

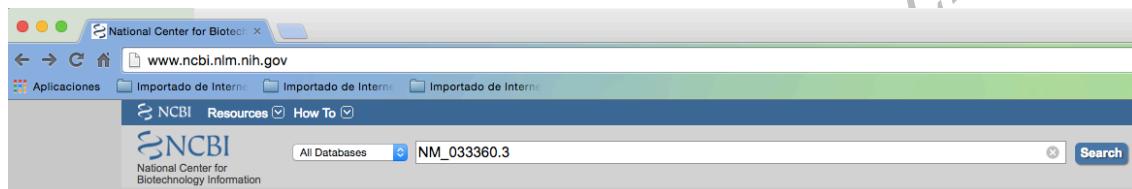
Taller del curso: Introducción al diseño de primers

Clase 7: ejercicios de diseño de primers

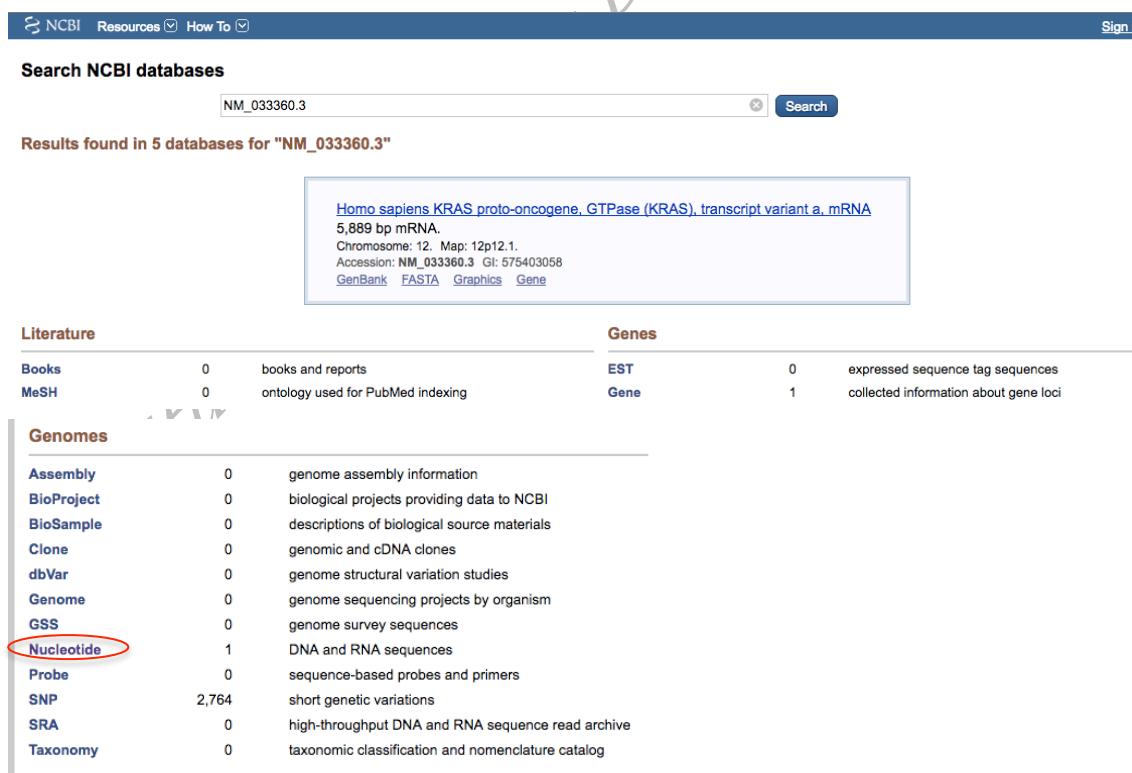
Caso 1: Primers para el gen K-RAS del exón 2 (humanos)

A.- Diseño de primers usando Primer 3 Plus:

1.- Entrar a la base de datos del NCBI www.ncbi.nlm.nih.gov/ y buscar el gen de interés, en nuestro caso buscaremos la secuencia del gen K-RAS que tiene como referencia en el GenBank el número NM_033360.3 Solo es necesario tipar el nombre y darle click en “search”



2.- Aparecerá la siguiente pantalla, de ella trabajaremos con “Nucleotide” por ello le damos click (que ofrece secuencias de ADN o ARN)



The screenshot shows the NCBI search results for the accession number NM_033360.3. At the top, there is a search bar with the query "NM_033360.3". Below the search bar, it says "Results found in 5 databases for 'NM_033360.3'".

In the center, there is a box containing information about the gene: "Homo sapiens KRAS proto-oncogene, GTPase (KRAS), transcript variant a, mRNA", "5,889 bp mRNA.", "Chromosome: 12. Map: 12p12.1.", and "Accession: NM_033360.3 GI: 575403058". Below this box, there are links for "GenBank", "FASTA", "Graphics", and "Gene".

On the left side, there are two main sections: "Literature" and "Genomes". Under "Literature", there are entries for "Books" (0 results) and "MeSH" (0 results). Under "Genomes", there are entries for "Assembly" (0 results), "BioProject" (0 results), "BioSample" (0 results), "Clone" (0 results), "dbVar" (0 results), "Genome" (0 results), "GSS" (0 results), "Nucleotide" (1 result), "Probe" (0 results), "SNP" (2,764 results), "SRA" (0 results), and "Taxonomy" (0 results). The "Nucleotide" entry is circled in red.

3.-Aparecerá la siguiente pantalla donde puedes observar si lo que tienes es una secuencia de ADN o ARN, esto es muy importante. Click en fasta para trabajar con la secuencia en ese formato.

Homo sapiens KRAS proto-oncogene, GTPase (KRAS), transcript variant a, mRNA

NCBI Reference Sequence: NM_033360.3

[FASTA](#) [Graphics](#)

Go to:

```

LOCUS      NM_033360          5889 bp   mRNA    linear    PRI 12-JUN-2016
DEFINITION Homo sapiens KRAS proto-oncogene, GTPase (KRAS), transcript variant
           a, mRNA.
ACCESSION  NM_033360
VERSION    NM_033360.3  GI:575403058
KEYWORDS   RefSeq.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Haplorrhini;
           Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 5889)
AUTHORS   Yun J, Mullarky E, Lu C, Bosch KN, Kavalier A, Rivera K, Roper J,
           Chio II, Giannopoulou EG, Rago C, Muley A, Asara JM, Paik J,
           Elemento O, Chen Z, Pappin DJ, Dow LE, Papadopoulos N, Gross SS and
           Cantley LC.
TITLE     Vitamin C selectively kills KRAS and BRAF mutant colorectal cancer
           cells by targeting GAPDH
JOURNAL   Science 350 (6266), 1391-1396 (2015)
PUBMED    26541605
REMARK    GeneRIF: Vitamin C is selectively toxic to cells with mutant KRAS
           or BRAF alleles.

```

4.- Copiamos la secuencia en un block de notas. Si observan el lado derecho verán que en la pantalla aparece "Pick Primers". Esa es otra forma de diseñar primers que pueden usar para comparar los obtenidos con Primer 3 Plus.

