recopilación de mutantes obtenidas por la condición C.**

	<b>Candidata</b>	<b><i>E. coli</i></b>	<b><i>P. aeruginosa</i></b>	<b>Fenotipo</b>	<b>Secuencia</b>	<b>Descripción</b>
23	Mut-C1	✓	✓	✓	-	Fenotipo wt
24	Mut-C2	✓	✓	✓	-	Fenotipo wt
25	Mut-C3	✓	✓	✓	-	Fenotipo wt
26	Mut-C4	✓	✓	✓	-	Fenotipo wt
27	Mut-C5	✓	✓	✓	-	Fenotipo wt
28	Mut-C6	✓	✓	✓	-	Fenotipo wt

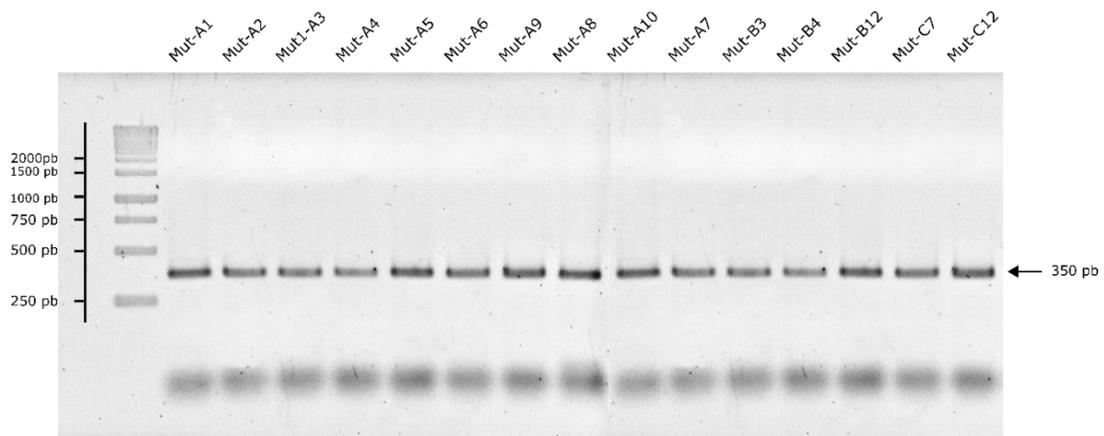
29	Mut-C7	✓	✓	✓	✓	Fenotipo wt
30	Mut-C8	✓	✓	✓	-	Fenotipo wt
31	Mut-C9	✓	✓	✓	-	Fenotipo wt
32	Mut-C10	✓	✓	✓	-	Fenotipo wt
34	Mut-C12	✓	✓	✓	✓	<b>Fenotipo mutante</b>
35	Mut-C13	✓	✓	✓	-	Fenotipo wt
36	Mut-C14	✓	✓	✓	-	Fenotipo wt
37	Mut-C15	✓	✓	✓	-	Fenotipo wt
38	Mut-C16	✓	✓	✓	-	Fenotipo wt
39	Mut-C17	✓	✓	✓	-	Fenotipo wt
40	Mut-C18	✓	✓	✓	-	Fenotipo wt
41	Mut-C19	✓	✓	✓	-	Fenotipo wt
42	Mut-C20	✓	✓	✓	-	Fenotipo wt
43	Mut-C21	✓	✓	✓	-	Fenotipo wt
44	Mut-C22	✓	✓	✓	-	Fenotipo wt

### 8.3 Secuenciación de candidatas

Una vez obtenida la biblioteca de candidatas, se prosiguió a realizar la secuenciación de 15 de ellas, para ello se empleó el kit BigDye™ Terminator v1.1 de ThermoFisher Scientific. La secuenciación de las candidatas se realizó a partir del producto de PCR de colonia de cada una de las versiones de *g2* en *Pseudomonas aeruginosa*, teniendo un control estricto de que cada PCR se realizara de la misma colonia de la que se realizó cada tapiz para probar el fenotipo antiCRISPR.

Para ello, cada colonia fue disuelta en 20  $\mu\text{L}$  de agua estéril y de dicha solución se tomó una alícuota para inocular medio líquido LB, y el resto se calentó a  $95^{\circ}\text{C}$  durante 5 minutos para lisar las células, y se procedió a realizar la PCR del DNA de la solución. La fenotipificación de cada colonia se realizó a partir de los cultivos obtenidos.

Los productos de PCR obtenidos se corrieron en geles de agarosa al 1% con el fin de determinar que no existiera ningún producto inespecífico (Fig. 15). Una vez confirmado esto, se purificaron los amplicones con el kit SAP-Exo y se prosiguió a realizar la reacción de secuenciación. Los electroferogramas obtenidos para cada una de las candidatas se pueden consultar en los anexos.



**Fig. 15. PCR de candidatas secuenciadas.** El gel muestra los amplicones obtenidos a partir de una PCR de colonia de las candidatas a secuenciar (*P. aeruginosa*). Cada carril representa cada una de las candidatas secuenciadas (nombres mostrados en la parte superior).

Las secuencias fueron analizadas y exportadas en formato fasta mediante el software Ridom TraceEdit. Posteriormente se analizaron los cambios a nivel de nucleótidos y aminoácidos mediante el software SeaView, empleando el alineador ClustalW integrado en el programa. Todas las secuencias fueron alineadas tomando como base la secuencia silvestre de g2. Posteriormente, el alineamiento nucleotídico fue visualizado en el software JalView. Las secuencias nucleotídicas obtenidas para las 15 candidatas se muestran a continuación, donde en azul se muestran los nucleótidos idénticos a la secuencia query (g2 wt), y en blanco todos los cambios:

<i>G2/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-A1/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-A2/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-A3/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-A4/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-A5/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-A6/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-A7/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-A8/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-A9/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-A10/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-B3/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-B4/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-B12/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-C7/1-204</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>Mut-C12/1-203</i>	1	ATGACTACGTTACCAGCATCGTTACCACTAACCTGATTTGGCGGATT
<i>G2/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-A1/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-A2/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-A3/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-A4/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-A5/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-A6/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-A7/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-A8/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-A9/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-A10/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-B3/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-B4/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-B12/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-C7/1-204</i>	144	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA
<i>Mut-C12/1-203</i>	143	GAAACGCTAAAGCGGCGCAGTTGAAAGACGGCGAGTGGCTTAATGTCTCGCATGAAGCCTGA

**Fig. 16. Alineamiento nucleotídico de las variantes de g2 secuenciadas.** En la imagen se muestra un alineamiento de las secuencias obtenidas para cada una de las candidatas, siendo el query g2. Cada cuadro en blanco representa un cambio en la secuencia analizada, respecto a la secuencia silvestre de g2.

Como se muestra en la figura 16, la mayoría de las secuencias poseen más de una mutación por molécula, sin embargo, algunas de estas dan lugar a cambios sinónimos, es decir, que ambos codones codifican el mismo aminoácido. Por lo tanto, para analizar los cambios no sinónimos en cada una de las secuencias, se realizó un alineamiento a nivel de aminoácidos. Para ello se empleó el alineamiento nucleotídico realizado con ClustalW, y se tradujeron *in silico* las secuencias, obteniéndose lo siguiente:

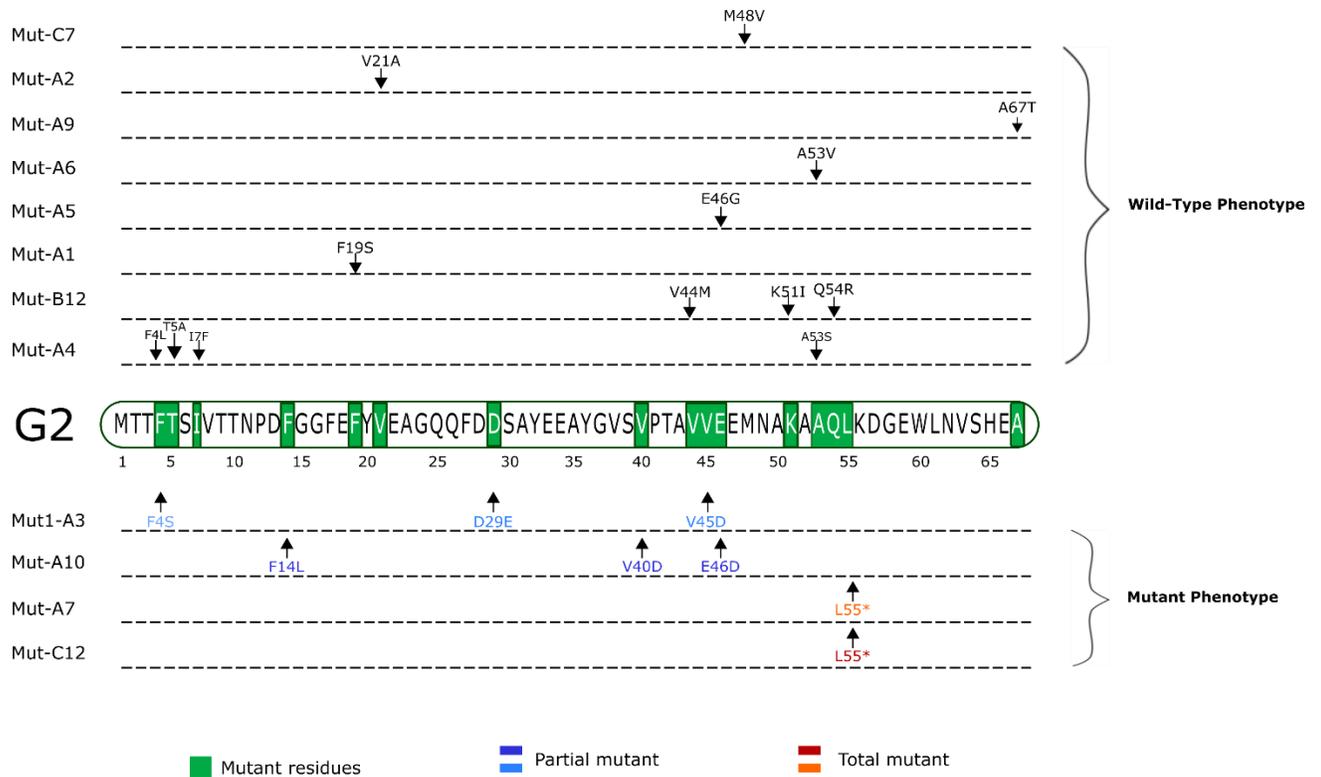


**Fig. 17. Alineamiento aminoacídico de las candidatas secuenciadas.** La imagen representa un alineamiento aminoacídico de las variantes de g2. El círculo azul muestra las candidatas con un fenotipo mutante parcial; el círculo rojo muestra las variantes con fenotipo mutante total. Los colores en el alineamiento representan un aminoácido distinto. La presencia de asteriscos denota la formación de un codón de paro en esa posición.

Cada color en el alineamiento de la figura 17 representa un aminoácido distinto, por lo que colores distintos a los de la secuencia query, denotan un cambio de aminoácido.

Con un círculo azul, se muestran las secuencias que dan origen a una proteína con fenotipo parcialmente mutante; mientras que aquellas con un círculo rojo representan proteínas con fenotipo totalmente mutante, es decir, que pierde por completo su función antiCRISPR.

Con el propósito de tener una lectura más clara de las mutaciones obtenidas en cada posición, y el efecto que produjo cada una de ellas, se realizó una figura donde se mapearon cada una de las mutaciones en cada una de las candidatas secuenciadas (Fig. 18).



**Fig. 18. Mapeo de candidatas secuenciadas sobre la secuencia de G2 silvestre.** En la imagen se muestran las posiciones en las que se obtuvieron cambios según las secuencias analizadas. La secuencia mostrada en el centro de la figura corresponde a la de G2 silvestre. Los recuadros en verde denotan los residuos que presentaron cambios en alguna de las secuencias. Las mutaciones mapeadas en la parte superior de la secuencia corresponden a fenotipos silvestres, mientras que las mapeadas en la parte inferior poseen fenotipos mutantes (en azul las parciales y en rojo las totales).

En el centro de la figura se muestra la secuencia silvestre de G2, donde en negro aparecen los aminoácidos que no presentaron ningún cambio en las secuencias, y en verde aquellos que si tuvieron. En la parte superior de la secuencia, se muestran mapeadas las mutantes que tuvieron un fenotipo silvestre, mostrándose en un renglón diferente las distintas candidatas (e.g. Mut-A4 señalada en el primer renglón arriba de G2, tuvo 4 cambios: F4L, T5A e I7F).

En la parte inferior de la secuencia, se mapearon las candidatas con un fenotipo mutante; en azul se muestran las que tuvieron un fenotipo mutante parcial, y en rojo aquellas con un fenotipo mutante total. De estas candidatas, las dos con fenotipo mutante parcial (Mut-A3 y Mut-A10) tuvieron 3 mutaciones en diferentes codones cada una. Las dos mutantes que tuvieron fenotipo mutante total presentaron un codón de paro prematuro en la lisina 55 (L55\*); el de Mut-A7 fue generado por el cambio del codón de lisina por uno de paro, y Mut-C12 tuvo la deleción del nucleótido 142 lo cual movió el marco de lectura, ocasionando la formación de un codón de paro en la misma posición que Mut-A7.

Adicionalmente, Mut-C12 presentó múltiples cambios a lo largo de la secuencia, originados por la deleción de un nucleótido (Fig. 16), sin embargo, dado que la formación del codón de paro prematuro confiere el fenotipo mutante total (observado en Mut-A10), no es posible atribuirle ningún efecto al resto de los cambios en la secuencia.

En la figura 18 sólo se muestran 12 de las 15 candidatas secuenciadas, esto debido a que 3 de ellas presentaron una secuencia silvestre.

Como ya se mencionó, durante este trabajo empleamos 3 condiciones de mutagénesis por PCR distintas, las cuales variaron en la concentración de  $MnCl_2$ ,  $MgCl_2$ , y/o número de ciclos de extensión (véase sección 7.7 de materiales y métodos).

Tras secuenciar 15 candidatas obtenidas por estas 3 condiciones, se determinó el número promedio de mutaciones por molécula de cada una de ellas, obteniéndose:

Condición A= 2.9 mutaciones/molécula  
Condición B= 1.3 mutaciones/molécula  
Condición C= 1.5 mutaciones/molécula

Por lo tanto, la tasa de error por posición (número de mutaciones por molécula/longitud del fragmento a mutagenizar) de cada una de las condiciones es:

Condición A= 0.014  
Condición B= 0.006  
Condición C= 0.007

La ecuación que dicta la probabilidad P de obtener k mutaciones por molécula en una secuencia de n pares de bases, mediante una condición de mutagénesis con una tasa de error por posición de  $\epsilon$ , se representa de la siguiente manera:

$$P(k, n, \epsilon) = \frac{n!}{(n - k)! * k!} * \epsilon^k (1 - \epsilon)^{n - k}$$

Donde:

k= número de mutaciones

n= longitud de la secuencia

$\epsilon$ = tasa de error por posición

Por lo que la probabilidad de obtener mutantes con una sola mutación por molécula para cada una de las condiciones es:

$$P_A(1, 204, 0.014) = 16.3\%$$

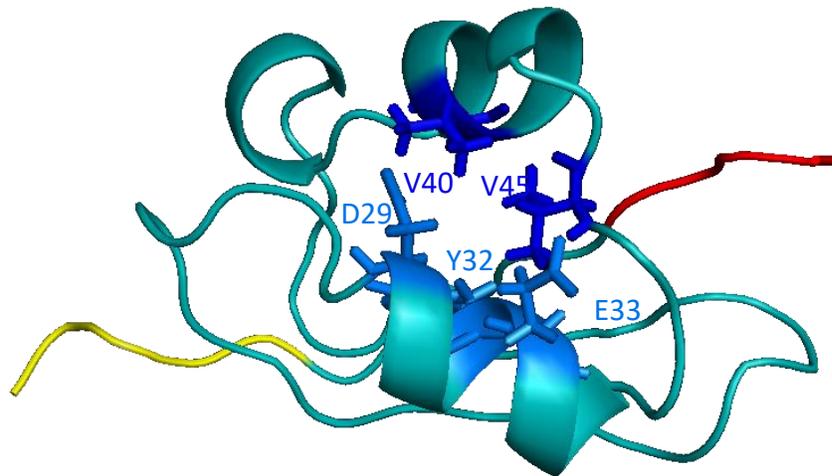
$$P_B(1, 204, 0.0063) = 35.6\%$$

$$P_C(1, 204, 0.0073) = 33.6\%$$

Debido a lo anterior, hace sentido el hecho de que la mayoría de las secuencias obtenidas contuviesen más de una mutación por molécula (Fig. 16).

