

**A premature stop codon in the CYP2C19 gene explains the unexpected sensitivity of vultures to diclofenac toxicity.**

Emmanuel Oluwasegun Adawaren<sup>1</sup>, Christiaan Labuschagne<sup>2</sup>, Aron Abera<sup>2</sup>, Vinny Naidoo<sup>1</sup>.

<sup>1</sup>Department of Paraclinical Science, Faculty of Veterinary Science, University of Pretoria, Gauteng, SOUTH AFRICA

<sup>2</sup>Iqaba Biotechnical Industries (Pty) 525 Justice Mahomed St, Muckleneuk, 0002Pretoria, Gauteng, SOUTH AFRICA

**Transcriptome Assembly**

Figure 1: Transcriptome Analysis for vulture kidney. Local Blast analysis against the Golden Eagle CYP2C19-like mRNA sequence, resulted in one similarity being identified in the Trinity aligned database. The result from the programme is presented in **A** below and the translation in **B** below, with the abnormal stop codon already identifiable

**A**

```
>TRINITY_DN10911_c0_g2_i3 len=1542 path=[4885:0-46 4914:47-401
      4912:402-523 4894:524-565 4895:566-589 4917:590-711
      4915:712-874 4913:875-981 4908:982-1366 4892:1367-1541]
      [-1, 4885, 4914, 4912, 4894, 4895, 4917, 4915, 4913,
      4908, 4892, -2]
      Length = 1542

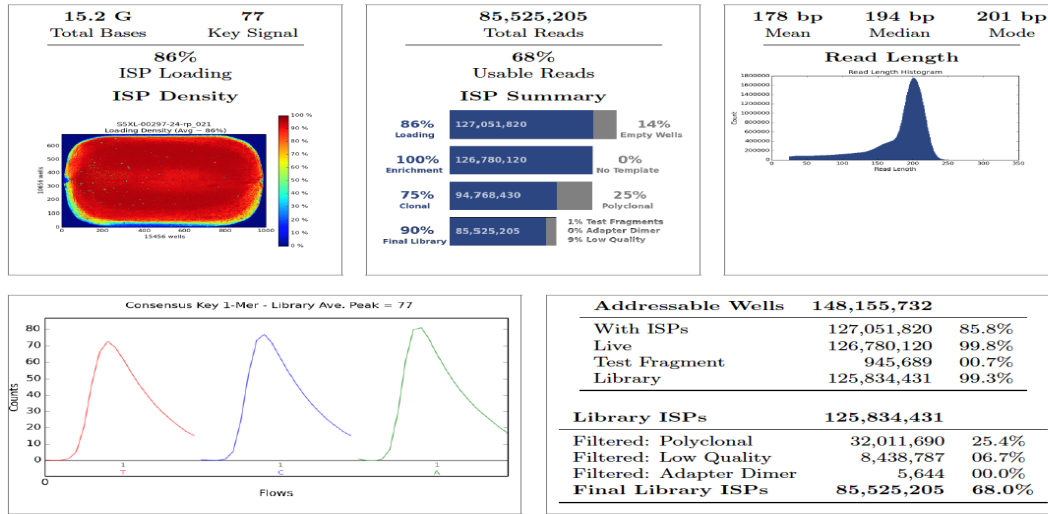
Score = 1132 bits (571), Expect = 0.0
Identities = 1083/1253 (86%), Gaps = 3/1253 (0%)
Strand = Plus / Minus
1      atgatgCGGTGaaagaagccttggttgatcacgCGGacaagtttgctGCCagaggacgca 60
61      tgccaataggGacagggctaacaatggattagggattgttttagcaacaacaaggagt 120
121     ggttacaagtccgTCGGtttgctctcagtactctgcgcaactttggaatggggaaaagga 180
181     gcattgaagagaggatccaggaggaaactgagtacttgctagaagagatcaacaaaaca 240
241     agggAACaccttttgaccaaccttcatgctgagctgtgctgtctccaatgtcatatgct 300
301     ccatcgtctttgggaaacgatatgactatgaagacaagaagttcctggccctgatgagca 360
361     acatgaacaacatctttgagctgatgaaactccccatggggacagctctaccagatgttct 420
421     caaagatcctggattacttgcccttttgGCCacacacaaaatattcacagaatttgatg 480
481     ctctaaaagcctttgCGtcagaggaggtgaagacgcaccaatcctccctagatcccagct 540
541     cccccaggatttcatcgactgcttccctcagaaaaatgcaggaggagaaagagcatcca 600
601     attccagtttccacatgaagaacctgataaccagcaccttcgacttgttcattactggaa 660
761     ctgagacaattagcaccactgcaagatatgggcttctgcttcttctcaaataccgaaga 720
721     tacaagagaaagttcaagaagagattgaccaggtagtgggatgatcacgaagaccttggtg 780
781     tggctcatcagaccagatgccctacacagatgcagtggtccatgaaatccagcgcttca 840
841     tctccctcatccccctgggtctccctcagctgtgaccaaagacaccggcttcagagagt 900
901     atgtcattcccaagggcaccacagtctttcccatcctcagttctgtcctccatgacagta 960
961     aagagtttccaaacccaaatgagttcaaccctggacatttcttgaatgacaatggcacct 1020
1021    ttaggaagagtgagttcttcatgcccttctcagcaggggaagcgaatatgcctggagagg 1080
1081    gcctggcagcatggagatattcttagtcataaccaccatcctgcagaactttaccttga 1140
1141    agcctgttgcgacccccaggaactcaacataaccccgcactgagtgggacaggcaacg 1200
1201    tacctcctgcctaccagctctgtgctctcccccgctgaaaagcacaaaacctc 1253
```