Homo sapiens KRAS proto-oncogene, GTPase (KRAS), transcript variant a, mRNA

NCBI Reference Sequence: NM_033360.3

[GenBank](#) [Graphics](#)

```
>gi|575403058|ref|NM_033360.3| Homo sapiens KRAS proto-oncogene, GTPase (KRAS),
transcript variant a, mRNA
TCCCTAGGCCGCGCCGCCGGCGCAGGCAGCACGGCGCCGCACTGGCGGCCGAAGGTGGCGCCG
CTCGGCCACTACTCCCGCCCCCGCCATTTCGACTGGGAGCGAGCCGAGGGCAGGCACTGAAGCGCCG
CGGGGCCAGAGCTCACGGCTCCAGCTGGCGAGAGAGGCGCTCTGAATAATCACTGAATAAACTT
GTGGTAGTTGGAGCTGGGGCTAGGCAAGCTGGCCTTGACGATACAGCTAATTGAGATCATTGGTGG
ACGAATATGATCCAACAAATAGAGGATTCTACAGGAACAACTAGTAATTGATGGAGAAACCTGTCTT
CGATATTCTCGACACACAGCTAACGACACTACAGTCAATCAGCCACCACTACATGAGCAGTGGGAG
GGCTTCTTGTGTATTGCCATAAAATAACTAAATCATTTGAAGATATTCACTTATAGAGAACAAA
TTAAAGACTTAAGGACTCTGAAGATGTAACCTATGGCTCTACTAGGAAATAATGTGATTGCTCTAG
AACAGTAGACACAAAAACAGGCTCAGGACTTAGCAAGAAGTTATGAAACATCAGCA
```

5.- Empezaremos a trabajar con Pick Primers del NCBI, para ello dar click en Pick Primers y luego trabajaremos con Primer 3 Plus.

The screenshot shows the NCBI sequence analysis interface. At the top, there are three dropdown menus: 'Change region shown', 'Customize view', and 'Analyze this sequence'. Under 'Analyze this sequence', the 'Run BLAST' option is visible, followed by 'Pick Primers', which is circled in red. Below these are 'Highlight Sequence Features' and 'Find in this Sequence'. In the bottom left corner, there is a small blue box containing the number '2'.

6.- Recordamos nuestros conceptos de diseño de primers para PCR en tiempo real y seleccionamos 50 a 200 bp como largo del amplicón. Cerciórate la especie (en este caso humana) y click en *Get Primers*

The screenshot shows the NCBI Primer-BLAST interface. In the 'PCR Template' section, the accession number NM_033360.3 is entered. Under 'Primer Parameters', the forward primer length is set to 70 bp, and the reverse primer length is set to 10 bp. The PCR product size is specified as 700 bp. The 'Primer melting temperatures (Tm)' section shows values: Min 57.0, Opt 60.0, Max 63.0, and Max Tm difference 3. In the 'Exon/intron selection' section, the exon junction span is set to 7 bp and the exon junction match is set to 4 bp. A note at the bottom states: "Note: Parameter values that differ from the default are highlighted in yellow".

7.- Luego de esperar un tiempo (no te preocupes que a veces demora) obtendrás una serie de datos, no solo un par de primers, recuerda que puedes elegir un forward de uno de los primers propuestos y un reverso de otro. Lo importante es que sean los mejores y que tengan una buena Tm aprox. 60, GC% 50, muy similar o igual Ta y considera el largo del primer (tal como está en las diapositivas y se explico en clase)

Primer pair 1							
	Sequence (5'->3')	Template strand	Length	Start	Stop	Tm	GC%
Forward primer	ACTGTTTGCCCCGAAATGGA	Plus	20	5463	5482	60.18	50.00
Reverse primer	CCTGCTGGTGACTGGCATTA	Minus	20	5816	5797	60.04	55.00
Product length	354						
Products on intended target							
>NM_033360.3 Homo sapiens KRAS proto-oncogene, GTPase (KRAS), transcript variant a, mRNA							
product length = 354							
Forward primer 1	ACTGTTTGCCCCGAAATGGA	20					
Template	5463					
Reverse primer 1	CCTGCTGGTGACTGGCATTA	20					
Template	5816					

8.- Para cerciorarte de los Ta puedes hallarlo con el siguiente link [http://tmcalculator.neb.com/#!/_](http://tmcalculator.neb.com/#!/) recuerda que esto es una aproximación el Ta se determina experimentalmente.

9.- Por otro lado, podemos trabajar con Primer 3 Plus disponible en el siguiente link <http://primer3plus.com/cgi-bin/dev/primer3plus.cgi> y copiar la secuencia en formato fasta que guardamos como indicamos en el ítem 4. En este caso estamos diseñando algo general, si queremos amplificar solo el codón 12 y 13 tenemos que indicarlo como el target, o si queremos excluir cierta secuencia podemos colocarla como excluded regions.

Obtendremos la siguiente pantalla:

The screenshot shows the Primer3Plus software interface. At the top, there are tabs for 'Main', 'General Settings', 'Advanced Settings', 'Internal Oligo', 'Penalty Weights', and 'Advanced Sequence'. The 'Advanced Settings' tab is currently selected. Below the tabs, there is a sequence input area containing a large DNA sequence. To the left of the sequence, there is a 'Mark selected region:' button followed by a set of selection tools. To the right of the sequence, there is a 'Save Sequence' button. Below the sequence input area, there are fields for 'Excluded Regions', 'Targets', 'Included Region', 'Primer overlap positions', and 'Pair OK Region List'. The main sequence area has a light blue background with some faint text.

10.- Click en Advanced settings, donde usaremos los mismos conceptos de largo de amplicón, Tm, etc que aprendimos en las clases teóricas y están en las diapositivas.

This screenshot shows the 'Advanced Settings' section of the Primer3Plus software. It includes fields for 'Primer Must Match 5 Prime', 'Internal Must Match 5 Prime', 'Primer Must Match 3 Prime', 'Internal Must Match 3 Prime', 'Left Primer Acronym' (F), 'Internal Oligo Acronym' (IN), 'Right Primer Acronym' (R), 'Primer Name Spacer', 'Product Tm' (Min: 60, Opt: 60, Max: 100), 'Product Size' (Min: 50, Opt: 100, Max: 200), and several checkboxes: 'Pick Anyway', 'Liberal Base', 'Do not treat ambiguity codes in libraries as consensus', 'Use Lowercase Masking', and 'Sequencing'.

11.- Click en Pick Primers

This screenshot shows the 'Pick Primers' section of the Primer3Plus software. It includes fields for 'Max Poly-X' (5), 'Table of thermodynamic parameters:' (SantaLucia 1998), 'Max #N's' (0), 'Salt correction formula:' (SantaLucia 1998), 'CG Clamp' (1), 'Use Thermodynamic Primer Alignment:', and a checkbox for 'Activates Settings Starting with TH:'.

12.- Se observará la siguiente pantalla que contiene los primers propuestos por el Software, nosotros debemos elegir los mejores de acuerdo a su Tm, Ta similar, contenido de GC, que tenga G o C en el 3'OH, etc.

Igual que con el otro software, también se puede elegir el Forward de uno y el reverso de otro. Pero deben de ser compatibles en cuanto a Tm y Ta.

Lamentablemente no es posible asegurar con un 100% de certeza que un primer diseñado mediante una herramienta de software va a ser completamente efectivo, pero sin duda le ayudará a aproximarse a la solución más óptima de una manera muy rápida.

The screenshot shows the Primer3Plus web interface. At the top, there are tabs for 'More...', 'Source Code', 'Help', and 'About'. Below the tabs, there's a back button and a checked checkbox for 'Pair 1: gi|575403058|ref|NM_033360.3| Homo sapiens K1'. The interface displays two primers: Left Primer 1 (Start: 3216, Length: 20 bp) and Right Primer 1 (Start: 3377, Length: 20 bp). Both primers have their sequences, GC content (GC), melting temperature (Tm), annealing temperature (Any), extension temperature (End), thermal stability (TB), hybridization probability (HP), 3' stability (3' Stab), and a penalty value. A 'Pair' section shows the product size (162 bp) and its properties. Below the primers, a table lists sequence segments from positions 1 to 251, each with its corresponding sequence. The background of the page has a watermark reading 'Instituto de Genética Barbara McClintock'.

13.- Elegir a los mejores primers para el siguiente paso que, es ver si forman o no homodímeros, heterodímeros y horquillas.

B.- Formación de estructuras secundarias que pueden formar los primers escogidos.

