UNIVERSIDAD ORT URUGUAY

Facultad de Ingeniería | Licenciatura en Bioinformática

Concurso de Bioinformática

Dirigido a	Estudiantes que en 2026 cursan 3.° de EMS, o que lo cursaron en 2025
Modalidad	Individual • 100 % remoto • Tutor docente opcional
Herramienta	Excel (no se requiere saber programar)
Entrega	Planilla Excel + video de hasta 5 minutos
Fecha límite	Lunes 7 de diciembre de 2026, 23:59 h
Premios	Becas del 80 %, 50 % y 50 % para la Licenciatura en Bioinformática

Contacto

¿Dudas sobre el concurso o los conceptos? Escribinos:

Nombre	Correo
Ignacio Ferres	ferres_i@ort.edu.uy
Josefina Vera	vera@ort.edu.uy
Felipe Machado	machado@ort.edu.uy

Índice

Parte I El concurso

1. De qué se trata

Parte II Conceptos de biología molecular

2. Lo que necesitás saber de biología molecular
3. Proteínas, aminoácidos y cómo compararlas

Parte III El desafío

4. El desafío paso a paso
5. Qué entregar y cómo te evalúan

Anexo Material de consulta

6. Tabla del código genético, glosario y checklist

Contacto

PARTE I

El concurso

1. De qué se trata

Vas a armar una planilla de Excel que traduzca cinco secuencias de ADN a proteínas, las compare con una proteína de referencia de *Escherichia coli* y diga cuál es la más parecida. La entrega vence el **lunes 7 de diciembre de 2026 a las 23:59**.

El problema

Una empresa biotecnológica usa una cepa de *E. coli* que produce una enzima clave para su proceso productivo. Al escalar la producción, la cepa empezó a fallar. El equipo de I+D tiene cinco secuencias de ADN de otras bacterias, sin función conocida, y necesita saber cuál podría reemplazar a la enzima original antes de gastar en ensayos de laboratorio.

Tu rol

Hacer de bioinformático: analizar las secuencias con una herramienta propia y proponer la mejor candidata. La bioinformática no reemplaza al laboratorio; genera una hipótesis bien fundamentada que después se prueba, y así ahorra tiempo y dinero.

PARTE II

Conceptos de biología molecular

2. Lo que necesitás saber de biología molecular

Toda la planilla se apoya en una idea: el ADN se lee de a tres letras, y cada grupo de tres indica un aminoácido de la proteína.

¿Qué es la bioinformática?

Es la unión de biología, informática y matemática para guardar, analizar e interpretar grandes volúmenes de datos biológicos: secuencias de ADN, secuencias de aminoácidos y estructuras 3D de proteínas. Algunos ejemplos:

- Seguir las variantes del coronavirus en 2020 secuenciando genomas virales.
- Encontrar genes de resistencia a la sequía para mejorar cultivos.
- Probar en la computadora si una molécula puede unirse a una proteína y servir como fármaco, antes de ir al laboratorio.

El ADN

El ADN está formado por dos hebras enrolladas en doble hélice. Cada hebra es una cadena de **nucleótidos**: un grupo fosfato, un azúcar (desoxirribosa) y una **base nitrogenada**. Hay cuatro bases: **A** (adenina), **T** (timina), **C** (citosina) y **G** (guanina).

Las dos hebras son complementarias: A se aparea con T y C con G. Por eso alcanza con conocer una hebra, y los archivos de secuencias (formato FASTA) muestran solo una. El genoma humano tiene unos 3.000 millones de pares de bases; el orden de esas letras es la información.

Para la computadora, una secuencia de ADN es simplemente un texto: `ATGGCTAAAGGT...`. Buscar a ojo una letra cambiada en millones es imposible; para eso existen las herramientas bioinformáticas.

Genes: las recetas

Pensá el ADN como un **recetario** y cada **gen** como una receta: un tramo de ADN con la información para fabricar una molécula, que en este desafío siempre es una proteína. Un gen tiene tres partes:

1. **Promotor**: donde se une la ARN polimerasa para empezar a copiar.
2. **Región codificante**: lo que después se traduce a proteína. Es la parte con la que vas a trabajar.
3. **Terminador**: donde termina la copia.