**B**

gatgcggtgaaagaagccttggttgatcacgcggacaagtttgctgccagaggacgcatg 60  
D A V K E A L V D H A D K F A A R G R M  
cfaataggagacagggctaacaatggattagggattgttttagcaacaacaaggagtgg 120  
P I G D R A N N G L G I V F S N N K E W  
ttacaagtccgctcggtttctctcagtactctgcgcaactttggaatggggaaaaggagc 180  
L Q V R R F A L S T L R N F G M G K R S  
attgaagagaggatccaggaggaaactgagtacttgctagaagagatcaacaaaacaag 240  
I E E R I Q E E T E Y L L E E I N K T K  
ggaacaccttttgacccaaccttcatgctgagctgtgctgtctccaatgtcatatgctcc 300  
G T P F D P T F M L S C A V S N V I C S  
atcgtctttgggaaacgatatgactatgaagacaagaagttcctggccctgatgagcaac 360  
I V F G K R Y D Y E D K K F L A L M S N  
atgaacaacatctttgagctgatgaactccccatggggacagctctaccagatgttctca 420  
M N N I F E L M N S P W G Q L Y Q M F S  
aagatcctggattacttgccctttggcccacacaacaaaatattcacagaatttgatgct 480  
K I L D Y L P F G P H N K I F T E F D A  
ctaaaagcctttgctcagaggaggtgaagacgcaccaatcctccctagatcccagctcc 540  
L K A F A S E E V K T H Q S S L D P S S  
ccccaggatttcatcagctgcttccctcagaaaaatgcaggaggagaaagagcatcccaat 600  
P Q D F I D C F L R K M Q E E K E H P N  
tccagtttccacatgaagaacctgataaccagcaccttcgacttgttcattactggaact 660  
S S F H M K N L I T S T F D L F I T G T  
gagacaattagcaccactgcaagatatgggcttctgcttcttctcaaatacccgaagata 720  
E T I S T T A R Y G L L L L L K Y P K I  
caagagaaagttcaagaagagattgaccaggtagtggaatgatcacgaagaccttggtg 780  
Q E K V Q E E I D Q V V G - S R R P C V  
gctcatcagaccagatgccttacacagatgcagtggtccatgaaatccagcgccttcatc 840  
A H Q T Q M P Y T D A V V H E I Q R F I  
tccctcatccccctgggtctccctcagctgtgaccaaaagacaccggcttcagagagtat 900  
S L I P L G L P H A V T K D T G F R E Y  
gtcattoccaaagggcaccacagctcttcccatcctcagttctgtcctccatgacagtaaa 960  
V I P K G T T V F P I L S S V L H D S K  
gagtttccaaacccaaatgagttcaaccctggacatttcttgaatgacaatggcaccttt 1020  
E F P N P N E F N P G H F L N D N G T F  
aggaagagtgagttcttcatgcccttctcagcaggaagcgaatatgccctggagagggc 1080  
R K S E F F M P F S A G K R I C P G E G  
ctggcacgcatggagatattcttagtcataaccaccatcctgcagaactttaccttgaag 1140  
L A R M E I F L V I T T I L Q N F T L K  
cctgttgctcagccccaggaactcaacataacccccgacactgagtgaggacaggcaacgta 1200  
P V V D P Q E L N I T P T L S G T G N V  
cctcctgctaccagctctgtgctctccccgctgaaaagcacaacacctc  
P P A Y Q L C A L P R - K A Q N L

Figure 2: Cape vulture next generation genome sequencing summary report

Run Summary



Barcode Name	Sample	Bases	$\geq Q20$	Reads	Mean Read Length
No barcode	none	547,156,409	423,328,218	2,913,204	188 bp
IonXpress_001	rp_021_1	14,662,043,279	11,293,276,080	82,602,222	178 bp

## Quantification of *CYP2C19* expression through Real-Time RT-qPCR

### **RT-qPCR Material and Methods**

Liver samples collected from the same bird used for the transcriptome analysis. RNA was isolated from *G. africanus* liver tissue samples preserved in RNAlater. Approximately 100 mg of tissue was homogenized in Eppendorf tubes using a mini glass pestle and liquid nitrogen. The RNA was subsequently isolated with Trizol according to the manufacturer's instruction (Invitrogen). RNA quality and quantity was assessed with an Xpose spectrophotometer (Trinean). The RNA was then DNase-1 treated (Thermo Scientific), and cDNA synthesized using a Revert-Aid cDNA synthesis kit as 10 µl reactions containing 1 µg of total RNA (Thermo Scientific). The expression of *CYP2C19* was assessed through Real-Time RT-qPCR using a Step-One Plus thermal cycler (Applied Biosystems). Beta-actin (*ACTB*) was applied as housekeeping gene. The PCRs were performed as 10 µl reactions containing 5 µl SYBR Green JumpStart Taq Readymix (Sigma), 1 µM of each primer and 4 µl of template cDNA. Each PCR programme included an enzyme activation step (95°C, 5 min) followed by 40 cycles including denaturing (95°C, 15 sec), annealing (58 to 62°C) and elongation (72°C, 30 sec) steps. All reactions were run in triplicate. Primers targeting *G. africanus CYP2C19* were designed using Geneious Prime (Biomatters) (Table 1).

Table 1. Sequences and characteristics of primers applied for Real-Time RT-qPCR and Sanger sequencing.

<b>Primer</b>	<b>Forward</b>	<b>Reverse</b>
Pair 1	CAGACGTGTGCTCTCCGAT	AAACTTGTCCGCGTGATCAA
Pair 2	CCTTCATGCTGAGCTGTGCT	CTGGTAGAGCTGTCCCATG
Pair 3	CAGAGGAGGTGAAGACGCAC	ACAAGTCGAAGGTGCTGGTT
Pair 4	GGCTCATCAGACCCAGATGC	GGAAAGACTGTGGTGCCCTT
Pair 5	TGCGTCAGAGGAGGTGAAGA	ACTGAGGATGGGAAAGACTG
<i>G. fulvus</i> -ACTB	CTATCCAGGCTGTGCTGTCC	TGAGGTAGTCTGTCAGGTCAGG

Suitable annealing temperatures per primer pair were determined through gradient PCRs using 10 ng of template per reaction. Melt-curve analysis was furthermore applied to confirm the amplification of a single product. The quality of primers was further assessed based on 7-point 2-fold serial dilution of cDNA (20 ng to 0.312 ng per reaction) applied to calculate PCR efficiency using the standard curve method. The PCR products representing two individuals were sequenced through bi-directional using Sanger sequencing by a commercial laboratory (Inqaba Biotec). The sequencing reads were aligned, and consensus sequences evaluated using NCBI Blast.

### **RT-qPCR Results**

The expression of *CYP2C19* could successfully be evaluated using two independent primer sets with PCR efficiency values being within the acceptable limits of 90 to 105% (Bustin et al. 2009). The fragments amplified matched the *CYP2C19* transcript and portion of the predicted open reading frames described in this paper (Figure 4).