Para este caso usaremos los primers de Kristensen et al. (2010).

F: 5'-gaataataaacttgtggtagttggagct-3'

R: 5'-atcgtaaggcactctgcctac-3'

1.- Ir al software OligoAnalyzer disponible en el siguiente link <https://www.idtdna.com/calc/analyzer> y copiar la secuencia del primer Forward en la pantalla input, de la siguiente forma y siempre en la dirección 5' a 3' Colocar la concentración ideal del primer o la que siempre se usa. En este caso dejaremos todo por default (así como está). Dar click en Analyze.



OligoAnalyzer 3.1

[Instructions](#) | [Definitions](#) | [Feedback](#)

Sequence 5'- GAA TAT AAA CTT GTG GTA GTT GGA GCT -3' <input type="button" value="Clear Sequence"/> <input type="button" value="Add To Order"/>	27 Bases Parameter sets SpecSheet (Default)  Target type DNA  Oligo Conc 0.25 μM Na⁺ Conc 50 mM Mg⁺⁺ Conc 0 mM dNTPs Conc 0 mM	<input style="background-color: #D9A319; color: white; font-weight: bold; padding: 5px; margin-bottom: 5px;" type="button" value="Analyze"/> <input type="button" value="Hairpin"/> <input type="button" value="Self-Dimer"/> <input type="button" value="Hetero-Dimer"/> <input type="button" value="NCBI Blast"/> <input type="button" value="Tm Mismatch"/>
---	---	--

2.- Dentro de los resultados podemos ver la secuencia, su reverso complementario, el largo, el contenido de GC% que para este caso no es bueno ya que solo llega a 37% y debería estar por 50%.

RESULTS

SEQUENCE	5'- GAA TAT AAA CTT GTG GTA GTT GGA GCT -3'
COMPLEMENT	5'- AGC TCC AAC TAC CAC AAG TTT ATA TTC -3'
LENGTH	27
GC CONTENT	37 %
MELT TEMP	55 °C
MOLECULAR WEIGHT	8393.5 g/mole
EXTINCTION COEFFICIENT	273300 L/(mole·cm)
nmole/OD ₂₆₀ :	3.66
μg/OD ₂₆₀ :	30.71

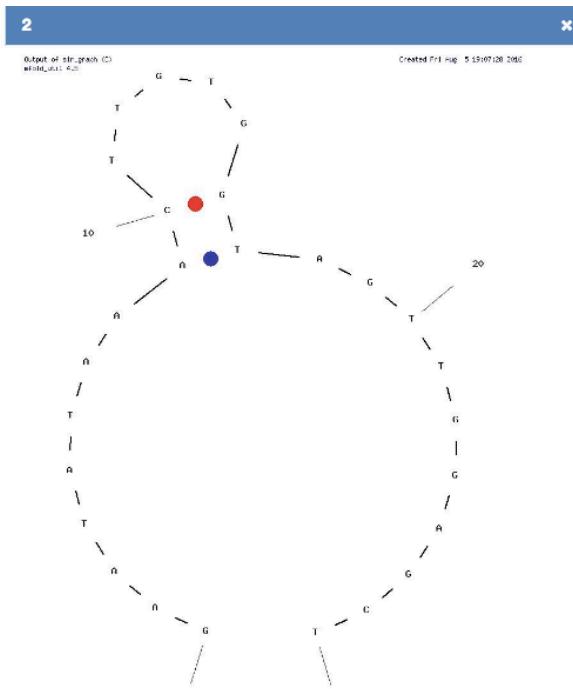
3.- Click en Hairpin para ver la formación de horquillas. Recuerden que un delta de G negativo significa que la reacción ocurre espontáneamente. Los Tm también son importantes pues mientras más pequeño, se requiere menor energía para romper esos puentes de hidrógeno, por tanto son mejores.¹

Structures

Structure Name	Image	ΔG(kcal.mole ⁻¹)	T _m (°C)	ΔH(kcal.mole ⁻¹)	ΔS(cal.K ⁻¹ mole ⁻¹)	Output
1		-0.62	32.4	-25.6	-83.79	<input type="button" value="Ct"/> <input type="button" value="Det"/>
2		0.19	20.6	-12.6	-42.89	<input type="button" value="Ct"/> <input type="button" value="Det"/>

4.- Otro punto a considerar es donde se produce la horquilla, entre los extremos o entre las zonas del medio. Además si son varios los puntos de unión o pocos. El

mejor escenario es que se produzcan entre las zonas medianas y con pocos puntos de unión. Demos click en la figura que está en la columna de Image



En este caso vemos que solo tiene dos puntos de unión y que su Tm es bajo. Además que se producen en el medio por lo que es buena señal. Así debemos evaluar para cada primer escogido.

5.- Click en Self-dimer para ver los homodímeros que se pueden formar. De la misma manera si se producen menos homodímeros y estos tienen menos uniones será mejor. Para ello debemos fijarnos en el delta de G y en cada uno de los resultados.

Homo-Dimer Analysis

The delta G is calculated by taking into account the longest stretch of complementary bases. These pairs of complementary bases are represented by a solid line. Dotted lines represent additional complementary bases for that dimer structure, but their presence does not impact calculated delta G values. Actual delta G values may vary based on presence of additional complementary bases. The Maximum Delta G value refers to the free energy of the oligo sequence binding to its perfect complement.

Dimer Sequence:

5'- GAATATAAACTTGTGGTAGTTGGAGCT -3'
Maximum Delta G: -45.61 kcal/mole

Delta G: -6.34 kcal/mole Base Pairs: 4

5' GAATATAAACTTGTGGTAGTTGGAGCT
|||||
3' TCGAGGTTGATGGTGTCAAATATAAG

Delta G: -4.88 kcal/mole Base Pairs: 4

5' GAATATAAACTTGTGGTAGTTGGAGCT
: |||| : :: :
3' TCGAGGTTGATGGTGTCAAATATAAG

6.- Click en HeteroDimer para calcular las posibles dímeros entre el primer y su reverso complementario y entre el primer Forward con el reverso. Click en

Calculate. Para analizar tener el mismo criterio que para el ítem anterior.

Hetero-Dimer Analysis

The delta G is calculated by taking into account the longest stretch of complementary bases. These pairs of complementary bases are represented by a solid line. Dotted lines represent additional complementary bases for that dimer structure, but their presence does not impact calculated delta G values. Actual delta G values may vary based on presence of additional complementary bases. The Maximum Delta G value refers to the free energy of the oligo sequence binding to its perfect complement.

Primary Sequence: 5'- GAATATAAACTTGTGGTAGTTGGAGCT -3'
Secondary Sequence: 5'- ATCGTCAAGGCACCTTGCCCTAC -3'

Maximum Delta G: -45.61 kcal/mole

Delta G: -5.5 kcal/mole **Base Pairs4**

5' GAATATAAACTTGTGGTAGTTGGAGCT
:: : ||| :
3' CATCCGTTCTCACCGAACCTGCTA

Delta G: -3.9 kcal/mole **Base Pairs4**

5' GAATATAAACTTGTGGTAGTTGGAGCT
||| :
3' CATCCGTTCTCACCGAACCTGCTA

C.- Especificidad de los primers

Si trabajamos con secuencias conocidas como es en el caso humano, podemos usar softwares que nos permitan ver si el primer que hemos elegido se ancla en otras regiones del genoma. Recuerden que los primers deben ser específicos para que no se produzca un falso positivo por amplificación de productos que NO son nuestro amplicón. Para ello podemos usar software como MFEprimer2.0 para verificar la especificidad en el ADN genómico y ARN mensajero/ADN copia e incluso hacer una electroforesis virtual.