El dogma central: ADN → ARN → proteína

La receta original no sale del recetario. La célula hace un borrador y cocina con él:

1. **Replicación**: el ADN se copia a sí mismo cuando la célula se divide.
2. **Transcripción**: la ARN polimerasa copia el gen a **ARN mensajero** (ARNm), el borrador. En el ARN la T se reemplaza por **U** (uracilo).
3. **Traducción**: en el ribosoma, el ARNm se lee y se arma la proteína, la “torta” final.



Tu planilla va a simular el último paso: leer la secuencia de a tres letras y escribir la proteína. Como vas a trabajar con ADN, vas a buscar **T** donde los libros de ARN dicen **U**: el codón de inicio AUG del ARN es **ATG** en tu secuencia.

Codones y código genético

Un **codón** es un grupo de tres nucleótidos seguidos. El ribosoma lee el ARNm de a un codón por vez, y cada codón corresponde a un aminoácido (por ejemplo, ATG → metionina, CCT → prolina). La tabla que dice qué codón va con qué aminoácido es el **código genético** (está completa en el Anexo). Tiene tres características:

- **Se lee en tripletes**, sin saltear ni superponer letras, desde el codón de inicio.
- **Es degenerado** (redundante): hay 64 codones y solo 20 aminoácidos, así que varios codones dan el mismo. La leucina, por ejemplo, tiene seis: CTT, CTC, CTA, CTG, TTA y TTG.
- **Es universal**: casi todos los seres vivos usan el mismo código. Lo que cambia entre especies es qué codones prefieren para cada aminoácido.

Inicio, stop y marco abierto de lectura (ORF)

- **Codón de inicio**: **ATG** (metionina). Ahí empieza la proteína.
- **Codones stop**: **TAA, TAG y TGA**. No codifican ningún aminoácido; indican que la proteína terminó.
- **ORF** (*open reading frame*, marco abierto de lectura): el tramo que va desde el codón de inicio hasta el primer stop en el mismo marco. Es la parte de la secuencia que realmente se traduce.

> **Ejemplo**
> En `CCATGGCTTGTAAATAGGC`, el ORF empieza en el primer ATG y se lee así: ATG · GCT · TGT · AAA · TAG. Eso da **M–A–C–K** y el TAG frena la traducción.

Mutaciones: por qué las proteínas parecidas no son idénticas

Las cinco secuencias del desafío son parecidas a la de referencia, pero con cambios. Esos cambios tienen efectos distintos:

Tipo de mutación	Qué pasa	Ejemplo
Silenciosa	Cambia el codón pero no el aminoácido	`GCT → GCC`: los dos son alanina
De cambio de sentido	Cambia un aminoácido; puede afectar o no la función	`GCT → ACT`: alanina → treonina
Sin sentido	Aparece un stop antes de tiempo y la proteína queda corta	`AAA → TAA`: lisina → stop
Inserción o deleción	Entra o sale una base y se corre el marco de lectura; todo lo que sigue se lee distinto	`ATG GCT TGT → ATG GCA TTG T…`

Un cambio de aminoácido importa más si ocurre en una zona clave (por ejemplo, el sitio activo de una enzima) o si el nuevo aminoácido tiene propiedades muy distintas. Por eso, además de contar coincidencias, conviene mirar las propiedades de las proteínas (sección 3).

3. Proteínas, aminoácidos y cómo compararlas

Una proteína es una cadena de aminoácidos; su secuencia y las propiedades de esos aminoácidos determinan su forma y su función.

¿Qué es una proteína?

Es una molécula formada por **aminoácidos** unidos en cadena por enlaces peptídicos (por eso se la llama cadena polipeptídica). La cadena se pliega en una forma 3D, y de esa forma depende lo que hace: las **enzimas**, como la del desafío, son proteínas que aceleran reacciones químicas.

