## Tutorial for using Case It to search DNA sequences

General sequence searching

Setting search parameters for the PCR procedure

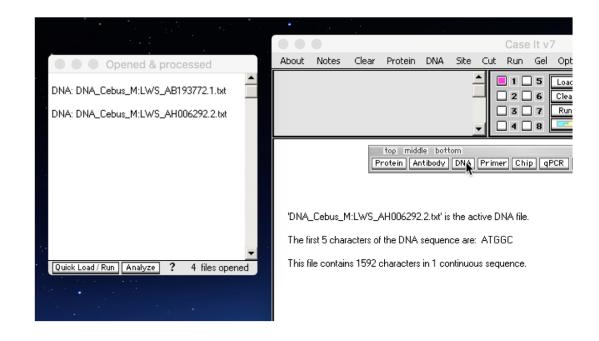
Setting search parameters for Southern and dot blots

Obtaining sequences for research purposes

The first part of this tutorial shows the basic procedure for using Case It v7 to search DNA sequences. Click the **DNA button** on the silver button bar, and double-click on these folders: **Case It V7 PC -> Cases -> Color Vision -> Sample Sequences**. Shift-click on the two sequences inside the **Sample Sequences** folder to highlight them, and then click the **Open** button:

							Case	e It v	7					
Abo	out Notes	Clear	Protein	DNA	Site	Cut	Run	Gel	Options	; Cases	Tut	orials		
					-			5	Load	1.0	r.	4	0 sites	
						-   0	2	6	Clear	[agaro	se]	#bases set	located	
							3	7	Run	60		100	0 sites	
							4	8		minutes	: set	match % set	cummulative	
	top middle bottom Protein Antibody DNA Primer Chip gPCR Probe Enzyme Cut DNA													
	$\bullet \circ \circ$	Sh	ift-click	to ope	en mu	ltiple	e DNA	A file	S					
L	.ook in:	🛅 Sar	nple seque	ences			- 4		<b>a</b> 🗈					
	DNA Ce	bus M:LW:	5 AB19377 5 AH00629	92.2.txt										▼
	File <u>n</u> ame:	1.	txt" "DNA i	Cebus M	1:LWS #	4H00(	6292.2	.txt"		Open k	나는			-
	Files of <u>typ</u> e	e: All	files					-	] _ (	Iancel				
	_		Open as <u>r</u> e	ead-only	/		_			_				

The two open files appear in the Opened & processed window. Click the DNA button again...



...and double-click through these folders: Case It v7 PC -> Cases -> Color Vision -> DNA for 180 277 285 locations. Shift-click to open all of the files in the DNA for 180 277 285 locations folder.

	- 11 - L											•		_
							Case	lt v	7					
About	Notes C	Ilear P	Protein	DNA	Site	Cut	Run	Gel	Options	Cases	Tuto	prials		
						▲ 🔽	1	] 5	Load	1.0 %		4	0 sites	
					1	-   C	2	6	Clear	[agarose	e]	#bases set	located	
							3	] 7	Run	60		100	0 sites	
						16	4	8		minutes 🗄	set	match % set	cummulative	
		t	op middl	e bott	om									
		Pro	tein An	tibody	DNA	Prime	er Chi	p qi	PCR Pro	e Enzyn	ne (	Cut DNA		
			Shift	-click	to op	oen m	nultipl	e DN	IA files					
	Look in:		DNA fo	or 180 2	277 28	5 locat	ions	•	<u> </u>		i= 1			
DNA_0		100.111	S tct S se					_			<u> </u>			
			S det A a											
The firs			S tac Y ty											
This file	DNA :	277 MW	/Sitte Fip	henyla	l. ex 5	.txt								
			S acc T ti			·····								
	DNA :	285 MW	/Sigce A (	alanine	ex 5.t	xt								Ŧ
Original													► <u></u> ∃ 1	
1	, E:l					e	• - I '		E h. 4"					Ŷ
1592 E	File <u>n</u> ame	;;	l'exe		35 MW	5 gcc i	H alanır	ne ex	5.txt"	<u>O</u> pe				
	Files of <u>ty</u>	ype:	All file:	s					•	Cano	cel			
					and on							_		
				ar as [t	au-on	γ								

## Click on the line DNA: DNA\_180\_LWS\_tct\_S\_serine\_ex\_3.txt, and then click the Quick Load / Run button and select Search sequences -> Add selected file to the 'Enter search sequence' field.

<u>Note</u>: The explanation of the filename is as follows: In Case It, DNA files must begin with the characters **DNA**. The number **180** refers to a location on Exon 3 (designated here as\_**ex\_3**) of the opsin gene, and the characters **tct** refer to the DNA codon associated with the amino acid **serine (S)**. One of these three characters would be at position 180 on exon 3 of the opsin gene. Color vision in primates and many other animals depend on which amino acid is associated with this and other positions on the opsin gene. For more information, visit the **Evo-Ed** site at <a href="http://www.evo-ed.org">http://www.evo-ed.org</a> and click on **Monkey Opsins**.

	· ·	1997 - 19	.1		1.1				
								Case	e It v
Opened & processed	About	Notes	Clear	Protein	DNA	Site	Cut	Run	Gel
DNA: DNA_Cebus_M:LWS_AB193772.1.txt								1 [ 2 [ 3 [	5 6 7
DNA: DNA_180_LWS_tct_S_serine_ex_3.txt			_	top midd	lle bott	torn	-	4	8
DNA: DNA_180_MWS_gct_A_alanine_ex_3.txt			P	rotein Ar	ntibody	DNA	Prim	er Chi	ip q
DNA: DNA_277_LWS_tac_Y_tyrosine_ex_5.txt									
DNA: DNA_277_MWS_ttc_F_phenylalex_5.txt									
DNA: DNA_285_LWS_acc_T_threonex_5.txt	'DNA	_180_LW	S_tct_S	_serine_ex	_3.txť i	is the a	ictive [	DNA file	е.
	Thef	irst 5 char	acters of	this sequ	ence ar	e: CC1	rgg		
	This f	ile contair	ns 30 ch	aracters in	a single	e seque	ence.		
Quick Load / Run Analyze ? 9 files opened									
DNA gel     Add selected file to the Electron sequences       Southern blot     Load well[s]       Protein gel     Search sequences after log			nce' field	5eque	ence û	. Û			A

The 30-character sequence in this file appears in the second field from bottom of the **Sequence Analysis** window (lower right in the screen shot below). Click the **Search** button and select **Search options** (or click the purple button to the left of the Search button).

About Notes Clear Protein DNA Site Cut Run Gel Options Cases Tutorials	Sequence analysis
Image: Control in the image: Contro	Options Search results field
top middle bottom Protein Antibody DNA Primer Chip gPCR Probe Enzyme Cut DNA	
'DNA_285_LWS_acc_T_threonex_5.txt' is the active DNA file. The first 5 characters of this sequence are: CTGGG This file contains 30 characters in a single sequence.	
■ Original order     Size     Sequence	
✓ Help ►	Range and length of selection
Genbank Lab Bench Data Screen ELISA Dot Blot 96-well PCR Chip qPCR Sequence analysis Opened Loaded Ge1 / blot Help	<b></b>

Drag over the search sequence to determine the position of the three lowercase characters. As you do so, the length is updated in the Range field just above the search string field, and the highlighted characters appear there. Note that the three lowercase letters are in positions 14 through 16, so enter those numbers in the Search Options box, as shown on the next slide...

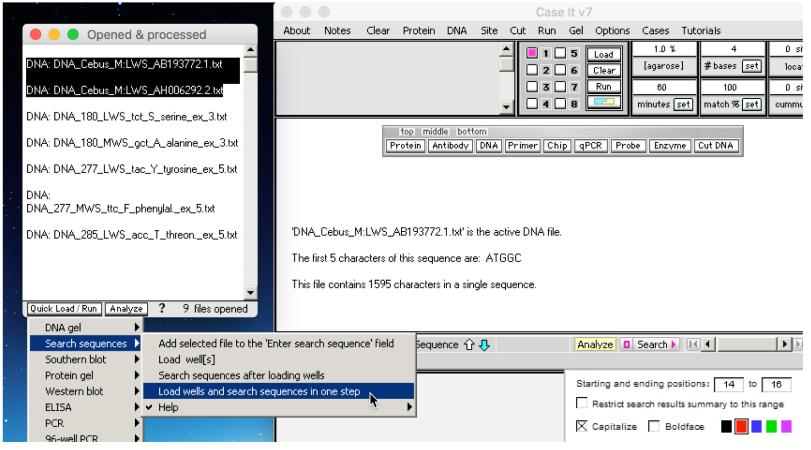
nce are: CTGGG a single sequence.				r
nce 순 🕂	Analyze  Search	1 ₩	Export field Range-[1,16] Length16	-
	Starting and ending positions: to to Restrict search results summary to this range			•
ISA DotBlot 96-wellPCR Ch	Close Apply settings to search sequence	▼. He]p	Enter replacement sequence	•

After entering 14 and 16 as the starting and ending positions, click the **Apply settings to search sequence** button. The default is to highlight the starting and ending characters in red, and capitalize any lowercase letters, but these settings can be changed as shown below.

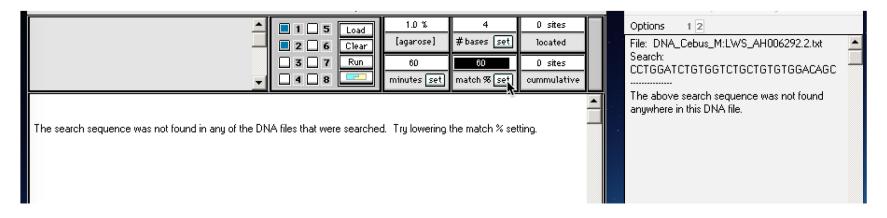
single sequence.	<b>•</b>		-
e 介 ሇ	Analyze Search K K A K A A A A A A A A A A A A A A A	Export field Range and length of selection	
	Starting and ending positions: 14 to 16 Restrict search results summary to this range Capitalize Boldface	CCTGGATCTGTGGGTCTGCTGTGGGACAGC	•
A DotBlot 96-wellPCR CI	hip gPCR Sequence analysis Opened Loaded Ge1 / blot Help	Enter replacement sequence	-

Highlight the two DNA sequences at the top of the **Opened & processed** window, then click the **Quick Load / Run button** of that window and select **Search sequences -> Load wells and search sequences in one step** as shown below. These sequences are from the Capuchin (genus Cebus), a New World monkey known for its tool use.