1.- Entrar a <http://biocompute.bmi.ac.cn/CZlab/MFEprimer-2.0/>

The screenshot shows the MFEprimer-2.0 web interface. At the top, there are two tabs: "Single Mode Check Dimer" and "Batch Mode Virtual Electrophoresis". Below the tabs, there is a dropdown menu for "Background database selection" set to "Zebrafish - RNA". The main area contains fields for "Enter the primer sequences (5' → 3')": "Sequence 1" and "Sequence 2". There are "Add +" and "Del -" buttons, and "Run" and "Reset" buttons. Below these fields are "Results filter settings [optional]:" with inputs for Tm (°C) From: 30 To: 80, PPC (%): From: 30, Size (bp): From: 50 To: 2000. The "Experimental settings [optional]:" section includes Concentration of monovalent cations (usually KCl, mM): 50.0, Concentration of dNTPs (mM): 0.25, Concentration of divalent cations (usually MgCl₂, mM): 1.5, and Annealing oligo concentration (nM): 50.0. On the right side, there are links to "www.mfeprimer.com", "Protocol (PDF)", "Database requires Download / Manual", "FAQ", "HowTo (中文, English)", "Thanks", and "You may interest" sections featuring "MPRIMER", "VizPrimer", "MPprimer", and "CeIRNAi".

2.- Colocar el organismo que se está estudiando (en nuestro caso humano), colocar la secuencia de ambos primers. Click en Run.

The screenshot shows the MFPrimer-2.0 software interface. At the top, it says "MFPrimer-2.0" and "A fast thermodynamics-based program for checking PCR primer specificity". There are tabs for "Single Match" and "Check Duplication". Below this, a section titled "Background database selection: (Multiple database selection)" has a dropdown menu set to "Human - RNA & Genomic". Under "Enter the primer sequences (5' → 3'):" there are two input fields: "Sequence 1" containing "gaatataaaacttggtagttggagct" and "Sequence 2" containing "atcgtaaggcactcttgccctac". Below the sequences are buttons for "Add +", "Del -", "Run" (which is circled in red), and "Reset". To the right of the "Run" button are "Example" and "Clear" buttons.

3.- Encontrarás los resultados de la secuencia de los primers, el largo, el contenido de GC, Tm, que tu ya conoces por los resultados previos y que podrás corroborar. Por otro lado, los amplicones potenciales. Recuerda que los primers deben ser específicos. Puedes colocar multialign si deseas realizar un alineamiento múltiple. También puedes ver el detalle de los amplicones, en que posición se están anclando los primers, algunos parámetros adicionales y como debes citarlo en tu paper.

A continuación la figura que se obtiene:

The screenshot shows the MFPrimer-2.0 Report. At the top, it says "MFPrimer-2.0 Report" and "Query". Below this, a table shows two primer sequences: Seq1 (GAATATAAACTTGTGGTAGTTGGAGCT) and Seq2 (ATCGTCAAGGCACTCTTGCCCTAC). The table includes columns for Query ID, Sequence (5' → 3'), Length, GC%, and Tm. A "View Hairpins & Dimers" button is also present. Below the table, a section titled "Descriptions of 3 potential amplicons" lists three amplicons with their details. At the bottom, there are buttons for "Select All", "ExportAmplicon", "MultiAlign", "Cluster & MultiAlign", "Virtual Electrophoresis", and "Reset".

Query ID	Sequence (5' → 3')	Length	GC%	Tm
Seq1	GAATATAAACTTGTGGTAGTTGGAGCT	27	37.0	60.0
Seq2	ATCGTCAAGGCACTCTTGCCCTAC	23	52.2	63.4

Amplicon details

[What's this?](#) [How to explain the result?](#)

1: Seq1 + Seq2 ==> gi|34485723|ref|NM_004985.3| Homo sapiens v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS), transcript variant b, mRNA

PPC = 88.7%, Size = 56 (bp), GC content = 48.2%
Fp: Tm = 60.0 (°C), ΔG = -25.4 (kcal/mol), 3ΔG = -3.4 (kcal/mol)
Rp: Tm = 63.4 (°C), ΔG = -25.4 (kcal/mol), 3ΔG = -2.5 (kcal/mol)
Binding sites: 188(27/27) ... 243(23/23)

>>>Seq1
1 27
5' GAATATAAACTTCTGGTACTTGGACCT 3'
::: 243
5' GAATATAAACTTCTGGTACTTGGACCTgg...ggcCTAGGCCAAGACTGCCCTTGACCAT 3'
188 :::::::::::::::::::::
3' CTAGGCCAAGACTGCCCTTGACCAT 5'
23 1
Seq2<<<

>Amp_1 Seq1 + Seq2 ==> gi|34485723|ref|NM_004985.3| Homo sapiens v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS), transcript variant b, mRNA
GAATATAAACTTCTGGTACTTGGACCTggcCTAGGCCAAGACTGCCCTTGACCAT

2: Seq1 + Seq2 ==> gi|34485724|ref|NM_033360.2| Homo sapiens v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS), transcript variant a, mRNA

PPC = 88.7%, Size = 56 (bp), GC content = 48.2%
Fp: Tm = 60.0 (°C), ΔG = -25.4 (kcal/mol), 3ΔG = -3.4 (kcal/mol)
Rp: Tm = 63.4 (°C), ΔG = -25.4 (kcal/mol), 3ΔG = -2.5 (kcal/mol)
Binding sites: 188(27/27) ... 243(23/23)

>>>Seq1
1 27
5' GAATATAAACTTCTGGTACTTGGACCT 3'
::: 243
5' GAATATAAACTTCTGGTACTTGGACCTgg...ggcCTAGGCCAAGACTGCCCTTGACCAT 3'
188 :::::::::::::::::::::
3' CTAGGCCAAGACTGCCCTTGACCAT 5'
23 1
Seq2<<<

Parameters

Database: Human.ma, Human.genomic	K value: 9
Degenerate sequence: No	PPC cutoff: 30.0 (%)
Size start: 50 (bp)	Size stop: 2000 (bp)
Tm start: 30.0 (°C)	Tm stop: 80.0 (°C)
Concentration of monovalent cations: 50.0 (mM)	Concentration of divalent cations: 1.5 (mM)
Concentration of annealing oligos: 50.0 (nM)	Concentration of dNTPs: 0.25 (mM)
Time used: 0 m 3 s	Date: 2016-08-06 09:07:56

Citation

Wubin Qu, Yang Zhou, Yanchun Zhang, Yiming Lu, Xiaolei Wang, Dongsheng Zhao, Yi Yang and Chenggang Zhang. (2012) MFEprimer-2.0: A fast thermodynamics-based program for checking PCR primer specificity. Nucleic Acids Res. (Web Server Issue): W205-8. DOI: 10.1093/nar/gks552.

Wubin Qu, Zhiyong Shen, Dongsheng Zhao, Yi Yang and Chenggang Zhang. (2009) MFEprimer: multiple factor evaluation of the specificity of PCR primers. Bioinformatics, 25(2), 276-278.

Recuerda que también puedes hacer un BLAST para alinear tus primers con el genoma de interés para saber si se ancla a otros sitios y pueda generar productos inespecíficos.

Caso 2: Primers para plantas que no tienen el genoma completo secuenciado pero se dispone de secuencias parciales de parientes cercanos

Muchas de las organismos están parcialmente secuenciados o no están secuenciados, por lo que, nos cuesta más su estudio y en muchos casos no es posible. Una de las estrategias que se propone si nos encontramos con estos problemas es realizar un BLAST (alineamiento de secuencias con secuencias de otras plantas muy relacionadas) o usar las secuencias PopSet (sequence sets from phylogenetic and population studies)

Lo que conviene es usar secuencias de organismos del mismo género y en el peor de los casos de la misma familia. El peligro de usar esta técnica, si no disponemos si quiera de secuencias parciales de nuestro organismo de interés es que los primers NO anclen.

Para este ejemplo usaremos al género Lupinus.