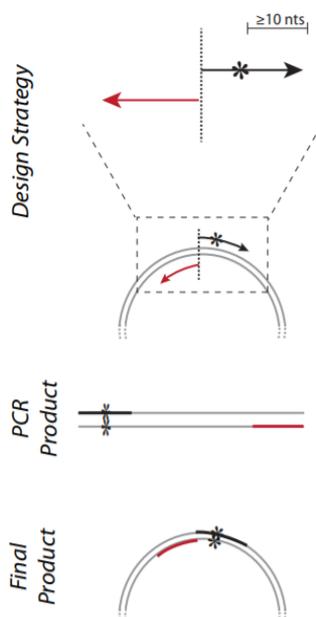
#### 8.4 Mutagénesis sitio dirigida

Con el fin de realizar un análisis más profundo de las mutaciones obtenidas, se realizó un modelamiento de la estructura tridimensional de la proteína antiCRISPR empleando la herramienta iTasser. La estructura predicha por dicho software se muestra a continuación (Fig. 19), donde en amarillo se muestra el extremo amino terminal de la proteína, y en rojo el carboxilo terminal. Adicionalmente se encuentran señalados los residuos que se hipotetiza se encuentran relacionados con el reconocimiento del sistema CRISPR-cas (véase Discusión).



**Fig. 19. Estructura tridimensional de G2.** La imagen es una predicción de la proteína G2 obtenida con iTasser. En color azul marino se muestran los aminoácidos encontrados en secuencias con fenotipo mutante parcial (V40 y V45). En azul celeste se aprecian los aminoácidos propuestos como parte de la superficie de interacción con el sistema CRISPR-cas. La sección amarilla de la proteína denota el extremo amino terminal, y en rojo se muestra el carboxilo terminal.

### A. Substitutions



Con el fin de evaluar dicha hipótesis, se siguió una estrategia de mutagénesis sitio dirigida, para lo cual se diseñaron 4 pares de oligonucleótidos empleando la herramienta NEBaseChanger. Los primers se diseñaron como se muestra en la figura 20. En el caso de las mutaciones por sustitución, éstas son creadas mediante el diseño de un primer mutagénico con un cambio en el centro (figura 20).

Se incluyen al menos 10 nucleótidos que sean complementarios al plásmido en el extremo 3' del primer. El extremo 5' del segundo primer debe comenzar en la siguiente base del extremo 5' del primer oligo, y continuar en dirección opuesta en la cadena complementaria.

**Fig. 20. Diseño de oligos para el kit Q5 Site-Directed Mutagenesis.** En la imagen se muestran en negro y rojo los oligos complementarios al vector. En asterisco se muestra el cambio contenido en la secuencia.

Este segundo primer puede ser 100% complementario a la secuencia del plásmido, o puede contener cambios si se desea. En nuestro caso, se diseñó un par de oligos que incluyeran mutaciones en las posiciones V40 y V45 (una en cada oligo), y un par de primers distintos para cada una de las 3 posiciones restantes (D29, Y32 y E33).

### Secuencia V40/45D

ATG ACT ACG TTC ACC AGC ATC GTT ACC ACT AAC CCT GAT TTC GGC GGA TTC GAG TTC TAC  
 GTT GAA GCG GGA CAG CAG TTC GAC GAT TCA GCA TAT GAA GAG GCA TAC GGT GTT AGC  
 GTT CCT ACC GCC GTG GAC GAG GAA ATG AAC GCT AAA GCG GCG CAG TTG AAA GAC GGC  
 GAG TGG CTT AAT GTC TCG CAT GAA GCC TGA

Nombre	Secuencia	Longitud	%GC	Tm	Ta*
V45D_forward	ACCGCCGTGGwCGAGGAAATG	21	62	68°C	66°C
V40D_reverse	AGGAwCGCTAACACCGTATG	20	50	65°C	

Donde w es igual a Adenina y Timina, lo cual da lugar a un codón de Valina y Aspartato, por lo cual, con este par de primers es posible obtener 4 combinaciones de secuencia: ambos residuos silvestres (V40V, V40V), la valina 40 silvestre y la valina 45 mutante (V40V, V45D), la valina 40 mutante y la valina 45 silvestre (V40D, V45V), y ambos aminoácidos mutantes (V40D, V45D).

#### Secuencia D29

ATG ACT ACG TTC ACC AGC ATC GTT ACC ACT AAC CCT GAT TTC GGC GGA TTC GAG TTC TAC  
GTT GAA GCG GGA CAG CAG TTC GAC GCT TCA GCA TAT GAA GAG GCA TAC GGT GTT AGC  
GTT CCT ACC GCC GTG GTC GAG GAA ATG AAC GCT AAA GCG GCG CAG TTG AAA GAC GGC  
GAG TGG CTT AAT GTC TCG CAT GAA GCC TGA

Nombre	Secuencia	Longitud	%GC	Tm	Ta*
D29_forward	CAGTTCGACGbTTCAGCATATGAAG	25	48	62°C	63°C
D29_reverse	CTGTCCCGCTTCAACGTA	18	56	65°C	

Donde b es igual a Citocina, Timina y Guanina, lo cual da lugar a un codón de Valina, Glicina y Alanina, por lo cual, con este par de primers es posible obtener 3 combinaciones de secuencia: D29V, D29G y D29A. Este diseño fue pensado con el fin de responder si lo importante para el reconocimiento del sistema CRISPR-cas es la posición del residuo, o las propiedades de éste, e.g. polaridad.

### Secuencia Y32

ATG ACT ACG TTC ACC AGC ATC GTT ACC ACT AAC CCT GAT TTC GGC GGA TTC GAG TTC TAC  
GTT GAA GCG GGA CAG CAG TTC GAC GAT TCA GCA GCT GAA GAG GCA TAC GGT GTT AGC  
GTT CCT ACC GCC GTG GTC GAG GAA ATG AAC GCT AAA GCG GCG CAG TTG AAA GAC GGC  
GAG TGG CTT AAT GTC TCG CAT GAA GCC TGA

Nombre	Secuencia	Long	%GC	Tm	Ta*
Y32_forward	CGATTCAGCAg <sub>n</sub> TGAAGAGGCATACGGTGTAG	33	52	65°C	66°C
Y32_reverse	TCGAACTGCTGTCCCGCT	18	61	69°C	

En el diseño de estos oligos se incluyeron 2 sustituciones, por lo que el codón TAT (correspondiente a tirosina) cambia por el condón GnT. N es igual a los 4 nucleótidos (Adenina, Citocina, Timina y Guanina), lo cual da lugar, en conjunto con el cambio de T por G a un codón de Aspartato, Valina, Glicina y Alanina por lo cual, con este par de primers es posible obtener 4 combinaciones de secuencia: Y32D, Y32V, Y32G y Y32A.

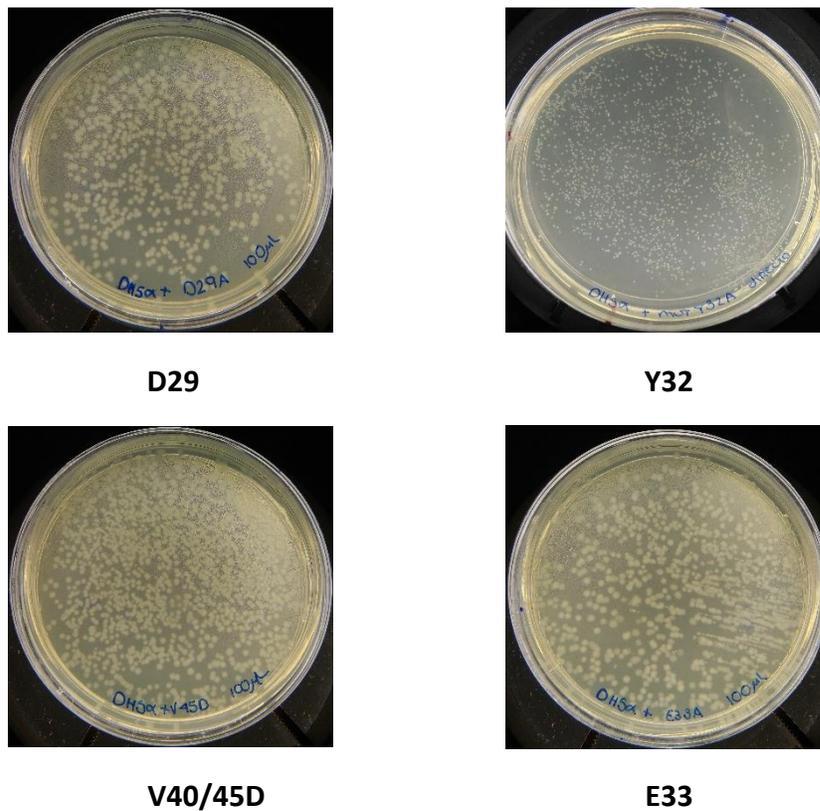
### Secuencia E33

ATG ACT ACG TTC ACC AGC ATC GTT ACC ACT AAC CCT GAT TTC GGC GGA TTC GAG TTC TAC  
GTT GAA GCG GGA CAG CAG TTC GAC GAT TCA GCA TAT GCA GAG GCA TAC GGT GTT AGC  
GTT CCT ACC GCC GTG GTC GAG GAA ATG AAC GCT AAA GCG GCG CAG TTG AAA GAC GGC  
GAG TGG CTT AAT GTC TCG CAT GAA GCC TGA

Nombre	Secuencia	Longitud	%GC	Tm	Ta*
E33_forward	GATTCAGCATATGbAGAGGCATACG	22	50	61°C	66°C
Y32_reverse	TCGAACTGCTGTCCCGCT	18	61	69°C	

Por la proximidad a la secuencia de Y32, se empleó el mismo oligo reverso. Adicionalmente, al igual que los primers D29A, se diseñaron con una b la cual es igual a Citocina, Timina y Guanina, lo cual da lugar a un codón de Valina, Glicina y Alanina por lo cual, con este par de primers es posible obtener 3 combinaciones de secuencia: E33V, E33G y E33A.

Se realizó la mutagénesis sitio-dirigida de los diferentes residuos propuestos como parte de la superficie de interacción de G2 con el sistema CRISPR-cas, empleando el kit Q5® Site-Directed Mutagenesis de NEB. Los vectores mutantes en cada uno de estos residuos se transformaron en las células quimiocompetentes DH5α proporcionadas en el kit.



**Fig. 21. DH5α transformadas con las diferentes mutantes sitio dirigidas.** En la imagen se muestran cajas representativas de la transformación de las distintas versiones de g2 en *E. coli*. Todas las cajas contienen LB con gentamicina (15 µg/mL).

Cabe mencionar que las imágenes anteriores solo son una muestra de las cajas obtenidas, puesto que la reacción de transformación fue dividida en 4 para las mutantes Y32 y V40/45D; las mutantes D29 y E33 se dividieron en 3. Esto debido al número de combinaciones de mutaciones que daban lugar cada uno de los oligos degenerados, por ejemplo, los oligos de Y32 dan lugar a 4 combinaciones: Y32A, Y32G, Y32V y Y32D, por lo tanto, con esto se trató de evitar la sobre representación de una sola clona.

A partir de cada una de estas cajas, se tomaron 3 colonias distintas y se realizó un cultivo pool en medio LB, a partir del cual se extrajo plásmido. Esto con el fin de aumentar la probabilidad de obtener todas las posibles mutaciones. Posteriormente, los pools de las mutantes Y32 y V40/45D (4 pools para cada una), se transformaron en *P. aeruginosa* quimiocompetentes, quedando pendiente la transformación de los pools de D29 y E33.

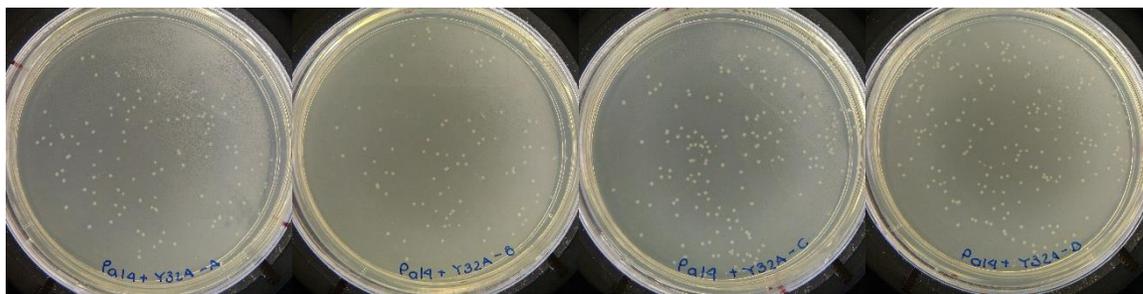


Pa14+V40/45D pool A

Pa14+V40/45D pool B

Pa14+V40/45D pool C

Pa14+V40/45D pool D



Pa14+Y32 pool A

Pa14+Y32 pool B

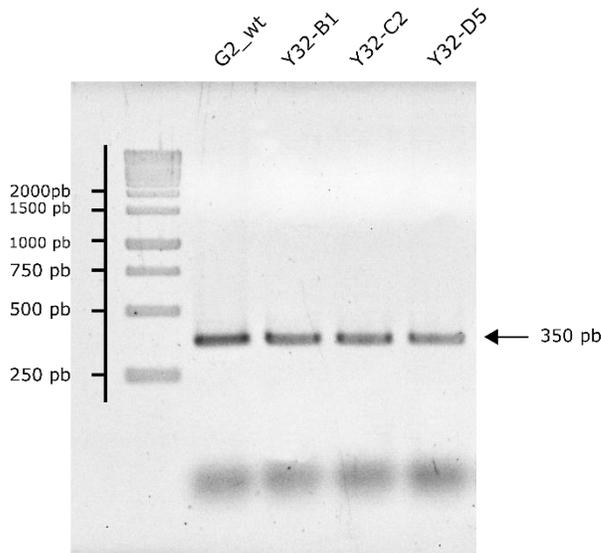
Pa14+Y32 pool C

Pa14+Y32 pool D

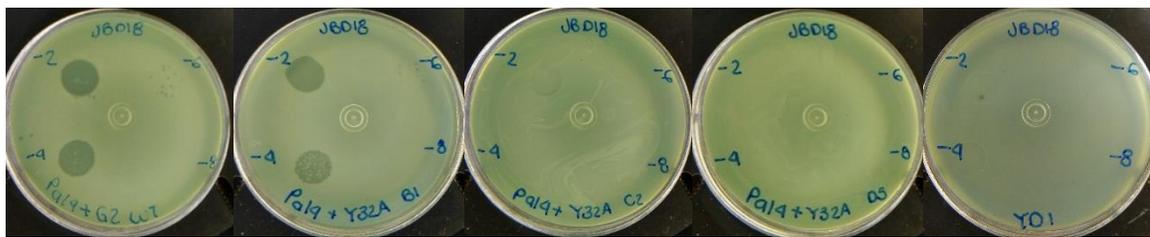
**Fig. 22. PA14 transformadas con los diferentes pools de cada mutante sitio dirigida.** En la imagen se muestran cajas representativas de la transformación de los distintos pools de plásmido obtenidos de las mutantes Y32 y V40/45D de G2 en *P. aeruginosa*. Todas las cajas contienen LB con gentamicina (50 µg/mL).

A partir de cada caja pool se tomaron 6 colonias (24 colonias en total para cada mutagénesis), y se realizaron PCR de cada una, tomando un respaldo de cada candidata para realizar cultivos de éstas y proceder a fenotipificar.

A continuación, se muestran las PCR obtenidas de 4 de las 24 colonias analizadas de las mutantes Y32 (Fig. 23), así como cajas del perfil de infección del fago sensible JBD18 sobre cada una de ellas (Fig. 24).



**Fig. 23. PCR de candidatas provenientes de mutagénesis Y32.** El gel de la izquierda muestra los amplicones obtenidos a partir de una PCR de colonia de las candidatas provenientes de la mutagénesis Y32 que se secuenciaron. En el primer carril se muestra el control positivo de PCR (g2 WT), y el resto de los carriles representan cada una de las candidatas secuenciadas (nombres mostrados en la parte superior).



**PA14+G2 WT      PA14+ G2 Y32-B1      PA14+ G2 Y32-C2      PA14+ G2 Y32-D5      PA14+ G2 Y32-D1**

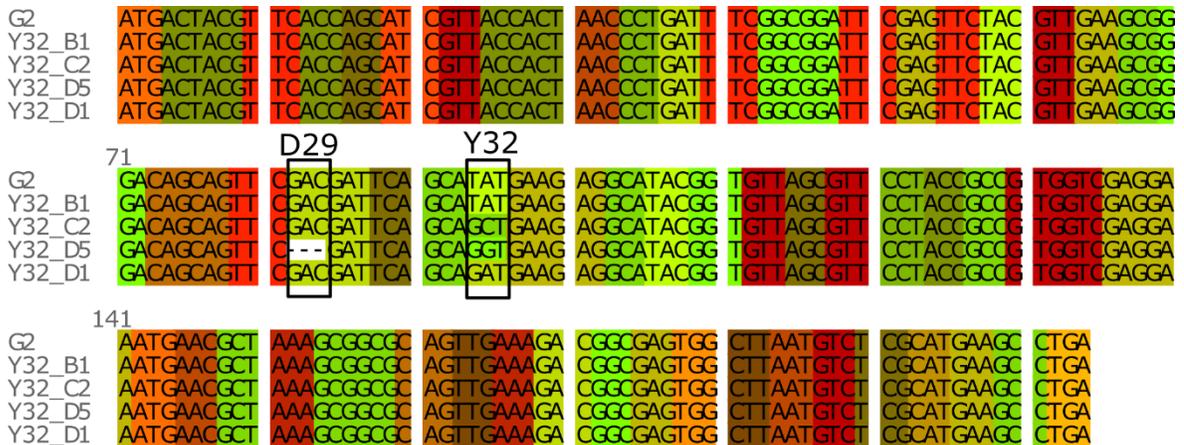
**Fig. 24. Fenotipificación de mutantes Y32.** En las imágenes se muestran los distintos fenotipos obtenidos en las mutantes Y32 secuencias. En el extremo izquierdo tenemos el control del antiCRISPR silvestre. En el resto de las imágenes se muestran las candidatas que se eligieron para su secuenciación, teniendo una para cada uno de los fenotipos encontrados: Y32-B1 muestra un fenotipo silvestre, Y32-C2 mutante parcial, Y32-D5 y Y32-D1 mutante total.

