

Conceptos básicos del análisis de variantes

Fundamentos y herramientas bioinformáticas para el análisis genómico

Lucía Spangenberg, Camila Simoes

15 de octubre de 2025



Disclaimers



Camila Simoes

Disclaimers



Camila Simoes

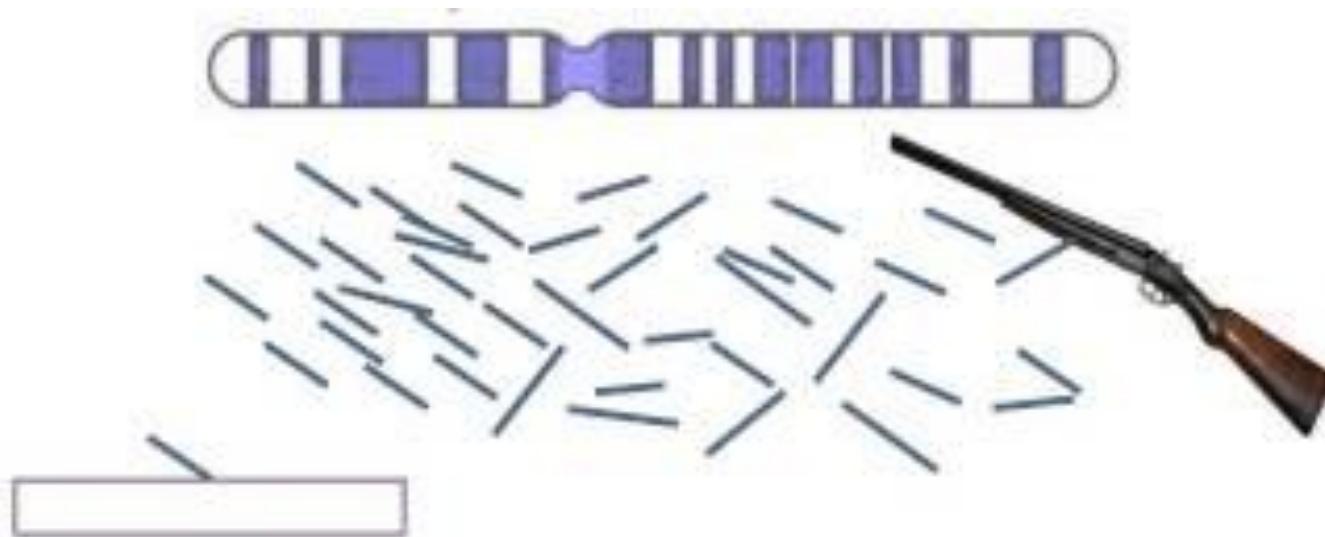


Disclaimers



Camila Simoes





[Link a video](#)



Millones de reads

¿Como sacar sentido a esos datos?

Germinal vs somático

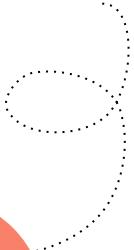
Variantes germinales heredadas

- Se producen en células germinales (espermáticas y ovocitos). Presentes en todas las células del organismo.
- Se obtienen de la sangre
- Se clasifican de acuerdo a su patogenicidad.
- En cáncer: asociadas a susceptibilidad. Eg. BRCA1/2

Variantes somáticas adquiridas

- Se producen en un tejido adulto. **No** son heredables.
- Se obtienen del tumor
- Se clasifican de acuerdo a su acciónabilidad. Eg. BRAF (inhibidores de BRAF)

Germinal

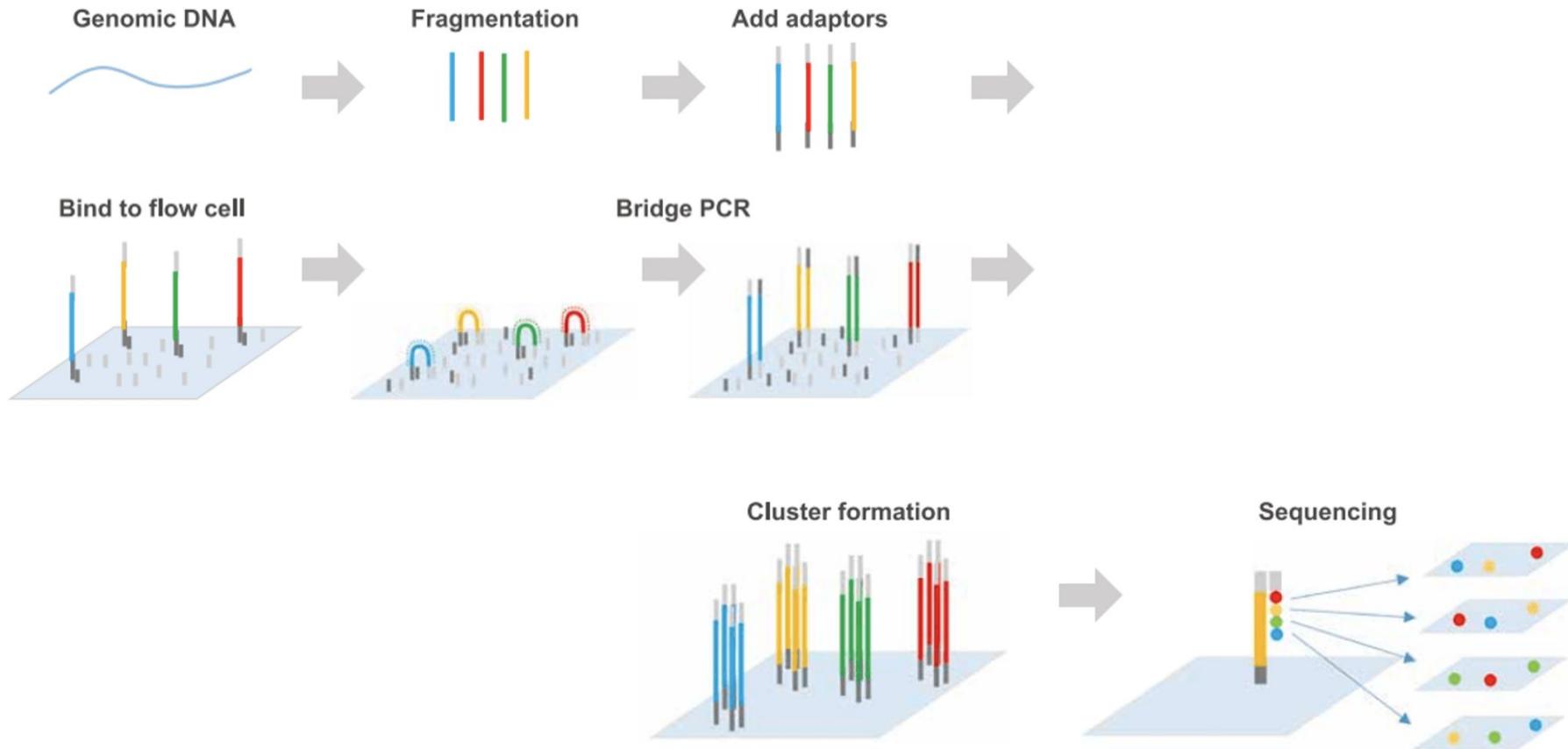


Variantes germinales heredadas

- Se producen en células germinales (espermáticas y ovocitos). Presentes en todas las células del organismo.
- Se obtienen de la sangre
- Se clasifican de acuerdo a su patogenicidad.
- En cáncer: asociadas a susceptibilidad. Eg. BRCA1/2

Variantes somáticas adquiridas

- Se producen en un tejido adulto. No son heredables.
- Se obtienen del tumor
- Se clasifican de acuerdo a su acciónabilidad. Eg. BRAF (inhibidores de BRAF)