En bioinformática cada aminoácido se escribe con **una letra**, así que una proteína también es un texto: `MAKGT...`.

Los 20 aminoácidos y sus grupos

Cada aminoácido tiene una cadena lateral distinta. Según cómo se comporta a pH 7 (el pH de referencia que vas a usar), se agrupan así:

Grupo	Qué significa	Aminoácidos (letra)
Hidrofóbicos (no polares)	Repelen el agua; suelen quedar en el interior de la proteína	Glicina (G), Alanina (A), Valina (V), Leucina (L), Isoleucina (I), Metionina (M), Fenilalanina (F), Triptófano (W), Prolina (P)
Polares sin carga	Se llevan bien con el agua pero no tienen carga neta	Serina (S), Treonina (T), Cisteína (C), Asparagina (N), Glutamina (Q), Tirosina (Y)
Con carga positiva (básicos)	Cada uno suma +1 a la carga	Lisina (K), Arginina (R), Histidina (H)*
Con carga negativa (ácidos)	Cada uno suma −1 a la carga	Aspartato (D), Glutamato (E)

**A pH 7 la histidina está casi siempre sin carga, por eso muchas calculadoras le asignan 0 o un valor muy chico (+0,1). Los libros no siempre coinciden en dónde ubicar la glicina, el triptófano o la tirosina: elegí una clasificación, dejá escrito cuál usaste y por qué.*

Propiedades que vas a calcular

Con la secuencia de aminoácidos podés describir cada proteína con algunos números:

- **Longitud**: cuántos aminoácidos tiene, contando desde la metionina inicial hasta el último antes del stop.
- **Composición**: cuántos aminoácidos de cada tipo hay, en número y en porcentaje del total.
- **Carga neta aproximada a pH 7**: positivos menos negativos.
- **Porcentaje de hidrofóbicos**: cuántos aminoácidos hidrofóbicos hay sobre el total.

La carga aproximada se calcula así:

$$\text{Carga} \approx (\text{nº de K} + \text{nº de R}) - (\text{nº de D} + \text{nº de E})$$

Si dos proteínas tienen función parecida, esperamos que estos valores también se parezcan. Una candidata mucho más corta (por un stop prematuro) o con la carga invertida es sospechosa, aunque comparta muchos aminoácidos.

Porcentaje de identidad

Es la medida principal para comparar dos secuencias: se ponen una debajo de la otra, posición por posición, y se cuenta cuántas coinciden.

$$\%\ \text{identidad} = (\text{posiciones que coinciden} \div \text{longitud del alineamiento}) \times 100$$

Ejemplo con dos proteínas cortas:

Posición	1	2	3	4	5	6	7	8
Referencia	M	K	T	A	Y	I	A	K
Candidata	M	K	S	A	Y	L	A	K
¿Coincide?	sí	sí	no	sí	sí	no	sí	sí

Coinciden 6 de 8 posiciones: 6 ÷ 8 × 100 = **75 % de identidad**.

Cuanto más alta la identidad con la referencia, más probable es que la proteína tenga la misma función. Así se anotan genes nuevos en la investigación real: se comparan con bases de datos de proteínas conocidas. Igual, es una hipótesis que después se confirma en el laboratorio.

Dos detalles para pensar en tu análisis

- **Longitudes distintas.** Si las proteínas no miden lo mismo, tenés que decidir sobre qué longitud dividís (la de la referencia, la más larga o solo las posiciones comparadas). Cualquiera puede ser válida si explicás tu elección.
- **No todas las diferencias pesan igual.** Cambiar I por L (dos hidrofóbicas parecidas) es menos grave que cambiar K por E (de positiva a negativa). Mirar esto suma al razonamiento bioinformático.

PARTE III

El desafío

4. El desafío paso a paso

Tu planilla tiene que hacer sola lo que hace una célula: pegar una secuencia de ADN en una celda y obtener la proteína, sus propiedades y la comparación con la referencia, sin pasos manuales.