Las 4 candidatas que se muestran en las figuras anteriores fueron secuenciadas y se encontró que la candidata con fenotipo silvestre (Y32-B1) de hecho tiene secuencia silvestre (Y32Y, fig. 25), a pesar de que los oligos fueron diseñados para que todas las secuencias obtenidas fueron mutantes, por lo cual es posible que no se haya degradado por completo el DNA molde durante la reacción con la enzima *DpnI*.

La mutante con fenotipo mutante parcial (Y32-C2), que permite la infección del fago sensible al sistema CRISPR-cas con aproximadamente 1000 veces menor eficiencia que el antiCRISPR silvestre, tuvo un cambio de la Tirosina por Alanina (Y32A, fig. 25).

Una de las mutantes con fenotipo mutante total (Y32-D5), en la cual no se observó ninguna placa en el tapiz, inesperadamente tuvo 2 modificaciones, el cambio de la Tirosina por Glicina (Y32G, fig. 25), y sorpresivamente presentó la delección de un aminoácido completo en la posición 29 (D29).

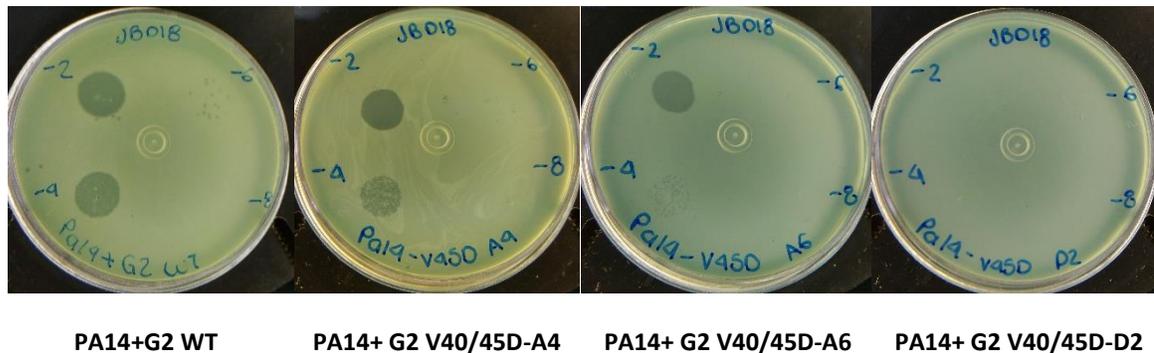
Por último, la mutante total restante, Y32-D1, tuvo un cambio de la Tirosina por un Aspartato (Y32D, fig. 25), lo cual nos confirma que la presencia de una Tirosina en esta posición es importante, ya que el cambio de esta por aminoácidos con distintas propiedades fisicoquímicas, da como resultado la pérdida de la función, en este caso una pérdida total.



**Fig. 25. Alineamiento nucleotídico de las candidatas secuenciadas pertenecientes a la mutagénesis sitio dirigida Y32.** En la imagen se muestra un alineamiento a nivel de nucleótidos, donde cada color representa un codón distinto. En cuadros negros se señalan las posiciones donde hubo cambios en la secuencia, y con letras negras se señala el aminoácido codificado por dichos codones.

Por otro lado, para la mutagénesis de las valinas ya encontradas por mutagénesis al azar, si bien aún no se cuenta con las secuencias, se fenotipificaron 24 candidatas. De estas se obtuvieron 14 con fenotipo silvestre, 2 con fenotipo mutante parcial y 8 mutantes totales.

En la figura 26, se muestra una imagen representativa de cada uno de estos fenotipos:



**Fig. 26. Fenotipificación de mutantes V40/45D.** En las imágenes se muestran los distintos fenotipos obtenidos en las mutantes V40/45D secuencias. En el extremo izquierdo tenemos el control del antiCRISPR silvestre. En el resto de las imágenes se muestran candidatas representativas de cada uno de los fenotipos encontrados: V40/45D-A4 muestra un fenotipo silvestre, V40/45D-A6 mutante parcial y V40/45D-D2 mutante total.

Las candidatas mostradas en la fig. 26 aún están en la espera de ser secuenciadas, pero por lo que hemos hipotetizado anteriormente, es posible que las candidatas con fenotipo silvestre sean silvestres en secuencia (por el diseño de los oligos de mutagénesis), las de fenotipo mutante parcial solo tengan el cambio de alguna de las Valinas por Aspartato, y la mutante total podría tener el cambio de ambas Valinas por Aspartato.

## 9. DISCUSION

El primer objetivo del presente proyecto fue la búsqueda de aminoácidos conservados en las distintas familias de proteínas antiCRISPR mediante un análisis de alineamiento múltiple de secuencias. Este objetivo se planteó para identificar residuos aminoacídicos posiblemente implicados en el reconocimiento del sistema CRISPR-cas. Sin embargo, existen puntos que limitan la comparación entre cada una de las familias, el primero de ellos es el bajo porcentaje de similitud entre ellas, siendo las más similares AcrF3 y AcrF4 con un porcentaje de identidad de 28%.

Por otro lado, no existe suficiente evidencia experimental sobre los componentes del sistema de vigilancia a las cuales se une cada una de las familias (solo se ha determinado el mecanismo de acción de 3 de las 10 familias que inhiben el sistema I-F), por lo tanto, resulta evidente que proteínas con diferentes blancos interactuarán de manera disímil con éstos, lo cual se traduce en la ausencia de aminoácidos conservados entre las distintas familias antiCRISPR.

La estrategia que consideramos más informativa para responder cuáles son los aminoácidos importantes para mantener la función antiCRISPR fue la mutagénesis al azar. La mutagénesis al azar permite probar estructural y funcionalmente regiones importantes dentro de un gen. La mutagénesis *in vitro*, acoplada a un método de selección, es especialmente útil cuando se desconocen las posiciones aminoacídicas significativamente importantes, como es el caso de G2, puesto que, como se mencionó anteriormente, mediante alineamientos de las diversas proteínas antiCRISPR hasta ahora reportadas, no se ha logrado identificar aminoácidos conservados entre éstas.

A pesar de que existen diversos métodos para introducir mutaciones aleatoriamente, se ha observado que uno de los más versátiles y simples es el uso de polimerasas poco fidedignas, que promuevan errores en el copiado de un gen a través de una reacción de PCR (Cadwell et al., 1992). La tasa de error de la polimerasa Taq es de 0.001-0.02% por nucleótido por pase de la polimerasa, dependiendo de las condiciones de la reacción (Eckert et al., 1990; Eckert et al., 1991). La fidelidad de la PCR se reduce mediante el incremento de la concentración de  $MgCl_2$ , la adición de  $MnCl_2$  a la mezcla de reacción, un desbalance en la concentración de dNTPS, aumento en la concentración de la Taq polimerasa, e incremento del número de ciclos de extensión (Leung et al., 1989; Cadwell et al., 1992). Sin embargo, como se mencionó anteriormente (Resultados 8.3), las probabilidades de obtener una mutación puntual por clona eran bajas, y aunado a esto, tuvimos una baja eficiencia de clonación del gen de interés (44 candidatas con inserto único de 1000 analizadas), por lo tanto, se decidió realizar un análisis exhaustivo de las secuencias ya obtenidas, con el fin de proponer residuos candidatos para cambiar por mutagénesis sitio dirigida.

El primer paso fue analizar las secuencias en su estructura primaria (Fig. 18). Debido a que las secuencias con un fenotipo mutante son más informativas, nos centramos en éstas, haciendo uso de la información proporcionada por las de fenotipo silvestre cuando fuera necesario.

La primer mutante analizada fue Mut-A3, la cual posee 3 cambios puntuales: F4S, D29E y V45D. En la primera mutación, tenemos el cambio de una Fenilalanina por una Serina; la Fenilalanina es un aminoácido no polar aromático, mientras que la Serina es un aminoácido polar no cargado, sin embargo, en la mutante Mut-A4 está presente un cambio en la misma posición (F4L), además de otros dos en la misma región (T5A, I7F), teniendo un fenotipo silvestre, por lo que mutaciones en esta región no parecen tener un efecto en la función del gen. La segunda mutación (D29E) fue el cambio de un Aspartato por un Glutamato, los cuales son aminoácidos polares de carga negativa, con una composición química muy similar entre ellos, por lo cual no resulta plausible que un cambio de uno por el otro dé lugar a cambios fenotípicos. Por último, tenemos el cambio de una Valina por un Aspartato (V45D), lo cual cambia un aminoácido no polar por uno polar con carga negativa, el cual parece ser el cambio más relevante en la secuencia, y que pudiera estar confiriendo el fenotipo observado.

La segunda mutante analizada fue Mut-A10, la cual, al igual que Mut-A3, posee un fenotipo mutante parcial, y 3 mutaciones en su secuencia: F14L, V40D y E46D. El primero es un cambio entre 2 aminoácidos no polares, Fenilalanina y Leucina, lo cual pareciera no tener un gran efecto sobre la proteína, ya que es poco probable que el cambio entre dos aminoácidos con propiedades fisicoquímicas similares induzca un cambio conformacional que pueda afectar la función.

La segunda mutación es el cambio de una Valina por un Aspartato, el cual es el mismo caso que se explicó para la mutante Mut-A3, pero en distinta posición (V40), por lo cual es plausible que esta mutación contribuya al fenotipo mutante.

Por último, se observó el cambio de glutamato 46 por aspartato, al igual que en la posición 29 (D29E) de Mut-A3, sin embargo, además de lo que ya se explicó para D29E, contamos con una secuencia con fenotipo silvestre con un cambio puntual del glutamato 46 por glicina, lo que nos confirma que un cambio de este aminoácido no tiene repercusión alguna sobre la funcionalidad del antiCRISPR.

Adicionalmente a lo ya mencionado, es oportuno aclarar que hasta este punto no se contaba con evidencia que nos señalara que el fenotipo parcial solo era conferido por una mutación puntual, y no por una combinación de varias.

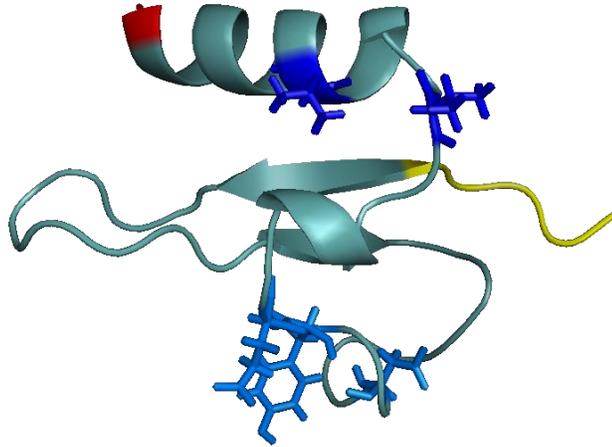
Para el caso de las mutantes totales, la primera de ellas, MutA-7, presentó un cambio que generó un codón de paro prematuro (L55\*), dando lugar a una proteína 13 aminoácidos más corta, lo cual nos dice que estos aminoácidos son importantes para el funcionamiento de la proteína, ya sea en el reconocimiento o en el plegamiento correcto de ésta.

La mutante Mut-C12 presenta múltiples mutaciones originadas por la delección de un nucleótido, lo que además provoca un codón de paro prematuro en la misma posición que la mutante anterior (L55\*), por lo que no es posible darles una lectura a los cambios de los demás residuos puesto que la generación del codón de paro da lugar a la pérdida total de la función.

Debido a que el análisis de las mutaciones en la estructura primaria fue informativo, pero no nos dio indicios claros de los sitios de interacción, se decidió realizar un análisis de las mutaciones en la estructura terciaria. Para ellos se mandó modelar la proteína al servidor de iTasser (Fig. 19).

En la figura 19 se muestran las Valinas 40 y 45, las cuales hipotetizamos se encuentran implicadas en la pérdida parcial de la función antiCRISPR (por lo que se explicó anteriormente), sin embargo, como se puede observar en rojo, el carboxilo terminal no se encuentra cerca de ellas, y según nuestros resultados, este segmento de la proteína es indispensable para el funcionamiento, lo que nos hizo plantearnos la posibilidad de que la pérdida de esos 13 aminoácidos afectara la estructura de la proteína, más que intervenir directamente con el reconocimiento del sistema CRISPR-cas.

Para explorar la hipótesis anterior, se modeló con la misma herramienta (iTasser) la mutante Mut-A7, obteniéndose la siguiente estructura:

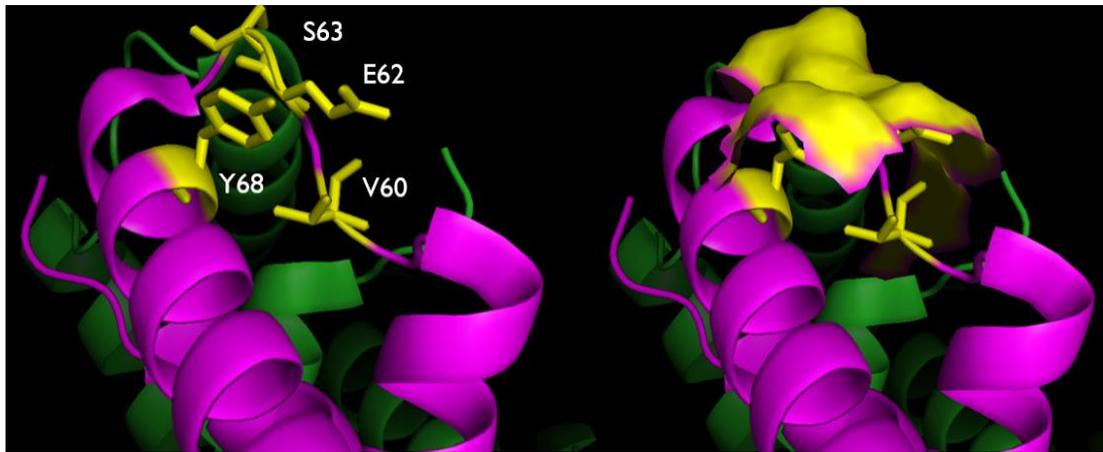


**Fig. 27. Estructura tridimensional de Mut-A7.** La imagen es una predicción de la proteína G2 en su versión mutante Mut-A7, obtenida con iTasser. Al igual que en la fig. 19, en color azul marino se muestran los aminoácidos encontrados en secuencias con fenotipo mutante parcial (V40 y V45). En azul celeste se aprecian los aminoácidos pertenecientes a la hélice alfa que originalmente se encontraba frente a V40 y V45. La sección amarilla de la proteína denota el extremo amino terminal, y en rojo se muestra el carboxilo terminal.

Como es posible observar, la proteína sufre una reestructuración, haciendo que la hélice alfa que antes se encontraba frente a las valinas (en azul marino), se aleje de ellas (en azul claro). Esto nos llevó a pensar si estos aminoácidos pudieran estar relacionados con el reconocimiento.

En busca de más argumentos que soportaran la hipótesis anterior, decidimos analizar las estructuras ya reportadas de antiCRISPR y su sitio de interacción con el sistema CRISPR-cas. La primera que analizamos fue la reportada por Xiaofei Wang y colaboradores (Xiaofei Wang et al., 2016). Ellos reportaron la estructura del antiCRISPR AcrF3, el cual se sabe que se une a la nucleasa Cas3. Ellos encontraron que AcrF3 forma un homodímero que interactúa con los distintos dominios de la Cas3, siendo el que llamó nuestra atención el dominio RecA2, esto debido a que contiene una Valina (V60) en su zona de interacción.

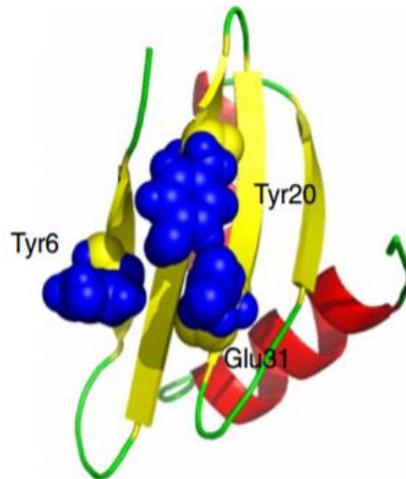
Esta Valina, en conjunto con la Tirosina 68, la Serina 63 y el Glutamato 62 proporcionan una superficie que permite la interacción (Fig. 28) con el dominio antes mencionado.