Fastq

Header

Fastq

Read

Fastq

Calidad

@A00155:405:HNMNHDSDXY:3:1101:1380:1125 2:N:0:GTTGGATT+TCTATGTC
AAATGGAGTTGAATGCATTTGTTTAAATCAAAGAGGCACATTGTTGTCATTCCTGATTGAATCCCTTCCATTCCACTTGTTCATAATCATTCCGATAAGTAGTCCAATTGCATGGATTGGTTCTAATC
+
:FFFFFFFFFFFFFFFFFF:FFFFFFFFFFFFFFFFFF:FFFFFFFFFFFFFFFFFF:FFFFFFFFFF:FFFF:FFFF:FFFFFFFFFF:FFFF:FF:FFFF:FFFF:FF:FFFF:FFFF:FFFF:
@A00155:405:HNMNHDSDXY:3:1101:3622:1125 2:N:0:GTTGGATT+TCTATGTC
TGGCCCCAGGGCCCATTTGGTACCCAGCTATGGGAGCCTCACTTCAGTCACTGTACTCCTCGTCCCTTCCGAGGGGAAGCCTTCCCTGTCCTCGTACCACTCTGGCTCTCCATGCATTCAA
+
FFFFFFFFFF:FFFFFFF,FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFF:FFFF
@A00155:405:HNMNHDSDXY:3:1101:4833:1125 2:N:0:GTTGGATT+TCTATGTC
TCTGCCCTTGTTGTCAGCCAGGGCATCTGGCTCGGAAGCTGGCAGTCTCTCATCTCTGTCTGATCAAATTCAACCCACATTAGCCAGGTGACATGGTGGCTCTCCCAGGGTGGAGG
+
:FFFFFFFFFF:FFF:F:FFFFFFFFFF:FFFF:FFFFFFFFFF:FFFF:FFFFFFFFFF:FFFF:FFFFFFFFFF:FFFF:FFFFFFFFFF:FFFF:FFFFFFFFFF:FFFF:FFFF
@A00155:405:HNMNHDSDXY:3:1101:5647:1125 2:N:0:GTTGGATT+TCTATGTC
ACCTAACTCCCTCTCCCGTGGTCATGCTGGCTTCTGGTATTAGCTTGTAGCAGAGCAGTAGAGGTGAGTGGTGTGCTCTCCAGACCCATCTCTCTGAAGACCTGCCAG
+
,F:F:,FFFFFFFFFF:FFFFF,F:FFFFF:F:FF:,F,FFFFFFFFFF,FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFF,FFFF:F:FFFF:FFFF,FFF:FFFFF,,FFF
@A00155:405:HNMNHDSDXY:3:1101:8865:1125 2:N:0:GTTGGATT+TCTATGTC
CGAGCGCTCGGTCAAGGAGAGGTACTTTGCACTGGAACTCTCTCTAGCTCCAGCAGCTGCTCGTGTGAAGGCAGTCGCCGCCGCGTCTGCCGCGCCGCTCGCCGGGG
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFF
@A00155:405:HNMNHDSDXY:3:1101:9842:1125 2:N:0:GTTGGATT+TCTATGTC
TAGAACATAGCTTTAGCAGGAACTGTTTACAGTCAGTTGACTCTGAAGTAAACACATATTAGCTGTGTGATAAGTTCTTCACATGGAAAGAGGGACCCATGTTATCAG
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFF
@A00155:405:HNMNHDSDXY:3:1101:14615:1125 2:N:0:GTTGGATT+TCTATGTC
TCAGGACTGCTGTCAGCGCTGCAATGTCAGGAACTGCTGGCATAGCCGTCTCCAGAGGGAGCAGACTGTGCCAGCACCCATCTCAGGCTCAGCCCCGGAACTCT
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFF
@A00155:405:HNMNHDSDXY:3:1101:15411:1125 2:N:0:GTTGGATT+TCTATGTC
GTCTCCCTGAGAACGTTAGCAGCGCGTACCTGGGCCAGAAAAGAACATTCTCAGGGTGGCGCAAATTCAAAGGCCAGAAATGCTAGAACATTCCAGAAACCTCT
+
FFFFFFFFFF:FFFFFFFFFF:FFF:F:FF,F::FF,F:FF,F:FF,F:FF:FFFFF:,FF:F:,FFFFF:,F:F:FF,::F,FF,::F,FF,F:FFF:,F:,FFFF,FFF,F:,F:FFF,
@A00155:405:HNMNHDSDXY:3:1101:19569:1125 2:N:0:GTTGGATT+TCTATGTC
ATTGTTAAATGGGAAGAGAGTATGCTGTGTTACAACCCATTAGCTTGGATTAGCACAACACATCAAACCATACCAGTGGTATGGCTAGGACCAACCCAGTAC
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF
@A00155:405:HNMNHDSDXY:3:1101:20238:1125 2:N:0:GTTGGATT+TCTATGTC
TCCTGACGTCACTGGCGCGACCCCCCGCCGCGCCAGCGCGCTGCGCCCAAAGCTAGGCCGCCAAAGCGCGAGGAGCGGGCGGGCGAGCGCGTGTCCGGATCT
+
FFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFF
@A00155:405:HNMNHDSDXY:3:1101:21522:1125 2:N:0:GTTGGATT+TCTATGTC
ATGGTACCGATTACCTGCCATTAAAGAACATCCACTGTCAGCTGGTAAGATATTCCAAAGGATAATCAAATTCCGACCATGGTAATGTGGCTCTGGAGTCT
+
FFFF:FFFF:FFFFFFFFFF:FFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFF
@A00155:405:HNMNHDSDXY:3:1101:21630:1125 2:N:0:GTTGGATT+TCTATGTC
AAATTAGGAAACCCAGAGAACAGAGAACAGAGGGAGGGGGTGGTGGAGCGCGAGGTAGCCAGTCAGCTGGATCTGGCTGGAAAGAGTGGAGTGGCG
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFF
@A00155:405:HNMNHDSDXY:3:1101:25337:1125 2:N:0:GTTGGATT+TCTATGTC
AAACAAATACACCTTACCTTCTTTTTTTTTTATTCTGAGATGGAGTCTGCTATGTCACCTGGCTGGAGTCCACCTGGTGTGCTGGAGTCT
+
FFFFFFFFFF:FFFFFFFFFF:FF:FF:FFFFFFFFFF:FFF:,,FFF:,,FFF:,,FF,FFFF,:FFF:FFF,:F:FF,:FFF,FF:,,FFF,FF:FF:FFF,FF:FFF:,,F:FFF,FFF
@A00155:405:HNMNHDSDXY:3:1101:27489:1125 2:N:0:GTTGGATT+TCTATGTC
ACCTGCACATTCTGCACATGTACCCAGAACATTAAAGTAAATTAAAAATAGCATTCTATAACTTCTACTGAATAATGAATAGAAGAATAATGAGATT
+
FFFFFFFFFF:FFFFFFFFFF:FF:FF:FFFFFFFFFF:FFF:,,FFF:,,FFF:,,FF,FFFF,:FFF:FFF,:F:FF,:FFF,FF:,,FFF,FF:FF:FFF,FF:FFF:,,F:FFF,FFF