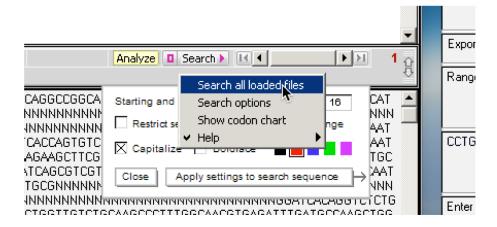
<u>Note</u>: For present purposes loading wells is not associated with running gels, but rather is a necessary step for the search procedure to work.



A message appears that the 'search sequence was not found in any of the DNA files that were searched. Try lowering the match % setting.' Type **60** into the **match %** box and click the set button...



...and then click the Search button on the gray divider bar, and select Search all loaded files to research the files at this new match % setting.



Results appear both in the white field of the **Data Screen** (as a summary of all results), and also in the **Search results** field of the **Sequence analysis** window (showing results for each file separately) Clicking the 1 or 2 at the top of the Sequence analysis window shows search results for the files loaded in either well 1 or well 2. There was one 70% match for the first file searched (AB193772.1), and both 70% and 60% matches for the second file searched (AH006292.2). Note that the gray divider bar has been dragged down to show all results – the 'drag handle' for the divider bar is on its extreme right at the cursor location in the screen shot below.

About Notes Clear Protein DNA Site Cut Run Gel Options Cases Tutorials	Sequence analysis
Image: Section of the section of th	Options 1 2 File: DNA_Cebus_M:LWS_AH006292.2.txt Search: CCTGGATCTGTGGGTCTGCTGTGGGACAGC
DNA_Cebus_M:LWS_AB193772.1.txt 1: 725 - 754 [ 725 - 754 ] 70.00 CCTGGATCTGTGGTCTGCTGTGGGACAGC CTCCTGGATCTGGTCTGCTGTGGGACGGC	ORIGINAL ORDER Fragment number, location of search sequence within a particular fragment, [cumulative location], and % match 
DNA_Cebus_M:LWS_AH006292.2.txt 1: 725 - 754 [ 725 - 754 ] 70.00 CCTGGATCTGTGGTCTGCTGTGGGACAGC CTCCTGGATCTGGGCTGCTGTGTGGACAGC	1: 727 - 756 [ 727 - 756 ] 60.00
DNA_Cebus_M:LWS_AH006292.2.txt 1: 727 - 756 [ 727 - 756 ] 60.00 CCTGGATCTGTGGTCTGCTGTGGGACAGC CCTGGATCTGGGCTGCTGTGTGGACAGCCC	Export field
■ Original order       Size       Sequence	

Change the value in the match% box from 60 to 70, then click the **set** button. Use the **Search** button as shown below to select **Search all loaded files**. This command can also be given from the Quick Load / Run button of the Opened & processed window.

<u>Note</u>: The gray divider bar has been returned to its default position in the screenshot below. You can do this either by dragging the bar up by its drag handle, or by using the Analyze button and select Divider->Return gray divider bar to its default position (this step not shown here). You would now need to scroll to see all of the contents of the white field, but to activate the scrollbar you would need to first click in the white field, otherwise the scrollbar may pop up to its top position when released.

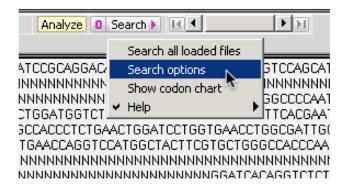
	•		7 Run 8 🖃	60 minutes set	70 match % set	2 sites cummulative	
DNA_Cebus_M:LWS_AB193772.1.txt							1
1: 725 - 754 [ 725 - 754 ] 70.00 CCTGGATCTGTGGTCTGCTGTGTGGACAGC CTCCTGGATCTGGTCTGCTGTGTGGACGGC							
DNA_Cebus_M:LWS_AH006292.2.txt							
1: 725 - 754 [ 725 - 754 ] 70.00 CCTGGATCTGTGGTCTGCTGTGTGGACAGC CTCCTGGATCTGGGCTGCTGTGTGGACAGC							
DNA_Cebus_M:LWS_AH006292.2.txt							
■Original order Size X Sequence 介 🕂			Analyze	🕽 Search 🕨 🔣		► E 1	
1					loaded files		Ŷ
ATGGCCCAGCAGTGGAGCCTCCAGAGGCTCG CTTCACCTACACCAACAGCAACTCCACCAGAGN NNNNNNNNNNNNNNNNNNNNNNNNNNNNN	INNN INNNI CACC	INNNNNN NNNNNNN AGTGTCT	INNNNNNN NNNNNNN GGATGGTC	C/ Search op IN Show codo N Y Help T	tions 🕅 on chart ►	STCCAGCAT INNNNNNN GGCCCCAAT TTCACGAAT GGCGATTGC	

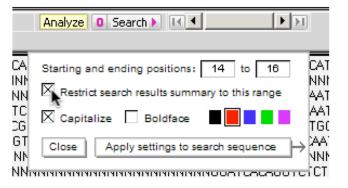
The files are searched again at the new setting of 70%, meaning that only hits with at least a 70% match are shown. Lines in the upper field of the Sequence analysis window represent hits, so if you click on one of these lines then the corresponding characters are highlighted in the file being searched, at the bottom of the Data Screen. In addition, those characters appear in the field at the bottom of the Sequence analysis window, so that you can compare them with the search string just above it. Note that in this case, the characters GCT were present in the file, whereas the characters in red just above it are TCT. What does this mean? Before explaining this, we'll use another Search Options feature to make it easier to see...

		1 🗌 5	Load	1.0 %	4	1 sites		Options	1 2	
		2 6	Clear	[agarose]	#bases set	located			ebus_M:LWS_AH00629	92.2.txt 📃 📥
		3 7	Run	60	70	1 sites		Search:	төтөөтстөстөтөтө	
		4 🗌 8		minutes set	match % set	cummulative			1010010100101010	
DNA_Cebus_M:LWS_AB193772.1.txt 1: 725 · 754 [ 725 · 754 ] 70.00 CCTGGATCTGTGGGTCTGCTGTGTGGACAGC CTCCTGGATCTGGTCTGCTGTGTGGACGGC DNA_Cebus_M:LWS_AH006292.2.txt 1: 725 · 754 [ 725 · 754 ] 70.00 CCTGGATCTGTGGTCTGCTGTGGGACAGC CTCCTGGATCTGGGCTGCTGTGTGGACAGC							· · ·	within a part location], an 	mber, location of search cular fragment, [cumulati	
							-			
						► E	-	Export field		-
Original order     Size Sequence      Original order		A	nalyze	Search 🕨 🔣			兌	Damas 11.00	20 June 14 20	•
1									)] Length30 TCTGGGCTGCTGTGTG	GACAGO
CTTCACCTACACCAACAGCAACTCCACCAGAGI								0.00100.1		- 01/01/02
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN TACCACATCGCTCCCAGATGGGTGTACCACCT										-
GGGCTGGTGCTGGCGGCCACCATGAAGTTCA							-	CCTGGATC	TGTGGTCTGCTGTGTG	GACAGC 🔺
TGACCTGGCAGAAACCGTCATCGCCAGCACTA									<b>I</b>	•
GTGTGTCCTGGAGGGCTACACGGTCTCCCTG								CTCCTGGA	тстобёстостотото	GACAGC 🔺
NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN GTCCCTGGCCATCATTTCCTGGGAGAGAGGTGG										
CCATCGTGGGAGTTGCCTT CTCCTGGATCTGU	GCTGC	CTGTGTGG	ACAGC	CGCCCATCTT	TGGTTGGAGC	AGNNNNNN				

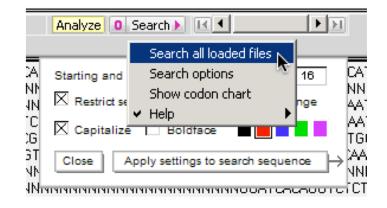
Since the Search Options box had been closed in the preceding slide, open it again using the **Search** button (or by clicking the button with a purple O to the left of the Search button).

Click the checkbox next to **Restrict search results summary to this range**, so that an **X** appears in this checkbox.





Then use the Search button and select **Search all loaded files**. This will re-search the files with this additional option checked.



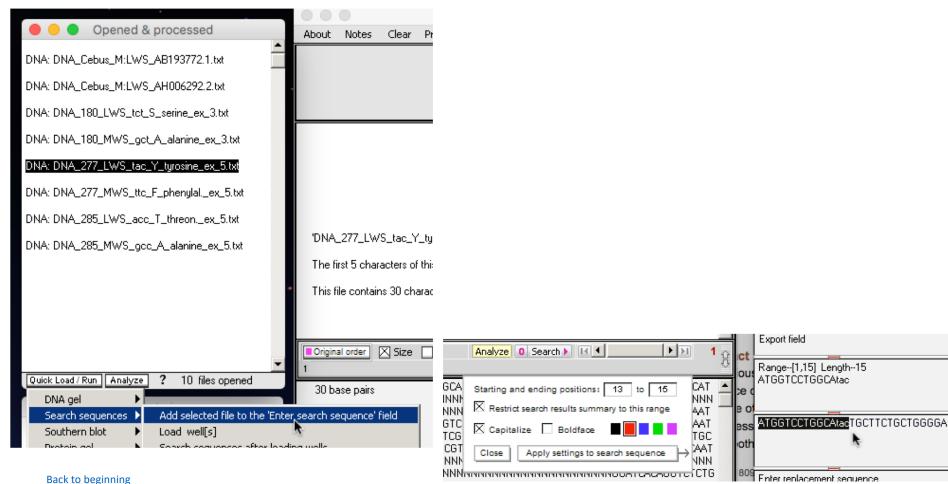
The summary of search results now shows only the three characters in the selected range (14 through 16), making it easier to see that TCT is present in the first file and GCT is present in the second file at location 180 in exon 3 of the opsin gene associated with color vision. Each of these triplets of letters represents a **codon**, so to determine which amino acids are associated with them we need a codon chart. Click the **Search** button and select **Show codon chart** as shown below.