**Fig. 28. Estructura tridimensional de AcrF3: Acercamiento del sitio de interacción con RecA2 de Cas3.** La imagen representa la estructura de AcrF3 obtenida mediante cristalización por Xiaofei Wang y cols. En morado se muestra uno de los monómeros que conforman el dímero de AcrF3; en verde se muestra el otro monómero. Los aminoácidos mostrados en amarillo pertenecen al sitio de interacción con el dominio RecA2 de la proteína Cas3. En el extremo derecho de la figura se muestra la superficie de interacción.

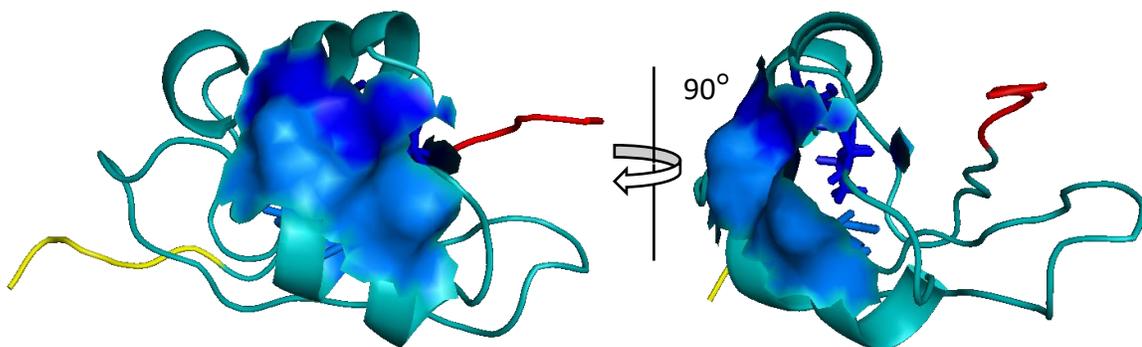
Posteriormente, se analizó la estructura del antiCRISPR AcrF1, reportada por el grupo del profesor Alan Davidson, en la cual encontraron que la superficie de interacción está compuesta por la Tirosina 6, la 20, y el Glutamato 31 (Fig. 29).

Esto, en conjunto con lo reportado por Xiaofei Wang, nos dio pistas de que tal vez las tirosinas y los aminoácidos de carga negativa pudieran estar asociados con la interacción con las proteínas Cas.



**Fig. 29. Estructura tridimensional de AcrF1.** La imagen representa la estructura de AcrF1 obtenida mediante RMN por Maxwell KL y cols. En amarillo se muestran las láminas b, en rojo las hélices alfa, y en verde los loops. Los aminoácidos mostrados en azul pertenecen al sitio de interacción con la proteína Csy3.

Encontrado lo anterior, analizamos los aminoácidos que se encontraban en la estructura de nuestra proteína, específicamente en la zona cercana a las Valinas 40 y 45 que se ve desplazada en ausencia del carboxilo terminal, y encontramos un Aspartato en la posición 29 (D29), una Tirosina en la 32 (Y32) y un Glutamato en 33 (E33), por lo cual nosotros hipotetizamos que son estos residuos los que pudieran interactuar con el sistema bacteriano, teniendo una superficie de interacción como se muestra a continuación:



**Fig. 30. Posible superficie de interacción propuesta para G2.** La imagen muestra la superficie de interacción propuesta para el antiCRISPR G2. En verde se muestra el core de la proteína; en el centro de ella se muestra graficada la densidad de electrónica de los aminoácidos propuestos como sitios de interacción (en azul) con el sistema CRISPR-cas. En la derecha se puede apreciar una vista de perfil de la superficie.

Para probar la hipótesis anterior, se diseñaron primers para mutar los aminoácidos antes mencionados (véase Métodos 7.23 y Resultados 8.4). Hasta el momento, hemos secuenciado 4 candidatas de la mutagénesis de la Tirosina 32 (Y32). Respecto a la mutante total que tuvo un cambio de Tirosina por Glicina y la delección del Glutamato 29 (mismo residuo planeado cambiar por mutagénesis sitio dirigida), es imposible atribuirle la pérdida de la función a alguno de los 2 cambios, aunque nos podría dar indicios de que ambos aminoácidos son importantes en la interacción de G2 con el sistema de defensa bacteriano.

Por otro lado, encontramos que el cambio de este aminoácido por una Alanina (Y32A), que es el cambio de un aminoácido polar por uno no polar, hace que la función de la proteína decrezca aproximadamente 1000 veces (Fig. 24). Para este resultado existen dos posibles explicaciones, la primera es que nuestra hipótesis sea correcta y que efectivamente la tirosina 32 interactúa directamente con el sistema CRISPR-cas; sin embargo, existe la posibilidad de que dicho fenotipo sea resultado de la desestabilización de la hélice alfa en la cual se encuentra este aminoácido.

Por otro lado, la mutante Y32-D1, que presentó el cambio de la Tirosina por un Aspartato (Y32D), que es el cambio de un aminoácido polar por otro polar, pero con carga negativa, induce la pérdida total de la función antiCRISPR, lo cual nos sugiere que este aminoácido está relacionado con el reconocimiento del sistema CRISPR-cas. Otro argumento que apoya esta hipótesis es que anteriormente Maxwell y colaboradores (Maxwell et al., 2016) reportaron una colección de mutantes sitio-dirigidas del antiCRISPR AcrF1, en la cual cuentan con mutantes con cambios en aminoácidos que conforman hélices alfa (entre ellos el cambio de la Tirosina 55 por una Alanina), en los cuales no observaron disminución de la función antiCRISPR.

De hecho, los únicos residuos que encontraron que un cambio en ellos induce una pérdida parcial o total de la función, fueron los que posteriormente se demostró que efectivamente interactúan con el sistema CRISPR-cas, por la obtención de su estructura en conjunto con el complejo Csy (Chowdhury et al., 2017). Lo cual nos sugiere que es altamente probable que la posición Y32 de nuestra proteína, forme parte de la superficie de interacción con el sistema CRISPR-cas.

A pesar de que aún no contamos con las secuencias de las mutantes V40/45, los fenotipos observados en éstas (Fig.26) indican que efectivamente, tal como se había hipotetizado, las valinas 40 y 45 son las responsables de los fenotipos mutantes parciales observados en las secuencias de la mutagénesis aleatoria (mut-A3 y mut-A10), y por lo tanto, es altamente probable que la superficie de interacción que propusimos sea la correcta, aunque es necesario hacer los ensayos de fenotipificación y secuenciación de los residuos D29 y E33.

Creemos que este resultado es altamente valioso ya que nos confirma que elegimos correctamente la estrategia para la búsqueda de aminoácidos implicados en la función de nuestra proteína, puesto que otras metodología, como la empleada por el grupo del profesor Alan Davidson, la cual consistió en la modificación de todos los aminoácidos polares del antiCRISPR AcrF1 por Alanina (aminoácido no polar), pueden perder información como la que nuestro grupo y el del profesor Zhu encontramos respecto a las valinas.

Es evidente que las valinas, al ser hidrofóbicas, no pueden contribuir a la unión con el sistema CRISPR-cas mediante puentes de hidrógeno, pero si pueden estabilizar la unión mediante la formación de bolsillos hidrofóbicos. De hecho, existen numerosos casos conocidos donde la unión de sustratos y cofactores a una enzima involucran interacciones con un área hidrofóbica expuesta, frecuentemente en “grietas” en la superficie (Lijnzaad et al. 1996). Por lo tanto, estrategias donde solo se consideran las posibles interacciones mediante los aminoácidos polares, parece no mostrar la totalidad de posibles interacciones.

Adicionalmente, nosotros hipotetizamos que el mecanismo mediante el cual G2 inhibe el sistema CRISPR-cas es mediante la unión con las proteínas del complejo CRISPR (Csy), o mediante la unión a la nucleasa Cas. Nuestros argumentos son que mediante análisis bioinformáticos no encontramos dominios de unión a DNA o RNA en nuestra proteína. Además, creemos que G2 no posee un “sitio activo” ya que hasta ahora no hemos podido encontrar una mutación puntual que dé lugar a un fenotipo mutante total, lo que nos sugiere que la inhibición se da por un conjunto de aminoácidos, i.e. superficie de interacción.

## 10. CONCLUSIONES

- No existe conservación de aminoácidos en la estructura primaria de las diferentes familias de proteínas antiCRISPR que inhiben el sistema CRISPR-cas I-F.
- La condición de mutagénesis al azar por PCR con una probabilidad mayor de obtener mutaciones en un solo nucleótido fue la B (35.6%).
- El ensayo de colony blot estandarizado para *Pseudomonas aeruginosa* mostró ser una alternativa eficiente para la detección de genes.
- Los resultados de fenotipificación mostraron que los últimos 13 aminoácidos del carboxilo terminal de G2 son indispensables para el funcionamiento de la proteína.
- El estudio sugiere que los aminoácidos Y32, V40 y V45 posiblemente forman parte del sitio de interacción con el sistema CRISPR-cas.

## 11. PERSPECTIVAS

- Secuenciar más candidatas Y32 con fenotipo mutante total para determinar que cambios puntuales dan lugar a ese fenotipo.
- Secuenciar las candidatas V40/45D con fenotipos mutantes para corroborar la intervención de las valinas 40 y 45 en la interacción con las proteínas bacterianas.
- Probar los fenotipos y secuenciar las variantes D29 y E33.
- Clonar y purificar g2 en el vector pET28a, para su posible cristalización.

## REFERENCIAS

1. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* 169:5429–5433.
2. Mojica, F. J., Diez-Villaseñor, C., Garcia-Martinez, J., and Soria, E. (2005). Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* 60:174–182.
3. Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315:1709–1712.
4. Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., Dickman, M. J., Makarova, K. S., Koonin, E. V., and van der Oost, J. (2008). Small CRISPR RNAs guide antiviral defence in prokaryotes. *Science* 321:960–964.
5. Marraffini, L. A., and Sontheimer, E. J. (2010). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genet.* 11:181–190.
6. Van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M., and Brouns, S. J. (2009). CRISPRbased adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.* 34:401–407.
7. Agari, Y., Sakamoto, K., Tamakoshi, M., Oshima, T., Kuramitsu, S., and Shinkai, A. (2010). Transcription profile of *Thermus thermophilus* CRISPR systems after phage infection. *J. Mol. Biol.* 395:270–281.
8. Godde, J. S., and Bickerton, A. (2006). The repetitive DNA elements called CRISPRs and their associated genes: Evidence of horizontal transfer among prokaryotes. *J. Mol. Evol.* 62:718–729.
9. Grissa, I., Vergnaud, G., and Pourcel, C. (2007a). The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics* 8:172.

10. Horvath, P., Romero, D. A., Cou<sup>^</sup>te<sup>^</sup>-Monvoisin, A. C., Richards, M., Deveau, H., Moineau, S., Boyaval, P., Fremaux, C., and Barrangou, R. (2008). Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* 190:1401–1412.
11. Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I., and Koonin, E. V. (2006). A putative RNA-interference-based immune system in prokaryotes: Computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct.* 1:7.
12. Sorek, R., Kunin, V., and Hugenholtz, P. (2008). CRISPR: A widespread system that provides acquired resistance against phages in bacteria and archaea. *Nat. Rev. Microbiol.* 6:181–186.
13. Jansen, R., Embden, J. D., Gastra, W., and Schouls, L. M. (2002b). Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* 43:1565–1575.
14. Deveau, H., Garneau, J. E., and Moineau, S. (2010). CRISPR/Cas system and its role in phagebacteria interactions. *Annu. Rev. Microbiol.* 64:475–493.
15. Lillestøl, R. K., Redder, P., Garrett, R. A., and Brugger, K. (2006). A putative viral defense mechanism in archaeal cells. *Archaea* 2:59–72.
16. Pourcel, C., Salvignol, G., and Vergnaud, G. (2005). CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* 151:653–663.
17. Hale, C., Kleppe, K., Terns, R. M., and Terns, M. P. (2008). Prokaryotic silencing (psi) RNAs in *Pyrococcus furiosus*. *RNA* 14:2572–2579.
18. Lillestøl, R. K., Shah, S. A., Bru<sup>¨</sup>gger, K., Redder, P., Phan, H., Christiansen, J., and Garrett, R. A. (2009). CRISPR families of the crenarchaeal genus *Sulfolobus*: Bidirectional transcription and dynamic properties. *Mol. Microbiol.* 72:259–272.
19. Haft, D. H., Selengut, J., Mongodin, E. F., and Nelson, K. E. (2005). A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* 1:e60.

20. Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S. M., Ma, W., and Doudna, J. A. (2009). Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure* 17:904–912.
21. Zegans, M. E., Wagner, J. C., Cady, K. C., Murphy, D. M., Hammond, J. H., and O'Toole, G. A. (2009). Interaction between bacteriophage DMS3 and host CRISPR region inhibits group behaviors of *Pseudomonas aeruginosa*. *J. Bacteriol.* 191:210–219.
22. E. Semenova, et al., Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence, *Proc. Natl. Acad. Sci. U. S. A.* 108 (25) (2011) 10098e10103.
23. Pawluk A., Bondy-Denomy J., Cheung V. H., Maxwell K. L. and Davidson A. R. (2014). New Group of Phage Anti-CRISPR Genes Inhibits the Type I-E CRISPR-Cas System of *Pseudomonas aeruginosa*. *mBio.* 5(2): e00896.
24. Cazares A., Mendoza-Hernández G. and Guarneros G. (2014). Core and accessory genome architecture in a group of *Pseudomonas aeruginosa* Mu-like phages. *BMC Genomics.* 15: 1146.
25. Cady K. C., Bondy-Denomy J., Heussler G. E., Davidson A. R. and O'Toole G. A. (2012). The CRISPR/Cas adaptive immune system of *Pseudomonas aeruginosa* mediates resistance to naturally occurring and engineered phages. *Journal of Bacteriology.* 194(21): 5728-5738.
26. Bondy-Denomy J., Pawluk A., Maxwell K. L. and Davidson A. R. (2013). Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature.* 493(7432): 429-432.
27. Bondy-Denomy J., Garcia B., Strum S., Du M., Rollins M. F., Hidalgo-Reyes Y., Wiedenheft B., Maxwell K. L. and Davidson A. R. (2015). Multiple mechanisms for CRISPR–Cas inhibition by anti-CRISPR proteins. *Nature.* 526(7571): 136-139.
28. De la Cruz, C. A. (2016). Caracterización del gen g2 del fago H70 de *Pseudomonas aeruginosa* como anti-CRISPR. Tesis de Maestría especialidad Genética y Biología Molecular, CINVESTAV.

29. Cadwell RC., and. Joyce G. (1992). Randomization of Genes by PCR Mutagenesis. *PCR Met. and Appl.* 2:28-33.
30. Eckert, K.A. and T.A. Kunkel. (1990). High fidelity DNA synthesis by the *Thermus aquaticus* DNA polymerase. *Nucleic Acids Res.* 18: 3739-3744.
31. Eckert, K.A. and T.A. Kunkel. (1991). DNA polymerase fidelity and the polymerase chain reaction. *PCR Methods Applic.* 1: 17-24.
32. Leung, D.W., E. Chen, and D.V. Goeddel. (1989). A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* 1: 11-15.
33. Wang X., Yao D., Xu J., Li A., Xu J., Fu P., Zhou Y and Zhu Y. (2016). Structural basis of Cas3 inhibition by the bacteriophage protein AcrF3. *Nat. Stru. Mol. Biol.* 23(9): 868-870.
34. Maxwell KL., Garcia B., Bondy-Denomy J., Bona D., Hidalgo-Reyes Y. and Davidson AR. (2016). The solution structure of an anti-CRISPR protein. *Nat. Com.* 7:13134.
35. Lijnzaad P., Berendsen H. and Argos P. A Method for Detecting Hydrophobic Patches on Protein Surfaces. *Proteins.* 26(2): 192-203.
36. Chowdhury S., Carte J., Rollins M., Golden S., Jackson R., Hoffmann C., Nosaka L., Bondy-Denomy J., Maxwell K., Davidson AR., Fischer E., Lander G., and Wiedenhef B. (2017). Structure Reveals Mechanisms of Viral Suppressors that Intercept a CRISPR RNA-Guided Surveillance Complex. *Cell.* 169, 47–57.

## ANEXOS

### ANEXO 1. Medios y soluciones.

#### 1.1 Medio LB (Medio Luria-Bertani)

##### Por litro:

Bacto-triptona	10 g
Extracto de levadura	5 g
NaCl	5 g

##### Para cajas:

Agar	15 g
------	------

Disolver los ingredientes y calentar para que la disolución sea completa. Esterilizar por autoclave.

#### 1.2 Medio SOC (Super Optimal broth with Catabolite repression)

##### Por litro:

Bacto-triptona	20 g
Extracto de levadura	5 g
NaCl	0.584 g
KCl	0.186 g
MgSO <sub>4</sub>	2.4 g

##### Ajustar pH 7.5

##### Esterilizar

Glucosa 20% estéril	20 mL
---------------------	-------

Disolver los ingredientes y calentar para que la disolución sea completa. Se ajusta el pH a 7.5 con una solución de NaOH 5N. Esterilizar por autoclave. Añadir la glucosa después de la esterilización.