Los materiales

La organización te entrega un archivo con seis secuencias en formato **FASTA**: la secuencia de referencia de *E. coli* y las cinco secuencias problema. Cada una tiene dos partes:

```
>Secuencia_problema_1
ATGAAACGTCTGGCT...
```

La línea que empieza con `>` es solo el **nombre** (un título); no es parte de la secuencia. Lo que copiás en tu planilla es lo que viene debajo.

Lo que tiene que hacer tu planilla

1. **Recibir la secuencia.** Una celda de entrada donde se pega el ADN. Conviene limpiarla: pasar todo a mayúsculas y quitar espacios o saltos de línea.
2. **Encontrar el inicio.** Ubicar el primer ATG y quedarse con la secuencia desde ahí.
3. **Separar en codones.** Cortar esa secuencia de a tres letras, idealmente un codón por fila.
4. **Traducir.** Buscar cada codón en una tabla del código genético (podés armarla en otra hoja con la tabla del Anexo) y escribir al lado su aminoácido.
5. **Cortar en el stop.** Detectar el primer codón stop y descartar todo lo que viene después.
6. **Armar la proteína.** Unir los aminoácidos en un solo texto, por ejemplo `MKTAYIAK...`.
7. **Calcular propiedades.** Longitud, composición, carga aproximada y porcentaje de hidrofóbicos (sección 3).
8. **Comparar.** Calcular el porcentaje de identidad de cada candidata contra la referencia y, si podés, de las candidatas entre sí (por ejemplo, en una tabla de 6 × 6).
9. **Responder.** Una celda o recuadro visible que diga qué secuencia elegiste y por qué.

Una buena prueba: pegá una secuencia nueva en la celda de entrada. Si todo se recalcula solo, tu herramienta está automatizada.

Pistas de Excel

No hay una única forma correcta. Estas funciones te pueden servir (nombres de Excel en español; en Uruguay los argumentos se separan con punto y coma):

Para…	Funciones útiles
Limpiar el texto	`MAYUSC`, `SUSTITUIR`, `ESPACIOS`
Encontrar el ATG	`HALLAR` o `ENCONTRAR`
Recortar texto	`EXTRAE`, `IZQUIERDA`, `DERECHA`, `LARGO`
Generar posiciones 1, 4, 7…	`SECUENCIA` (Excel 365) o `FILA`
Traducir cada codón	`BUSCARV` o `BUSCARX` contra tu tabla del código genético
Encontrar el primer stop	`COINCIDIR`, `SI`, `SI.ERROR`
Unir los aminoácidos	`CONCAT` o `UNIRCADENAS`

Para…	Funciones útiles
Contar aminoácidos	`CONTAR.SI` sobre la columna, o `LARGO` menos `LARGO(SUSTITUIR(...))` sobre el texto
Comparar posición por posición	`EXTRAE` de a una letra y `SUMAPRODUCTO`

Podés apoyarte en tutoriales o herramientas de IA para entender cómo funciona una fórmula, pero el trabajo tiene que ser tuyo y en el video vas a tener que explicar qué hace cada parte. Un docente de química, biología o informática de tu liceo también puede acompañarte como tutor.

Cómo elegir la mejor candidata

La identidad con la referencia es el criterio principal, pero no el único. Antes de decidir, mirá también:

- ¿La proteína tiene una longitud parecida a la de referencia, o se cortó antes por un stop?
- ¿Su carga y su porcentaje de hidrofóbicos se parecen a los de la referencia?
- Donde hay diferencias, ¿los aminoácidos cambiados son de propiedades parecidas?

> **Desafío integrador**
> Tu respuesta es más fuerte si combina varios de estos argumentos y explica por qué descartaste a las otras cuatro.

5. Qué entregar y cómo te evalúan

Entregás dos cosas, una planilla y un video de hasta 5 minutos, antes del **lunes 7 de diciembre de 2026 a las 23:59**.

1. La planilla Excel

- Permite ingresar secuencias de ADN y calcula automáticamente la información de las proteínas.
- Muestra los análisis de las cinco secuencias y la comparación con la referencia.
- Incluye una **respuesta explícita**: cuál de las cinco secuencias se parece más a la referencia y por qué.