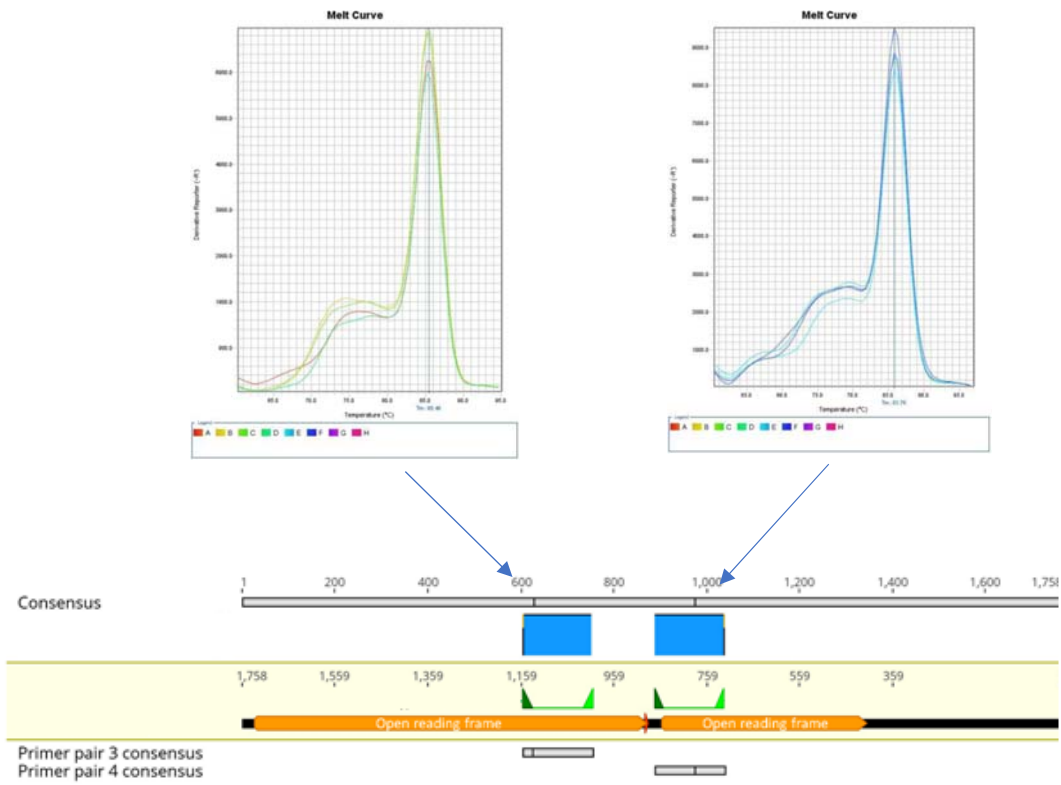


Figure 3: Visual illustration of the sequenced products of two RT-qPCR primer pairs in reference to the *Gyps coprotheres* CYP2C19 transcript identified through RNA.

Pair_3	-----	0
Pair_4	-----	0
Vulture	ATGGAGCTCCTGGGAGCAGCCACTGTTGTCCTCCTGGTTTGCATGCTTGCTGCTATCC	60
Pair_3	-----	0
Pair_4	-----	0
Vulture	TTTCGAGCATGGAGAGGGAGGTCTGGAAAGGGGAAGATGCCTCCAGGACCGGCTCCCCTT	120
Pair_3	-----	0
Pair_4	-----	0
Vulture	CCCATCCTAGGTAATGTGCTGCAAGTCAAACCAAAGAACTTGGCCAAAACCTGCAGAAAG	180
Pair_3	-----	0
Pair_4	-----	0
Vulture	CTCAGTGAAGAGTATGGACCAGTGTTCACAGTGCACCTTGGGCTCTGACCCAGTGGTGGTG	240
Pair_3	-----	0
Pair_4	-----	0
Vulture	CTGCACGGACATGATGCGGTGAAAGAAGCCTTGGTTGATCACGCGGACAAGTTTGCTGCC	300
Pair_3	-----	0
Pair_4	-----	0
Vulture	AGAGGACGCATGCCAATAGGAGACAGGGCTAACAAATGGATTATGGATTGTTTTAGCAAC	360
Pair_3	-----	0
Pair_4	-----	0
Vulture	AACAAGGAGTGGTTACAAGTCCGGCGGTTGCTCTCAGTACTCTGCGCAACTTTGGAATG	420
Pair_3	-----	0
Pair_4	-----	0
Vulture	GGGAAAAGGAGCATTGAAGAGAGGATCCAGGAGGAACTGAGTACTTGCTAGAAGAGATC	480
Pair_3	-----	0
Pair_4	-----	0
Vulture	AACAAAACAAAGGGAACACCTTTTGACCAACCTTCATGCTGAGCTGTGCTGTCTCCAAT	540
Pair_3	-----	0
Pair_4	-----	0
Vulture	GTCATATGCTCCATCGTCTTTGGGAAACGATATGACTATGAAGACAAGAAGTTCCTGGCC	600
Pair_3	-----	0
Pair_4	-----	0
Vulture	CCGATGAGCAACATGAACAACATCTTTGAGCTGATGAACTCCCCATGGGGACAGCTCTAC	660
Pair_3	-----	0
Pair_4	-----	0
Vulture	CAGATGTTCTCAAAGATCCTGGATTACTTGCTTTTGGCCACACAACAAAATATCCACA	720
Pair_3	-----CAGAGGAGGTGAAGACGCACCAAGCCTCCCTA	32
Pair_4	-----	0
Vulture	GAATTCGATGCTCTAAAAGCCTTTGCGTCAGAGGAGGTGAAGATGCACCAAGCCTCCCTA	780
Pair_3	GATCCCAGCTCCCCCAGGATTTTCATCGACTGCTTCCTCAGAAAAATGCAGGAGGAGAAA	92
Pair_4	-----	0
Vulture	GATCCCAGCTCCCCCAGGATTTTCATCGACTGCTTCCTCAGAAAAATGCAGGAGGAGAAA	840
Pair_3	GAGCATCCCAATTCCAGTTTCCAC-----	116
Pair_4	-----	0
Vulture	GAGCATCCCAATTCCAGTTTCCACATGAAGAACCTGATAACCAGCACCTTCGACTTGTTT	900

Pair_3	-----	116
Pair_4	-----	0
Vulture	ATTACTAGAACTGAGACAATTAGCACCAGTGTAAAGATATGGGCTCCTGCTTCTTCTCAA	960
Pair_3	-----	116
Pair_4	-----	0
Vulture	TACCCGAAGATACAAGAGAAAGTTCAAGAAGAGATTGACCAGGTAGTGGGATGATCACGA	1020
Pair_3	-----	116
Pair_4	-----GGCTCATCAGACCCAGATGCCCTACACAGATGCAGTGGTCCATGAAATC	49
Vulture	AGACCTTGTGTGGCTCATCAGACCCAGATGCCCTACACAGATGCAGTGGTCCATGAAATC	1080
Pair_3	-----	116
Pair_4	CAGCGCTTCATCTCCCTCATCCCCTGGGTCTCCCTCACACTGTGACCAAAGACACCGGC	109
Vulture	CAGCGCTTCATCTCCCTCATCCCCTGGGTCTCCCTCACACTGTGACCAAAGACACCGGC	1140
Pair_3	-----	116
Pair_4	TTCAGAGAGTATGTCATTCCCAAGGGCACCACAGTCTTTCC-----	150
Vulture	TTCAGAGAGTATGTCATTCCCAAGGGCACCACAGTCTTTCCCATCCTCAGTTCTGTCTC	1200
Pair_3	-----	116
Pair_4	-----	150
Vulture	CATGACAGTAAAGAGTTTCCAAACCCAATGAGTTCAACCCTGGACATTTCTTGAATGAC	1260
Pair_3	-----	116
Pair_4	-----	150
Vulture	AATGGCACCTTTAGGAANAGTGAGTTCTTCATGCCCTTCTCAGCAGGAAGCGAATATGC	1320
Pair_3	-----	116
Pair_4	-----	150
Vulture	CCTGGAGAGGGCCTGGCAGCATGGAGATATTCTTAGTCATAACCACCATCCTGCAGAAC	1380
Pair_3	-----	116
Pair_4	-----	150
Vulture	TTTACCTTGAAGCCTGTTGTGCGACCCCCAGGAACTCAACATAACCCCGACACTGAGTGGG	1440
Pair_3	-----	116
Pair_4	-----	150
Vulture	ACAGGCAACGTACCTCCTGCCTACCAGCTCTGTGCTCTCCCCGCTGA	1488