1.- Entrar al NCBI, colocar el género de interés en este caso *Lupinus* y buscar en las secuencias POPSET

Search NCBI databases

Help

lupinus

Search

Results found in 25 databases for "lupinus"

Literature	Genes
Books 11 books and reports	EST 10,352 expressed sequence tag sequences
MeSH 7 ontology used for PubMed indexing	Gene 47,463 collected information about gene loci
NLM Catalog 4 books, journals and more in the NLM Collections	GEO DataSets 65 functional genomics studies
PubMed 1,564 scientific & medical abstracts/citations	GEO Profiles 0 gene expression and molecular abundance profiles
PubMed Central 2,842 full-text journal articles	HomoloGene 0 homologous gene sets for selected organisms
Health	PopSet 236 sequence sets from phylogenetic and population studies
ClinVar 0 human variations of clinical significance	

2.- Supongamos que el gen que deseo estudiar es LEGCYC1B disponible en <https://www.ncbi.nlm.nih.gov/popset/342325892>

Debo cerciorarme que la secuencia (no importa parcial) este en el conjunto de secuencias POPSET.

En la parte de abajo de la ventana mostrada podemos ver los alineamientos múltiples y en el lado derecho dice RUN BLAST, al cual le damos click

Genistae Cycloidea-like protein group 1B-like (LEGCYC1B) gene, partial sequence.

PopSet: 342325892

[GenBank](#) [FASTA](#)

[Go to:](#)

Study Details

Increased rates of diversification prior to invasion of the Andes reveal ecological correlates of replicate continental radiations in *Lupinus* (Leguminosae)

Drummond,C.S., Eastwood,R.J., Miotto,S.T.S. and Hughes,C.E.

[Go to:](#)

Sequences in this data set

HM562704.1	Lupinus micranthus voucher FHO:C.E.Hughes 1982 Cycloidea-like protein group 1B (LEGACYC1B) gene, partial cds
HM562703.1	UNVERIFIED: Lupinus albus voucher FHO:C.E.Hughes 1970 Cycloidea-like protein group 1B-like (LEGACYC1B) gene, partial sequence
HM562702.1	Lupinus angustifolius voucher FHO:C.E.Hughes 1972 Cycloidea-like protein group 1B (LEGACYC1B) gene, partial cds
HM562701.1	Lupinus hispanicus voucher FHO:Pannell 1 Cycloidea-like protein group 1B (LEGACYC1B) gene, partial cds
HM562700.1	Lupinus arboreus voucher FHO:cult. UK Cycloidea-like protein group 1B (LEGACYC1B) gene, partial cds
HM562699.1	Lupinus chachas voucher FHO:C.E.Hughes 2027 Cycloidea-like protein group 1B (LEGACYC1B) gene, partial cds
HM562698.1	Lupinus solanagrorum voucher FHO:C.E.Hughes 2014 Cycloidea-like protein group 1B (LEGACYC1B) gene, partial cds
HM562697.1	Lupinus sp. 1-CEH voucher FHO:C.E.Hughes 2118 Cycloidea-like protein group 1B (LEGACYC1B) gene, partial cds
HM562696.1	Lupinus mutabilis voucher FHO:C.E.Hughes 2291 Cycloidea-like protein group 1B (LEGACYC1B) gene, partial cds
HM562695.1	Lupinus sp. 1-CEH voucher FHO:C.E.Hughes 2107 Cycloidea-like protein group 1B (LEGACYC1B) gene, partial cds

3.- Click en run blast alignment (puede demorar un poco)

Help

Send to:

Analyze this data set

Run BLAST alignment

Tree View

Related information

BioCollections

Nucleotide

4.- Obtendremos el alineamiento de las secuencias

NIH U.S. National Library of Medicine NCBI National Center for Biotechnology Information Sign in to NCBI

BLAST® > blastn suite-2sequences > RID-ZY5E0U84114 Home Recent Results Saved Strategies Help

BLAST Results

Edit and Resubmit Save Search Strategies > Formatting options > Download You Tube How to read this page Blast report description

Blast 2 sequences

Job title: gj|342325892 (631 letters)

RID ZY5E0U84114 (Expires on 11-06 20:44 pm)

Query ID HM562673.1 Subject ID 32 subjects

Description Genista umbellata voucher FHO:C.E.Hughes 1980 Cycloidea-like Description [See details](#)

Molecule type dna Molecule type dna

Subject Length 20916 Program BLASTN 2.7.1+ [Citation](#)

Other reports: [Search Summary](#) [Distance tree of results](#) [MSA viewer](#)

Graphic Summary

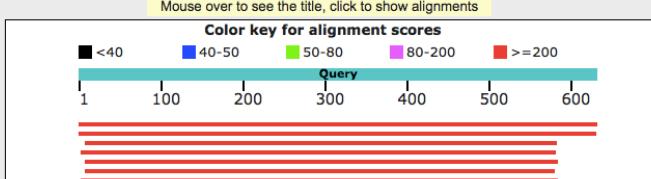
Distribution of the top 32 Blast Hits on 32 subject sequences [?](#)

Mouse over to see the title, click to show alignments

Color key for alignment scores

<40	40-50	50-80	80-200	>=200
-----	-------	-------	--------	-------

Query 1 100 200 300 400 500 600



Sequences producing significant alignments:

Select: All None Selected: 0

	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	UNVERIFIED: Genista umbellata voucher FHO:C.E.Hughes 1980 Cycloidea-like protein group 1B-like (LEGCYC1B) gene, partial cds	1119	1119	100%	0.0	100%	HM562673.1
<input type="checkbox"/>	UNVERIFIED: Sparium junceum voucher FHO:C.E.Hughes 2190 Cycloidea-like protein group 1B-like (LEGCYC1B) gene, partial cds	982	982	99%	0.0	94%	HM562674.1
<input type="checkbox"/>	UNVERIFIED: Lupinus albus voucher FHO:C.E.Hughes 1970 Cycloidea-like protein group 1B-like (LEGCYC1B) gene, partial cds	814	814	90%	0.0	90%	HM562703.1
<input type="checkbox"/>	Lupinus hispanicus voucher FHO:Pannell 1 Cycloidea-like protein group 1B (LEGCYC1B) gene, partial cds	812	812	91%	0.0	90%	HM562701.1
<input type="checkbox"/>	Lupinus chachas voucher FHO:C.E.Hughes 2027 Cycloidea-like protein group 1B (LEGCYC1B) gene, partial cds	805	805	91%	0.0	90%	HM562699.1
<input type="checkbox"/>	Lupinus angustifolius voucher FHO:C.E.Hughes 1972 Cycloidea-like protein group 1B (LEGCYC1B) gene, partial cds	805	805	90%	0.0	90%	HM562702.1
<input type="checkbox"/>	UNVERIFIED: Lupinus huarenensis voucher FHO:C.E.Hughes 2241 Cycloidea-like protein group 1B-like (LEGCYC1B) gene, partial cds	803	803	92%	0.0	89%	HM562679.1
<input type="checkbox"/>	Lupinus sp. 1-CEH voucher FHO:C.E.Hughes 2118 Cycloidea-like protein group 1B (LEGCYC1B) gene, partial cds	801	801	91%	0.0	90%	HM562697.1
<input type="checkbox"/>	Lupinus solanagromus voucher FHO:C.E.Hughes 2014 Cycloidea-like protein group 1B (LEGCYC1B) gene, partial cds	800	800	91%	0.0	90%	HM562698.1
<input type="checkbox"/>	UNVERIFIED: Lupinus microphyllus voucher FHO:C.E.Hughes 2272 Cycloidea-like protein group 1B-like (LEGCYC1B) gene, partial cds	796	796	92%	0.0	89%	HM562677.1
<input type="checkbox"/>	Lupinus sarmentosus voucher FHO:C.E.Hughes 2015 Cycloidea-like protein group 1B (LEGCYC1B) gene, partial cds	796	796	91%	0.0	89%	HM562678.1
<input type="checkbox"/>	Lupinus mutabilis voucher FHO:C.E.Hughes 2291 Cycloidea-like protein group 1B (LEGCYC1B) gene, partial cds	796	796	91%	0.0	89%	HM562696.1
<input type="checkbox"/>	Lupinus arboreus voucher FHO:cult. UK Cycloidea-like protein group 1B (LEGCYC1B) gene, partial cds	796	796	92%	0.0	90%	HM562700.1

Como en cualquier BLAST, se observa el porcentaje de identidad, el E-value, etc. Se espera un E-value mucho menor, pero en este caso lo usaremos, ya que es meramente un ejemplo. En la parte de abajo de nuestros resultados está la ventana de alineamientos y se muestran que tan similares son las secuencias. Como lo que quiero es buscar genes que amplifiquen para este gen en mi organismo de interés (alguna especie del género Lupinus) puedo diseñar primers para las regiones conservadas.