DNA_Cebus_M:LWS_AB193772.1.txt	<b>_</b>
1: 725 - 754 [ 725 - 754 ] 70.00 TCT TCT TCT	
DNA_Cebus_M:LWS_AH006292.2.txt	
1: 725 - 754 [ 725 - 754 ] 70.00 TCT GCT	-
Original order 🗌 Size 🛛 Sequence 🔂 🕂	Analyze 0 Search K 4 N 1
1	Search all loaded files
ATGGCCCAGCAGTGGAGCCTCCAGAGGCTCGCAGGCCGGCA CTTCACCTACACCAACAGCAACTCCACCAGAGNNNNNNNN NNNNNNNNNNNNNNNNNNNNNN	Starting and Search options 16 CAT ANN AAT AAT AAT TGC

The codon chart indicates that **TCT** is the codon associated with the amino acid **Serine (S)**, whereas **GCT** is associated with **Alanine (A)**, so the individual monkeys from which these samples were taken had **S** and **A**, respectively, at the **180** position on exon 3 of the opsin gene. This has significance in terms of the wavelengths of colors that these two individuals could detect (see See <u>Color Vision Polymorphism in Wild Capuchins and Spider Monkeys in Costa</u> <u>Rica</u>, also referenced on Slide 18).

About Notes Clear Protein DN/ DNA codon chart RNA codon chart IUPAC notation Second base											
				Т			С			Α	
			TTT	F	Phenyl- alanine	TCT	S	Serine	TAT	Y	Tyrosi
DNA_Cebus_M:LWS_AB193772.1.txt		Т	TTC	F		TCC	S		TAC	Y	
1: 725 - 754 [ 725 - 754 ] 70.00			TTA	L	Leucine	TCA	S		TAA		stop
			TTG	L		TCG	S		TAG		stop
			CTT	L		CCT	Р	Proline	CAT	н	Histidi
DNA_Cebus_M:LWS_AH006292.2.txt		С	CTC	L		CCC	Р		CAC	н	
1: 725 - 754 [ 725 - 754 ] 70.00 TCT			CTA	L		CCA	Р		CAA	Q	Gluta- mine
GCT	base		CTG	L		CCG	Р		CAG	Q	
	First b		ATT	Ι	Iso- leucine	ACT	Т	Threo- nine	AAT	N	Aspara gine
		A	ATC	Ι		ACC	Т		AAC	Ν	
Original order     Size      Sequence			ATA	Ι		ACA	Т		AAA	К	Lysine
1 ATGGCCCAGCAGTGGAGCCTCCAGAGG			ATG	М	Methio- nine	ACG	Т		AAG	K	
CTTCACCTACACCAACAGCAACTCCACC			GTT	v	Valine	GCT	Α	Alanine	GAT	D	Aspart acid
TACCACATCGCTCCCAGATGGGTGTAC		G	GTC	v		GCC	Α		GAC	D	

In the Opened & processed window, click on the line DNA: DNA\_277\_LWS\_tac\_Y\_tyrosine\_ex\_5.txt and use the Quick Load/Run button to select Search sequences -> Add selected file to the 'Enter search sequence' field. The search sequence associated with that file will appear in the second field from the bottom in the Sequence analysis window, as shown at right. Drag to determine the position of the lowercase letters tac (13 through 15), open the Search options box, enter these numbers as the starting and ending positions, and finally click Apply settings to search sequence.



There were 90% and 80% matches, respectively, for the two files. An examination of the codon chart (not shown here) reveals that TAC is the codon for tyrosine (Y), and TTC is the codon for phenylalanine (F). These represent the amino acids at location 277 of exon 5 of the opsin gene, another important location for color vision in primates and other animals. If you repeat this procedure for location 285 on exon 5, using the search string for **DNA**; **DNA\_285\_LWS\_acc\_T\_threon\_ex\_5.txt**, the amino acids present at this location can be determined for both files.

Verify that for the monkey associated with AB193772.1, the amino acids at positions 180, 277, and 285 are S Y, and T, respectively. See <u>Color Vision</u> <u>Polymorphism in Wild Capuchins and Spider Monkeys in Costa Rica</u>, the source of this data, for information relating to its significance. Question: What are the comparable amino acids at these positions in humans with full color vision, and humans that are colorblind?

DNA_Cebus_M:LWS_AB193772.1.txt		ORIGINAL ORDER Fragment number, location of search sequence
1: 1217 - 1246 [ 1217 - 1246 ] 90.00 TAC TAC		within a particular fragment, [cumulative location], and % match
DNA_Cebus_M:LWS_AH006292.2.txt		1: 1217 - 1246 [ 1217 - 1246 ] 80.00
1: 1217 - 1246 [ 1217 - 1246 ] 80.00 TAC TTC		
	-	Export field
Original order     Size   Sequence	Analyze 0 Search 🕨 💶 🕨 원 1 ਨੂ	
1	<u> </u>	Range-[1,30] Length30
GTACCCCGGGGTGCAGCCTTACATGATCGTCCTCATGATCAC	Starting and ending positions: 13 to 15	
CTCCAAGTGTGGCTGGCCATCCGAGCTNNNNNNNNNNNN NNNNNNNNNNNNNNNNNNNNNNNN	Restrict search results summary to this range TCC	
GAATCCACCCAGAAGGCAGAGAAGGAAGTGACACGCATGGT( CTACGCCTTCTTCGCATGCTTTGCTGCTGCCAACCCTGGCTA	Capitalize 🗌 Boldface 🔳 📕 📕 📕 🗱 🕅	ATGGTCCTGGCATACTGCTTCTGCTGGGGA
GCCAAAAGTGCCACTATCTACAACCCCATTATCTATGTCTTTA	Close Apply settings to search sequence NNN	
NNTTTCGAAACTGCATCTTGCAGCTTTTTGGGAAGAAGGTTG	กรอกายอย่ายายการายายเกอยอยยายเกกกระหม่AGG	

<u>Next, we'll demonstrate the enhanced PCR feature of Case It v7</u>. Earlier versions of Case It required a 100% match of forward and reverse primers, but Version 7 allows you to set the match percent at any level. As an example, we'll use DNA sequences from a variety of New World and Old World monkeys, obtained from GenBank. Before beginning, use the Clear menu to clear all existing data from Case It v7.

About Notes Clear Protein DNA Site Cut Run Gel C								Case	elt v	7
	About	Notes	Clear	Protein	DNA	Site	Cut	Run	Gel	Op
Other options			Cle	ar everyth	ning? 🕨	Ye	5	1	] 5 [	Loa
			Oth	ner options	; 🕨		T	2	6	Clea

Click the **DNA** button and navigate through the folders **Case It v7 PC -> Cases -> Color Vision -> Sample sequences**. Shift-click to select the two DNA Cebus files used earlier, and click the **Open** button. Then click the **DNA** button again, select the other two files in the folder, and click the **Open** button. Note that you cannot shift click to open all four files simultaneously, since the second two files contain multiple DNA sequences each.

Shift-click to open multiple DNA files		Shift-click to open multiple DNA files
Look in: 📄 Sample sequences 💽 🔬 🖾 📃	top middle bottom Protein Antibody DNA	Look in: 📄 Sample sequences 💽 👍 🛄 🛅 🧮
DNA Cebus MLW5 AB193772.1.txt DNA Cebus MLW5 AH006292.2.txt DNA ex 5 NWMs.txt DNA ex 5 OWMs.txt	Clear Protein DNA Site	DNA Cebus MLW5 AB193772.1.txt DNA Cebus MLW5 AH006292.2.txt DNA ex 5 NWMs.txt DNA ex 5 OWMs.txt
File name: 1.txt" "DNA Cebus MLW5 AH006292.2.txt"		File name: DNA ex 5 NWMs.txt" "DNA ex 5 OWMs.txt"
Files of type: All files Cancel		Files of type: All files Cancel
Open as read-only		Open as read-only

Navigate to the primers for exons 3 and 5 folder and open Primer for exon 5.txt. It will appear as a line at the bottom of the Opened & processed window after the DNA sequences opened previously. The blue lines indicate files that contain multiple DNA sequences, indicated by the underscore preceding each filename. The Cebus files contain one sequence per file.

Shift-click to open multiple primer files		Case
Look in: 📄 primers for exons 3 and 5 💽 🔥 🔛 🔚	Opened & processed	About Notes Clear Protein DNA Site Cut Run
Primer for exon 3.txt	DNA: DNA_Cebus_MLWS_AB193772.1.txt	Filename: Primer for exon 5.txt
Primer for exon 5.txt Primers for exons 3 and 5.txt	DNA: DNA_Cebus_MLWS_AH006292.2.txt	FP: 5'-> AGAGTCCGAATCCACCCAGA <-3'
	DNA+: DNA_ex_5_NWMs.txt	RP: 5'-> tgccggttcatgaagacat <-3'
	DNA:.>AH012145.2_Saguinus_labiatu	
	DNA:.>AH012134.2_Leontopithecus_	
File name: Primer for exon 5.txt Open	DNA:.>AH006290.2_Cebus_olivaceus	
Files of type: All files Cancel	DNA:.>AB193796.1_Ateles_geoffroy	
Open as read-only	DNA:.>DNA_KP867090_Cacajao_calvu	'DNA_Cebus_MLWS_AB193772.1.txt' is the active DNA file.
	DNA+: DNA_ex_5_0WMs.txt	The first 5 characters of this sequence are: ATGGC
	DNA:.>XM_023198523.1_Piliocolobus	This file contains 1595 characters in a single sequence.
	DNA:.>XR_979216.1_Macaca_nemestr	
	DNA:.>XM_012002287.1_Mandrillus_	Original order X Size Sequence 介 ♣
	DNA:.>XM_010373544.1_Rhinopithec	1
	DNA:.>XM_007993132.1_Chlorocebus	1595 base pairs
	PRIMERS: Primer for exon 5.txt	
	Quick Load / Run Analyze ? 13 files opened	

Set the match % setting to 70, meaning that hits will occur for any match of 70% or higher (for both forward and reverse primers).