Figure 4: Alignment of the primer pair 3 and 4 and the predicted CYP2C19

### RT-qPCR References

Bustin, S.A., Benes, V., Garson, J.A., Hellems, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L. and Vandesompele, J., 2009. The MIQE Guidelines: Minimum Information for Publication of Quantitative Real-Time PCR experiments. *Clinical chemistry*, 55, pp.611-662.

**NextGen Genome Analysis**

```
Genome ----- 0
Eagle_CYP2C19 CATAGCTAACATAGCCCTTCGGTATATAAACTCGGGCTCCGGCAGCCAGTCACACCCTCC 60

Genome ----- 0
Eagle_CYP2C19 TGAGCCACAGAAGCAAATGGACTTCTCGGGACCAGCCACTGTTGTCTCTCGTTTGCAT 120

Genome -----NNNNNNNNNNNNNNNN-NNN 19
Eagle_CYP2C19 TGCTTGCTGCTATCCTTCACAGCATGGAGGGGAGGTCTGAAAGGGGAAGATGCCTCC 180

Genome NNNN--NNNNNNNNNN-NNNNNNNNNNNN-NNNNNNNNNNNNNN-NNNNNNNNNNNN- 73
Eagle_CYP2C19 AGGACCGGCTCCCTTCCCATCCTAGGTAATGTGCTGCAAGTGAACCAAAGATTTGGC 240

Genome NNNNNNN-NNNNNNNNNNNNNN-N--NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 128
Eagle_CYP2C19 CAAAACCTCCAGAAGGTAAGTCTACTTTCTTCTTACTTTCTTCTGCAGGCAAGAAA 300

Genome NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 188
Eagle_CYP2C19 TGTGTGGGACCTGTGCATTGGGATAACCTGTGGCCATAGCTTGATAAGTGGCAATCCCAG 360

Genome NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 248
Eagle_CYP2C19 AGAAGCAGAAAAGTAATGGGGTCTCAGCTGCTCCCCTCACTGCTGTTGCCATTGTTGT 420
*****

Genome CTAGGGCCTGATACCTGGGAACCTCTCCACAACCACATTTATCATGAGAGCACCAAATC 308
Eagle_CYP2C19 CCAGGGCCTGATAGCTGGGAACCTCTCCACAACCACATTTATCATGACAGGACCAAATC 480
* ***** *

Genome TTGACATTGCTGCTGTGTGAGGCTGCACTGCCACAGCGGATGCAGCAGCAAGGCCACT 368
Eagle_CYP2C19 TTGACGTTGCTGCTGTGTGAGGCTGCACTGCCACAGTGGGATGCAGCAGCAAGGCCACT 540
*****

Genome TTTCAACTCCTCCTCCATCTGCTCCATGGCCATGTTCTCTCTTGCAATTTGACAGCTC 428
Eagle_CYP2C19 TTTCAACTCCTCCTCCATCTGCTCCATGGCCATGTTCTCTCTTGCAATTTGACAGCTC 600
*****

Genome AGTGAAGAGTATGGACCAGTGTTCACAGTGCACCTGGGCTCTGACCCAGTGGTGGTGTG 488
Eagle_CYP2C19 AGTGAAGAGTATGGACCAGTGTTCACAGTGCACCTGGGCTCTGACCCAGTGGTGGTGTG 660
*****

Genome CACGGACATGATGCGGTGAAAGAAGCCTTGTTGATCACGCNNNNNNNNNNNNNNNNNN 548
Eagle_CYP2C19 CACGGACATGATGCGGTGAAAGAAGCCTTGTTGATCACGCAGAAGAGTTGCTGCCAGA 720
*****

Genome NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 608
Eagle_CYP2C19 GGACACATGCCAGTAGGAGACAGGGCTAACAAATGGATTAGGTATGTTTGCTGATTGAGCT 780

Genome NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 668
Eagle_CYP2C19 CCTTCTGAGGAGAAGGGAGGGCTGAGGGTGTGTTCCGGCAGATTAGAGCTGGAGTGTACG 840

Genome NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 728
Eagle_CYP2C19 TGCACGAGGAAGACTGCGGGGAGTGGGTGGTCCACCTGGAGAAGGAAGGCAGAGT 900
*****

Genome GGGCTGTAATTGCTGGCTTCCACCACTCAAAGCGGGTAGCAGAGAAATGGAGCCGATTTC 788
Eagle_CYP2C19 GGGCTGTAATTGCTGGCTTCCACCACTCAAAGAGGGTAGCAGAGAAATGGAGCCGATTTC 960
*****
```



Genome ATCTCAGCGATGGGCAGTGAAAGAGCAAGCGGCAACAGTCACCAGCTGCAGCAAGGGAAA 848  
Eagle\_CYP2C19 ATCTCAGCGATGGGCAATGAAAGGGCAAGCAGCAACAGTCACCAGCTGCAACAAGGGAAA 1020  
\*\*\*\*\*

Genome TTATGTGCGGATATGAGTAATGAATTGTTACAGGGAGGTTGGGTATGTGGTGAGAAGG 908  
Eagle\_CYP2C19 TTATGTGCGGATATGAGTAATGAATTGTTACAGGGAGGTTGGGTATGTGGTGAGAAGG 1080  
\*\*\*\*\*

Genome CCAAAGCAGGCAGTTGGACTGCAGGACCTCTAGAGATTGCCTCCAAACCACAGCACCT 968  
Eagle\_CYP2C19 CCAAAGCAGGCAGTTGGACCGCAGGACCTCTAGAGATTGCCTCCAAACCACAGCACTCT 1140  
\*\*\*\*\*

Genome ATGACACTACCACATCAGTGCTCTCCAGGGACAGGATTCACTGTCACCATGAGCCTCAG 1028  
Eagle\_CYP2C19 CTGACACTACCACATCAGTGCTCTCCAGGGACAGGATTCACTGTCACATGTGAGCCTCAG 1200  
\*\*\*\*\*