No hace falta que sea compleja, pero sí que sea clara: celdas de entrada marcadas, títulos y resultados fáciles de encontrar.

2. El video (máximo 5 minutos)

1. Cómo hiciste el análisis.
2. Qué resultados obtuviste y qué tuviste en cuenta para comparar las proteínas.
3. Por qué la secuencia elegida es la mejor candidata.
4. Qué mejoras le podrías agregar a tu herramienta.
5. **Último minuto**: qué te motiva a participar y a estudiar la Licenciatura en Bioinformática.

No necesitás cámara profesional: podés grabar tu pantalla (por ejemplo, con Zoom) mientras mostrás la planilla. Explicá la idea general de cada fórmula (“esta celda busca el ATG y corta la secuencia desde ahí”), no cada detalle técnico.

Criterios de evaluación

Criterio	Peso	Qué se mira
Razonamiento bioinformático	30 %	Que entiendas los conceptos y justifiques la elección con datos
Implementación técnica en Excel	25 %	Que la planilla funcione, esté automatizada y sea clara
Comunicación y presentación	25 %	Que el video explique bien el proceso y los resultados

Criterio	Peso	Qué se mira
Motivación	20 %	Lo que contás en el último minuto del video

Quién puede participar

- Estudiantes que en 2026 cursan **3.° de EMS**, o que lo cursaron en 2025, de cualquier departamento del país. El concurso es 100 % remoto.
- Todas las orientaciones, **excepto** Creativo Artístico, General y Ciencias Sociales y Humanidades – UCA Derecho y Ciencias Políticas. También bachilleratos equivalentes de planes anteriores.
- La participación es individual; un docente del liceo puede acompañarte como tutor, pero no es obligatorio.
- El proyecto tiene que hacerse durante el período del concurso.
- No pueden participar familiares directos de funcionarios de la Facultad de Ingeniería de la Universidad ORT Uruguay.

Premios

Los tres mejores puntajes ganan becas para la **Licenciatura en Bioinformática** de la Facultad de Ingeniería de la Universidad ORT Uruguay, con inicio en marzo de 2027:

Puesto	Beca
1.°	80 %
2.°	50 %
3.°	50 %

Las becas son solo para esa carrera, se usan hasta el inicio del ciclo de marzo de 2027, y los ganadores no pueden haber cursado otra carrera en ORT. Los resultados se avisan por correo a todos los inscriptos y se publican en la web del concurso.

ANEXO

Material de consulta

6. Tabla del código genético (en ADN)

Esta es la tabla que tu planilla tiene que usar para traducir. Está escrita con **T** (ADN), no con U (ARN). La fila indica la primera base, la columna la segunda, y cada celda trae el codón completo. Podés copiarla a una hoja aparte con dos columnas (codón y aminoácido) para buscar desde ahí.

1.ª base	2.ª base: T	2.ª base: C	2.ª base: A	2.ª base: G
T	`TTT` Phe (F)	`TCT` Ser (S)	`TAT` Tyr (Y)	`TGT` Cys (C)
	`TTC` Phe (F)	`TCC` Ser (S)	`TAC` Tyr (Y)	`TGC` Cys (C)
	`TTA` Leu (L)	`TCA` Ser (S)	`TAA` **Stop**	`TGA` **Stop**
	`TTG` Leu (L)	`TCG` Ser (S)	`TAG` **Stop**	`TGG` Trp (W)
C	`CTT` Leu (L)	`CCT` Pro (P)	`CAT` His (H)	`CGT` Arg (R)
	`CTC` Leu (L)	`CCC` Pro (P)	`CAC` His (H)	`CGC` Arg (R)
	`CTA` Leu (L)	`CCA` Pro (P)	`CAA` Gln (Q)	`CGA` Arg (R)
	`CTG` Leu (L)	`CCG` Pro (P)	`CAG` Gln (Q)	`CGG` Arg (R)
A	`ATT` Ile (I)	`ACT` Thr (T)	`AAT` Asn (N)	`AGT` Ser (S)
	`ATC` Ile (I)	`ACC` Thr (T)	`AAC` Asn (N)	`AGC` Ser (S)
	`ATA` Ile (I)	`ACA` Thr (T)	`AAA` Lys (K)	`AGA` Arg (R)
	`ATG` **Met (M) · inicio**	`ACG` Thr (T)	`AAG` Lys (K)	`AGG` Arg (R)
G	`GTT` Val (V)	`GCT` Ala (A)	`GAT` Asp (D)	`GGT` Gly (G)
	`GTC` Val (V)	`GCC` Ala (A)	`GAC` Asp (D)	`GGC` Gly (G)
	`GTA` Val (V)	`GCA` Ala (A)	`GAA` Glu (E)	`GGA` Gly (G)
	`GTG` Val (V)	`GCG` Ala (A)	`GAG` Glu (E)	`GGG` Gly (G)