Opened & processed									Case	lt v7	,				
	<b></b>		Notes	Clear	Protein	DNA	Site	Cut	Run	Gel	Options	Cases	Tuto	orials	
DNA: DNA_Cebus_MLWS_AB193772.1.txt		Filename	: Primer f	or exon	5.txt		ŀ		1	5	Load	1.0 %		4	0 site
DNA: DNA_Cebus_MLWS_AH006292.2.txt							(- 3'		2		Clear	[agaros	e]	#bases set	locate
DNA+: DNA_ex_5_NWMs.txt		RP: 5'	> tgeogg	ottcatgaa	agacat <-	3'	Ŀ		3		Run	6D minutes [	set	70 match % set	D site cummula
DNA: >AH012145.2. Saouinus Tabiatu															

Shift click to select all the DNA files, making sure not to select the **Primers** file at the bottom, and then use the **Run** button and select **Run PCR**.

Opened & processed	Case It v7	
	About Notes Clear Protein DNA Site Cut Run Gel Options Cases Tutorials	
DNA: DNA_Cebus_MLWS_AB193772.1.txt		sites
DNA: DNA_Cebus_MLWS_AH006292.2.txt	FP: 5'-> AGAGTCCGAATCCACCCAGA <-3"	ated
DNA+: DNA_ex_5_NWMs.txt	RP:         5' -> tgccggttcatgaagacat <- 3'	sites at
DNA:.>AH012145.2_Saguinus_labiatu	Run PCR Bun Southern or Western blot	
DNA:.>AH012134.2_Leontopithecus_		
DNA:.>AH006290.2_Cebus_olivaceus	Note: Run menu above has more options	
DNA:.>AB193796.1_Ateles_geoffroy		
DNA:.>DNA_KP867090_Cacajao_calvu	'DNA_Cebus_MLW/S_AB193772.1.txt' is the active DNA file.	
DNA+: DNA_ex_5_0WMs.txt	The first 5 characters of this sequence are: ATGGC	
DNA:.>XM_023198523.1_Piliocolobus	This file contains 1595 characters in a single sequence.	
DNA:.>XR_979216.1_Macaca_nemestr		
DNA:.>XM_012002287.1_Mandrillus_	Original order ∑ Size _ Sequence 介 ↓     Analyze 0 Search ▶ K ▲	Ы
DNA:.>XM_010373544.1_Rhinopithec	1	
DNA:.>XM_007993132.1_Chlorocebus	1595 base pairs	
PRIMERS: Primer for exon 5.txt		

New lines appear in the Opened & processed window, each representing a PCR product (designated by the arrow preceding the filename). To see a summary of results shown the percentage for each hit, **right-click** on the white upper field and select **PCR results -> Put match % in the upper field**.

<u>Note</u>: most fields in Case It have associated pop-up menus, activated by right-clicking.

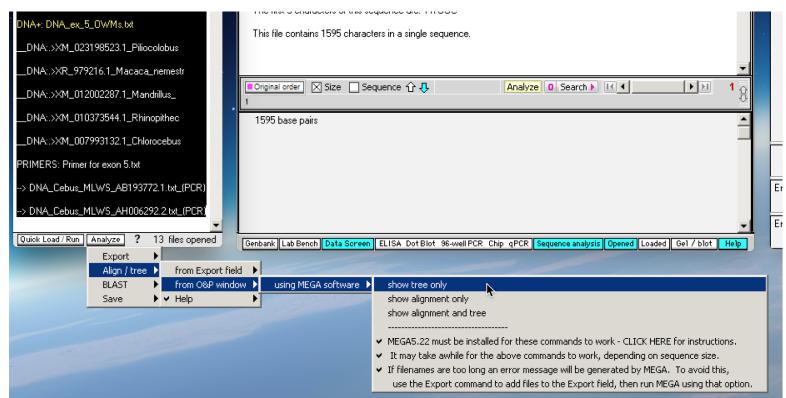
Opened & processed	Case It v7	
	About Notes Clear Protein DNA Site Cut Run Gel Options Cases Tutorials	
DNA:.>XR_979216.1_Macaca_nemestr	Filename: Primer for exon 5.txt	Optic
DNA:.>XM_012002287.1_Mandrillus_	FP:     5' -> AGAGTCCGAATCCACCCAGA <- 3'	Sear
DNA:.>XM_010373544.1_Rhinopithec	RP: 5'-> tgccggttcatgaagacat <-3'	
DNA:.>XM_007993132.1_Chlorocebus		
PRIMERS: Primer for exon 5.txt		
> DNA_Cebus_MLWS_AB193772.1.txt_(PCR)	Сору	
> DNA_Cebus_MLWS_AH006292.2.txt_(PCR)	Paste 🕨	
> >AH012145.2_Saguinus_labiatu_(PCR)	'>XM_007993132.1_Chlorocebus_(PCR)' is the active DNA file. Add Replace ►	
> >AH012134.2_Leontopithecus(PCR)	The first 5 characters of this sequence are: AGAGT Save ► PCR results ► Put match % results into upper field	d I
> >AH006290.2_Cebus_olivaceus_(PCR)	This file contains 223 characters in a single sequence.     Blot results	•
> >AB193796.1_Ateles_geoffroy_(PCR)		
> >DNA_KP867090_Cacajao_calvu_(PCR)	In Original order Size Sequence ↑ ↓ Analyze Search ► K ▲ ► ► ► 1 ↔	Expo
> >XM_023198523.1_Piliocolobus_(PCR)	1	Ranç
> >XR_979216.1_Macaca_nemestr_(PCR)	222 base pairs	
> >XM_012002287.1_Mandrillus(PCR)		Enter
> >XM_010373544.1_Rhinopithec_(PCR)		
> >XM_007993132.1_Chlorocebus_(PCR)		Enter

One of the matches is shown below, with a 95% match for the forward primer and a 94.74 match for the reverse primer. You would need to scroll in the field to see all of the results, so right-click to **Save** them to the Notepad of Case It. You can also **Copy** the results to the clipboard, for pasting into another application.

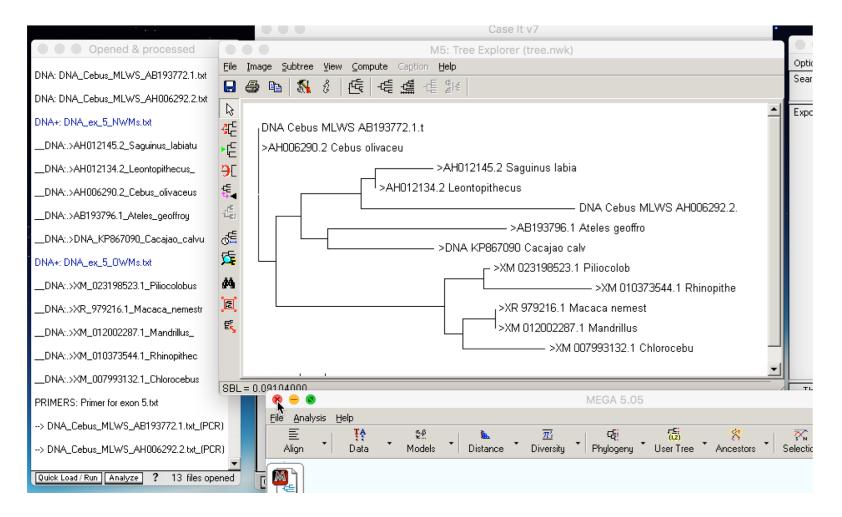
Opened & processed	About Notes Clear Protein DNA Site Cut Run Gel Options Cases Tutorials	
DNA:.>XM_023198523.1_Piliocolobus		Option
DNA:.>XR_979216.1_Macaca_nemestr	FP: 5' -> AGAGTCCGAATCCACCCAGA <- 3'	Search
DNA:.>XM_012002287.1_Mandrillus_		Export
DNA:.>XM_010373544.1_Rhinopithec	>XM_007993132.1_Chlorocebus	
DNA:.>XM_007993132.1_Chlorocebus	[FP match % , RP match %] ["95.00 FP", "94.74 RP"]	
PRIMERS: Primer for exon 5.txt	AGAGTCCGAATCCACCCAGA~FP	
> DNA_Cebus_MLWS_AB193772.1.txt_(PCR)	AGAGTCTGAATCCACCCAGADNA Copy Paste	
> DNA_Cebus_MLWS_AH006292.2.txt_(PCR)	ATGTCTTCATGAACCGGCARP Add  ATGTCTTTATGAACCGGCADNA Replace	
> >AH012145.2_Saguinus_labiatu_(PCR)	Save   Entire contents of upper field to Notepad	
> >AH012134.2_Leontopithecus(PCR)	PCR product 1 length = 222     PCR results >     Save all sequences in upper field of main window       Blot results >     Save a specific sequence in lower field of main window	
> >AH006290.2_Cebus_olivaceus_(PCR)		
> >AB193796.1_Ateles_geoffroy_(PCR)	<ul> <li>✓ IMPORTANT: Rename files every time you save them</li> <li>Important order</li> <li>Size Sequence ↑ ↓</li> <li>Ana</li> <li>Ido not replace an existing file with a file of the same name</li> </ul>	ie]
> >DNA_KP867090_Cacajao_calvu_(PCR)	1	
> >XM_023198523.1_Piliocolobus_(PCR)	222 base pairs	
> >XR_979216.1_Macaca_nemestr_(PCR)		Range
> >XM_012002287.1_Mandrillus(PCR)		<b>F</b> :

To create a phylogenetic tree of this data with one click, **shift-click** to highlight all of the PCR products, then use the **Analyze** button and select **Align/tree -> from O&P window -> using MEGA software -> show tree only**...

<u>Note</u>: This assumes that you have separately downloaded the MEGA executable and put it in the MEGA folder of Case It – for details, see the tutorial **Installing MEBA** software for use with Case It.



...and in a few seconds the tree appears. Case It also copy selected PCR products to the the clipboard and open two web-based sites for aligning sequences and building trees, the **MABL** site and the **MAFFT** site. See the **Bioinformatics** tutorial for details.



Before demonstrating set search parameters for Southern and dot blots, use the Clear menu and select **Clear everything? -> Yes**.