Genome TTCATCTCACACCCAGCACATCCAAGCCAGGCTTCCCTCTTCTCTGCAGGGATTGTTT 1088  
Eagle\_CYP2C19 TTCATCTCACACCCAGCACATCCAAGCCAGGCTTCCCTCTTCTCTGCAGGGATTGTTT 1260  
\*\*\*\*\*

Genome TTAGCAACAACAAGGAGTGGTTACAAGTCCGGCGGTTTGCTCTCAGTACTCTGCGCAACT 1148  
Eagle\_CYP2C19 TTAGCAACAACAAGGAGTGGTTAGAAGTCCGGCGGTTTGCTCTCAGTACTCTGCGCAACT 1320  
\*\*\*\*\*

Genome TTGGAATGGGGAAAAGGAGCATTGAAGAGAGGATCCAGGAGGAAACTGAGTACTTGCTAG 1208  
Eagle\_CYP2C19 TTGGAATGGGGAAAAGGAGCATTGAAGAGAGGATCCAGGAGGAAACTGAGTACTTGATGG 1380  
\*\*\*\*\*

Genome AAGAGATCAACAAAACAAGGGTATGGTTCACTTTTAATACAGTTATGTTTGCAAAAA 1268  
Eagle\_CYP2C19 AAGAGATCAACAAAACAAGGGTATGGTTCACTTTTAATACAGTTATGTTTGCAAAAA 1440  
\*\*\*\*\*

Genome AAGACTATATTAAGGAAGCCCATATCTTTAGCATAGAATTGAAGTTCCCGTGTGTTGGT 1328  
Eagle\_CYP2C19 AGGACTATATTAAGGAAACCCATGCTTTAGCATAGAATTGAAGTTCCCGTGTGTTGGT 1500  
\* \*\*\*\*\*

Genome CCATGCACCAGCTACCCCTTATTACAAGCAGTAAATAATTCCTGATCACGTACTATTC 1388  
Eagle\_CYP2C19 CCATGCACCAGCTACCCCTCATTACAAGCAGTAAATAATTCCTGATCACGTACTATTC 1560  
\*\*\*\*\*

Genome TTTTTGGAGGACCTTTCTGTGTTACTCAGAGACTGTGATCTGATATTTCCACTAACAACC 1448  
Eagle\_CYP2C19 TTTTTGGAGGACCTTTCTGTGTTACTCAGAGACTGTGATCTGATATTTCCACTAACAACC 1620  
\*\*\*\*\*

Genome AGAGCTCAGGTNN 1508  
Eagle\_CYP2C19 AGAGCTCAGGTACCCACAGCAGCCATAAGTAGCCATAGTTGGGGTTGGTTTTTTTAT 1680  
\*\*\*\*\*

Genome NNNNTGTAACCTTTTCTGAAATCACAGGAACACCTTTTGACCCAACCTTCATGCTGAGCTG 1568  
Eagle\_CYP2C19 GGTTTGTAACTTTTCTGAAATCACAGGAACACCTTTTGACCCAACCTTCATGCTGAGCTG 1740  
\*\*\*\*\*

Genome TGCTGTCTCCAATGTCATATGCTCCATCGTCTTTGGGAAACGATATGACTATGAAGACAA 1628  
Eagle\_CYP2C19 TGCTGTCTCCAATGTCATATGCTCCATCATCTTTGGGAAACGATATGACTATGAAGACAA 1800  
\*\*\*\*\*

Genome GAAGTTCCTGCCCCGATGAGCAACATGAACAACATCTTTGAGCTGATGAACCTCCCATG 1688

Eagle\_CYP2C19 GAAGTTCCTGGCCCTGATGAGCAACATGAACAACACCTTTGAGATGATGAACTCCCATG 1860  
\*\*\*\*\*

Genome GGGACAGGTACATTGAGTAGCCATATCGCAGTGGCAACATTCCTCCAGGGGAGCAGGAGG 1748  
Eagle\_CYP2C19 GGGACAGGTACATTGAGTAGCCATATTCGAGTGGCAACATTCCTCCAGGGGAGCAGGAGG 1920  
\*\*\*\*\*

Genome CCTCCCTCCCAATGAAGTCCCATTTCAGTCTTCTAACTAGGACCCCATCAATGCTTTGCT 1808  
Eagle\_CYP2C19 CCTCCCTCCCAATGAAGTCCCATTTCAGTCTTCTAACTAGGACCCCATCAATGCTTTGCT 1980  
\*\*\*\*\*

Genome GCACGGGAGGGAATTCATTCAAGAGCAAGAAGCACCTCACCAATGGAGTTGGCTTTGCC 1868  
Eagle\_CYP2C19 GCACAGGAGGGAATTCATTCAAGAGCAAGAAGCACCTCACCAATGCAGTTGGCTTTGCC 2040  
\*\*\*\*

Genome AGCTAAGTGCAGCTTCCCCTTAGTCTGAACACCCTTCTTCCCACCTCTAAGGTTCCCTA 1928  
Eagle\_CYP2C19 AGCTAAGTGCAGCTTCCCCTTAGTCTGAACACCCTTCTTCCCACCTCTAAGGTTCCCTA 2100  
\*\*\*\*\*

Genome CTACTACTAAGCAAGCTCCAACCACCTGTGAGACTGCCAGCCTCATAGCCCTTCTGCT 1988  
Eagle\_CYP2C19 CTACTACTAAGCAAGCTCCAACCACCTGTGAGACTGCCAGCCTCATAGCCCTTCTGCT 2160  
\*\*\*\*\*

Genome TGGCTCAGTTGAGAGGTTGTGGCCTTCTAGGCTTTTTGCTAATATCAAAGGATGGGGC 2048  
Eagle\_CYP2C19 TGGCTCAGTTGAGAGGTTGTGGCCTTCTAGGCTTTTTGCTAATATCAAAGGATGGGGC 2220  
\*\*\*\*\*

Genome AATGGTCGTGCTATGACCTTCTGTTTCTTCCCTTCCAGCTCTACCAGATGTTCTCAA 2108  
Eagle\_CYP2C19 AATGGTCATGCTAGGGACCTTCTGTTTCTTCCCTTCCAGCTCTACCAGATGTTCCCAA 2280  
\*\*\*\*\*

Genome GATCTGGATTACTTGCCTGGCCACANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2168  
Eagle\_CYP2C19 GATCTGGATTACTTGCCTGGCCACACAACAAAATATTACAGAATTTGATGCTCTAAA 2340  
\*\*\*\*