Fastq

40 Millones

@A00155:405:HNMMHDSXY:3:1101:1380:1125 2:N:0:GTTGGATT+TCTATGTC
AAATGGAGTTGAATGCATTTGTTTAAATCAAAGAGGCACATTGGTGTCTGCTTACCTGGATTGAATCCCTTCCATTCCACTCTGTTCATATACTCATTCCGATAAGTAGTCAATTGATGGATTGGTTCTAATC
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:3622:1125 2:N:0:GTTGGATT+TCTATGTC
TGGCCCAGGGCCCATTTGGTACCCAGCTGGGAGCCTCACTTCACTGCAGTACTCTCGTCCCTTCCGTAGGGGAAGCCTTCCCCCTCTCGTACCACTCTGGCTCTCCATGCATTCAAACGTAGGTGCGT
+
FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:4833:1125 2:N:0:GTTGGATT+TCTATGTC
TCTGCCCTTGTTGTCAGCCAGGGCATCTGGCTCGGCAAGTGGCAGCTCCCTCATCTCTGTCATGATCAAATTCAACCCACATTAGCCAGGTGACATGGTGGCTCTCCCAAGGGTGGAGGCTGTCAGGCTGCTGCA
+
:FFFFFFF:FFF:FF:FFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:5647:1125 2:N:0:GTTGGATT+TCTATGTC
ACCCATACTCCCTCTCCCGTGGTATGCTGCTTCACTTGTGTTAGATCTGTCAGAGCAGTAGAGGTGAGTGGTGTGATGCCAGCCCATCTCTCTGAAGACTGCCACGGCAAGCTGGAGCG
+
,F:F:,FFFFFF:FFFFFF,FFFF,F:FFFFFF:F:F:,F,FFFFFFF,FFFFFF,FFFFFF,FFFFFF,FFFFFF,FFFFFF,FFFFFF,FFFFFF,FFFFFF,FFFFFF,FFFF,FFF:,F:FFFFFF:FFFF,FFF:FFFFFF,,FFF
@A00155:405:HNMMHDSXY:3:1101:8865:1125 2:N:0:GTTGGATT+TCTATGTC
CGAGCGCTCGGTCAAGGAGAGGTACTTTGCACTGGAACTCTCTACTGCTCAGCAGTCTGCTGGTAAGGCAGTCGCCGCCGCGTTCTGCCGTAGAGTGGTGTGCTGCCGCGCCGCTGCTGGGGGTCTCC
+
FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:9842:1125 2:N:0:GTTGGATT+TCTATGTC
TAGAACATGCTTATGCAAGGAACTGTTTACAGTCAGTTGACTCTGAAGTAAACACATATTAGCTGTCATGTTCACTGGAAAGAGGGACCCATGTTATCAGACAGGAATCAGTTGATATGGG
+
FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:14615:1125 2:N:0:GTTGGATT+TCTATGTC
TCAGGACTGCTGTCAGCGCTGCAATGTCAGGAACTGCTGGCATAGCCGTTCCAGAGGGAGCAGACTGTGCCAGCACCCATCTCAGGCTCAGCCCGGAATCTGCCCTCTGGGGCTGCGCACCCATAG
+
FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:15411:1125 2:N:0:GTTGGATT+TCTATGTC
GTCTCCCTGAGAACGTTAGCAGCGCGTACCTGGGCCAGAAAAGAACATCTCAGGGTGGCGCAAATTCAAAGGCCAGAAATGCTAGAACATCTCAGAAACCTCTAGAACACTGGAAATGCCCTCAGGGTCC
+
FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:19569:1125 2:N:0:GTTGGATT+TCTATGTC
ATTGTTAAATGGGAAGAGATGCTGTGTTACAACCCCTAGTCAGCACAACACATCAAACCATACATTGGTATGGCTAGGACCCAACCCAGTATCAGAAATGAAGGAACGTACCCCTACAGGCCACAGCC
+
FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:20238:1125 2:N:0:GTTGGATT+TCTATGTC
TCCTGACGTCACTGGCGCGACCCCCCGCCGCGCCAGCGCGCTGCCAACAGCTAGGCCGCCAACAGCGCAGGGAGCAGCGCGTGTCCGGATCTCGCAGTGTCCGGCCCCCTCCCTGCC
+
FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:21522:1125 2:N:0:GTTGGATT+TCTATGTC
ATGGTACCGATTACCTGCCATTAAAGAACATCCACTGTCAGCTGGTAAGATATTCCAAAGGATAATCAAATTCCGACCATGGTAAATGTCCTGGAGTCTTGAGCCACAATCTGGTCAGAACATGCC
+
FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:21630:1125 2:N:0:GTTGGATT+TCTATGTC
AAATTCAGGAAACCCGAGAACAGAGAACAGAGGGAGGGTGGTGCAGCGCGAGGTGAGCCAGTCAGTGGATTCTGGCCTGGAAAGAAGTGGAGTGGCGTAGGACCTCTCCCGAGCAACGCCA
+
FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:25337:1125 2:N:0:GTTGGATT+TCTATGTC
AAACAAATATCACCTTATCACTTCTTTTTTTTTTATCTGAGATGGAGTCTGCTATGTCACCCAGGCTGGACTGCACTGGCGCATGCCCTGCTCCCGAGATCAAGCATTCTCTGCC
+
FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMMHDSXY:3:1101:27489:1125 2:N:0:GTTGGATT+TCTATGTC
ACCTGCACATTCTGCACATGTACCCAGAACATTAAGTAAATTTAAAAATAGCATTTACTGAAATAATGAATAGAAGAATAATGAGATTACTTTATGACTAACTACAAATTATTAGCTGCACAT

Fastq



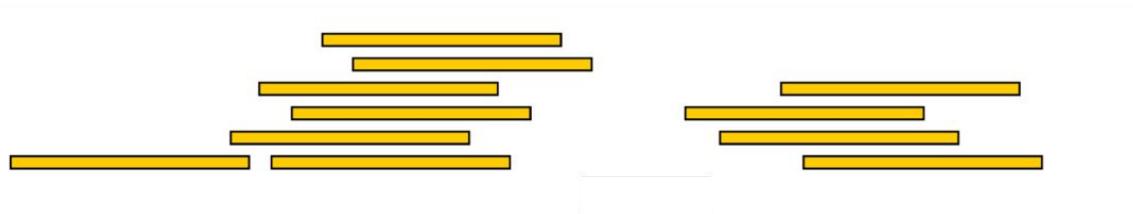
1 genoma humano 30x 300Gb



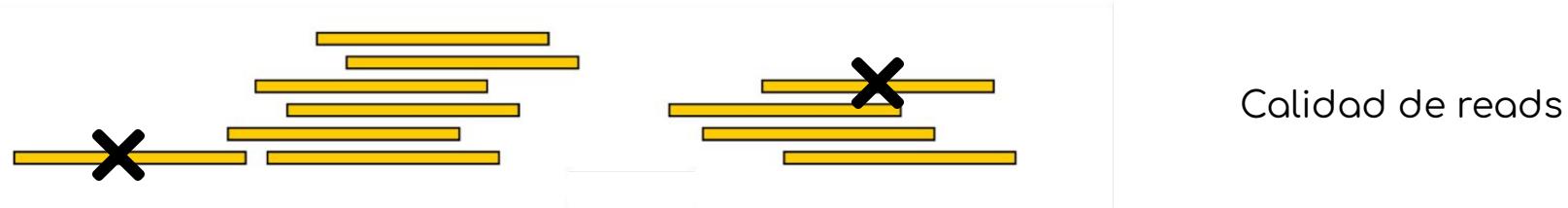
1 exoma humano 100x 5Gb

@A00155:405:HNMNHD\$XY:3:1101:1380:1125 2:N:0:GTTGGATT+TCTATGTC
AAATGGAGTTGAATGCATTTGTTTAATCAAAGAGGCCACATTGTTGTCATTCCTGATTGAATCCCTTCCATTCCACTCTGTTCATATACTCATTCCGATAAGTAGTCAATTGATGGATTGGTTCTAATC
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:3622:1125 2:N:0:GTTGGATT+TCTATGTC
TGGCCCCAGGGCCCATTTGGTACCCAGGTACTGGGAGCCTCACTTCACTGCACTGTAECTCTCGTCCCTTCCGTAGGGGAAGCCTTCCCCTGTCCTCGTACCACTCTGGCTCTCCATGCATTCAA
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:4833:1125 2:N:0:GTTGGATT+TCTATGTC
TCTGCCCTTGTTGTCAGCCAGGGCATCTGGCTCGGCAAGTGGCAGCTCCCTCATCTCTGTCATGATCAAATTCAACCCACATTAGCCAGGTGACATGGTGGCTCTCCCAGGGTGGAGGTCTG
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:5647:1125 2:N:0:GTTGGATT+TCTATGTC
ACCCATAATTCCCCCTCTCCCGTGGTCATGCTGGCTTCTGGTATAGCTTGTAGAGCAGAGTAGAGGTGAGTGGTGTGATGCCAGCCCATCTCTCTGAAGACCTGCCAGGA
+
,F:F:,FFFFFFF:FFFFFFF,FFFFF,F:FFFFFFF:F:FF:,F,FFFFFFF,FFFFFFF,FFFFFFF,FFFFFFF,FFFFFFF,FFFFFFF,FFFFFFF,FFFFFFF,FFFFFFF,FFFFFFF,,FFF
@A00155:405:HNMNHD\$XY:3:1101:8865:1125 2:N:0:GTTGGATT+TCTATGTC
CGAGCGCTCGGTCAAGGAGAGGTACTTTGCACTGGAACTCTCTCTGACTCCAGCAGCTGCTCGTGGTAAGGCAGTCGCCGCCGCGTTCTGCCGTAGCGTGTGCTGCCGCGCCGCT
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:9842:1125 2:N:0:GTTGGATT+TCTATGTC
TAGAACATAGCTTTAGCAGGAACCTGTTTACAGTCAGTTGACTCTGAAGGTAAACACATATTAGCTGTCGTGATAAGTTCTTCACATGGAAAGAGGGACCCATGTTATCAGACAGGA
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:14615:1125 2:N:0:GTTGGATT+TCTATGTC
TCAGGACTGCTGTCAGCGCTGCAATGTCAGGAACTGCTGGCATAGCCGTTCTCCAGAGGGAGCAGACTGTGCCAGCACCCATCTCAGGCTCAGCCCCGGAATCTGCC
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:15411:1125 2:N:0:GTTGGATT+TCTATGTC
GTCTCCCTGAGAACGTTAGCAGCGCGTTTACCTGGGCCAGAAAAGAACATTCTCAGGGTGGCGCAAATTCAAAGGCCAGAAATGCTAGAACATCTCAGAAACCT
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:19569:1125 2:N:0:GTTGGATT+TCTATGTC
ATTGTTAAATGGGAAGAGAGTATGCTGTGTTACAACCCCTAGGTGATTGACACACATCAAACCATACATTGGTATGGCTAGGACCAACCCAGTACAGAACATGAAG
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:20238:1125 2:N:0:GTTGGATT+TCTATGTC
TCCTGACGTCACTGGCGCGACCCCCCGCCGCGCCAGCGCGCTGCGCCCAAAGCTAGGCCGCGCCAAAGCGCGAGGAGCGGGCGGGCGAGCGCGTGTCC
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:21522:1125 2:N:0:GTTGGATT+TCTATGTC
ATGGTACCGATTACCTGCCATTCTAAAGAACATCCACTGCACTGTCAGCTGGTAAGATATTCCAAAGGATAATCAAATTCCGACCATGGTAAATG
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:21630:1125 2:N:0:GTTGGATT+TCTATGTC
AAATTAGGAAACCCGAGAACAGAGAACAGAGGGAGGGGTGGTGCAGCGCGAGGTGACCCAGTCAACTGGTACCTGGCCTGGAAAGAAGTGGAGTGGCG
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:25337:1125 2:N:0:GTTGGATT+TCTATGTC
AACAAATATCACCTTATCACTTCTTTTTTTTATCTCAGAGTGGAGTCTGCTATGTCACCCAGGCTGGACTGCACTGGCGCAGTCCGCTACTGCAC
+
:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:27489:1125 2:N:0:GTTGGATT+TCTATGTC
ACCTGCACATTCTGCACATGTACCCAGAACATTAAAGTAAATTAAAAATAGCATTCTACTGATAACTTCTACTGATAATAATGAATAGAAGAATAAT
:FFFFFFF:FFFFFFF:FFFFFFF:
@A00155:405:HNMNHD\$XY:3:1101:27489:1125 2:N:0:GTTGGATT+TCTATGTC
ACCTGCACATTCTGCACATGTACCCAGAACATTAAAGTAAATTAAAAATAGCATTCTACTGATAACTTCTACTGATAATAATGAATAGAAGAATAAT
:FFFFFFF:FFFFFFF:FFFFFFF:

Pipeline bioinformático



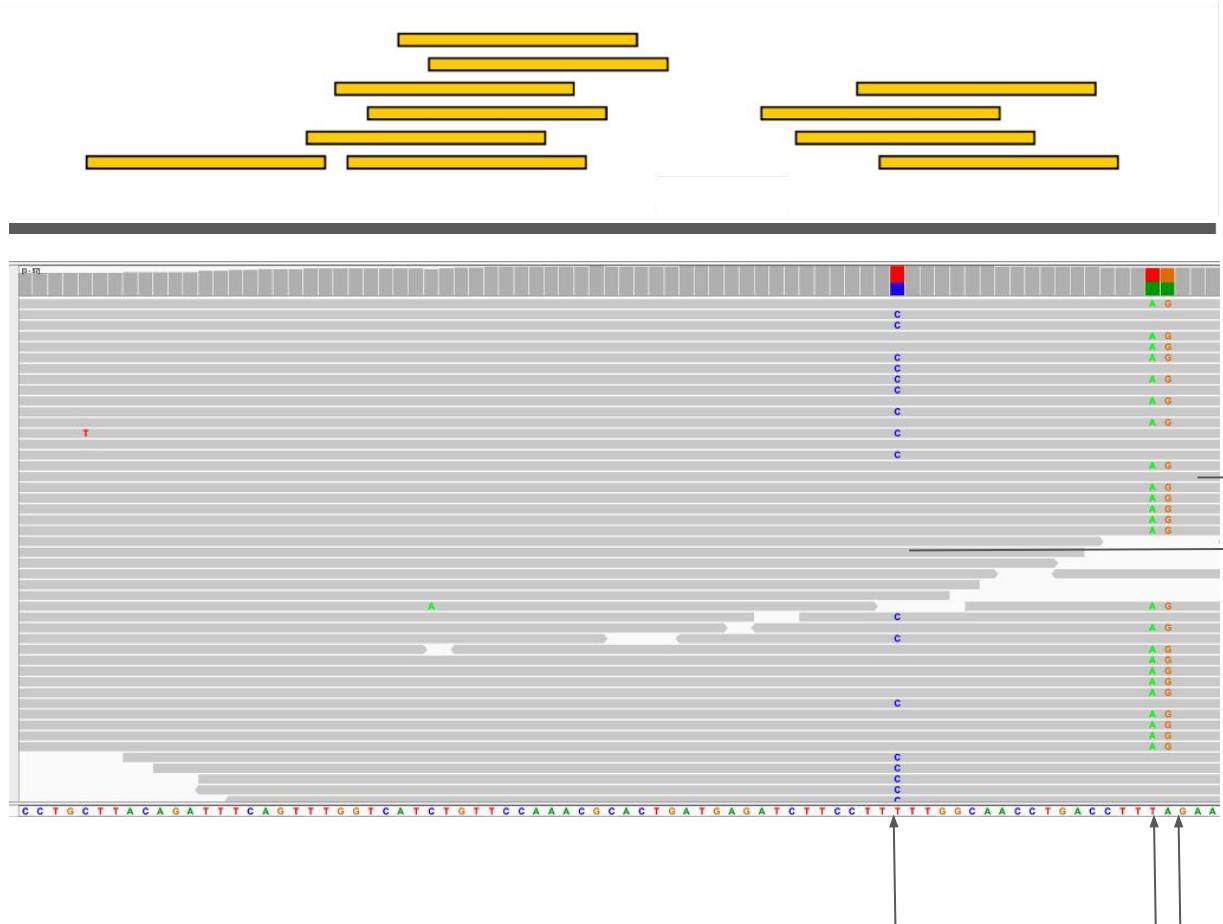
Pipeline bioinformático



Pipeline bioinformático



Pipeline bioinformático

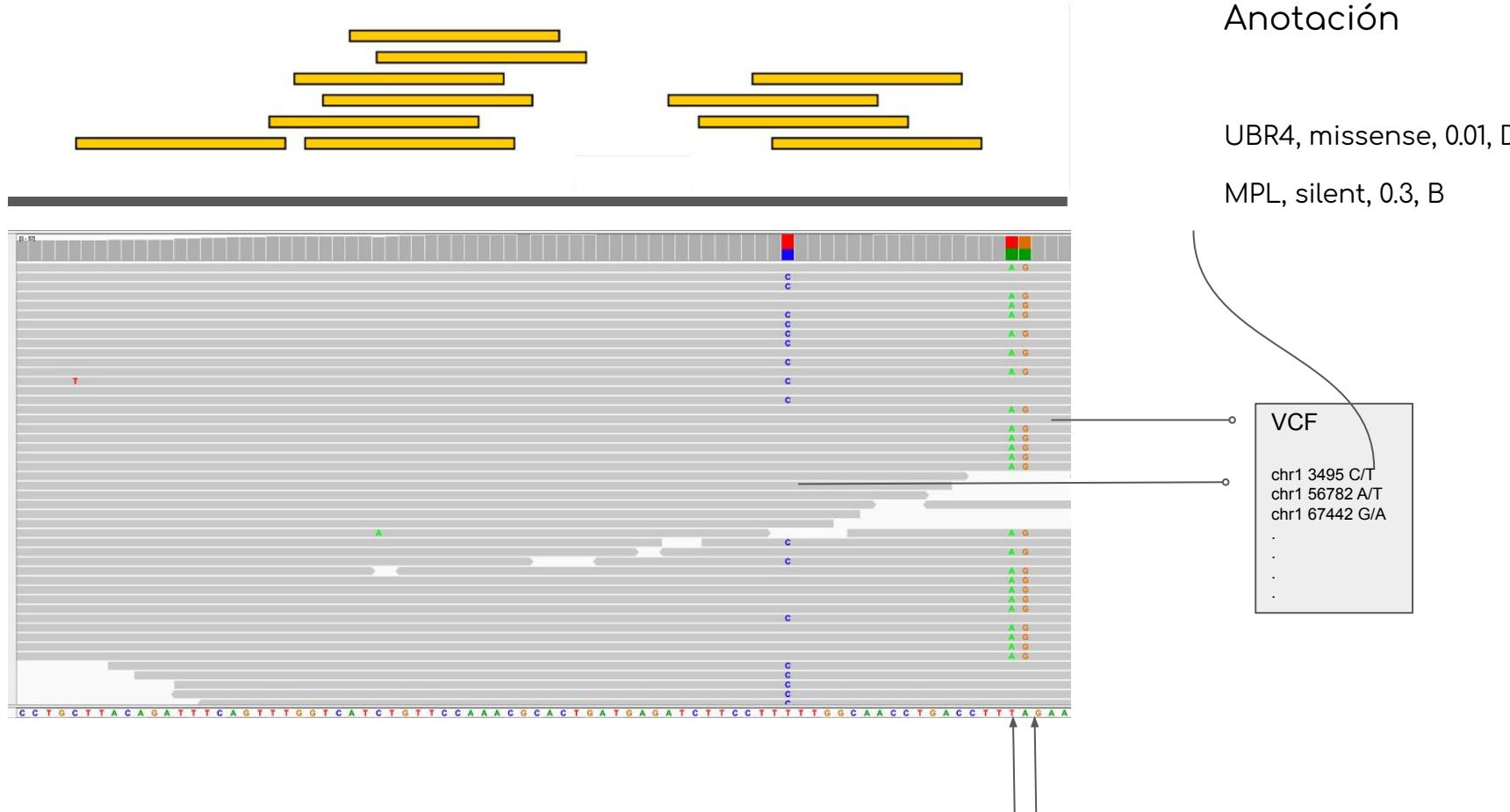


Llamado de variantes

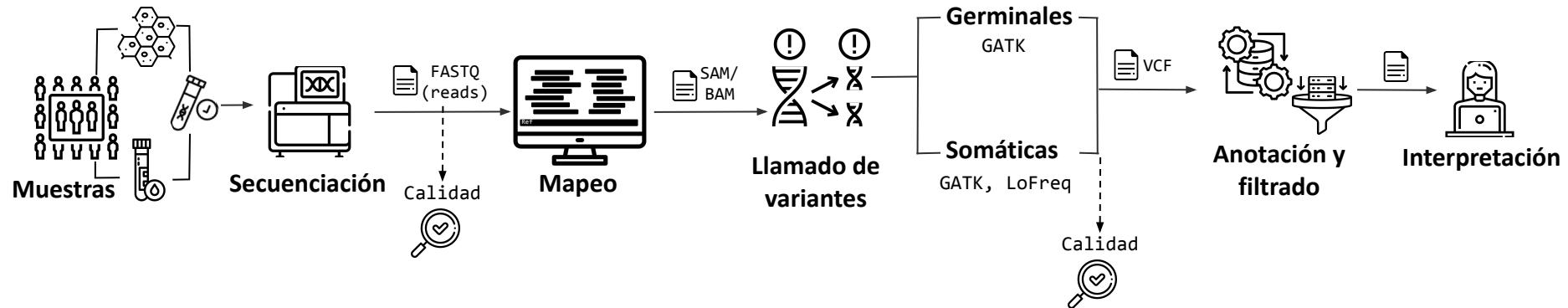
VCF

chr1 3495 C/T
chr1 56782 A/T
chr1 67442 G/A
. . .

Pipeline bioinformático



Pipeline bioinformático

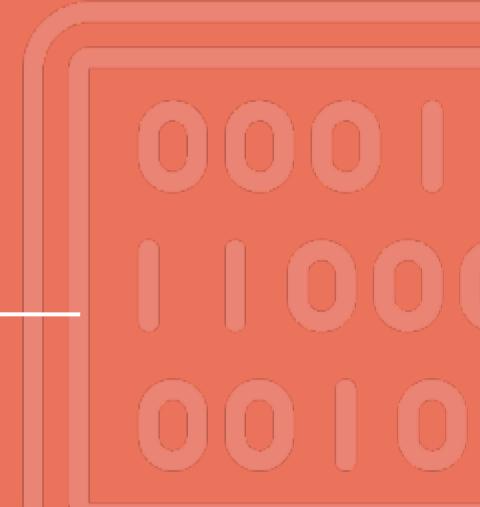


Pipeline bioinformático





Calidad Concepto #1



000110010010

GACCGATGACCCGGTTCAGGCTTCACCACTAGTGGAACGCCGCGTCTCGAACTTAACGGCGACC
TAAGGTTGACGACGGACCCAGCAGTGTGCTAATCTCAGCGCTCCGCTGACCCCTCAGCAAAGGGCTTG
GCTCAATCTCGTCCAGCCATTGACCATCGTGAGGGGTTGCTCTGTTATCGTGCCGAGCAGCTTG
AAAAACGAAATCGAGGCCATCTGGGGCCCCGATTACCGACGCTCTAGCCGCCACTCGGACATCAGA
CCAACCTGGGGTCCGCATCGCTCCGCCGGCAGCAGAAGCCGACGACACTACCGTGCCGCCCTCGA
ATCCTGCTACCACATGCCAGACACCACAACGACAACGAGGAGATTGATGACAGCGCTGCCAACGG
CGATAACCAGCACAGTTGGCAAGTTACTTCACCGAGCGCCGACAATACCGATTCCGTAACCGCTG
GGTAACCAGCTTAACCGTCGCTACACCTTGATACGTTGTTACGGCCCTCCAACCGTTCGCGCA
CGCCGCCGCTTGGCGATCGCAGAAGCACCCGCCGCGTTACAACCCCCCTGTTCATCTGGGGCGAGTCC
GGTCTCGGAAGACACACCTGCTACACGCCAGGCAACTATGCCAACGGTTCCGGGAATGCCGG
AAAATATGCTCCACCGAGGAATTCAACCAACGACTTCATTAACTCGCTCCGCCATGACCGCAAGGTCGC
CTCAACCGCAGTACCGCAGTAGACGTTGCTGCTGACCCAGATCCAATTCAAGGCAAAGAG
GTATTCAAGAGGAGTTCTCCACACCTCAACCCGAGCTGCTGAGACCGCATGCCATCTGGCGCAAGAAAGC
ACTGACGTACAACCAACCCGAGCTGGAGACCCGATCGCAAGCAATATCCGTGAACTCGAGG