						Case	e It v	7
About Notes	Clear	Protein	DNA	Site	Cut	Run	Gel	Ор
	Cle	ar everyth	ning? 🕨	Ye	5 N	1	5	Loa
	Oth	her options	5 🕨		Tř	2	6	Clea

Click on the DNA button of the silver button bar, then navigate through the folders **Case It v7 PC -> Cases -> Test sequences -> DNA for digestion -> Probe example using sickle cell** and open the file **DNA sickle cell control mutated.gen**.

						Case	lt v	7				
	About Notes	Clear Prot	in DNA	Site	Cut			Options	Cases T	utorials		
	nultin la DNA Gla				1	1	] 5 [	Load	1.0 %	4	0 sites	
Shift-click to open n		es i		-	_   ā	2		Clear	[agarose]	#bases set	located	
Look in: 📄 Probe example using sick	e cell 💌 👍					3	] 7 [	Run	60	100	0 sites	
DNA sickle cell control mutated.gen					-10	4	8		minutes se	t match % set	cummulative	
Enzyme MstII (cctNagg)			2	ttom								
Probe for sickle cell 90 percent match.TX	(T		2	DNA	Prime	er Chij	p qF	PCR Pro	be Enzyme	Cut DNA		
1			_									
File name: DNA sickle cell control mu	utated.gen	<u>O</u> pe	n									
Files of type: All files	•	Can										
Open as <u>r</u> ead-only												
												-
	Original order	🛾 Size 📃 Se	quence (	Ъ 🗘			Ar	nalyze 0	Search 🕨	<u>I</u>	► ►	슈
												-

Before running a Southern blot, we need the restriction enzyme Mst II to digest the DNA sequence. Click the **button with the blue/yellow symbol** (located under the Run button) and select **M** -> **Mst II**.

<u>Note</u>: This file could also be accessed by clicking the Enzyme button, then selecting that file from the currently active folder.

00001077	
About Notes Clear Protein DNA Site Cut Run Gel Options Case:	s Tutorials
▲ ■ 1 5 Load 1.0 2 6 Clear A → 3 7 Run B → 4 8 - C →	%     4     0 sites       *]     # bases set     located       70     0 sites       set     match % set     cummulative
top middle bottom Protein Antibody DNA Primer Chip qPCR F E ► H ►	ne Cut DNA
'DNA_sickle_cell_control_mutated.gen' is the active DNA file.       M         The first 5 characters of the DNA sequence are: ccttt       M         This file contains 95006 characters in 2 separate numbered sequences - s       S         Click on (or drag over) the numbers below to see information for each sequence       T	ACGT 1907 GATC 0140 ACGCGT
Click on (of diag over) the numbers below to see information for each sequence         Imoriginal order       ⊠ Size         Size       Sequence          1       2	I <mark>Ч++++</mark> i Miul Leoeo∀ T TAA II+++
47499 base pairs	Second S
	9949100 Mst II

Click the Cut DNA button on the silver button bar and select with a single enzyme.



The digested file appears as a new line in the Opened & processed window, designated by a \*\*\* prefix and a + Mst II suffix. Click this line and then click the Load button to load the digested sequence into well one.



Click on the **Probe** button of the silver button bar and open the file **Probe for sickle cell 90 percent match**.

									Case	e It v	7					
		About	Notes	Clear	Protein	DNA	Site	Cut	Run	Gel	Options	Cases	Tutoria	als		
· · ·						_		1	1	] 5	Load	1.0 %		4	0 sites	
	Shift-click to open	multiple	e probe	files					2	_	Clear	[agaros	e] #	bases set	located	
Look in:	Probe example using sid	ckle cell	- 🎄						3	7	Run	60		100	0 sites	
	control mutated.gen							∙∥ ¤	4	8		minutes (	set m	atch % set	cummulative	
Enzyme MstII	-	ТХТ									PCR Pro	obe Enzyi	me Cut	DNA		
							s the ac	tive D	NA file.							
1						_	ice are:	ccttt								
File <u>n</u> ame:	Probe for sickle cell 90	percent n	natch.TXT		<u>O</u> pen	R.	eparate	e numb	pered se	equer	nces - see	below.				

Change the number in the match% field to **70**, then click the **Set match** % button.

<u>Note</u>: Settings below 50% can cause the program to temporarily freeze and may give unpredictable results.

	· · ·					÷			•						
							Case	lt v	7						
About	Notes	Clear	Protein	DNA	Site	Cut	Run	Gel	Options	s Cases	Tub	orials			
Probe: F	robe for s	ickle cel	l 90 perce	nt mato	h.TX	<b>-</b> [ (	1	5	Load	1.0 %	,	4		0 sites	
aagggg	axxxagtad	aggggg	atgggxgaa	aggoga	atcac		2		Clear	[agaros	:e]	#bases	set	located	
						<b>[</b> ] (	3	7	Run	60		70		0 sites	
						- [	4	8		minutes	set	match 9	ð set	cummulative	
		P		lle boti htibody		Prim	er Chip	- q	PCR Pr	obe Enzy	me	Cut DNA	 ] • н	et match % elp	

Click the Run button and select Run Southern or Western blot.

-

		Case It v	7				
About Notes Clear F	Protein DNA Site Cu	it Run Gel	Options	Cases Tub	orials		
Probe: Probe for sickle cell 9	0 percent match. TX 📥		Load	1.0 %	4	0 sites	
aaggggaxxxagtacagggggatg	gggxgaaaggcgatcac	2 6	Clear	[agarose]	#bases set	located	
		3 7	Run	60	70	0 sites	
	<b>_</b>	4 8	Run D	NA or protein ç	jel	ative	
to	op middle bottom		Run P	CR			
	tein Antibody DNA Pri	mer Chip gl	Run 1	Southern or We	estern blot 🔥		
					Ú.		
			Note:	Run menu abo	ve has more op	tions	

A Southern blot will appear with a single fragment showing, representing a fragment that bound to the probe at this match % setting. Note that there are 38 migration distances showing in the table to the right, indicating that other fragments are not visible since they did not bind to the probe.

θ														
	0		25	<b>← → ←</b>	50	Diff.display	75		100	Mig	ration d	listance	s for la	ne 1
	1				1					1	2	3	4	5
										26.50	26.50	29.00	29.00	29.50
	2									6	7	8	9	10
	3									29.50	30.75		31.50	31.50
										11	12	13	14	15
	4								_				36.50	
	5									16	17	18	19	20
									_				49.25	
	6									21	2	23	24	_25
	7												59.25	
										26	_ 27	28	29	
	8												62.25	
	1.0 %	agarose								31	32	33	34	35
Q	runtir	me = 60											76.50	91.75
Ŀ	100	D volts								36	37	38		
										91.75	96.50	96.50		
	60	Set	Run Re	load		Options Data	Photo	Clear H	lide	14 4	► FI	fragn	nents 1	- 38
							1							

Click on the fragment to turn it from black to red, and note that the fragment was the 20<sup>th</sup> one on the gel prior to running the blot procedure.

	) 25	<b>€ </b> ≯ → €	50	Diff.display	75	100	Mig	ration d	listance	s for la	ne 1
1			1				1	2	3	4	5
2			k				26.50	26.50 7	29.00 8	29.00 9	29.50
з							29.50	30.75 12	30.75 13	31.50 14	31.50 15
4								32.25	36.50	36.50	38.25
5							16 38.25	17 43.50	18 43.50	19 49.25	20 49.25

To see the original gel upon which the blot is based, click on the **Run** button and select **Timed run**.

<u>-</u>	1.0 % agarose runtime = 60 100 volts			31         32         33         34         35           75.50         75.50         76.50         76.50         91.75           36         37         38           91.75         96.50         96.50
	60 Set		Photo Clear Hide	fragments 1 - 38
	Large → small	Timed run Hour run	Analyze 0 Search 🕨	KII ► 20 g

"Ghost fragments" will appear on the gel, that would not be visible in reality without staining.



To stain the fragments, click the blue button (or the orange or black buttons).



All fragments in the lane will turn blue, indicating that they have been stained.



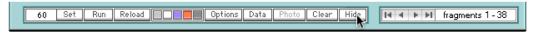
To re-run the blot to see which fragments were bound to the probe, use the **Run** button and select **Southern or Western blot**...

	60 Set	Run Reload C	ptions Data Photo Clear Hide 🛛 🖬 🖌 🕨 fragments 1 -	. 38
		Timed run		
	Large → small	Hour run	Analyze 0 Search M 4	<b>20</b>
Ľ	2345678	Southern or Western Blot 📐	20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38	$\leq$
	1401 L			

...and the fragment bound to the probe appears.

1		1			1	2	3	4	5
Ľ				2	26.50	26.50	29.00	29.00	29.50

Temporarily hide the blot by clicking the Hide button.



	Сору	Þ	
	Paste	Þ	
	Add	۶	
	Replace	۶I	
'DNA_sickle_cell_control_mutated.gen + M	Save	۱,	NA file.
	PCR results		
The first 5 characters of this file are: tcagg	Blot results	М	Put match % results into upper field

**Right-click** on the upper field and select **Blot results -> Put match % results into upper field**.

Actual match percentages (90%) appear in the field, indicating that both fragments 19 and 20 were bound to the probe. Note that the second and fourth lines are duplicates of the first and third lines, so they can be ignored. These results can be copied or saved by right-clicking.

			90.0 = % match for fragment 19 in lane 1
		Сору	20.0 – % match for hagment 12 in fane 1
	•	Paste	90.0 = % match for fragment 19 in lane 1
	•	Add	
	×	Replace	90.0 = % match for fragment 20 in lane 1
contents of upper field to Notepad		Save	90.0 – % match for fragment 20 in Jane 1
all sequences in upper field of main wind	•	PCR results	20.0 – 20 match for hagment 20 m lane f
a specific sequence in lower field of main	•	Blot results	
all sequences in upper field of main wind		Paste Add Replace Save PCR results	

Since only one band appeared after the Southern blot, fragments 19 and 20 must be at the same location, which can be verified by looking at the migration distances and clicking on the boxes associated with fragments 19 and 20.