Alignments

Query	Sequence	Score
HM562673	ATCTTC-TAATTCAACCTTYYTTCCCTTCTAACCTGAAATGCTTCTYCAAGCA	59
HM562674-.....YY..T.....	59
HM562703-.....T.....T.....	51
HM562701N.....T.....T.....	79
HM562699T.....T.....T.....	80
HM562702T.....T.....T.....	57
HM562679-G.....T.....T.....	85
HM562697T.....T.....T.....	87
HM562698T.....T.....T.....	51
HM562677-G.....T.....T.....	76
HM562678T.....T.....T.....	58
HM562696T.....T.....T.....	87
HM562700N.....T.....T.....	74
HM562688T.....T.....T.....	86
HM562692T.....T.....T.....	86
HM562693T.....T.....T.....	86
HM562694T.....T.....T.....	86
HM562704T.....T.....T.....	51
HM562681T.....T.....T.....	52
HM562683T.....T.....T.....	62
HM562687T.....T.....T.....	59
HM562690T.....T.....T.....	78
HM562691T.....T.....T.....	71
HM562695T.....T.....T.....	44
HM562676T.....T.....T.....	50
HM562689T.....T.....T.....	86
HM562684T.....T.....T.....	56
HM562685T.....T.....T.....	49
HM562686T.....T.....T.....	60
HM562682T.....T.....T.....	44
HM562675T.....T.....T.....	35
HM562680T.....T.....T.....	13

Si se dan cuenta los puntos indican que son la misma base (coinciden y se alinean), por otro lado, podemos ver que para la secuencia HM562674 difiere con respecto a la query (original que tomaron de referencia) en dos Y donde había citosina (revisar nomenclatura IUPAC descrita en clases) y en una T. Lo que muestro es

solo una parte del alineamiento, como verán se muestra por segmentos. Por tanto, me conviene diseñar primers para las regiones conservadas.

Esta podría ser la región para mi Forward de la base 32 a la 52 (recuadro rojo)

Query	1	ATCTTC-TAATTCAACCCCTYTTCCCTTCTTAACCTGAAAATGCTCTYCAAGCA	59
HM562673	1	59
HM562674	1 YY.. T.	59
HM562703	1 T.	51
HM562701	22 N.. T.	79
HM562699	29 T. T	80
HM562702	6 T.	57
HM562679	29 -G T. T	85
HM562697	36 T. T	87
HM562698	1 T. T	51
HM562677	20 -G. T. T. T	76
HM562678	7 T. T	58
HM562696	36 T. T	87
HM562700	17 N.. T. T	74
HM562688	35 T. T	86
HM562692	35 T. T	86
HM562693	35 T. T	86
HM562694	35 T. T	86
HM562704	1 T.	51
HM562681	2 T. T	52
HM562683	11 T. T	62
HM562687	8 T. T	59
HM562690	27 T. T	78
HM562691	20 T. T	71
HM562695	1 T. T	44
HM562676	1 T. T	50
HM562689	35 T. T	86
HM562684	6 T. T	56
HM562685	1 T. T	49
HM562686	9 T. T	60
HM562682	1 T. T	44
HM562675	1 T. T	35
HM562680	12 T.	13

Y esta región donde se unirá mi reverse de la base 391 a la 415 (recuadro rojo)

Query	363	TCGATCTACAAGACATGCTAGGGTTGRCAAAGCAAGCAACACCCTGAGTGSTATTC	422
HM562673	363	422
HM562674	369	..T..... A	428
HM562703	370 A	429
HM562701	392 A	451
HM562699	393 A	452
HM562702	376 T. A	435
HM562679	404 A	463
HM562697	400 A	459
HM562698	364 A	423
HM562677	392 A	451
HM562678	374 A	433
HM562696	403 A	462
HM562700	384 A	443
HM562688	405 A	464
HM562692	405 A	464
HM562693	405 A	464
HM562694	405 A	464
HM562704	367 A	426
HM562681	371 A	430
HM562683	381 A	440
HM562687	378 A	437
HM562690	397 A	456
HM562691	390 A	449
HM562695	357 A	416
HM562676	366 A	425
HM562689	405 A	464
HM562684	375 A	434
HM562685	368 A	427
HM562686	379 A.	438
HM562682	363 A	422
HM562675	351 A	410
HM562680	326 A	385

Para ello, podemos usar la secuencia Query, ya que se supone que la especie para la cual elegimos diseñar estos primers NO tiene la secuencia disponible en la base de datos, pero está emparentada a las que se muestran, es decir son del mismo género. Para hacerlo debo ir a Primer 3 Plus y especificar los segmentos que deseo.

La secuencia Query está disponible en <https://www.ncbi.nlm.nih.gov/nuccore/HM562673.1?report=fasta> que este caso pertenece a una especie de la misma familia pero no del mismo género, pero es una región que se ha conservado y nos puede servir para amplificar el segmento de interés.

UNVERIFIED: Genista umbellata voucher FHO:C.E.Hughes 1980 Cycloidea-like protein group 1B-like (LEGCYC1B) gene, partial sequence

GenBank: HM562673.1

[GenBank](#) [Graphics](#) [PopSet](#)

```
>HM562673.1 UNVERIFIED: Genista umbellata voucher FHO:C.E.Hughes 1980 Cycloidea-like protein group 1B-like (LEGCYC1B) gene, partial sequence
ATCTCTTAATTCTCATACCCCTTCTTCCTTAAACCTGAAAATGCTCTYCAAGCAACAACACCAC
ACCYTCTCATGATCCACTTGCTCATGTTCCGTACAATTACCAAGTCATCATCATATTCTATAACACAC
CTATAATCCAAGAAAACACTGCCAATTGGCTGTTCTGCTGCTGCAACGCCAAAACATGACCCCTAT
TATAACTGGTGGYGGTGCATCATCACTATGGGCTTCTCTTGCTCACAAAAGAAACCCAGCCAAAAAA
GATAGGCACAGCAAATTACACTGCTCAGGGCTTGAGGGACGGGAGGGTGAGGCTTCGATCCAGATCG
CVCGAAAGTTTTCGATCTACAAGACATGCTAGGGTTTGRCAAAGCAACCAACACCCCTGAGTGGSTAT
CAACAAAGCTCAAGAGAGCAATTAAGCAGCTAGCTAGAAAGAACATCAATGTCAGTGAAGGCTGATACT
AAGAGCTTCTCCTCATCTCGGATTGTGAGGACTGCAATGAAGTTATTTCAGGGATCAACAAATGAACAGC
AAGATATCACCATTGCTGATCATGCTAAACATTACAACAAAGGGTTAGATTCAAATCCTGTGAA
A
```

Por otro lado, si nos fijamos bien en los segmentos que tenemos disponibles, estos son muy limitados para el diseño de primers, por ejemplo la región para el Forward es solo de 21 bases (lo mismo que el largo de un primer) y la del reverse es de 25 bases, por lo que, no hay mucho de que escoger, para ello nosotros podemos ir tomando regiones de 18 a 25 pb probarlas en Primer3 y luego analizarlas en Oligoanalyzer para elegir a las mejores. En estos casos es mucho más complicado poder obtener primers que no formen subproductos, pero se eligen a los mejores.