TCGGTCCCCGACGATGTCCTCGAACTCATGCCAGCTATCGAACGCAATATCCGTGAACTCGAGG
GGCGTGTACCGGGTCAACCGCGTTGCCATCGCAAGAGCTCGGGGCCGCAAGGCCGAGCACTGCCAGTCAC
CTGCTTCGGCATCTGATGCCGACGCCATCGCAAGAGCTCGGGGCCGCAAGGCCGAGCACTGCCAGTCAC
GCCGAATCTCGACACTACCGTCGAAGAGCTCGGGGCCGCAAGGCCGAGCACTGCCAGTCAC
GACAGATTGCGA
CCGTGATCACACAACCGTCACTGTACGCCAACGCAAGATCCTGCGAGATGGCCGAGGCCGTCAGGTC
TTGATCACGTCAAAGAACTCACCACTCGCATCCGTCAGCGCTCCAAGCGTAGCACGGCGTTCTTCC
AACCGTTCTAAAAAAACTCTCTCCAGGTACACCAAGTACAGAGATTGGCTGTGAGTGTGCT
TGCACAAACCGCGCACAGACTCATACAGTCCCAGGTTCCGTTACAACCCACGCCATCCCCACCG
ACCCAAACACACACCCACAGTCATGCCACCGTCATCCACAACCTCGACCGACGTCGACCTGCACCAAGA
CCAGACTGCCCCAAACTGCACACCCCTCTAATACTGTTACCGAGATTCTCGTCGTTGTTCTGGAAA
ACAGCGCTGGGGATCGTCGCTGGATACCACCCGCTAACACTGGCTCGCGGTGGTCAGAGGTCAAT
ATCAACTTCAAGTTGACGTGAGAACGCTTACCGTTGTTGACTGCTGTTGCCGCCGTCGTTGG
TCACCGCGTCATGGGCATTGCTCGTGGCAGTCCCCACGCTAGCGGGGCCGCTAGCCACGGGATCGAACTC
ATCGTGAGGTGAAAGGGCGCAATGGACGCCGCTACGACAAGAGTTGGCCTCACCGACTTGACGTTG
TGCTACGAGAGTCTTCGCCATGCCGTTGCGTGGCTAAAAATCTGCCAGCCAGGCCGCCGCGTGC
GGTCTCCGGCGTTGTTGACCGGCTCGGACAACGGCTGACGATTCCGGATTGACTACCGAGGTT
CGCCGAGGCCAGGTTGGCGCTGAAATTGTTCTCCTGGAAGCGTTAGTTCTGGCCGATTGTTGT

ASCII

66 B
67 C
68 D
69 E
70 F
71 G
72 H
73 I
74 J
75 K

Calidad

$$70-33=37$$

Phred Quality Score	Error	Accuracy (1 - Error)
10	1/10 = 10%	90%
20	1/100 = 1%	99%
30	1/1000 = 0.1%	99.9%
40	1/10000 = 0.01%	99.99%
50	1/100000 = 0.001%	99.999%
60	1/1000000 = 0.0001%	99.9999%