### **1.3 Buffer TMG**

#### **Por litro:**

NaCl	11.7 g
Gelatina	1 g
Tris-HCl 1M pH 8.0	1 mL
MgSO <sub>4</sub> 1M	0.5 mL

#### **Ajustar pH 7.4**

Disolver los ingredientes y calentar para que la disolución sea completa. Se ajusta el pH a 7.4. Esterilizar por autoclave o filtración.

### **1.4 Solución Tfb I**

#### **Por litro:**

KOAc 5M	6 mL
RbCl <sub>2</sub>	12 g
CaCl <sub>2</sub> 1M	10 mL
MgCl <sub>2</sub> *4H <sub>2</sub> O 2M	25 mL
Glicerol 100%	150 mL

#### **Ajustar pH 5.8**

Disolver los ingredientes. Se ajusta el pH a 5.8 con ácido acético. Esterilizar por autoclave o filtración.

### 1.5 Solución Tfb II

#### Por litro:

MOPS	2 g
RbCl <sub>2</sub> 0.5M	20 mL
CaCl <sub>2</sub> 1M	75 mL
Glicerol 100%	150 mL

#### Ajustar pH 6.8

Disolver los ingredientes. Se ajusta el pH a 6.8 con KOH o NaOH. Esterilizar por autoclave o filtración.

### 1.6 Buffer SB 20x

#### Por litro:

NaOH	8 g
Ácido Bórico	45 g

Disolver los ingredientes. Disolver en relación 1:20 en agua para preparar buffer SB 1x.

### 1.7 TG Salt

#### Por litro:

MgCl <sub>2</sub>	0.57 g
CaCl <sub>2</sub>	8.3 g
Glicerol 100%	150 mL

#### Ajustar pH 6.8

Disolver los ingredientes. Ajustar pH a 6.8 y esterilizar por filtración.

### **1.8 Solución I (Colony Blot)**

NaOH	0.5 M
NaCl	1.5 M

### **1.9 Solución II (Colony Blot)**

Tris-HCl pH 7.2	1 M
-----------------	-----

### **1.10 Solución III (Colony Blot)**

Tris-HCl	0.5 M
NaCl	1.5 M

### **1.11 Solución de pre-hibridación (Colony Blot)**

#### **Para 10 mL:**

SDS 10%	1 mL
NaCl	0.58 g

Calentar la solución para que el SDS se mezcle por completo.