Por ejemplo elijamos para el forward la región de la 32 a la 49 (CTTAACCCTGAAAATGCT : 18bp) y para el reverse las 18bp empezando por la base 391 (CAAAGCAAGCAACACCCCT : 18bp).

Recuerda que podemos ir cambiando y colocar por ejemplo un largo diferente o tomando en el rango establecido las diferentes bases, con tal de elegir al mejor par posible.

5.- Entrar a Primer 3 Plus <https://primer3plus.com/cgi-bin/dev/primer3plus.cgi> y colocar en la parte de abajo los primers que estamos proponiendo.

El Forward será CTTAACCCCTGAAAATGCT

Mientras que el reverse será el reverso complementario de la secuencia que reconoce es decir AGGGTGTTGCTTGCTTTG

<input checked="" type="checkbox"/> Pick left primer	<input type="checkbox"/> Pick hybridization probe	<input checked="" type="checkbox"/> Pick right primer
or use left primer below.	(internal oligo) or use oligo below.	or use right primer below (5'->3' on opposite strand).
CTTAACCCCTGAAAATGCT		AGGGTGTGCTTGCTTTC

© by A. Untergasser : [Contact](#) : [Datenschutzerklärung](#) : [Impressum](#)

Damos click a Pick Primers y obtenemos lo siguiente:

<input checked="" type="checkbox"/> Pair 1: HM562673.1 UNVERIFIED: Genista umbellata v1
Left Primer 1: CTTAACCCCTGAAAATGCT
Start: 32 Length: 18 bp Tm: 49.6 C GC: 38.9 % Any: 0.0 End: 0.0 TB: 7.0 HP: 0.0 3' Stab: 3.8 Penalty: 12.371
Problems: Temperature too low;
Right Primer 1: AGGGTGTGCTTGCTTTC
Start: 408 Length: 18 bp Tm: 56.1 C GC: 50.0 % Any: 0.0 End: 0.0 TB: 9.0 HP: 0.0 3' Stab: 2.8 Penalty: 5.902
Problems: Temperature too low;
Pair: Product Size: 377 bp Any: 0.0 End: 0.0 TB: 16.0 Penalty: 18.274

Donde se muestran las características del primer F y R en este caso podemos descartar esta combinación, ya que ambos tienen un Tm que difiere en más de 5°C. Por otro lado, se muestra la posición que reconocen los primers.

Send to Primer3Manager					
Reset Form					
1	ATCTTCTAAT	TCATACCCCTT	YTTTCCCTTT	CCTTAACCCCT	GAAAATGCTT
51	CTYCAAGCAA	CAACAACCAC	ACCYTTCTTC	ATGATCCACT	TGCTCATGTT
101	CCGTACAATT	TACCAAGTCA	TCATCATATT	CATAACACAC	CTATAATCCA
151	AGAAACACTG	CCCAATTGTTG	CTGTTCTGC	TGCTGCTGCA	ACGCCAAAAC
201	ATGACCCCTAT	TATAAGTGGT	GGYGGTGCTC	ATCATCACTA	TGGGCTTTCT
251	TCTTGCTCA	CAAAGAAACC	AGCCAAAAAA	GATAGGCACA	GCAAAATTAA
301	CACTGCTCAG	GGCTTGAGGG	ACCGGAGGGT	GAGGCTTTCG	ATCCAGATCG
351	CVCGAAAGTT	TTTCGATCTA	CAAGACATGC	TAGGGTTTGR	CAAAGCAAGC
401	AACACCCCTTG	AGTGGSTATT	CAACAAGTCC	AAGAGAGCAA	TTAAGGAGCT
451	AGCTAGAACG	AAGAACATCA	ATGTCAGTGA	AGGTGATACT	AAGAGCTTCT
501	CCTCATCTTC	GGATTGTGAG	GAUTGCAATG	AAGTTATTTC	AGGGATCAAC
551	AATGAACAGC	AAGATATCAC	CATCGCTGAT	CATGATGCTG	TAAACTTACA
601	ACAACAAGGG	TTAGATTCAA	ATGCTGTGAA	A	

De esta forma vamos eligiendo al par correcto y luego procedemos al análisis con Oligoanalyzer.

Caso 3: Primers SSR para gramíneas

Los microsatélites son marcadores moleculares codominantes, secuencias repetidas en tandem muy usados en estudios de variabilidad y diversidad genética tanto en plantas como en animales y humanos.

Para poder diseñar primers que amplifiquen microsatélites de interés primero debemos encontrar el microsatélite que buscamos en la secuencia de interés.

1.- Eliges la secuencia en la cual deseas testear si hay microsatélites, para luego realizar tu estudios de variabilidad o diversidad genética. Recuerda que también puedes usar esta herramienta para otras especies, ya que el algoritmo te permite encontrar cualquier SSR en el formato fasta de acuerdo a las características que indicaste. En esta oportunidad trabajaremos con la secuencia de maíz (*Zea mays* cultivar B73 chromosome 7, B73 RefGen_v4, whole genome shotgun sequence) disponible en el siguiente link

https://www.ncbi.nlm.nih.gov/nuccore/NC_024465.2?report=fasta&from=113218229&to=113228563

2.- Entrar a SSRTTools de Gramene, esta es una página especializada de estas plantas, si la revisan podrán encontrar que ya hay una lista de primers probados para ellas pero si nuestro objeto de interés fuera una planta que no se ha estudiado mucho, podemos usar esta herramienta disponible en <http://archive.gramene.org/db/markers/ssrtool>

The screenshot shows the GRAMENE Ssrtool website. The URL in the address bar is archive.gramene.org/db/markers/ssrtool. The page has a green header with the GRAMENE logo and the word "Ssrtool". Below the header is a navigation menu with links: Search, Genomes, Species, Download, Resources, About, Help, and Feedback. Under the menu, there are links to Markers Home, Markers Search, Browse Map Sets, SSR Markers Resource, Microarray Probes, Help, Tutorial, and FAQs. The main content area is titled "SSRIT - Simple Sequence Repeat Identification Tool". It contains text about the tool's purpose, source code availability, citation information, and browser requirements. A note at the bottom states: "Note: Netscape 2.0 or greater, or IE 4.0 or greater required."

3.- Colocar los parámetros que estamos buscando como “seleccionar el máximo largo del motivo”. Por ejemplo, si pongo dímero es porque quiero que el software elija una secuencia de dos nucleótidos que se repitan n veces. En la letra b coloco por ejemplo 3 que sería el mínimo de veces que quiero que se repita este dímero. Tal como aparece en la siguiente figura:

- 1) Select search parameters
 - a) Select the maximum motif-length group you wish to find. For example, if you want to search for all SSRs up to and including pentamers (meaning, you'd like to search for dimers, trimers, tetramers, and pentamers), you should select 'pentamer' from the drop-down menu.
 - b) Enter the minimum number of repeats you will allow. Entering a '5,' for example, will match SSRs with five or more motif repeats, such as ag-5 ('agagagagag').

4.- Copio y pego mi secuencia fasta de interés en el recuadro del software, de la siguiente forma y le doy click a FIND SSRs:

2) Paste/Enter your sequence of interest into the textarea

The sequence(s) must be in FASTA format - meaning, there must be a title line with a '>' at the beginning for each sequence.