ASCII

66 B
67 C
68 D
69 E
70 F
71 G
72 H
73 I
74 J
75 K

Calidad

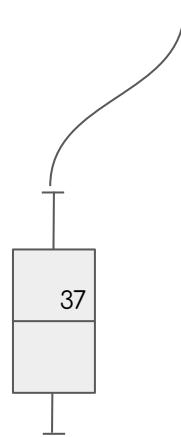
70-33=37

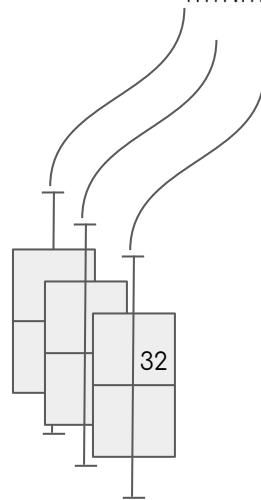
Phred Quality Score	Error	Accuracy (1 - Error)
10	1/10 = 10%	90%
20	1/100 = 1%	99%
30	1/1000 = 0.1%	99.9%
40	1/10000 = 0.01%	99.99%
50	1/100000 = 0.001%	99.999%
60	1/1000000 = 0.0001%	99.9999%

$$Q = -10 \log E$$

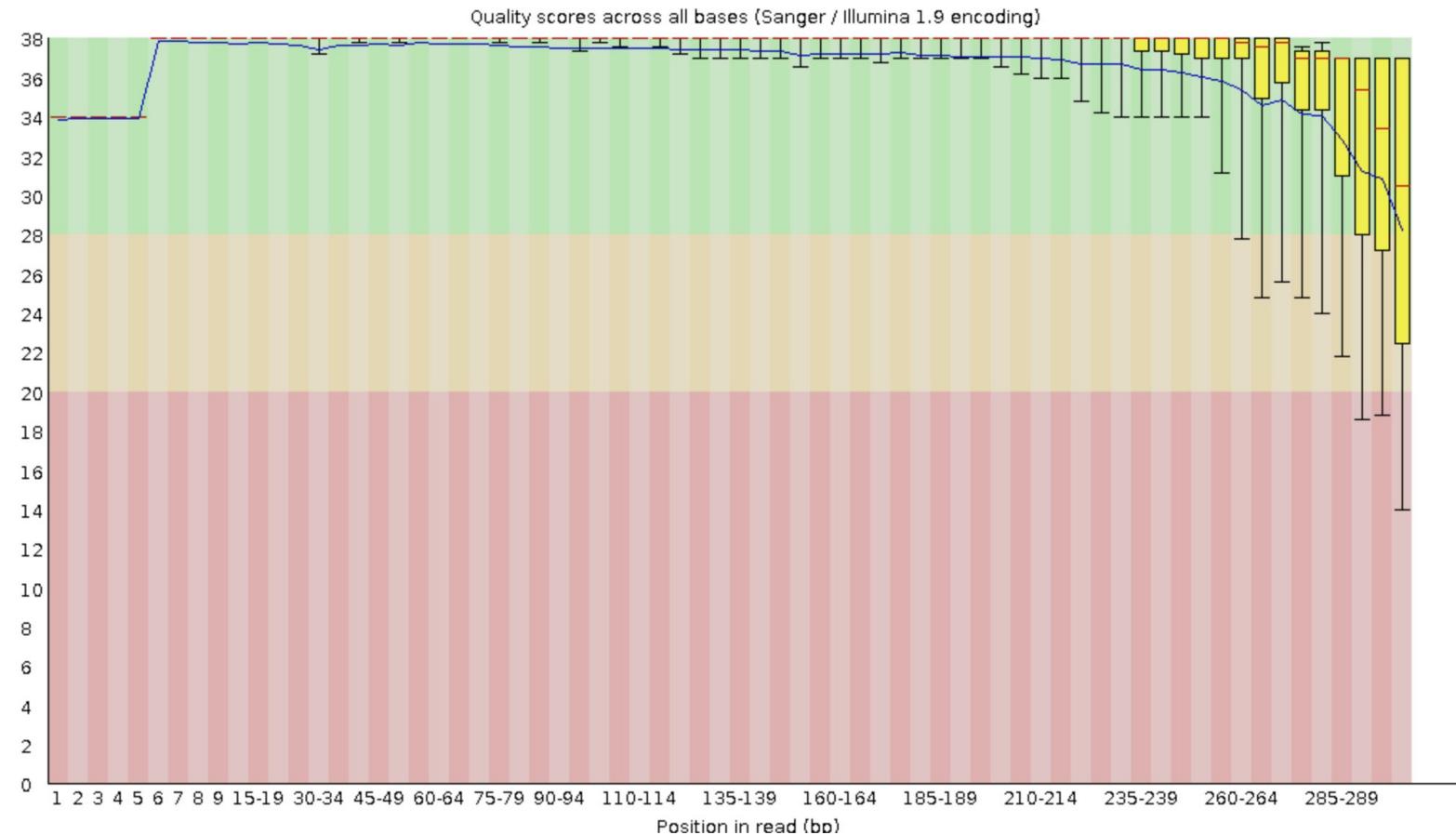
$$E = 10^{-\left(\frac{Q}{10}\right)}$$

@A00155:405:HNMNHDSDXY:3:1101:1380:1125 2:N:0:GTGGATT+TCTATGTC
AAATGGAGTTGAATGCATTTGTTTAATCAAAGAGGCACATTGGTGTCTGCTTACCTGGATTGAATCCCTTCCATTCCACTTGTTCATAATCATTCCGATAAGTAGTCAATTGATGGATTGGTTCTAATC
+
:FFFFFFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:
@A00155:405:HNMNHDSDXY:3:1101:3622:1125 2:N:0:GTGGATT+TCTATGTC
TGGCCCAGGGCCCATTTGGTACCCAGCTGGGAGCCTACTTCACTGCACTGTACTCCTCGTCCCTTCCGAGGGGAAGCCTTCCCCCTGTCCTCGTACCACTCTGGCTCTCCATGCATTCAA
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:
@A00155:405:HNMNHDSDXY:3:1101:4833:1125 2:N:0:GTGGATT+TCTATGTC
TCTGCCCTTGTTGTCAGCCAGGGCATCTGGCTCGGAAGCTGGCAGCTCTCATCTCTGTCTGATCAAATTCAACCCACATTAGCCAGGTGACATGGTGGCTCTCCAGGGTGGAGGCTGTC
+
:FFFFFFFFFF:FFF:F:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:
@A00155:405:HNMNHDSDXY:3:1101:5647:1125 2:N:0:GTGGATT+TCTATGTC
ACCCATACTCCCTCTCCCGTGGTCACTGGCTTCTGGTATTAGCTTGTAGAGCAGAGTAGGGTCACTGGTGTGATGCCAGACCCATCTCTCTGAAGACTGCCACGGCAAGGCTGGAGCG
+
,F:F:FFFFFFFFFF:FFFFF:FFFFF:FFF:F:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFF:FFFFF:,FFF
@A00155:405:HNMNHDSDXY:3:1101:8865:1125 2:N:0:GTGGATT+TCTATGTC
CGAGCGCTCGGTCAAGGAGAGGTACTTTGCACTGGAACTCTCTACTGCTCAGCAGCTGCTCGTGGTAAGGCAGTCGCCGCCGGTCTGCCGTAGACGTGGTGTGCCGCCGGCCTG
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:
@A00155:405:HNMNHDSDXY:3:1101:9842:1125 2:N:0:GTGGATT+TCTATGTC
TAGAACATAGCTTTAGCAGGAACCTGTTTACAGTCAGTTGACTCTGAAGTAAACACATATTAGCTGCTGTGATAAGTTCTTCACATGGAAAGAGGGACCCATGTTATCAGACAGGA
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:
@A00155:405:HNMNHDSDXY:3:1101:14615:1125 2:N:0:GTGGATT+TCTATGTC
TCAGGACTGCTGACTCGGCTGCACTGTCAGGAACTGCTGGCATAGCCGTTCTCCAGAGGGAGCAGACTGTGCCAGCACCCATCTCAGGCTCAGCCCGGAATCTGCCCTCTGGGG
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:
@A00155:405:HNMNHDSDXY:3:1101:15411:1125 2:N:0:GTGGATT+TCTATGTC
GTCTCCCTGAGAACGTTAGCAGCGCGTTTACCTGGGCCAGAAAAGAACATTCTCAGGGTGGCGCAAATTCAAAGGCCAGAAATGCTAGAACATCTCAGAACCTCTAGAA
+
FFFFFFFFFF:FFFFFFFFFF:FFF:FFF:F:FFF:F:FFF:F:FFF:F:FFF:F:FFF:F:FFF:F:FFF:F:FFF:F:FFF:F:FFF:F:FFF:F:FFF:F:FFF:F
@A00155:405:HNMNHDSDXY:3:1101:19569:1125 2:N:0:GTGGATT+TCTATGTC
ATTGTTAAATGGGAAGAGATGCTGTGTTACAACCCCTTAGGTTACAGCACAACACATCAAACCATACCAGTGGTATGGCTAGGACCAACCCAGTACAGAAATGAAGGA
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:
@A00155:405:HNMNHDSDXY:3:1101:20238:1125 2:N:0:GTGGATT+TCTATGTC
TCCTGACGTCACTGGCGCGACCCCCCGCCGCGCCAGCGCGCTGCCCAAGCTAGGCCGCCCAAGCGCGAGGAGCGGGCGGGCGAGCGCGTGTCCGGATCTCGCAG
+
FFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:
@A00155:405:HNMNHDSDXY:3:1101:21522:1125 2:N:0:GTGGATT+TCTATGTC
ATGGTACCGATTACCTGCCATTAAAGAACATCCACTGTCAGCTGGTAAGATATTCCAAAGGATAATCAAATTCCGACCATGGTAATGTGGCTCTGGAGTCTTGAG
+
FFFF:FFFF:FFFFFFFFFF:FFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:
@A00155:405:HNMNHDSDXY:3:1101:21630:1125 2:N:0:GTGGATT+TCTATGTC
AAATTCAGGAAACCCGAGAACAGAGAACAGAGGGAGGGGGTGGTGCAGCGCGAGGTCACTGGATTCTGGCTGGAAAGAGTGGAGTGGCGCTAGGACCTCTCC
+
FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:FFFFFFFFFF:
@A00155:405:HNMNHDSDXY:3:1101:25337:1125 2:N:0:GTGGATT+TCTATGTC
AAACAAATACACCTCTACTTCT
+
FFFFFFFFFF:FFFFFFFFFF:FFF:FFF:FFFFFFFFFF:FFF
@A00155:405:HNMNHDSDXY:3:1101:27489:1125 2:N:0:GTGGATT+TCTATGTC
ACCTGCACATTCTGCACATGTATCCAGAACATTAAAGTATAATTAAAAATAGCATTCTATAACTTCTACTGAATAATGAATAAGAATAATGAGAATAAGATTCTT
ACTAATCAAATTAGTCTGACAT