En rosado, los codones stop; en verde, el codón de inicio ATG.

Glosario

Término	Qué es
Aminoácido	Unidad que forma las proteínas. Hay 20, cada uno con una letra.
ARN mensajero (ARNm)	Copia de trabajo de un gen; lleva la información al ribosoma. Usa U en lugar de T.
Base nitrogenada	La “letra” de cada nucleótido: A, T, C o G (en ARN, U en vez de T).
Carga neta	Suma de las cargas de los aminoácidos de una proteína a un pH dado.
Código genético	Tabla que indica qué aminoácido corresponde a cada codón.
Codón	Grupo de tres nucleótidos que corresponde a un aminoácido o a una señal de stop.
Codón de inicio	ATG; marca dónde empieza la proteína y codifica metionina.
Codón stop	TAA, TAG o TGA; marca dónde termina la proteína.
Dogma central	El flujo de información ADN → ARN → proteína.
Enzima	Proteína que acelera una reacción química.
FASTA	Formato de texto para secuencias: una línea de nombre que empieza con “>” y debajo la secuencia.

Término	Qué es
Gen	Tramo de ADN con la información para producir una proteína (o un ARN).
Hidrofóbico	Que repele el agua.
Identidad de secuencia	Porcentaje de posiciones iguales entre dos secuencias alineadas.
Mutación	Cambio en la secuencia de ADN.
Nucleótido	Unidad del ADN: fosfato, azúcar y base nitrogenada.
ORF (marco abierto de lectura)	Tramo desde un codón de inicio hasta el primer stop en el mismo marco.
Proteína	Cadena de aminoácidos plegada que cumple una función en la célula.
Ribosoma	Estructura de la célula que lee el ARNm y arma la proteína.
Traducción	Paso de ARNm a proteína en el ribosoma.
Transcripción	Copia de un gen de ADN a ARN mensajero.

Checklist antes de entregar

- ☐ Mi planilla tiene una celda de entrada clara para pegar la secuencia de ADN.
- ☐ Al pegar una secuencia nueva, todo se recalcula solo.
- ☐ Encuentra el primer ATG y corta en el primer codón stop.
- ☐ Muestra la proteína de cada una de las cinco secuencias y de la referencia.
- ☐ Calcula longitud, composición, carga aproximada y porcentaje de hidrofóbicos.
- ☐ Calcula el porcentaje de identidad contra la referencia.
- ☐ Tiene una respuesta explícita: qué secuencia elijo y por qué.
- ☐ Mi video dura 5 minutos o menos y el último minuto cuenta mi motivación.
- ☐ En el video explico cómo hice el análisis, mis resultados, mi elección y qué mejoraría.
- ☐ Envié todo antes del lunes 7 de diciembre de 2026 a las 23:59.

Contacto

Si tenés dudas sobre el concurso, los conceptos o tu planilla, escribinos:

Nombre	Correo
Ignacio Ferres	ferres_i@ort.edu.uy
Josefina Vera	vera@ort.edu.uy
Felipe Machado	machado@ort.edu.uy