Genome NGCCTTTGCGTCAGAGGAGGTGAAGATGCACCAAGCCTCCCTAGATCCCAGCTCCCCCA 2228  
Eagle\_CYP2C19 AGCCTTTGTATCAGAGGAGGTGAAGATGCACCAAGCCTCCCTAGATCCCAGCTCCCCCA 2400  
\*\*\*\*\*

Genome GGATTTTCATCGACTGCTTCTCAGAAAAATGCAGGAGGTAAGGAGACAGCATGCAGCCCC 2288  
Eagle\_CYP2C19 GGATTTTCATCGACTGCTTCTCAGAAAAATGCAGGAGGTAAGGAGACAACATGCAGCCCC 2460  
\*\*\*\*\*

Genome AGCGCATCCGAAACATATCTACACTGAGCCCTTCCCTCTCCTAAGTAACAGAGACATT 2348  
Eagle\_CYP2C19 AGCACATCCGAAACATATCTACACCGAGCCCTTCCCTCTCCTAAGTAACAGAGACATT 2520  
\*\*\*

Genome GGGATGACCCCTTCCCAGGCNGCCGCTGTGCAGTGGGAAGACACCAATGATACCCAG 2408  
Eagle\_CYP2C19 GGGATGACCCCTTCCCAGGCNGCTGGCTGTGCAGTGGGAAGACACCAACGATACCCAG 2580  
\*\*\*\*\*

Genome ATATTGCCTGTGCTAGAGCTTATAGCTGTGTCTGCACATTACGAGCCACTACCTGCA 2468  
Eagle\_CYP2C19 ATATTGCCTGTGCTAGAGCTTATAGCTGTGTCTGCACATTACGAGCCACTACCTGCA 2640  
\*\*\*\*\*

Genome GCTGCTTGGACAGCCAGACTTTCATCTGCTCTACCAAAACACAGGGACATTTGCAGTCT 2528  
Eagle\_CYP2C19 GCTGCTTGGACAGCCAGACTTTCATCTGCTCTACCAAAACACAGGGACATTTGCAGTCT 2700

```

*****
Genome          GGTCTTGCAGTCTTTCCACCAGAAAGCAGTTGGCACAGCAAGGGCAAACAGGCAGCTATG 2588
Eagle_CYP2C19  GGTCTTGCAGTCTTTCCACCAGAAAGCAGTTGGCACAGCAAGGGCAAACAGGCAGCTATG 2760
*****

Genome          TACCATTTCAGTGCACAGCTTGAAGAAGGTTGCTCTGGGGTTCAGCTTACTGCACAAAGA 2648
Eagle_CYP2C19  TACCATTTCAGTGCACAGCTTGAAGAAGGTTGCTCTGGGGTTCAGCTTACTGCACAAAGA 2820
*****

Genome          TAGCATTGCAATGTCCCAGATGGCTGCTGCACCTTAGCAGAGAAATGCAACAGCTCTAGG 2708
Eagle_CYP2C19  TAGCATTGCAATGTCCCAGATGGCTGCTGCACCTTAGCAGAGAAATGCAACAGCTCTAGG 2880
*****

Genome          CAAAGAGGAAGGCCTGGGATGACACAAGAAAGGGCCACTTTTCTTAGCTCAGCCAGAGAA 2768
Eagle_CYP2C19  CAAAGAGGAAGGCCTGGGATGACACAAGAAAGGGCCACTTTTCTTAGCTCAGCCAGAGAA 2940
*****

Genome          ACAGCCCAGCCTTTCCCTAGGAGAAAGAGCATCCAATTCCAGTTTCCACATGAAGAAC 2828
Eagle_CYP2C19  ACAGCCCAGCCTTTCCCTAGGAGAAAGAGCATCCAATTCCAGTTTCCACATGAAGAAC 3000
*****

Genome          CTGATAACCAGCACCTTCGACTTGTTTACTTAGAAGTGAACAATTAGCACCAGCTGTA 2888
Eagle_CYP2C19  CTGATAACCAGCACCTTCGACTTGTTTACTTAGAAGTGAACAATTAGCACCAGCTGTA 3060
*****

Genome          AGATATGGGCTCTGCTTCTTCTCAAATACCGAAGATACAAGGTACCAGTCTGTGACCT 2948
Eagle_CYP2C19  AGATATGGGCTCTGCTTCTTCTCAAATACCGAAGATACAAGGTACCAGTCTGTGACCT 3120
*****

Genome          CTTCTCTCTAGCTATTCCCCTGGGCTGGGCATGTGTCTAACACCACATCCTTCTCCAA 3008
Eagle_CYP2C19  CTTCTCTCTAGCTATTCCCCTGGGCTGGGCATGTGTCTAACACCACATCCTTCTCCAA 3180
*****

Genome          CTTTTCTCTCTCTCTCATATTTACCCAAAAGGGCAAGGTCTTGACCATCATTC 3068
Eagle_CYP2C19  CTTTTCTCTCTCTCTCTCATATTTACCCAAAAGGGCAAGGTCTTGACCATCATTC 3240
*****

Genome          CCAGGGTGGCCGTTAGAGATCCCTCAGCTTCACTGGCATCATCATCAGCTGCACTCCCCT 3128
Eagle_CYP2C19  CCAGGGTGGCCGTTAGAGATCCCTCAGCTTCACTGGCATCATCATCAGCTGCACTCCCCT 3300
*****

Genome          CACAGCCAGGGACGCGACTGTGTACACATCTTTATCTCCAGTTATACGCACATTAGCCA 3188
Eagle_CYP2C19  CACAGCCAGGGACGCGACTGTGTCCACATCTTTATCTCCAGTTATACGCACATTAGCCA 3360
*****

Genome          AGTTACCAGGAGGACATTTTGCATGGGCCAAAGTGGAGCTATGTTCTAGTCAACACCA 3248
Eagle_CYP2C19  AGTTACCAGGAGGACATTTTGCATGGGCCAAAGTGGAGCTATGTTCTAGTCAACACCA 3420
***

Genome          AAACCAACATCAGTGCTCCTGTCTGTCTGTGGACAGCTTAGGGAAACCTGTTTA 3308
Eagle_CYP2C19  AAACCAACATCAATGCTCCTGTCTGTCTGTGGACAGCTTAGGGAAACCTGTTTA 3480
*****

Genome          GGTACGTTTGATTTTATTTACCCATTTTATGGTTTGTGCCTTTCATGACAGAGAAAGTTC 3368
Eagle_CYP2C19  GGTACGTTTGATTTTATTTACCCATTTTATGGTTTGTGCCTTTCATGACAGAGAAAGTTC 3540
**