1		1	2	3	4	5
	1	26.50	26.50	29.00 8	29.00 9	29.50 10
3		29.50	30.75	-	-	
		11	12	13	14	15
4		32.25	32.25	36.50	36.50 19	20
5		38.25	43.50	43.50	49.25	49.25
E F		21	- 22	23	24	25

To make sure that the probe match routine is working properly, set the match % to **91**...

3 7 Run 4 8	60 minutes set	91 match % set	0 sites cummulative	
	·	Se V He	et match %	F

...and no match will occur since 91 is higher than 90, the actual percentage required for a match with this probe and sequence.

1		1	2	3	4	5
		26.50	26.50	29.00	29.00	29.50

Set the match % back to 90...

_ 3 _ 7 Run _ 4 _ 8	60 minutes set	90 match % set	0 sites cummulative	
		Se V H	et match %	F

...and the black band reappears, indicating that the procedure is working properly.

1			1	2	3	4	5
Ľ			26.50	26.50	29.00	29.00	29.50

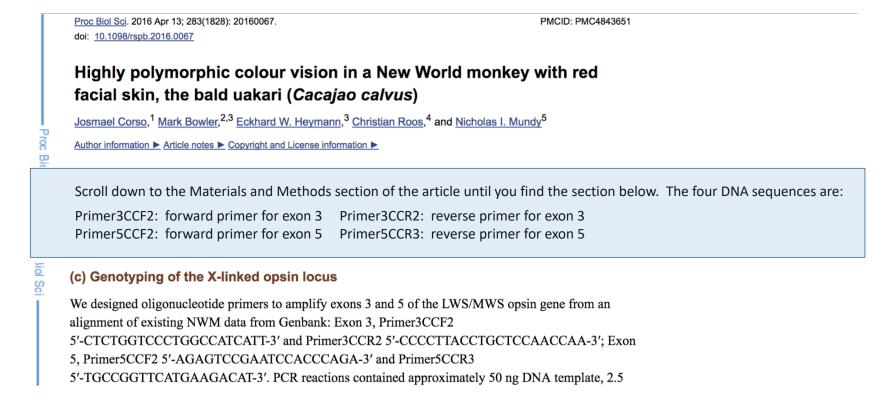
The previous probe searches assumed that there was a perfect match between the first 4 characters of the probe sequence and the hit site on the DNA sequence being searched, as shown in the **# bases** field below. This was done to speed up searching for existing Case It cases, so that is the default setting for the software. The value can be reduced down to zero, which is more realistic if you are using your own probe for research or case development purposes, but search times may increase substantially as described in the **Help** message below.

<u>Important</u>: The **# bases** box only applies to the Southern blot and dot blot routines. Values in this box have no effect on the general search and PCR routines described earlier in this tutorial, as those routines always search from the beginning to the end of sequences one character at a time, for greater realism and usefulness for research purposes. [Lower **match %** settings will slow down these routines, depending on the size of sequences being searched.]

4 0 sites Options		Quick help	Tutorials	Options	
#bases set located Search re	sults field 📃	This is the Data	Screen		
90 Set number of matching bases	<b>_</b>	This is the Data Scieen.			
match % V Help	Enter the number of bases that must match at the beginning of the probe sequence in order for the probe to bind to that site, and then click the 'set' button. Note: The default value of 4 gives faster results, but will miss matches if there is variation between the first 4 bases of the probe and the first 4 bases of the sequence to be searched. Setting the value to zero will detect any match equal to or greater than a particular setting, but may increase search time considerably,				ns a tions s ca lore It we nd al
_	depending on the length of the DNA sequ	uence[s] being s	earched.		

## Obtaining sequences for research purposes

A good place to start a research project is to find an article with published forward and reverse primers, for example the one below that analyzed exon 3 and exon 5 of the opsin gene, a gene involved with color vision in primates and many other animals (click here for a link to the full article.)



Nucleotides at position 180 on exon 3 and positions 277 and 285 on exon 5 help determine whether bald uakaris have dichromatic or trichromatic color vision, a characteristic with significant evolutionary and ecological consequences. [See the <u>Monkey Opsin section of the Evo-Ed website</u> for a detailed explanation of the molecular biology of color vision, as it relates to the evolution and ecology of New World and Old World primates.]

The first step is to create a Primer file that can be read by Case It. Highlight the **forward primer sequence for exon 5** and right-click to copy it to your computer's clipboard. <u>Note</u>: Although Case It v7 is capable of multiplex PCR, we will only use the primers for exon 5 for this example.

## (c) Genotyping of the X-linked opsin locus

We designed oligonucleotide primers to amplify exons 3 and 5 of the LWS/MWS opsin gene from an alignment of existing NWM data from Genbank: Exon 3, Primer3CCF2 5'-CTCTGGTCCCTGGCCATCATT-3' and Primer3CCR2 5'-CCCCTTACCTGCTCCAACCAA-3'; Exon 5, Primer5CCF2 5'-AGAGTCCGAATCCACCCAGA 5'-TGCCGGTTCATGAAGACAT-3'. PCR reactions contain

Click the **Notes** menu of Case It v7 and select **Show**, which causes a simple **Notepad** text editor to appear.



**Right-click** in the Notepad and select **Paste -> clipboard into Notepad at cursor location**.



The forward primer for exon 5 is now the first line in the Notepad.

		Notepad.dir	
	AGAGTCCGAATCCACCCAGA		
			-
Back to beginning			

Go back to the article and select and copy the **reverse primer** sequence for exon 5.

alignment of existing NWM data from Genbank: Exon 3, Primer3CCF2	
5'-CTCTGGTCCCTGGCCATCATT-3' and Primer3CCR2 5'-CCCCTTACCTGCTCCA	ACCAA-3'; Exon
5, Primer5CCF2 5'-AGAGTCCGAATCCACCCAGA-3' and Primer5CCR3	
5'-TGCCGGTTCATGAAGACAT-2' PCP reactions contained annrovimately 50 ng DN	IA template, 2.5
mM each dNTP, 1× PCR buffer, 1.	0.25 µM each

Position your cursor at the beginning of the second line of the Notepad, then **Paste** in the reverse primer sequence so that the forward primer is line 1 of the file, and the reverse primer is line 2.

Notepad.dir	
AGAGTCCGAATCCACCCAGA	-
Сору	_
Paste 🕨 clipboard into Notepad at cursor location	
Convert  Cipboard into Notepad replacing all contents of upper field	
Save	

Additional forward and reverse primers could be pasted in, for example the FP and RP for exon 3, but we will not do this here. The general rule is to add each primer set as additional lines in the file (e.g. lines 3 and 4 for the next set, lines 4 and 5 for the set after that, and so forth.) See the STD case for an example of how multiplex PCR works in Case It v7.

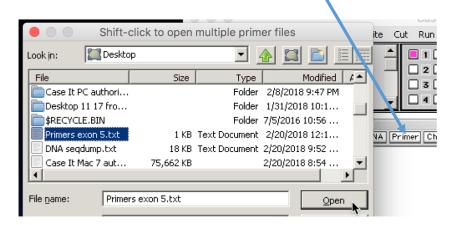
	Notepad.dir	
AGAGTCCGAATCCACCCAGA TGCCGGTTCATGAAGACAT		
		<b>_</b>

Use the **Save** command of the **Notes** menu, and save the file after renaming it **Primers exon 5** to the desktop or another location on your computer. Note that files containing primers must always include the word **Primers** at the beginning of the filename.

							Case	e It v7	7
About	Notes	Clear	Protein	DNA	Site	Cut	Run	Gel	Options
	Shov Oper						1	5	Load . Clear
	Save Foot		•				3	7	Run

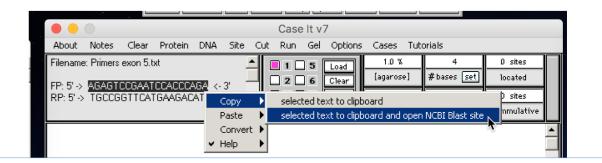
	• • •		Notepad	
	Save in:	esktop	🖸 🖾 🗹	
	====	E.BIN	🛅 Case It v7 PC	
1	2017 tax	es	📄 Desktop 11 17 from Windows 7	
	Barbersh	юр	🛅 Desktop 12 17 18	
1	Case It		🛅 Desktop 2 17 18	
i	Case It F	C authoring 10 25 13	🛅 Test sequences PC	
	Case It v	/607	📃 DNA seqdump.txt	
	File <u>n</u> ame:	Primers exon 5	▼ <u>S</u> a	ve

Click the <u>Primer</u> button on the silver button bar and open the **Primer exon 5.txt** file that you just saved. It will appear in the Opened & processed window, and also in the <u>gray field</u> of the main window of Case It.





We now need to obtain DNA sequences that contain the forward primer sequence for exon 5 of the opsin gene. **Double-click** on the forward primer in the gray field to highlight it, then **right-click** in the field and select **Copy -> selected text to clipboard and open NCBI Blast site**.



Your default web browser will automatically open to the BLASTN page of the NCBI web site. **Right-click** in the **Enter Query Sequence** box on this page, then **paste** the forward primer sequence into the box. **You may have to paste more than once** for this to work.

blastn <u>blastp</u> <u>blastx</u> <u>tblastn</u> <u>tblastx</u>	BLASTN programs search	nucleotic	le databases using a nucleo
Enter accession number(s), gi(s), or FASTA seque	ence(s) 😡	<u>Clear</u>	Query subrange 😡
Undo			From
Cut			То
Paste			

Scroll down on the page and click the **BLAST** button...



Search database Nucleotide collection (nr/nt) using Blastn (Optimize for somewhat simil

Once the page appears showing results, scroll down on the page and select the first 10 hits (we are doing this strictly to illustrate the procedure; in reality you would want to carefully examine hits and select those most appropriate for the hypothesis you are testing).