### **1.12 Solución de lavado I (Colony Blot)**

#### **Para 10 mL:**

SSC 25x	0.8 mL
---------	--------

### **1.13 Solución de lavado II (Colony Blot)**

#### **Para 20 mL:**

SSC 25x	1.6 mL
SDS 10%	2 mL

### **1.14 Solución de lavado III (Colony Blot)**

**Para 10 mL:**

SSC 25x                      0.4 mL

### **1.15 Solución de lavado IV (Colony Blot)**

**Para 10 mL:**

SSC 25x                      0.2 mL

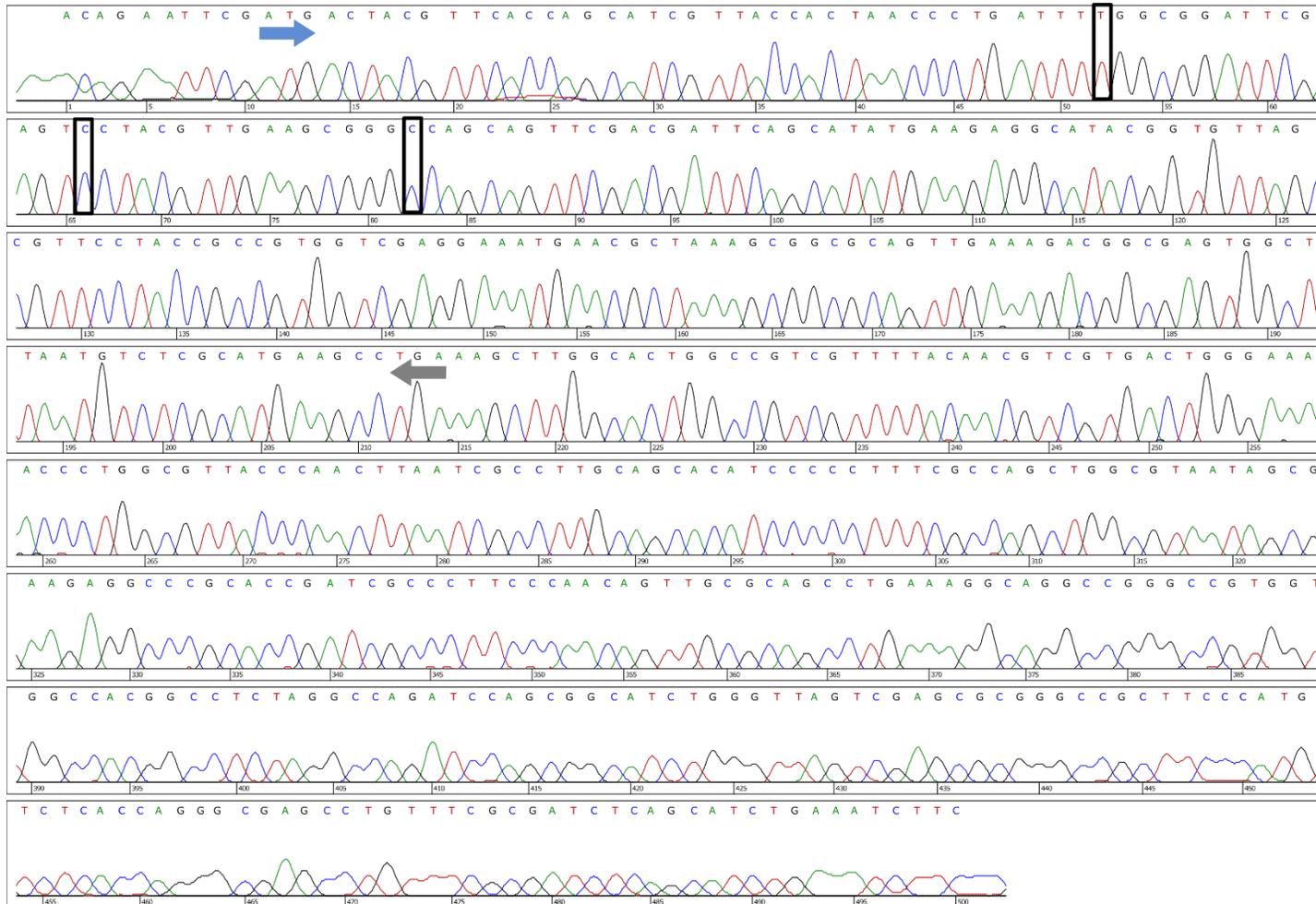
## ANEXO 2. Electroferogramas

Mut-A1

■ Start codon

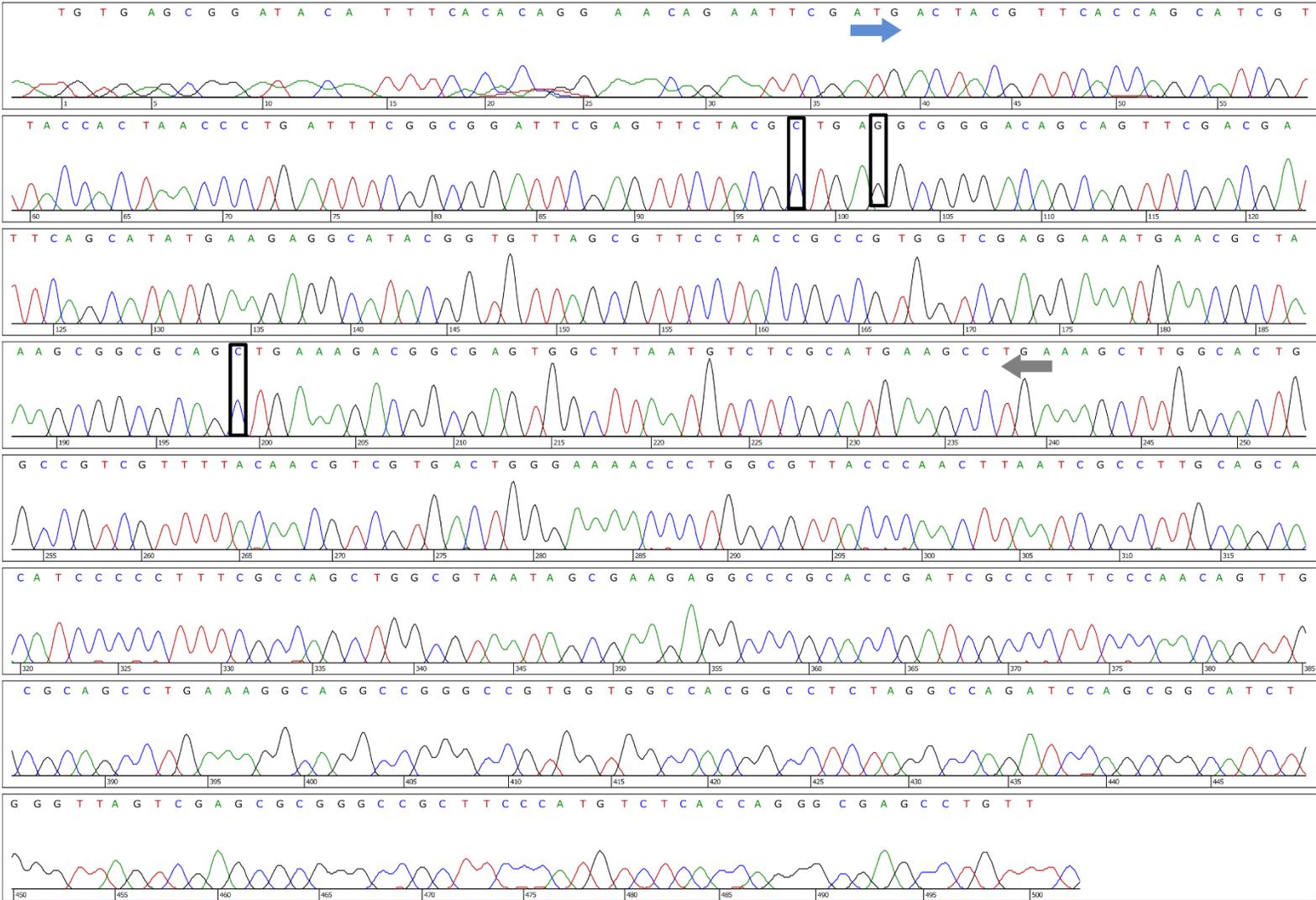
■ Stop codon

□ Mutation



# Mut-A2

■ Start codon    ■ Stop codon    □ Mutation

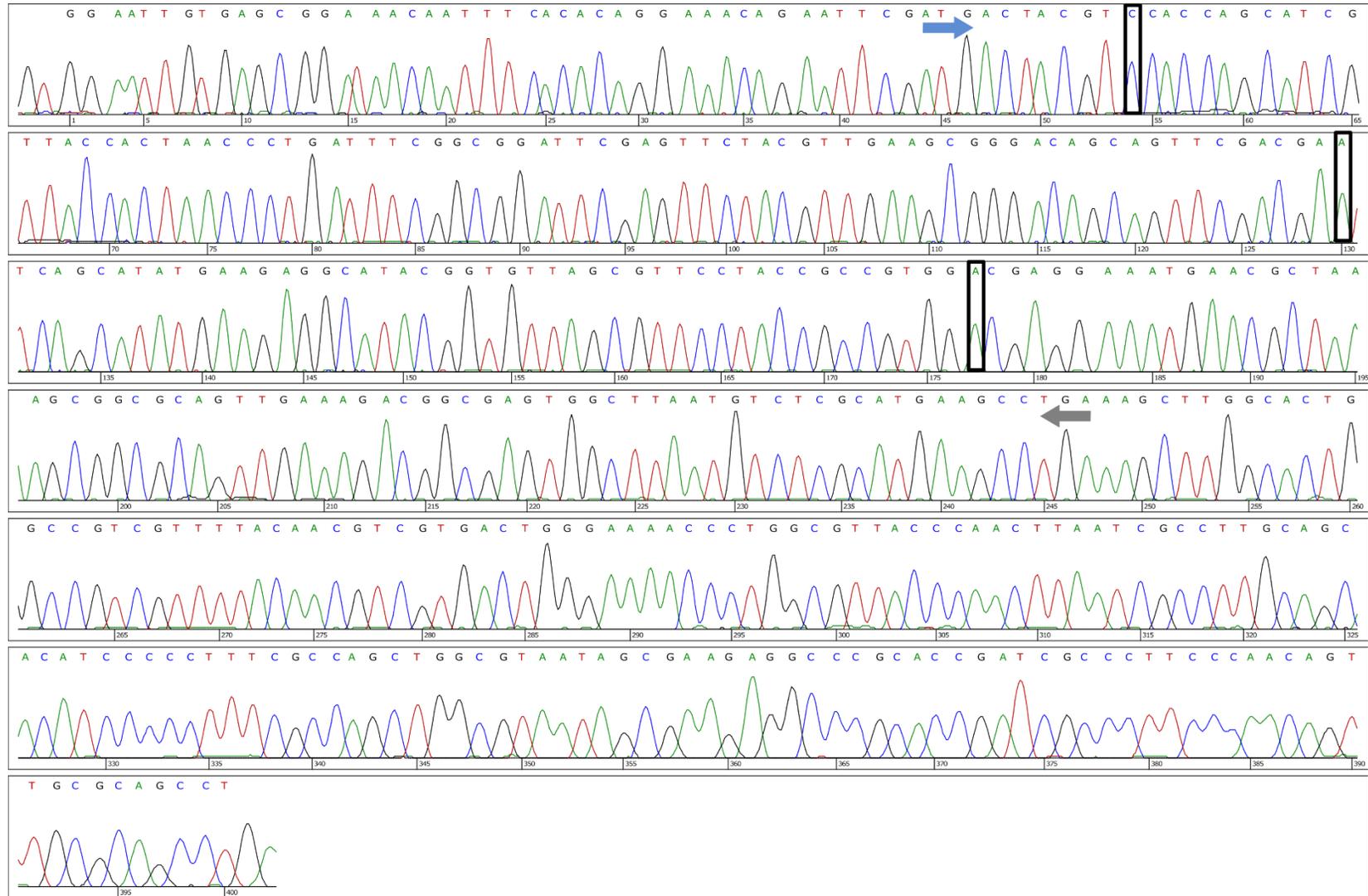


# Mut-A3

■ Start codon

■ Stop codon