```

Genome	AAGAAGAGATTGACCAGGTAGTGGGATGATCACGAAGACCTTGTGTGGCTCATCAGACCC	3428
Eagle_CYP2C19	AAGAAGAAATTGACCAGGTAGTGGGACGATCACGAAGACCTTGTGTGGCTCATCAGACTC	3600
	*****	
Genome	AGATGCCCTACACAGATGCAGTGGTCCATGAAATCCAGCGCTTCATCTCCCTCATCCCC	3488
Eagle_CYP2C19	AGATGCCCTACACAGATGCAGTGGTCCATGAAATCCAGCGCTTCATCTCCCTGTCCCC	3660
	*****	
Genome	TGGGTCTCCCTCACACTGTGACCAAAGACACCGCTTCAGAGAGTATGTCATTCCCAAGG	3548
Eagle_CYP2C19	TGGGTCTCCCTCACACTGTGACCAAAGACACCGCTTCAGAGAGTACGTCATTCCCAAGG	3720
	*****	
Genome	TTAGTGGGCAAGCTTCTCAAGTTCAACAACAAGGTGGCAGCATGTTTGCTCTTAGGCTG	3608
Eagle_CYP2C19	TTAGTGGGCAAGCTTCTCAAGTTCAACAACAAGGTGGCAGCATGTTTGCTCTTAGGCTG	3780
	*****	
Genome	TCAGTTACAAGCCAGAAGCTACAATACTGTGGCCTCACCTCGTTGTCACCCAGCCCTTC	3668
Eagle_CYP2C19	TCAGTTACAAGCCAGAAGCTACAATACTGTGGCCTGACCTCGTTGTCACCCAGCCCTTC	3840
	*****	
Genome	GTGAAGGTTTTATCCTCAGGCTCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	3728
Eagle_CYP2C19	ATGAAGGTTTTATCCTCAGGCTCCTGTAGCTCACTTGAATCCCTTGTGTCATTCTGTG	3900
	*****	
Genome	NN	3788
Eagle_CYP2C19	GACCGACTGGAACAGTAGCTCGCCAACACCTTGTCTACTTGCTCCGCAGCTTTACCT	3960
Genome	NN	3848
Eagle_CYP2C19	ATCTGCAGAAACAGTGCCTTAAACAAAAAACAGATCAGTGACTCCTCACACAAAAG	4020
Genome	NN	3908
Eagle_CYP2C19	GCAGCGAAGCAGTGGAACTGCTCACCACATCCAGGAGAGCCTGTACACATGCTTGGATG	4080
	*****	
Genome	CAAAGTCAGTTGGTGGTTATTGAGCACAGAAAGCCCTGAGCCACAAAGTCTTGAGGCT	3968
Eagle_CYP2C19	CAAAGTCAGTTGGTGGTTATTGAGCACAGAAAGCCCTGAGCCACAAAGTCTTGAGGCT	4140
	*****	
Genome	ATGTTTGCAATTCAAGGGAACAAGCAAACCTGTTTGCTCTGTTCTCACACACTTCTTGAG	4028
Eagle_CYP2C19	ATGTTTGCAATTCAAGGGAACAAGCAAACCTGTTTGCTCTGTTCTCACACACTTCTTGAA	4200
	*****	
Genome	GCAGTCACTGCTGGCCACTGCCAGAGGCAGGCTGATGGGCTAGAGCAACCCACGGCCTGG	4088
Eagle_CYP2C19	GCAGTCACTGCTGGCCACTGCCAGAGGCAGGCTGATGGGCTAGAGCAACCCACGGCCTGG	4260
	*****	
Genome	CCTGGCCCATGGCAGCGATGCTCATGTTACATCCTGTCTTCTGGGTGACCGGTACAGCA	4148
Eagle_CYP2C19	CCTGGCCCATGGCAGCGATGGTTCATGTTACATCCTGTCTTCTGGGTGACCGGTACAGCA	4320
	*****	
Genome	GGGGAACAGCAGCACACAGGCTGTTGGCGTTTTGATTTTCAGCTGCCTCTGTAATAA	4208
Eagle_CYP2C19	GGGGAACAGCAGCACACAGGCTGTTGGTGTGTTTTGATTTTCACCTGCCTCTGTAATAA	4380
	*****	

Genome	AAGGCAGGACCCAGACAACCTGCTCAGCTTCCAGCAGAGGGCTCTTGTGTCATCACTGCTG	4268
Eagle_CYP2C19	AAGGCAGGACCCAGAGAACCTGCTCAGCTTCCAGCAGAGGGCTCTTGTGTCATCACTGCTG	4440
	*****	
Genome	GATTATTTATCCCACCCCATCAATGACTGGTCAGGAAGAGGAAGGAAGAGGATCAT	4328
Eagle_CYP2C19	GATTATTTATCCCACCCCATCAATGGCTGGTCAGGAAGAGGAAGGAAGAGGATCGT	4500
	*****	
Genome	GGCAGGAGAAATTTGTCAGGAGAGGGGAAGAGAGCTTCATTTCCACAACCTTTGTGTGCTT	4388
Eagle_CYP2C19	GGCAGGAGAAATTTGTCAGGAGAGGGGAAGAGAGCTTCATTTCCACAACCTTTGTGTGCTT	4560
	*****	
Genome	TCGCTTTTGGGGCTCACAGAAAAGAAGAAACGTAGTGCATCTACAAGTTNTTTATTGTG	4448
Eagle_CYP2C19	TCGCTTTTGGGGCTCACAGAAAAGAAGAAACGTAGTGCATCTACAAGTTTATTGTTGTG	4620
	*****	
Genome	AAACACTAAGAAAGAGGAAGAATTAGGTGGGGTGTCCCCAGATGGCCAGAGCAAGGTGAG	4508
Eagle_CYP2C19	AAACACTAAGAAAGAGGAAGAATTAGGTGGGGTGTCCCCAGATGACCAGAGCAAGGTGAG	4680
	*****	
Genome	GATAAACGGTCTGAAAGTGAGAGGGAGCATTAGGAAAAGCAGCCAACAACCTGTCCCTGC	4568
Eagle_CYP2C19	GATAAATGGTCTGAAAGTGAGAGGGAGCATTAGGAAAAGCAGCCAACAACCTGTCCCTGC	4740
	*****	
Genome	TGTGACATCTAGCAGGGTCTCCCTAAAGCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	4628
Eagle_CYP2C19	TGTGACATCTAGCAGGGTCTCCCTAAAGCTGCCAGCAGAAGGGTCCCCAAAGGCA	4800
	*****	
Genome	AACAAAAGGCACCTGGTAGGGCTGTGCTGGGAAGCTCAGTCTTAACCTGGCAGGCAGTCGG	4688
Eagle_CYP2C19	AACAAAAGGCACCTGGTAGGGCTGTGCTGGGAAGCTCAGTCTTAACCTGGCAGGCAGTTGG	4860
	*****	
Genome	ATTAAGTGATCTAACTTGTCTCCTTCATGTTGTGCTCTTCTTCTTGTGTTACCTCAGG	4748
Eagle_CYP2C19	ATTAAGTGATCTAACTTGTCTCCTTCATGTTGTGCTCTTCTTCTTGTGTTACTTCAGG	4920
	*****	
Genome	GCACCACAGTCTTTCCATCCTCAGTCTGTCTCCATGACAGTAAAGAGTTTCCAACC	4808
Eagle_CYP2C19	GCACCACAGTCTTTCCATCCTCAGTCTGTCTCCATGACAGTAAAGAGTTTCCAACC	4980
	*****	
Genome	CAAATGAGTTCAACCCTGGACATTTCTTGAATGACAATGGCACCTTTAGGAAGAGTGAGT	4868
Eagle_CYP2C19	CAAATGAGTTCAACCCTGGACATTTCTTGAATGACAATGGCACCTTTAGGAAGAGTGAGT	5040
	*****	
Genome	TCTTCATGCCCTTCTCAGCAGGTAAGCACAGGCTTTCTCTCCGTCAACGTCTCTCTG	4928
Eagle_CYP2C19	TCTTCATGCCCTTCTCAGCAGGTAAGCACAGGCTTTCTCTCTGTCGACGTCTCTCTG	5100
	*****	
Genome	GTGACCACCTGCCCTCCCTATAACACACTCTCCCTTCCCAGCATGGTYCCCTGGT	4988
Eagle_CYP2C19	GTGACCACCTGCCCTCCCTATGACACACTCTCCCTTCCCAGCATGGTTCCCTGGT	5160
	*****	
Genome	GCTTCCTTGGGACTCACTGCTCTGANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	5048
Eagle_CYP2C19	GCTTCCTTGGGACTCACTGCTCTGGATAACTCCTCCCTGCTGCTCCCTTCTCCAGCTG	5220
	*****	