Alignments Bownload v GenBank Graphics Distance tree of results			
Description	Max score	Total score	Quer cove
Leontopithecus chrysomelas 563nm X-linked opsin genes, partial cds	42.1	42.1	100%
2 Callithrix pygmaea 563nm X-linked opsin genes, partial cds	42.1	42.1	100%
Callimico goeldii 563nm X-linked opsin genes, partial cds	42.1	42.1	1009
Callimico goeldii 543nm X-linked opsin genes, partial cds	42.1	42.1	1009
Leontopithecus chrysomelas 556nm X-linked opsin genes, partial cds	42.1	42.1	1009
Callithrix jacchus 556nm X-linked opsin genes, partial cds	42.1	42.1	1009
Callithrix geoffroyi 563nm X-linked opsin genes, partial cds	42.1	42.1	1009
Saguinus oedipus 563nm X-linked opsin gene, partial cds	42.1	42.1	1009
Saguinus midas 556nm X-linked opsin gene, partial cds	42.1	42.1	1009
Saguinus fuscicollis 563nm X-linked opsin gene, partial cds	42.1	42.1	1009

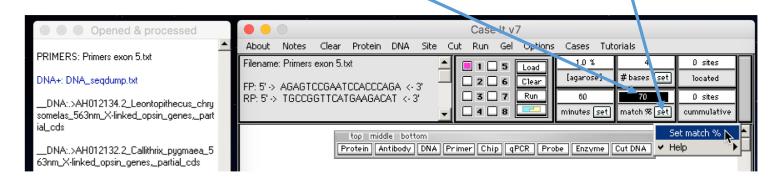
Click the **Download** button using the default setting (download the complete FASTA sequence). Click the **Continue** button, select **Save File**, and click **OK**. You will then be able to save the file to your desktop or another location on your computer.

●FASTA (complete sequence) ●FASTA (cliqued sequences)	You have chosen to open: seqdump.txt which is: TXT file from: https://www.ncbi.nlm.nih.gov What should Firefox do with this file?
<ul> <li>FASTA (complete sequence)</li> <li>FASTA (aligned sequences)</li> <li>GenBank (complete sequence)</li> <li>Hit Table (text)</li> <li>Hit Table (CSV)</li> </ul>	which is: TXT file from: https://www.ncbi.nlm.nih.gov
FASTA (aligned sequences)       nm X-linked opsin genes, pa         GenBank (complete sequence)       ad opsin genes, partial cds         Hit Table (text)       opsin genes, partial cds         Hit Table (CSV)       opsin genes, partial cds	from: https://www.ncbi.nlm.nih.gov
GenBank (complete sequence)     ad opsin genes, partial cds       Hit Table (text)     opsin genes, partial cds       Hit Table (CSV)     opsin genes, partial cds	What should Firefox do with this file?
Hit Table (text)     opsin genes, partial cds       Hit Table (CSV)     opsin genes, partial cds	
opsin genes, partial cds	Open with TextEdit (default)
Text	• Save File
XML nm X-linked opsin genes, pa	Do this automatically for files like this from now on.
ASN.1     opsin genes, partial cds       Continue     Cancel       d opsin genes, partial cds	Cancel OK N

The filename will be saved by the NCBI site with the name **seqdump.txt**. After locating the downloaded file, **change the name of the file**, making sure to begin the filename with the letters **DNA**. Otherwise, Case It v7 will not recognize it as a DNA file. In this example, the file has simply been name **DNA seqdump.txt**, but you should give it a more descriptive name that retains the DNA prefix.



Use the **DNA** button on the silver button bar to open the file. Enter <u>70</u> in the **match** % field and click the <u>Set</u> button.



Shift-click in the Opened & processed window to select all 10 lines (only 3 of which are shown here), then click the Run button and select Run PCR.

Opened & processed	Case It v7
	About Notes Clear Protein DNA Site Cut Run Gel Options Cases Tutorials
PRIMERS: Primers exon 5.txt	Filename: Primers exon 5.txt 🔺 🔲 1 🗆 5 🛛 Load 🔰 1.0 % 4 🛛 Sit
DNA+: DNA_seqdump.txt	FP: 5' -> AGAGTCCGAATCCACCCAGA <- 3'
DNA:.>AH012134.2_Leontopithecus_chry	RP: 5'-> TGCCGGTTCATGAAGACAT <-3' 3 3 7 Run 60 70 0 sit
somelas_563nm_X-linked_opsin_genes,_part	Run DNA or protein gel
ial_cds	top middle bottom Run PCR
DNA:.>AH012132.2_Callithrix_pugmaea_5	Protein Antibody DNA Primer Chip gr Run Southern or Western blot
63nm_X-linked_opsin_genes,_partial_cds	
DNA:.>AH012130.2_Callimico_goeldii_563	Note: Run menu above has more options
nm_X-linked_opsin_genes_partial_cds	

PCR products appear in the O&P windows, designated by arrows at the beginning of each filename. **Right-click** on the large white field and select **PCR results -> Put match % results into upper field**.

	top middle bottom Protein Antibody DNA Primer Chip gPCR Probe Enzyme Cut DNA
Opened & processed	Case It v7
<ul> <li>&gt;&gt;&gt;AH012134.2_Leontopithecus_chrysomelas</li> <li>_563nm_X-linked_opsin_genes_partial_cds_</li> <li>(PCR)</li> <li>-&gt;</li> <li>&gt;&gt;AH012132.2_Callithrix_pygmaea_563nm_X-linked_opsin_genes_partial_cds_(PCR)</li> </ul>	About       Notes       Clear       Protein       DNA       Site       Cut       Run       Gel       Options       Cases       Tutorials         Filename:       Primers exon 5.txt       Image: Strategy and the strategy and
> >AH012130.2_Callimico_goeldi_563nm_X-lin ked_opsin_genes,_partial_cds_(PCR) > >AH012129.2_Callimico_goeldi_543nm_X-lin ked_opsin_genes,_partial_cds_(PCR) > >AH012121.2_Leontopithecus_chrysomelas _556nm_X-linked_opsin_genes,_partial_cds_ (PCR)	Copy Paste Open Add '>AH012141.2_Saguinu Replace Save The first 5 characters of PCR results Put match % results into upper field Blot results Y Help This file contains 223 characters in a continuous sequence.
> ->AH012119.2_Callithrix_jacchus_556nm_X-li nked_opsin_genespartiaLcds_(PCR)	■ Original order       X Size       Sequence
> >AH012117.2_Callithrix_geoffroyi_563nm_X-li nked_opsin_genes,_partiaL_cds_(PCR)	222 base pairs
> >AH012149.2_Saguinus_oedipus_563nm_X- inked_opsin_gene,_partiaLcds_(PCR) >	
Back to beginning	Genbank Lab Bench Data Screen ELISA Dot Blot 96-well PCR Chip qPCR Sequence analysis Opened Loaded Ge1 / blot Help

For the first PCR product, there was a 100% match for the forward primer (FP), and a 94.74% match for the reverse primer. Note that the forward primer in the white field is the same as the forward primer in the gray field, whereas the reverse primer (RP) in the white field is the **reverse complement** of the reverse primer sequence in the gray field. When Case It v7 opens primer files, it automatically converts the reverse primer into its reverse complementary sequence. To see why this is necessary, click the **Tutorials** menu and watch the video tutorial called **Show PCR primer sites**.

Case it v7      About Notes Clear Protein DNA Site Cut Run Gel Options Cases Tutorials      About Notes Clear Protein DNA Site Cut Run Gel Options Cases Tutorials      About Notes Clear Protein DNA Site Cut Run Gel Options Cases Tutorials      Filename: Primers exon 5.txt     FP: 5' -> AGAGTCCGAATCCACCCAGA <- 3'     Rev 5' -> TGCCGGTTCATGAAGACAT <- 3'     AH012134.2_Leontopithecus_chrysomelas_563nm     X-linked_opsin_genes_partial_cds (PCR)      AH012134.2_Leontopithecus_chrysomelas_563nm     X-linked_opsin_genes_partial_cds
inked_opsin_gene_partial_cds DNA::>AH012141.2_Saguinus_fuscicollis_563nm X-linked_opsin_gene_partial_cds > >AH012134.2_Leontopithecus_chrysomelas_563nm
DNA::>AH012141.2_Saguinus_fuscicollis_563nm        NA::>AH012141.2_Saguinus_fuscicollis_563nm        X-inked_opsin_gene_partial_cds        >        >         >AH012134.2_Leontopithecus_chrysomelas_563nm         >AH012134.2_Leontopithecus_chrysomelas_563nm
DNA:.>AH012141.2_Saguinus_fuscicollis_563nm X-linked_opsin_genepartial_cds FP: 5' > AGAGTCCGAATCCACCCAGA <-3' > >AH012134.2_Leontopithecus_chrysomelas_563nm >AH012134.2_Leontopithecus_chrysomela
X-linked_opsin_gene_partial_cds       FP: 5'-> AGAG ICUGAA ICUACULAGA <-3'
> >AH012134.2_Leontopithecus_chrysomelas_563nm >AH01212134.2_Leontopithecus_chrysomelas_563nm >AH012134.2_Leont
>AH012134.2_Leontopithecus_chrysomelas_563nm
E ' E ' E B ADUUZIJA Z LEODIODINECUS CONSODEJAS OBJODINECUS DESDI DENES, DANIAL COS
[FP match % , RP match %]
> >AH012132.2_Callithrix_pugmaea_563nm_X-linked_ ["100.00 FP", "94.74 RP"]
opsin_genes_partial_cds_(PCR) AGAGTCCGAATCCACCCAGAFP
AGAGTCCGAATCCACCCAGADNA
-> >AH012130.2_Callimico_goeldi_563nm_X·linked_o
psin_genes_partial_cds_(PCR) ATGTCTTCATGAACCGGCA-RP
> >AH012129.2_Callimico_goeldii_543nm_X-linked_o
psin_genes,_partial_cds_(PCR) PCR product 1 length = 222

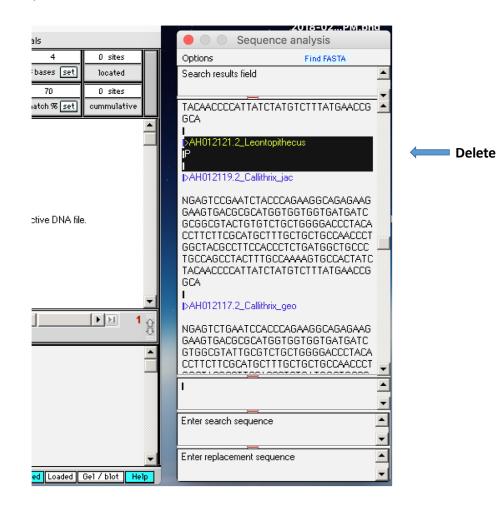
You can experiment with re-running PCR with different match percent settings, to find an optimal setting for particular combinations of primers and sequences being searched during the PCR procedure. Settings that are **too high** may miss FP and/or RP sites, whereas settings that are **too low** may match with spurious FP and RP sites, rendering the results meaningless.