□ Mutation

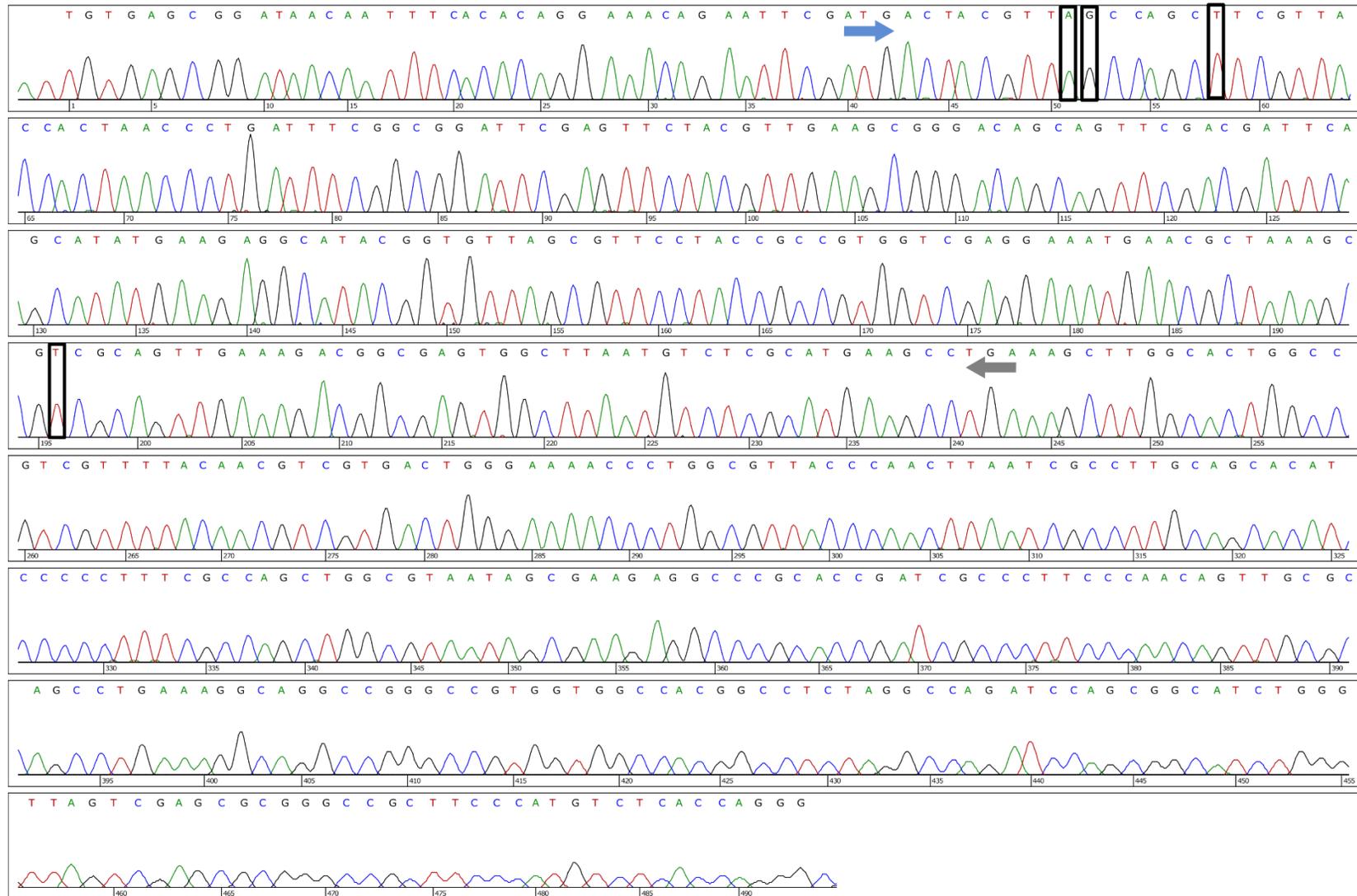


# Mut-A4

■ Start codon

■ Stop codon

□ Mutation

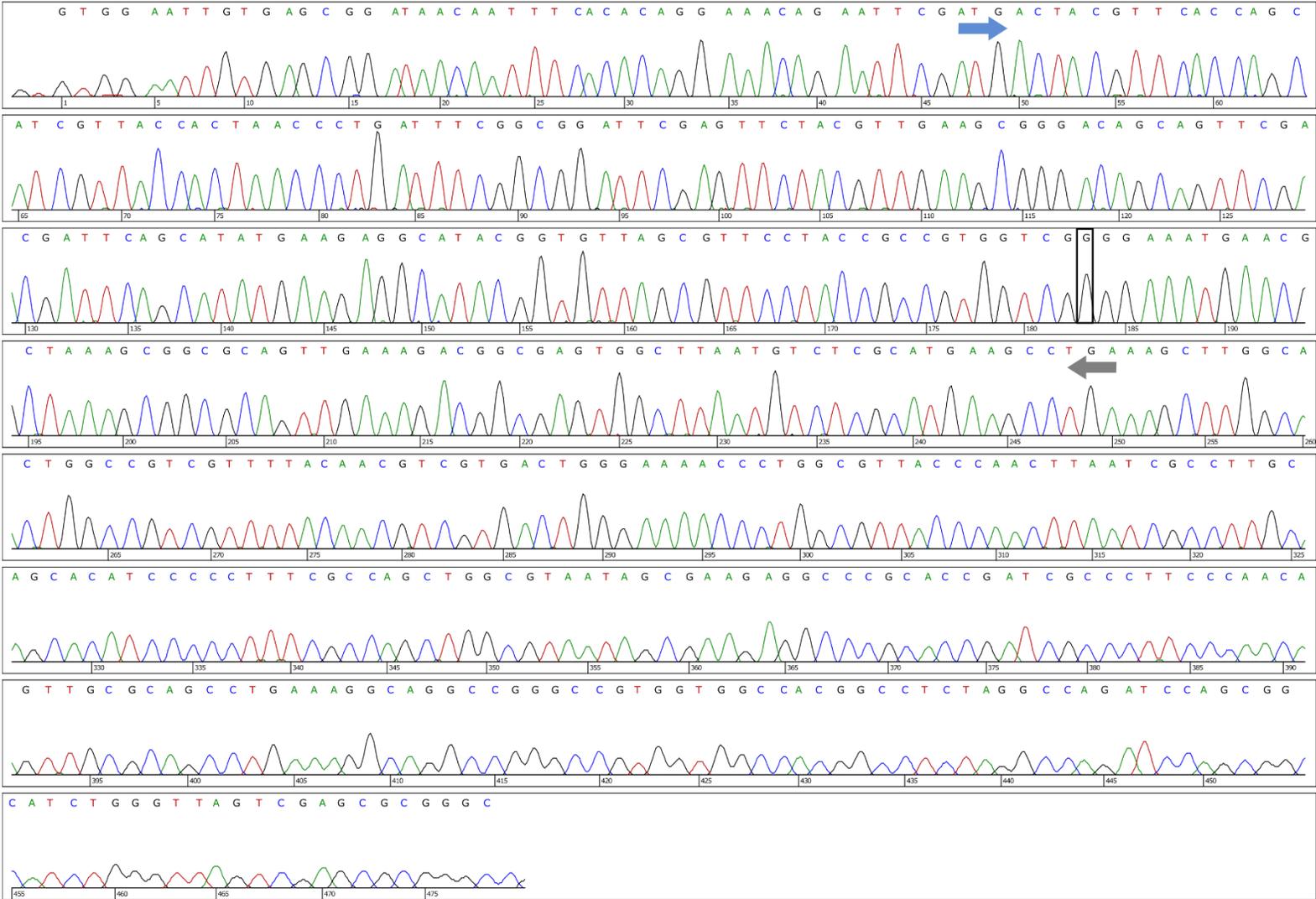


# Mut-A5

■ Start codon

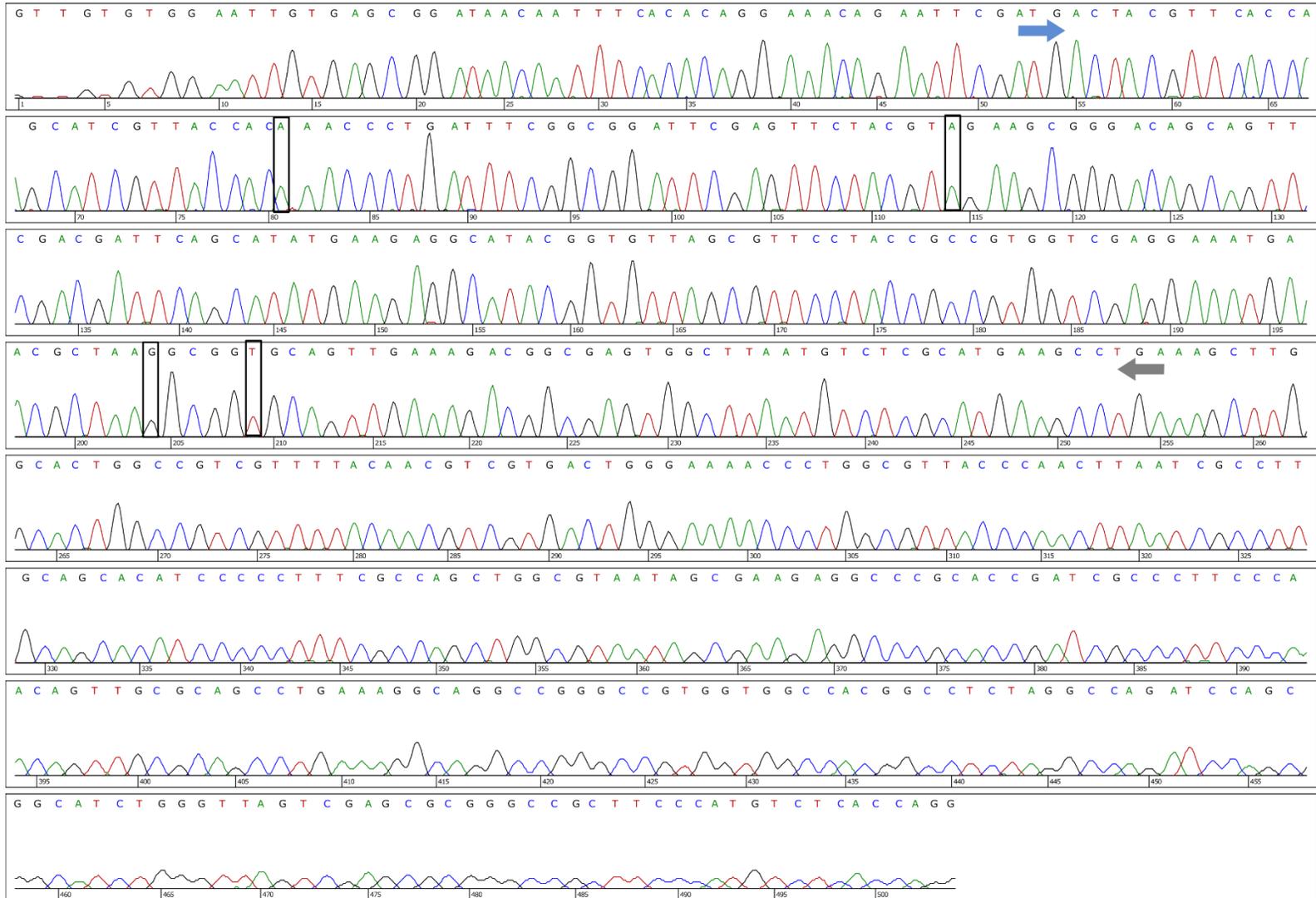
■ Stop codon

□ Mutation



# Mut-A6

■ Start codon    
 ■ Stop codon    
 □ Mutation

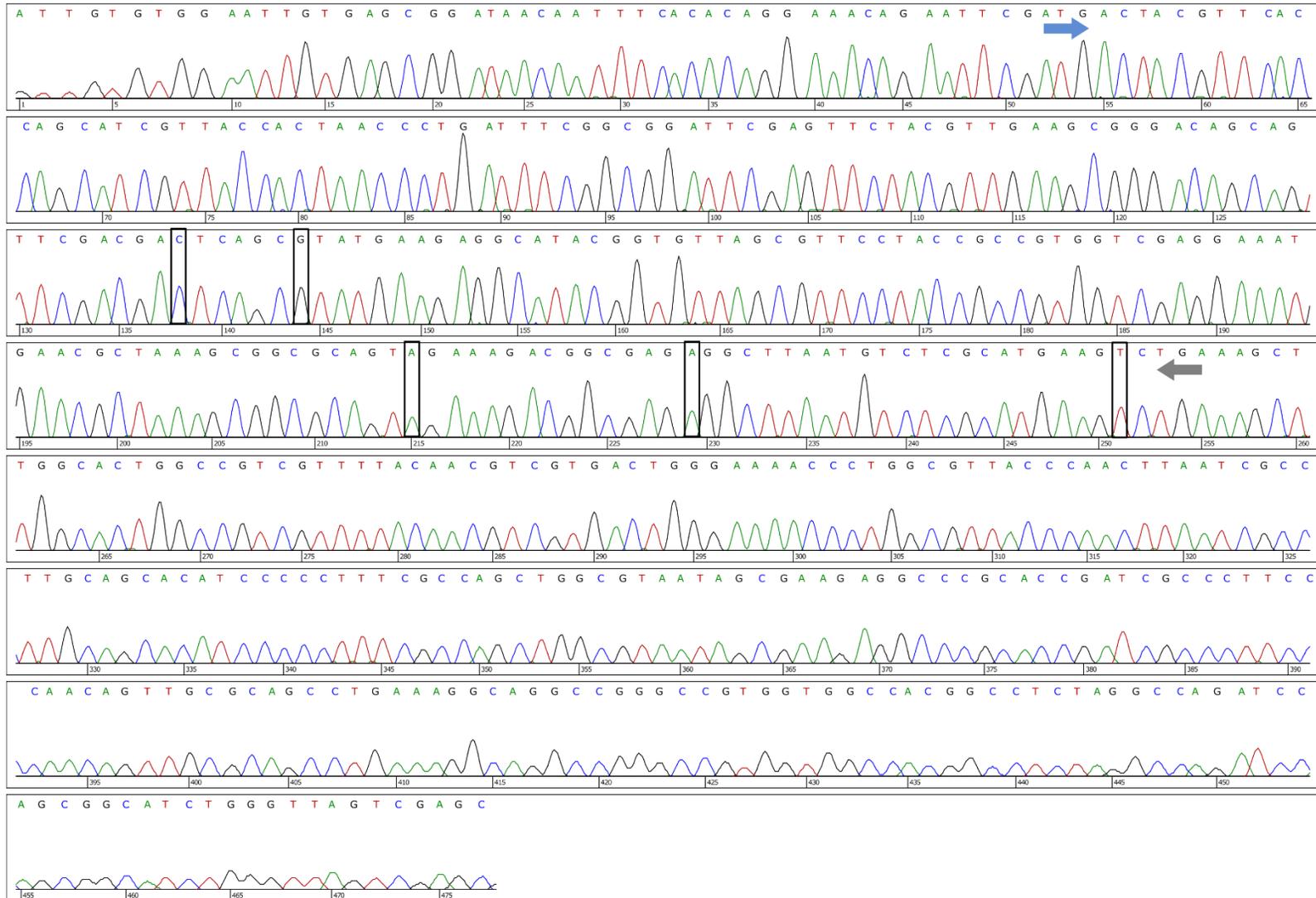


# Mut-A7

■ Start codon

■ Stop codon

□ Mutation

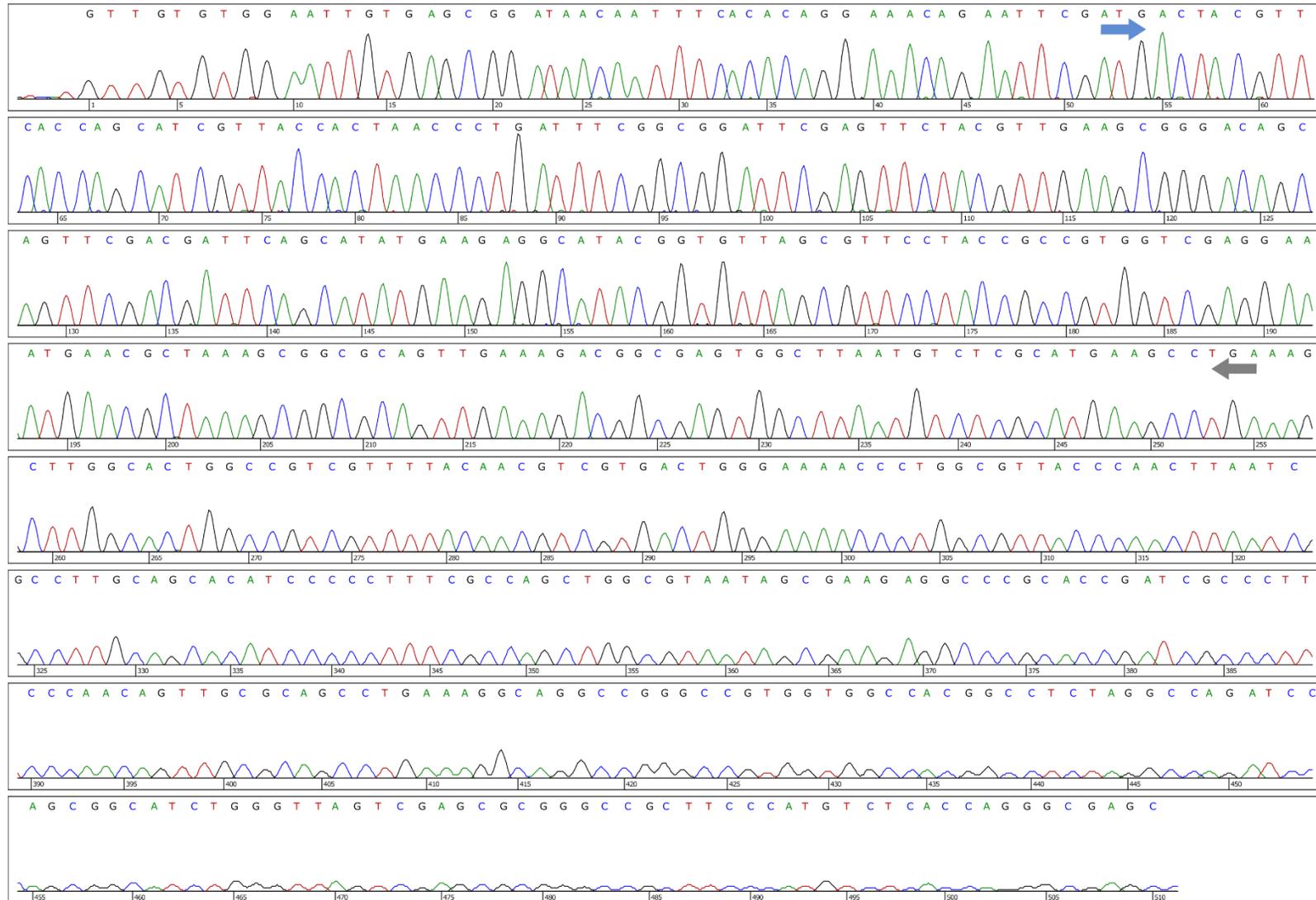


# Mut-A8

■ Start codon

■ Stop codon

□ Mutation

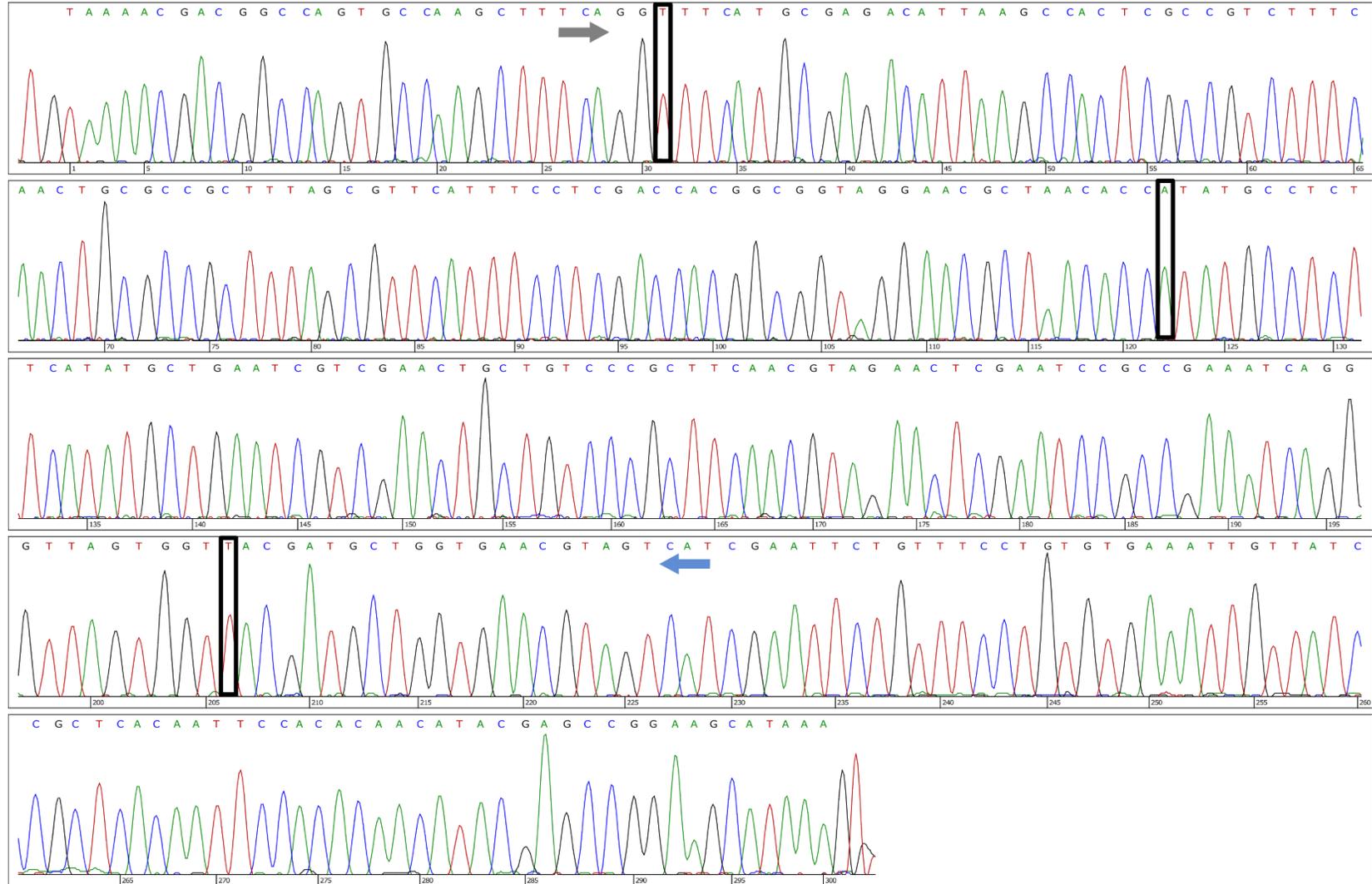


# Mut-A9

■ Start codon

■ Stop codon

□ Mutation

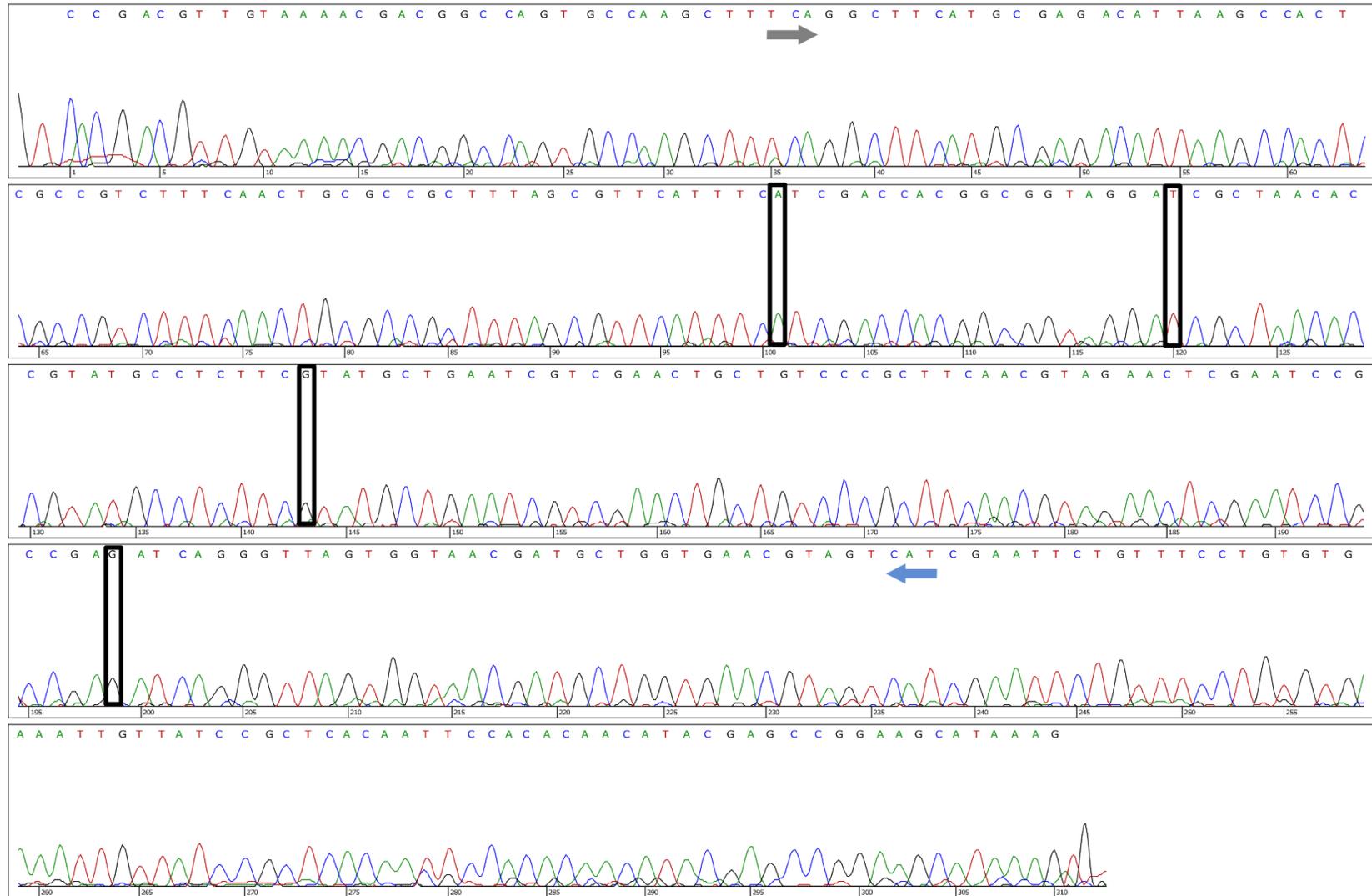


# Mut-A10

■ Start codon

■ Stop codon

□ Mutation

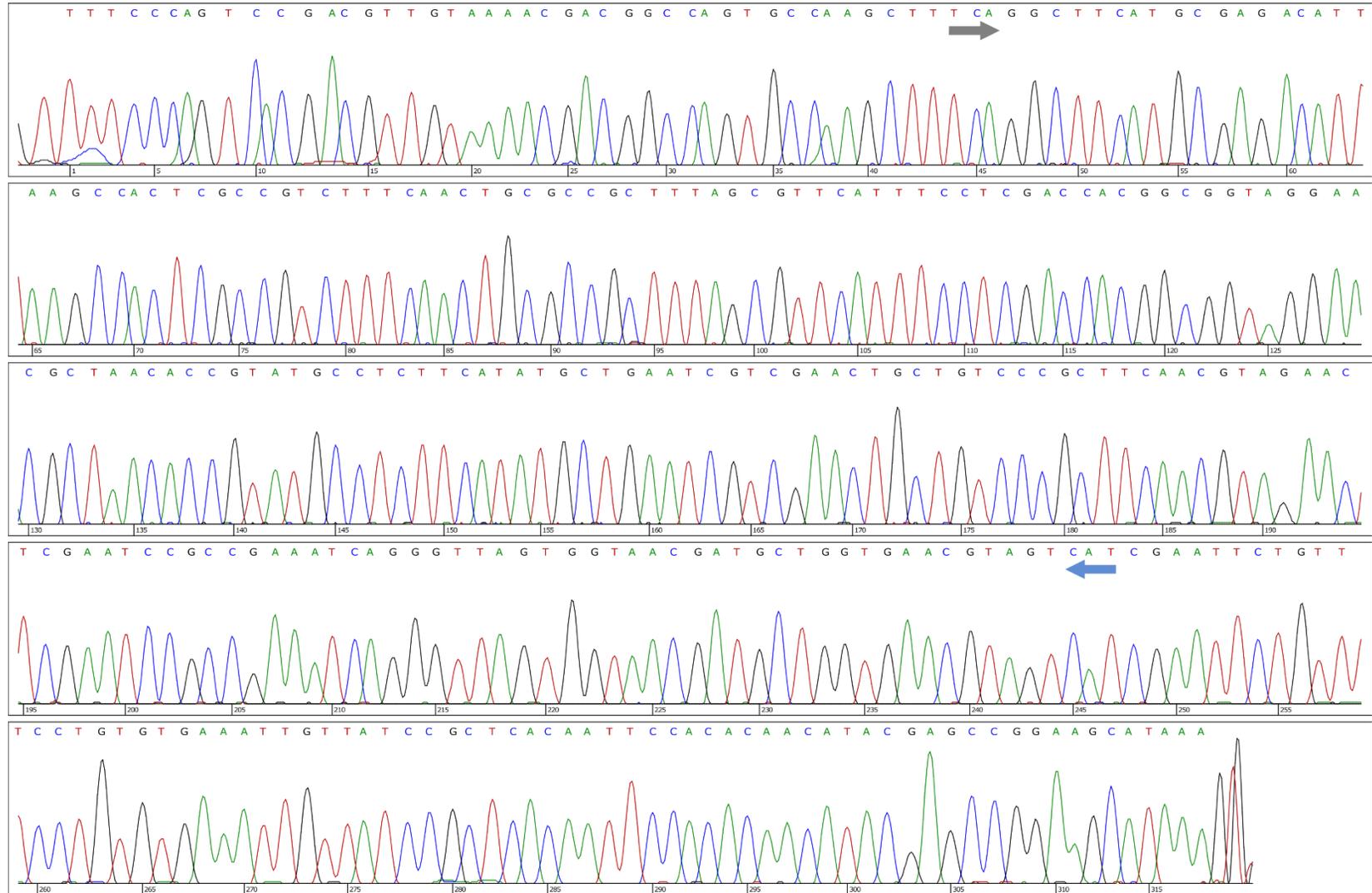