Genome	NN 5108
Eagle_CYP2C19	CATCTCCCATCTGTTCAGGCCTGGTGGCTTCTCAGACCTCCTACTAGGAGCTGCCAACG 5280
Genome	NN 5168
Eagle_CYP2C19	TAACGACATGAATGGGACCCACTGACACTTCTAAACACCTGTCACCTTCCCATAACAAC 5340
Genome	NNNNNNNNNNNNNNNNNNNNNNAACCTGTCTTTCTGTGGTAGCTCTGCTTTTCCCCAGCAC 5228
Eagle_CYP2C19	CACCAGAGTCATCTTAAACACCTGTCTTTCTGTGGTAGCTCCGCTTTTCCCCAGCAC 5400
	*****
Genome	TCTTAAGTACCTCTCAGGCTGCATGCAACCACTCACGTACCAATAGCTCCAACCCCTCTGC 5288
Eagle_CYP2C19	TCTTAAGTACCTCTCAGGCTGCATGCAACCACTCACGTACCAATAGCTCCAACCCCTCTGC 5460
	*****
Genome	TTTCTCACCCGTGCTTTATGTACCTTGCTCTCTCATCNNNNNNNNNNNNNNNNNNNN 5348
Eagle_CYP2C19	TTTCTCACCCGTGCTTTATGTACCTTGCTCTCTCATTTGCCTCCTGTGATGCCCCC 5520
	*****
Genome	NN 5408
Eagle_CYP2C19	TCCATGTTAAGACTTCTACTCCTCCTTCGCACACCAACCAAACTGCTAAGGGTCTTT 5580
Genome	NNNNNNNNNNNNNNNNNCTTTGCAATGTCTTAGTAATCCTGACAGTCATCTCACTGC 5468
Eagle_CYP2C19	TTCTGCAACAACCTTCTCTGCAATGTCTTAGTAATCCTGACAGTCACCTCACTGC 5640
	*****
Genome	TGCGGCTGTATGCCGTAATAAAGCTGTGCAACACATCTGCTTAGTACAAACTTAGTCT 5528
Eagle_CYP2C19	TGCGGCTGTATGCCGTAATAAAGCTGTGCAACACATCTGCTTAGTACAAACTTAGTCT 5700
	*****
Genome	ACTCAGCAACACCTGTTGAGTGTATTTGCTGTTTTCACTTCCAGGGAAGCGAATATGCC 5588
Eagle_CYP2C19	ACTCAGCAACACCTGTTGAGTGTATTTGCTGTTTTCACTTCCAGGGAAGCGAATATGCC 5760
	*****
Genome	CTGGAGAGGGCCTGGCACGCATGGAGATATCTTAGTCATAACCACCATCCTGCAGAACT 5648
Eagle_CYP2C19	CTGGAGAGGGCCTGGCACGCATGGAGATATCTTAGTCATAACCACCATCCTGCAGAACT 5820
	*****
Genome	TTACCTTGAAGCCTGTTGTGACCCCGAGAACTCAACATAACCCCGACACTGAGTGGGA 5708
Eagle_CYP2C19	TTACCTTGAAGCCTGTTGTGACCCCGAGAACTCAACATAACCCCGATACTGAGTGGGA 5880
	*****
Genome	CAGGCAACGTACCTCCTGCCTACCAGCTCTGTGCTCCTCCN----- 5755
Eagle_CYP2C19	CAAGCAACGTACCTCCTGCCTACCAGCTCTGTGCTCCTCCCGCTGAAAAGCACAAAACC 5940
	** *****
Genome	----- 5755
Eagle_CYP2C19	TCTCTCCATGGTATCCTGAGCCTGGCTACTCCTTTACATGTCCCTTACTAAAACCAACAG 6000
Genome	----- 5755
Eagle_CYP2C19	CAGAAGCATTGGTACCTTTGCAAGCCTCCAGGAAGGCAACCCAAACACAGAGTAGAT 6060
Genome	----- 5755
Eagle_CYP2C19	ACATGCACAGAGTAGACACCTTCAAAGCCTCCAGAACTGCTCCTACTACATTGCAAGGAG 6120

```

Genome ----- 5755
Eagle_CYP2C19 GCCCTGGGTGACATTGCAGCCTGTGCACTGACGCTCTGTCCACAGGTTTCATGGAGCGCAG 6180

Genome ----- 5755
Eagle_CYP2C19 TGACTGCATCAAACAGCTGATTTTCTTGTAACAGGGCTCGCATGGCTGCCCGTGACCTG 6240

Genome ----- 5755
Eagle_CYP2C19 CTTCTCCCACAGCACGTGTCCCACCAACGGCAGCACGCTTGCAGATAGATAGCGCTGCT 6300

Genome ----- 5755
Eagle_CYP2C19 TGTGCGTATGTTAATGACCCAGTAAAGAAAAATCTGTTTATGAGGA 6346

```

**Figure 5: Alignment of the next generation vulture genome sequencing results to the Eagle CYP2C19 gene (Chromosome 11) following alignment using HISAT2.**