FOR EXAMPLE,

```
>seq1
agagatttagatcgatcgctctctctctctcgatcgagatcgat
ggccatcatcatcatcatatcgatcgatcgatcgatcgatcgat
agaatagatatcgctatagagatcgagagatcgagagatgaga
>seq2
agagatagaaataatgagatagcggggggggggggcgctatacgcgctcg
gagagagatctctctctttagatcgatcgactcgatcgatata
agactcactcactcactcactcagcogat
```

Paste/Enter your sequence(s) here:

```
>NC_024465.2:113218229-113228563 Zea mays cultivar B73
chromosome 7, B73 RefGen_v4, whole genome shotgun sequence
TAAGCGCTGCACGATCACCTTCCGCCGAGCTGCATCCATCCCCAC
GATCCGCGCCAGGGCGCCGCAA
CGCCCTGCCGCTCGGTGGCCACTGACGTGTCACCCCACTCCCTTC
TTCTGTACCCCATCTCCCTTC
CCGCCACTTCTCCGCGCCATTTCCTTGCCCCTCGCCTCCACCTACT
AAATAATCCCCCAGCCAGCA
ATTCTCCCGAAAAAAAGACAAAACCTAAACCCGCGGAGCCGGGGTG
GAGAGCAGCCATTCTCGCTC
```

5.- Obtendré una lista de los microsatélites que estarían presentes, según los parámetros indicados. Por ejm el primer resultado nos indica que la secuencia tiene un motivo dímero de CA (citosina y adenina) que se repite 3 veces y te indica que inicia en la base 2261 y termina en la base 2266 (hay otro similar en la posición 4568 al 4573).

SSRs found in your sequence(s)

Sequence	Motif	No.of Repeats	SSR start	SSR end	SeqLength
NC_024465.2:113218229-113228563-1	ca	3	2261	2266	10386
NC_024465.2:113218229-113228563-2	tg	3	3233	3238	10386
NC_024465.2:113218229-113228563-3	tc	3	3865	3870	10386
NC_024465.2:113218229-113228563-4	at	3	3894	3899	10386
NC_024465.2:113218229-113228563-5	tc	3	4101	4106	10386
NC_024465.2:113218229-113228563-6	ta	3	4445	4450	10386
NC_024465.2:113218229-113228563-7	ca	3	4568	4573	10386
NC_024465.2:113218229-113228563-8	ta	3	4620	4625	10386
NC_024465.2:113218229-113228563-9	at	4	6445	6452	10386
NC_024465.2:113218229-113228563-10	tg	3	6637	6642	10386

6.- Ahora vamos a suponer que me interesa solo ese dímero de CA repetido 3 veces y voy a diseñar primer para que solo se amplifique esa región, para ver si entre cultivares puedo observar una diferencia en el largo de repeticiones (más corto a más largo). Para ello entro a Primer 3 Plus <https://primer3plus.com/cgi-bin/dev/primer3plus.cgi> y colocaré mi secuencia FASTA con algunas restricciones. Primero como quiero que amplifique la región que va de la base 2261 y termina en la base 2266, coloco en TARGETS 2261,6 Puedo modificar los ajustes según lo explicado en clase o puedo dejar todo por defecto y darle pick primers (dependerá del objeto de estudio).

The screenshot shows the Primer3Plus web interface. At the top, there are tabs for 'More...', 'Source Code', 'Help', and 'About'. Below that, there are 'Load server settings:' and 'Task:' dropdowns. The 'Task:' dropdown is set to 'generic'. To the right, there is a note: 'Select primer pairs to detect the given template sequence. Optionally targets and included/excluded regions can be specified.' Below these are several tabs: 'Main' (selected), 'General Settings', 'Advanced Settings', 'Internal Oligo', 'Penalty Weights', and 'Advanced Sequence'. Under the 'Main' tab, there is a 'Sequence Id:' input field and a note 'Paste template sequence below'. Below this is a large text area containing a FASTA sequence: >NC_024465.2:113218229-113228563 Zea mays cultivar B73 chromosome 7, B73 RefGen_v4, whole genome shotgun sequence. The sequence starts with TAAAGCGCTGCAAGCATCACCTTCCGGCGGAGGTGCATCCACGATCCGGCGCAGGGCGCCCAA and continues for several lines. Below the sequence, there are fields for 'Excluded Regions:', 'Targets:', 'Included Region:', 'Primer overlap positions:', and 'Pair OK Region List:'. There are also buttons for 'Mark selected region:' (with arrows and braces), 'Save Sequence', and 'Upload File'.

7.- Obtengo lo siguiente:

The screenshot shows the results page of Primer3Plus. At the top, there is a note: 'News: Try out our new tool: [Wiley-DNA-Editor](#) - A DNA/Plasmid editor running in your browser!' Below that, there are tabs for 'More...', 'Source Code', 'Help', and 'About'. The main area displays the results of the primer search. It shows a pair of primers: Left Primer 1 (AACAGAGGAATGGTGGAAA) and Right Primer 1 (TACCCATCTCCTTACGCCGA). Both primers have a length of 20 bp, a GC content of 50.0%, and a 3' Stab of 4.0. The penalties are 0.482 for the left primer and 0.107 for the right primer. The pair has a product size of 561 bp. There are buttons for 'Send to Primer3Manager' and 'Reset Form'.

Donde se muestra el resultado de varios pares de primers. Tal como se explico en clase, ver que se cumplen las reglas de diseño de primers. Además se muestra donde anclan ambos primers y la región que delimité.

1751	ATTGACAATT	GATAAACAGA	GGCAATGGTG	GGAATTCAAA	TCACAGCACA
1801	TGGATAAGGT	TCTGTTTTTC	AAGGTACATT	TATATTTGC	CACAAGTTAC
1851	TTGGTACTTA	AATGCAACTA	TTTTATTTAG	TGATTACAAT	AATAAAATAT
1901	GAATTTGAAT	AGATGGGCAA	ATTCTACGAA	CTATTTGAGA	TGGATGCACA
1951	TGTTGGAGCA	AAGGATCTTG	ATCTTCAGTA	CATGAAGGCA	CAAATATCTT
2001	CTCACATTAA	TGCTTATTGA	TTCTTCTGCT	ACTAAGCAAT	CCAGATGATC
2051	TGCCATTCT	CGTATTATCT	AGAACTCACT	GTCTGTATTC	AATCATGTTA
2101	TCTTAGGGTG	AACAGCCCCA	CTGTGGATTG	CCAGAGAAGA	ATTTATCAGT
2151	GAATTTGGAG	AAGTTAGCTG	AGAAGGTAAA	CTTATTTAAT	GTTTTCCTC
2201	ATTAAATTTC	ACACATAATT	GCATTTAGC	AATAACATTA	ATTGCATTAC
2251	AGGGTTATCG	AGTTCTGGTT	GTCGAGCAAA	CTGAAACACC	TGACCAGCTT
2301	GAACTTCGGC	GTAAGGAGAT	GGGTATAAAG	GACAAGGTAT	GCTTAGATTT

Si nos damos cuenta se subrayó otra región (muy cercana pero no es el microsatélite que buscamos) es la región AGTTCT que NO corresponde a lo que indicamos. Sin embargo, subrayé la región que correspondería al motivo estudiado. Este error se puede deber al algoritmo que usa el programa SSRTools para contar las bases.

Ahora solo faltaría usar Oligoanalyzer para ver si los primers elegidos generan horquillas, homodímeros o heterodímeros y elegir a los mejores. Recuerda que este es solo un ejemplo, trabajando con algunos plantas como maíz que ya cuenta con una base especializada es más fácil encontrar SRR ya probados y adquirirlos para nuestro estudio. Por otro lado, si trabajaremos con plantas no estudiadas previamente podemos usar esta herramienta y buscar genes que anteriormente se hayan reportado que presenten microsatélites y usarlo en nuestra especie de interés. Este software detecta cualquier tipo de microsatélites sin importar la especie en estudio.