# Mut-B3

■ Start codon

■ Stop codon

□ Mutation

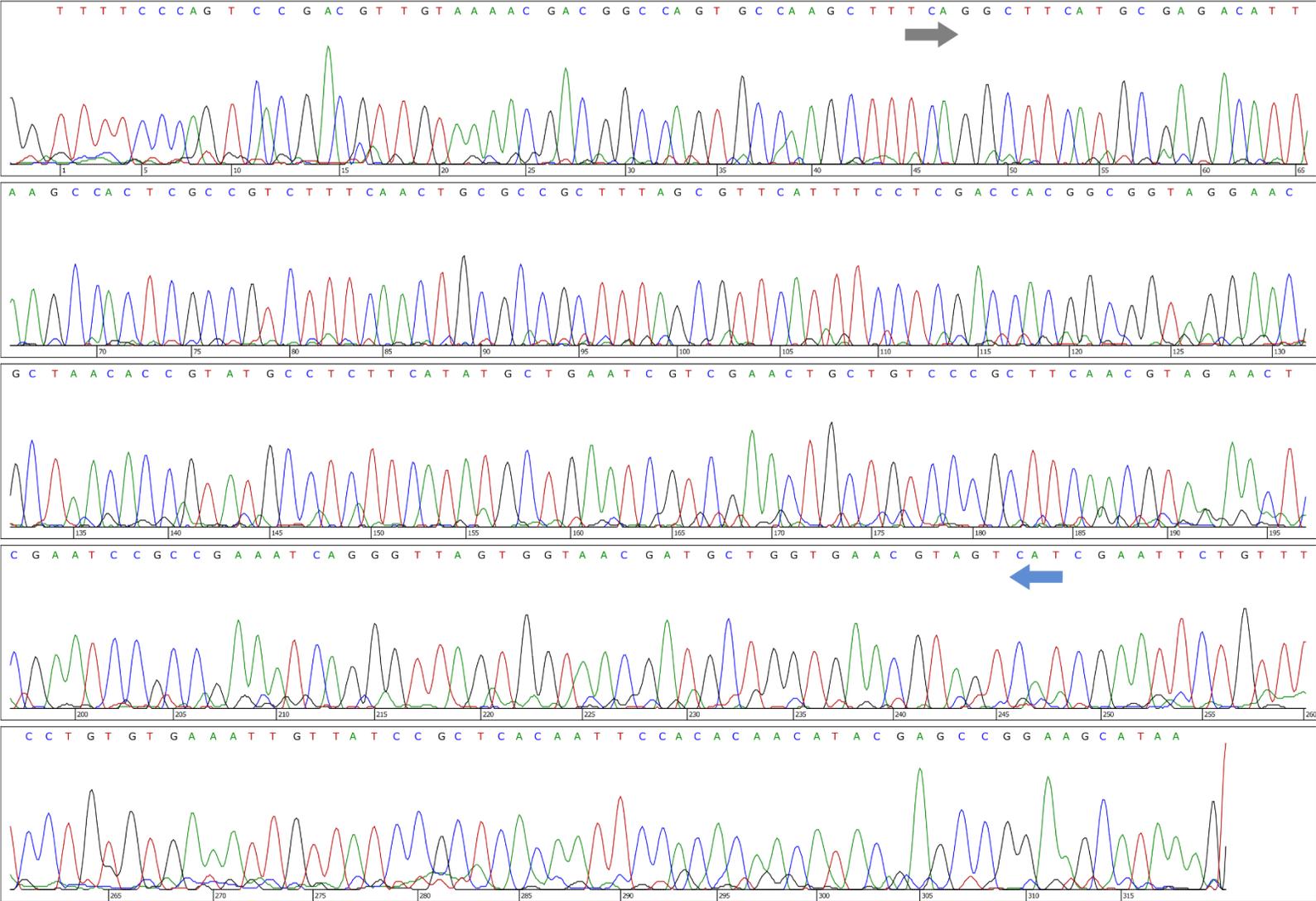


# Mut-B4

■ Start codon

■ Stop codon

□ Mutation

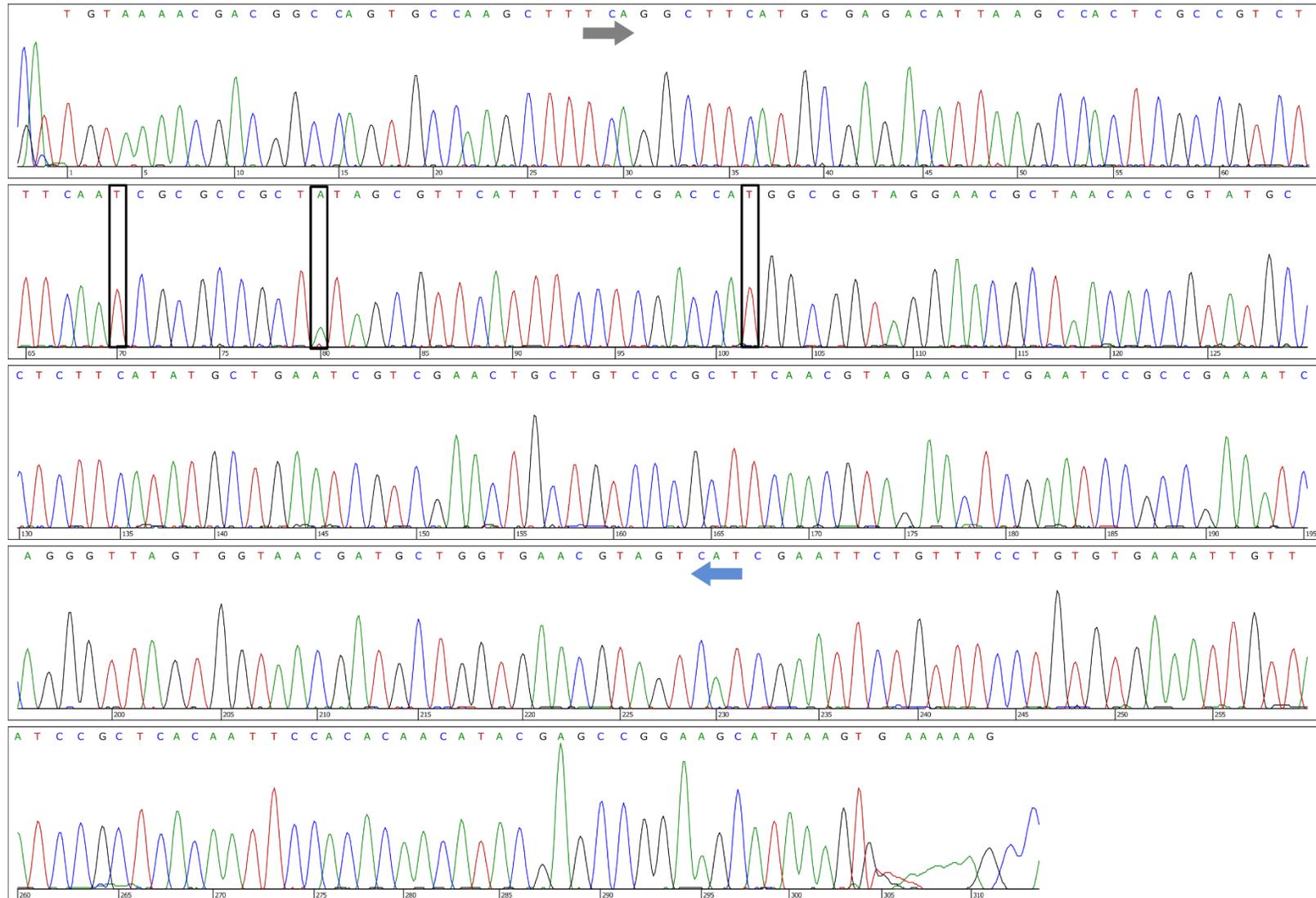


# Mut-B12

■ Start codon

■ Stop codon

□ Mutation

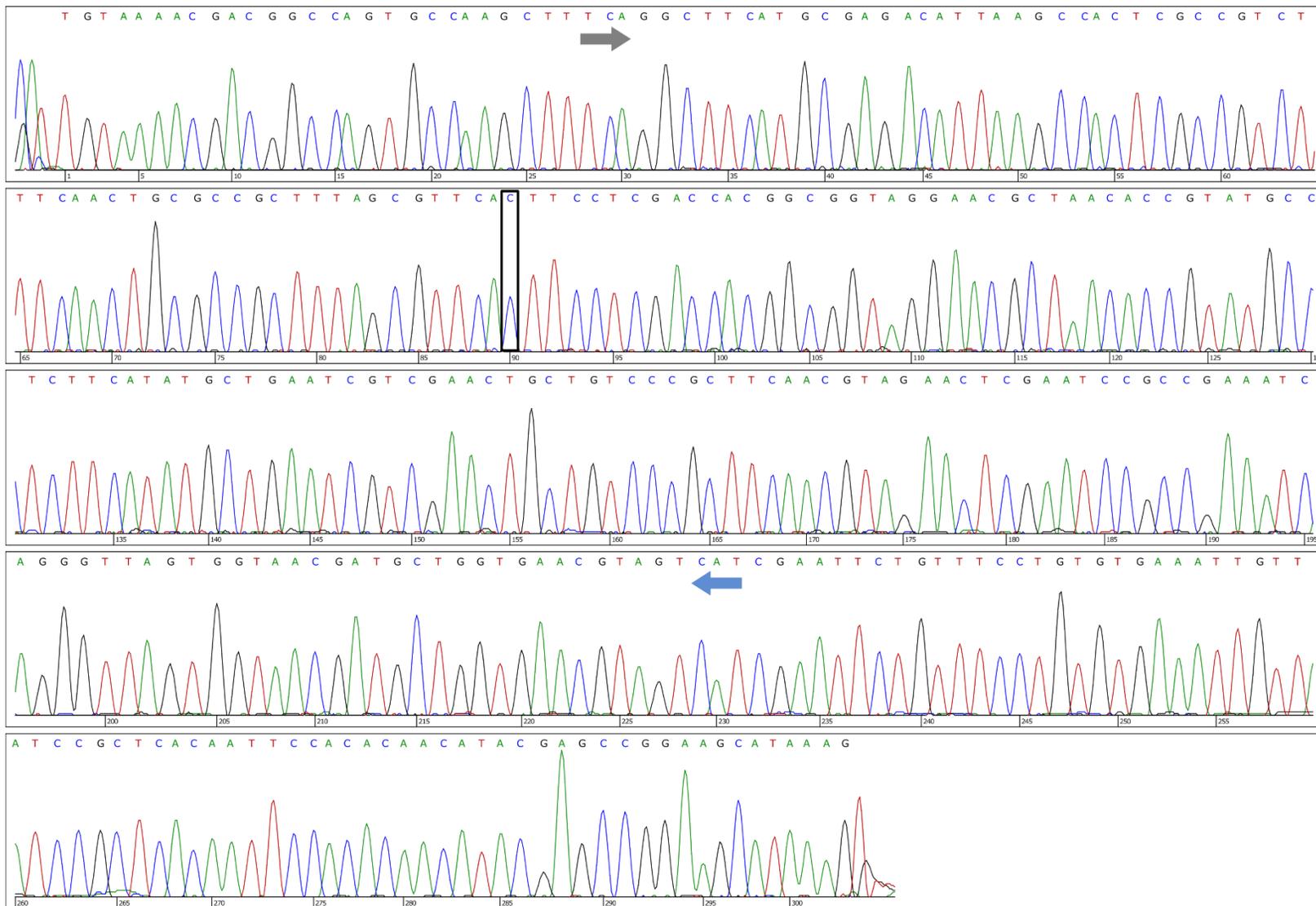


# Mut-C7

■ Start codon

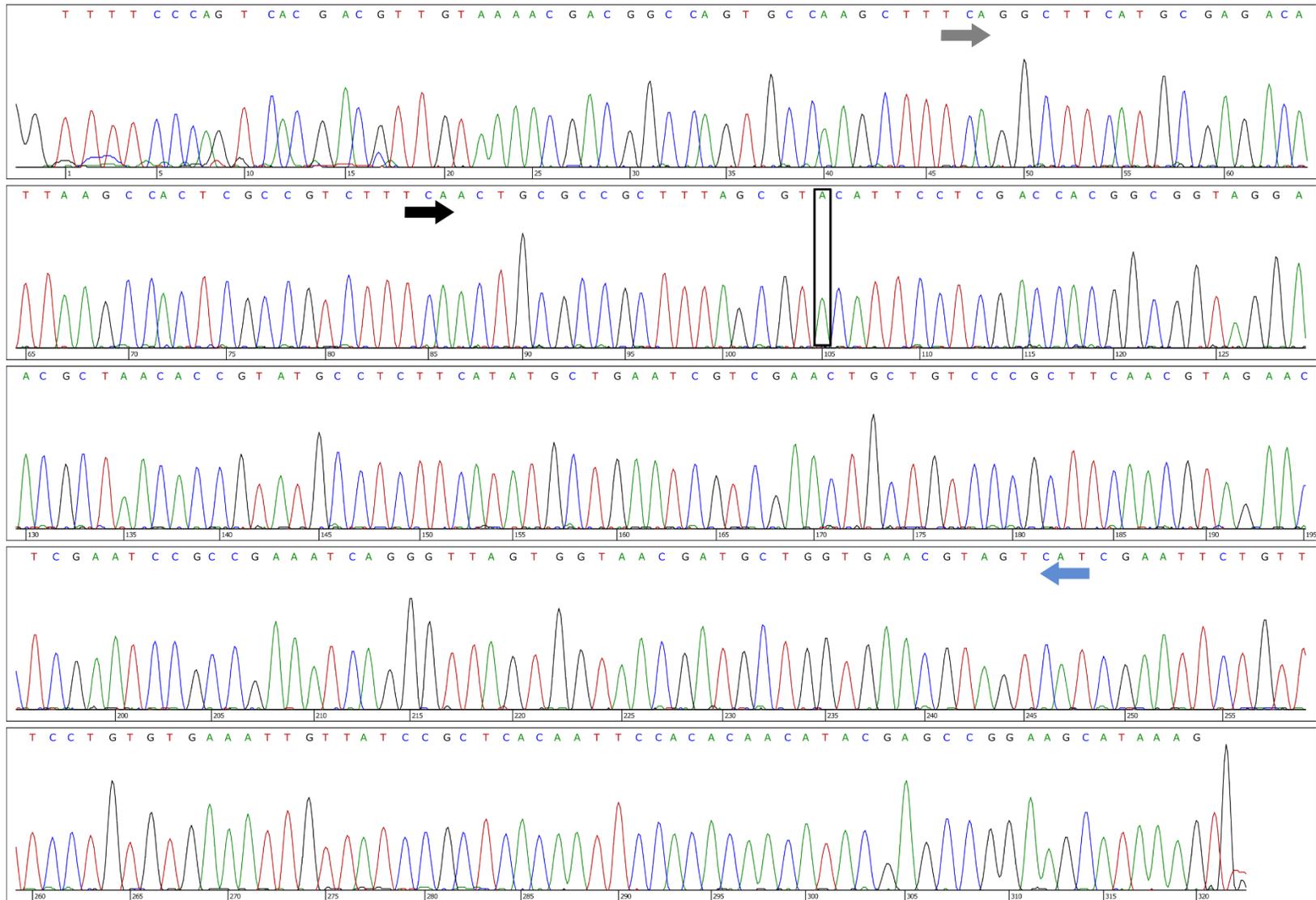
■ Stop codon

□ Mutation



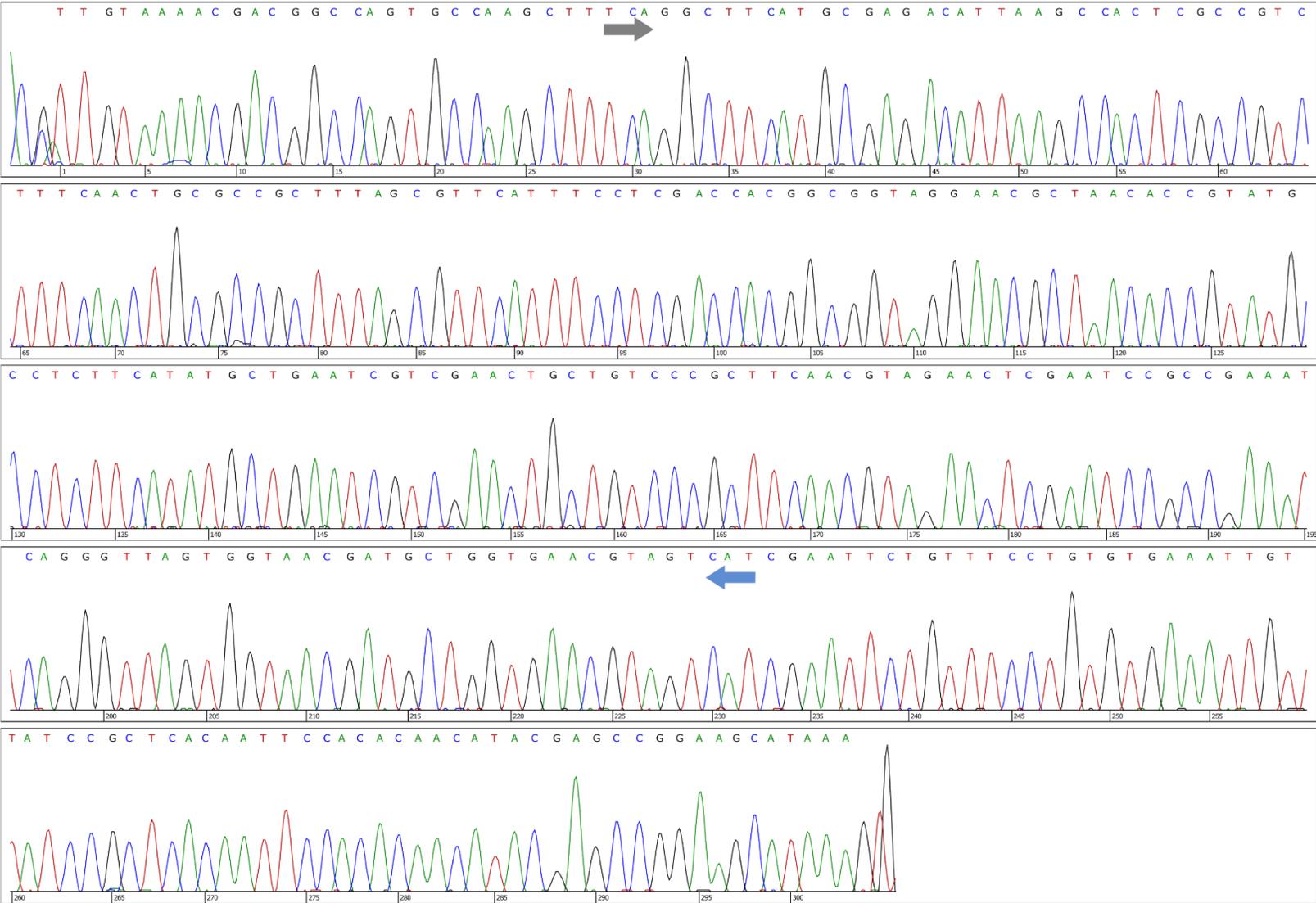
# Mut-C12

■ Start codon   
 ■ Original Stop codon   
 ■ New Generated Stop codon   
 □ Mutation



# Y32-B1

■ Start codon    ■ Stop codon    □ Mutation

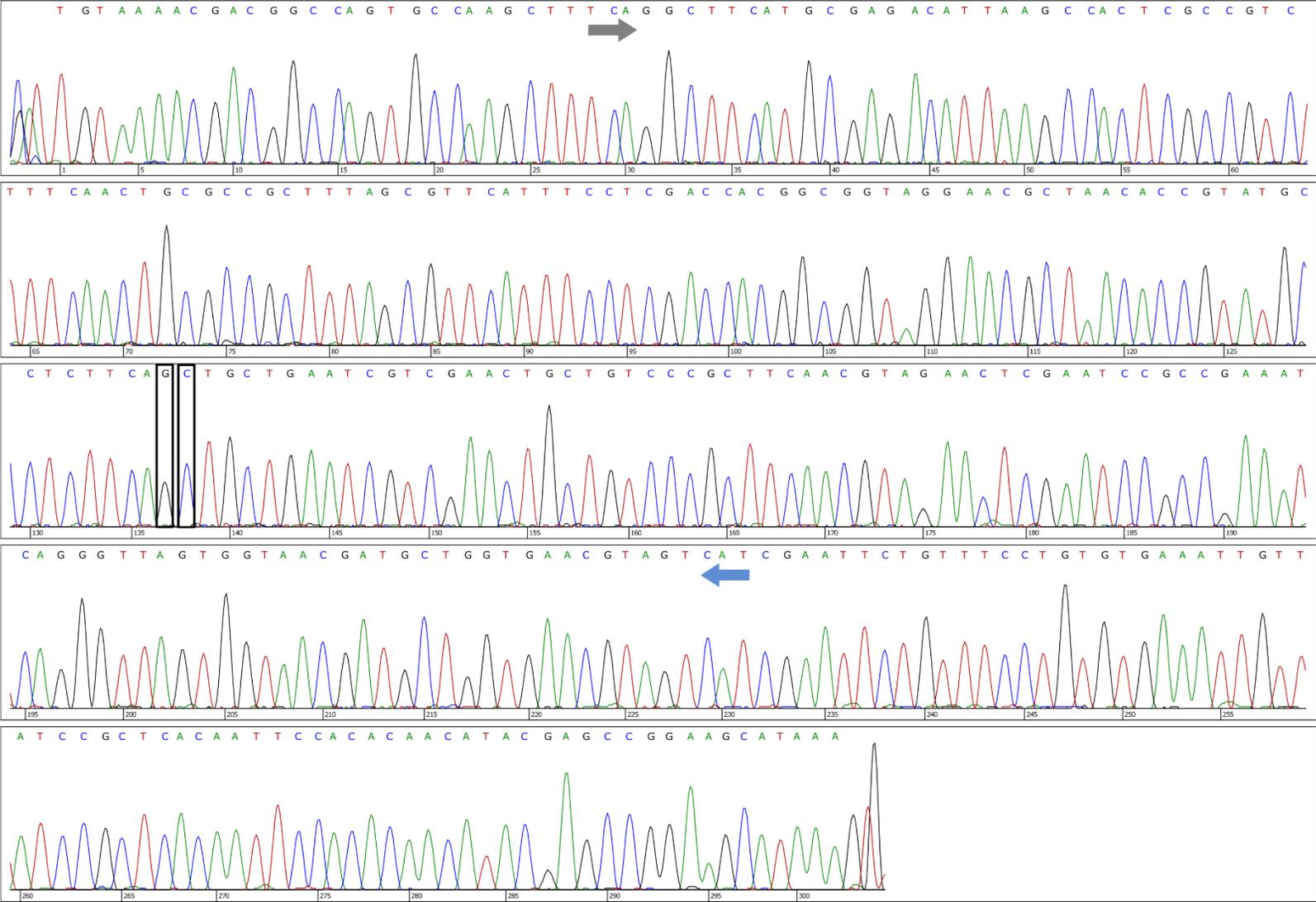


Y32-C2

■ Start codon

■ Stop codon

□ Mutation



Y32-D5

■ Start codon

■ Stop codon

□ Mutation