<u>Note</u>: The default match % setting is 100% because existing cases were originally developed with primers that matched exactly, and earlier versions of Case It did not have the capability of finding locations with less than a 100% match. With Version 7, it is now possible to use any % match setting, making the simulation more useful for research purposes.

These PCR products can now be analyzed via multiple alignment and tree-building. First, **shift-click** to highlight all PCR products (lines preceded by arrows) in the **Opened & processed** window, then click the **Analyze** button at the bottom of that window and select **Export -> Add selected file(s)**...

	Case It v7
Opened & processed	About Notes Clear Protein DNA Site Cut Run Gel Options Cases 1
<ul> <li></li></ul>	Filename: Primers exon 5.txt         FP: 5' -> AGAGTCCGAATCCACCCAGA <- 3'         RP: 5' -> TGCCGGTTCATGAAGACAT <- 3'
d_opsin_genes_partial_cds_(PCR) > >AH012129.2_Callimico_goeldii_543nm_X-linte d_opsin_genes_partial_cds_(PCR) > >AH012121.2_Leontopithecus_chrysomelus_55	
6nm_X-linked_opsin_genes,_partiaLcds_PCR) ···	'>AH012132.2_Callithrix_pygmaea_563nm_X-linked_opsin_genes,_partial_cds_(PCR)' is I The first 5 characters of this sequence are: NGAGT
>AH012119.2_Callithrix_jacchus_556nh_X-link ed_opsin_genes,_partial_cds_(PCR)	This file contains 223 characters in a single sequence.
<ul> <li>&gt;AH012117.2_Callithrix_geoffroyi_563nm_X-link</li> <li>ed_opsin_genes_partial_cds_(PCr)</li> <li>-&gt;</li> <li>&gt;AH012149.2_Saguinus_oedipus_563nm_X-lin</li> </ul>	
<pre>&gt;AH012147.2_Saguinus_oedupus_osonm_&lt;-iin ked_opsin_gene_partial_cds_PCR) &gt;AH012147.2_Saguinus_migas_556nm_X-linke</pre>	222 base pairs
d_opsin_gene,_partial_cds_(PCR) > >AH012141.2_Saguinus_tuscicollis_563nm_X-li nked_opsin_gene,_partit_Lcds_(PCR)	
Quick Load / Run Analyze ? 11 files opened	Genbank Lab Bench Data Screen ELISA Dot Blot 96-well PCR Chip qPCR Sequence analysis
Align / tree 🕨 Replace conten	e[s] in the O&P window to contents of the Export field [in the S.A. window] ts of the Export field with selected file[s] in the O&P window of Export Field to the desktop

The PCR products have been added to the Export field of the Sequence Analysis window. Scroll through the Export field to verify that PCR products were generated for all files except **>AH012134.2\_Leontopithecus**. Select the lines shown highlighted below and delete them by selecting them and hitting the **Delete** key of your keyboard.



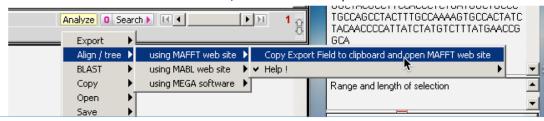
As an aside, it is important to verify that we really have isolated a region associated with exon 5 of the opsin gene. One way to do this is to click on lines in the **Opened & processed** window representing PCR products (lines preceded by arrows), to see if they are all about the same size.

For this example, all PCR products have 222 characters, suggesting that the match percent setting (70%) was about right for isolating exon 5.

<u>Note</u>: If there is a slight discrepancy in the number of characters in the white field above compared to the gray field below (e.g. 223 vs 222), go by the number in the gray field.

> >AH012130.2_Callimico_goeldii_563nm_X-linke d_opsin_genes_partial_cds_(PCR) > >AH012129.2_Callimico_goeldii_543nm_X-linke d_opsin_genes_partial_cds_(PCR)	'>AH012129.2_Callimico_goeldii_543nm_X-linked_opsin_genes,_partiaL_cds_(PCR)' is the active DNA file. The first 5 characters of this sequence are: AGAGT This file contains 223 characters in a single sequence.	•
> >AH012121.2_Leontopithecus_chrysomelas_55 6nm_X-linked_opsin_genes_partial_cds_(PCR)	Original order     X Size     Sequence     Analyze     Search     X       1     1	1 ☆
	222 base pairs	<b></b>

Previously (slides 24 and 25) we used MEGA software to build a tree, but for this example we'll click on the **Analyze** button and select **Align / tree -> using MAFFT web site -> Copy Export Field to clipboard and open MAFFT web site**.



Note: This particular web site is more reliable than the MABL site, which can be unresponsive at times.

Your default browser will automatically open to the first page of the MAFFT site. **Paste** contents of the clipboard into the Input field.

MAFFT version 7

Multiple alignment program for amino acid or nucleotide sequences

Download version Mac OS X	For a large number of short sequences, try an experimental service (2017/Jul).
<u>Windows</u> Linux	Multiple sequence alignment and NJ / UPGMA phylogeny
Source	
Online version	
Alignment	Input:
mafftadd	Paste protein or DNA sequences in fasta format. <u>Example</u>
Merge	
Phylogeny	CCAGCCTACTTTGCCAAAAGTGCCACTATCTACAACCCCATTATCTATGTCTTTATGAACCGGCA
<u>Rough tree</u>	
Merits / limitations	CCAGCCTACTTTGCCAAAAGTGCCACTATCTACAACCCCATTATCTATGTCTTTATGAACCGGCA
Algorithms	
Tips	
Benchmarks	CCAGCCTACTTTGCCAAAAGTGCCACTATCTACAACCCCGTTATCTATGTCTTTATGAACCGGCA
<u>Feedback</u>	
	CCAGCCTACTTTGCCAAAAGTGCCACTATCTACAACCCCGTTATCTATGTCTTTATGAACCGGCA
	CCAGCCTACTTTGCCAAAAGTGCCACTATCTACAACCCCATTATCTATGTCTTTATGAACCGGCA

Scroll down further on that web page and click the **Submit** button. Note that we will be using the default settings to build the tree.

MAFFT version	on 7 program for amino acid or nucleotide sequences
Download version Mac OS X Windows	Notify when finished (optional; recommended when submitting large data): Email address:
Linux Source Online version	Submit Reset

On the next page that appears, click **Phylogenetic tree**.

	GUIDANCE2 computes the residue-wise confidence scores and extracts well-aligned residues.
LAST hits (score>39) between the top	Refine dataset
sequence and the others.	Phylogenetic tree

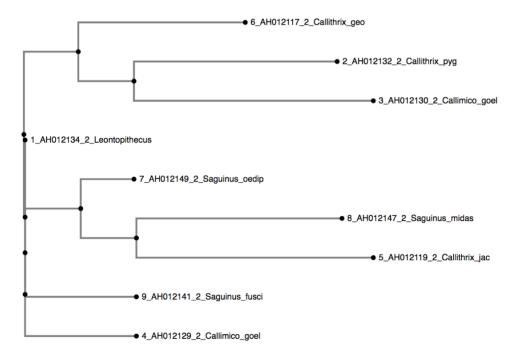
On the page that appears after that, click Go.

sequence and the others.	NJ or UPGMA tree (β)
Open all plots	9 sequences, 222 total sites, 221 gap-free sites, 221 <u>conserved sites</u>
Numero and Annual Annua	Gol Reset

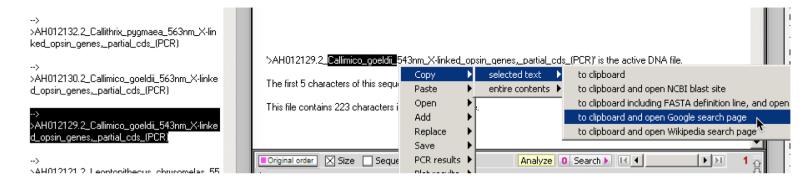
Finally, click **View tree on Phylo.io** (or one of the other options on that page).

sequence and the others.	Result (Phylo.io 1.0.0)
others.	Phylo.io runs on any modern browser.
Open all plots	Thylo. Io fulls of any modern browser.
Trachill - 18	View tree on Phylo.io

...and the tree appears. Whether or not this particular tree has any significance depends a number of factors, including whether the default settings were appropriate for this analysis and whether the BLAST results selected were appropriate for the hypothesis being tested. The tree is shown here simply to demonstrate the procedure for tree-building using the MAFFT web site.



There are other features of Case It v7 that are useful for gathering information for research. For example, you can copy text (such as a scientific name) from the large white field of the Data Screen and automatically open the search page of **Google** or **Wikipedia**, then paste the scientific name and search. This is a quick way to display information about that species, when gathering information for testing hypotheses on ecological, behavioral, evolutionary, or other topics. **This concludes the tutorial**.



#### Showing results for *Callimico Goeldii* Search instead for Callimico\_goeldii

#### Callimico goeldii (Callimico, Goeldi's Marmoset, Goeldi's Monkey ... www.iucnredlist.org/details/3564/0 •

Range Description: **Callimico goeldii** occurs in the upper Amazon from the Rio Caquetá in Colombia, south through the Peruvian Amazon and the extreme western Amazon of Brazil into the Pando region of northern Bolivia (Hernández-Camacho and Barriga-Bonilla 1966; Hernández-Camacho and Cooper 1976; ...

Taxonomy · Assessment Information · Geographic Range · Population

### Goeldi's marmoset - Wikipedia

https://en.wikipedia.org/wiki/Goeldi%27s marmoset ▼



# Goeldi's marmoset

The Goeldi's marmoset or Goeldi's monkey is a small, South American New World monkey that lives in the upper Amazon basin region of Bolivia, Brazil, Colombia, Ecuador, and Peru. Wikipedia