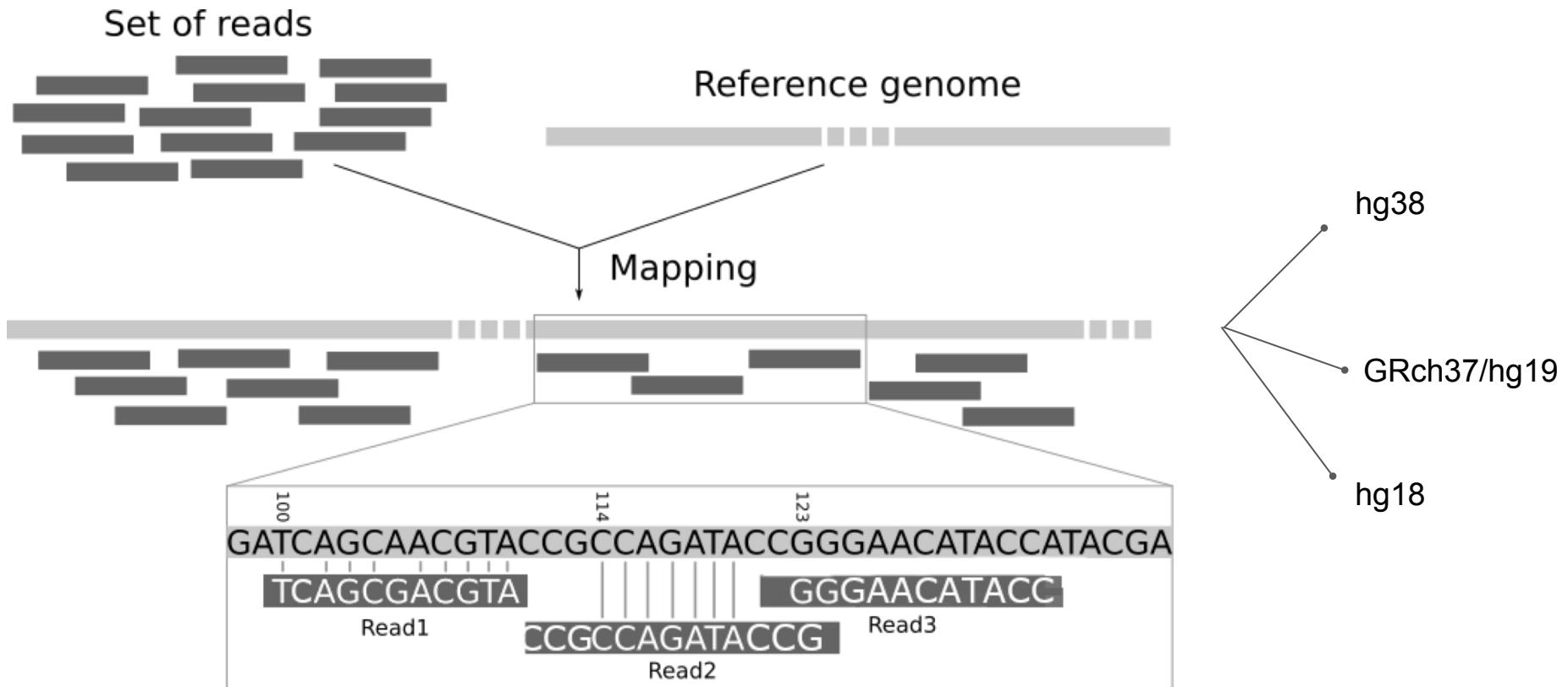
FastQC



Pipeline bioinformático



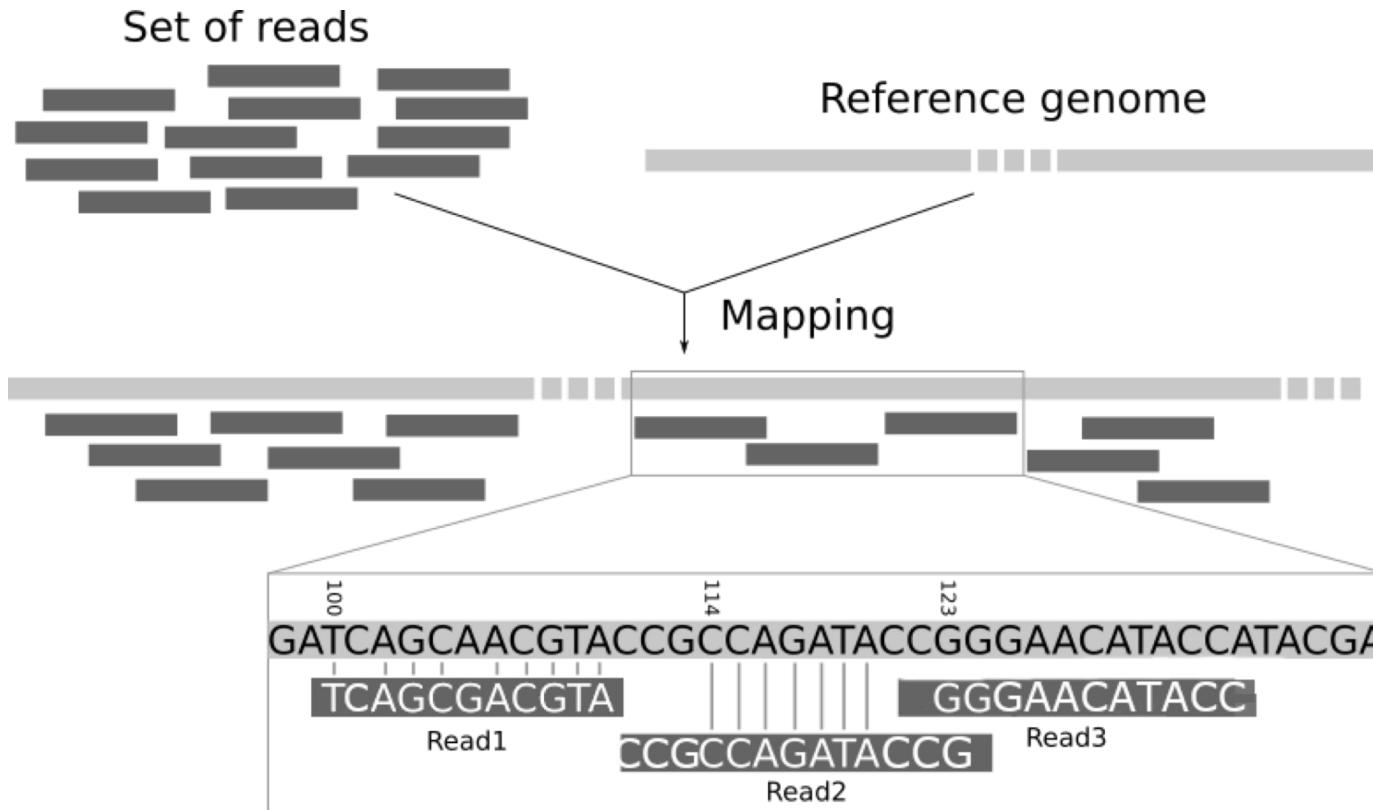
Mapeo



Mapeo

Problemas

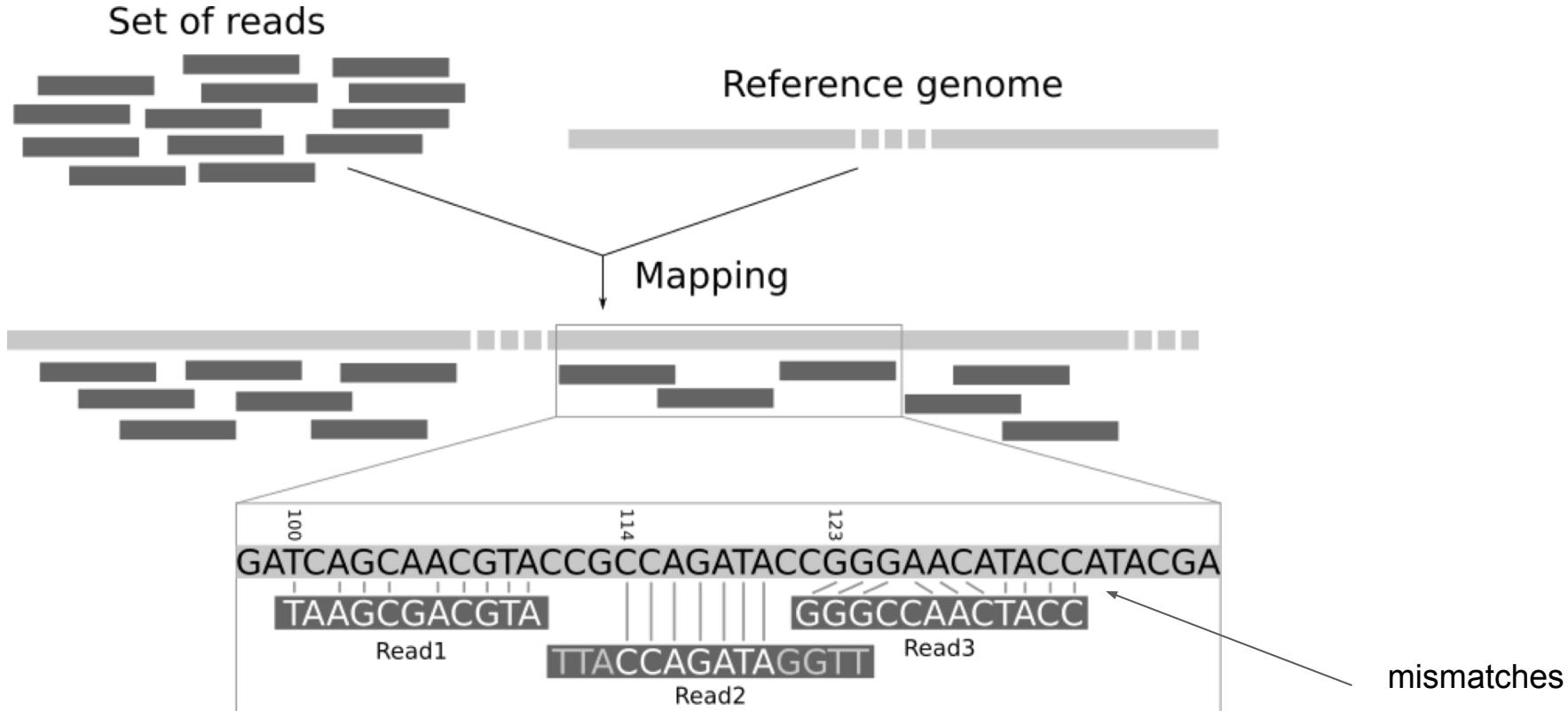
1. Genomas enormes y millones de reads
2. Reads son cortos
3. Regiones repetitivas
4. Mismatches (¿SNPs o errores?)



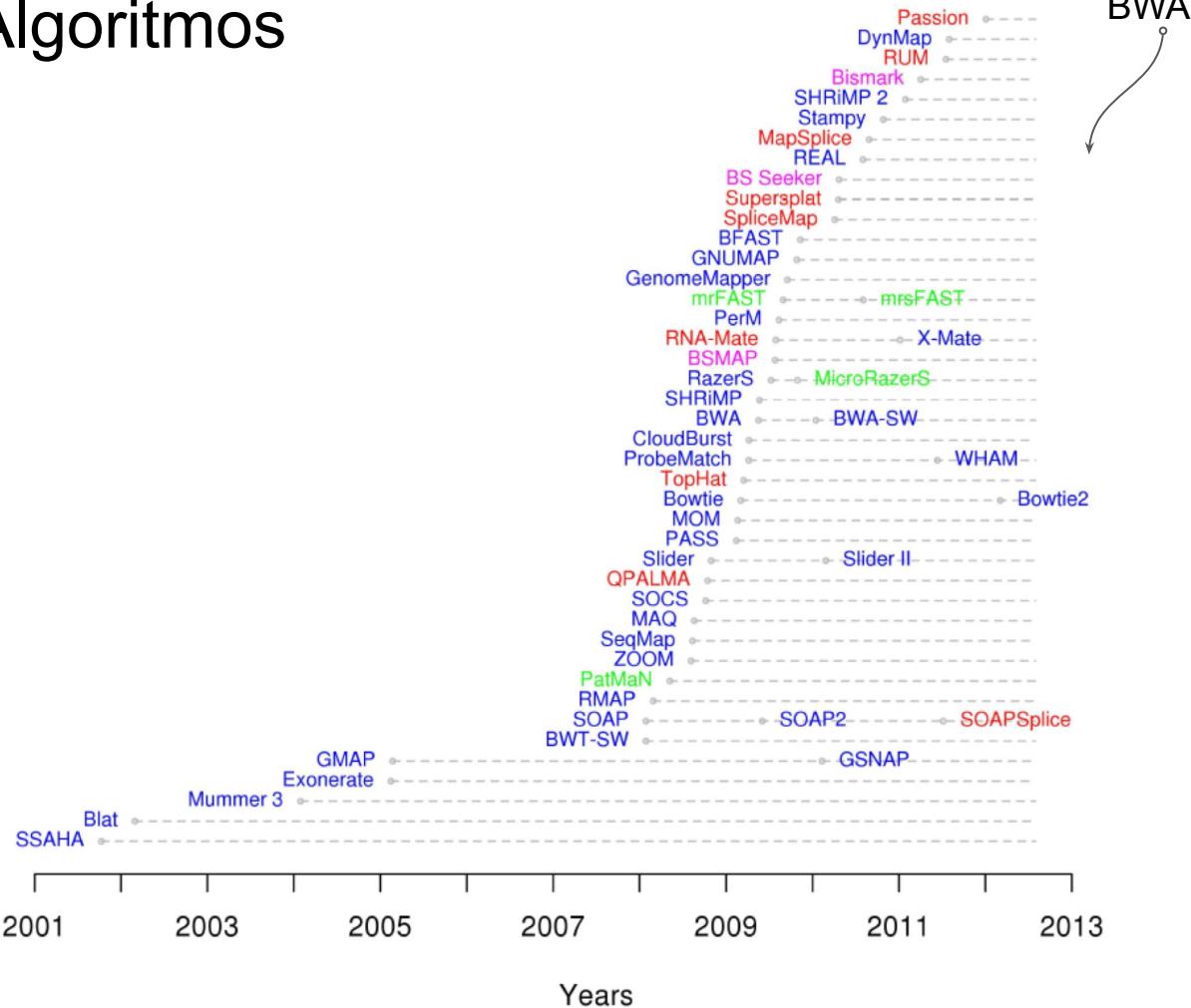
Mapeo

Problemas

1. Genomas enormes y millones de reads
2. Reads son cortos
3. Regiones repetitivas
4. Mismatches (¿SNPs o errores?)

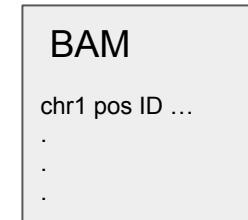
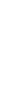


Algoritmos



BWA

- Unique mapped
- Unique mapped con mismatches
- Multiple mapped
- Multiple mapped con mismatches

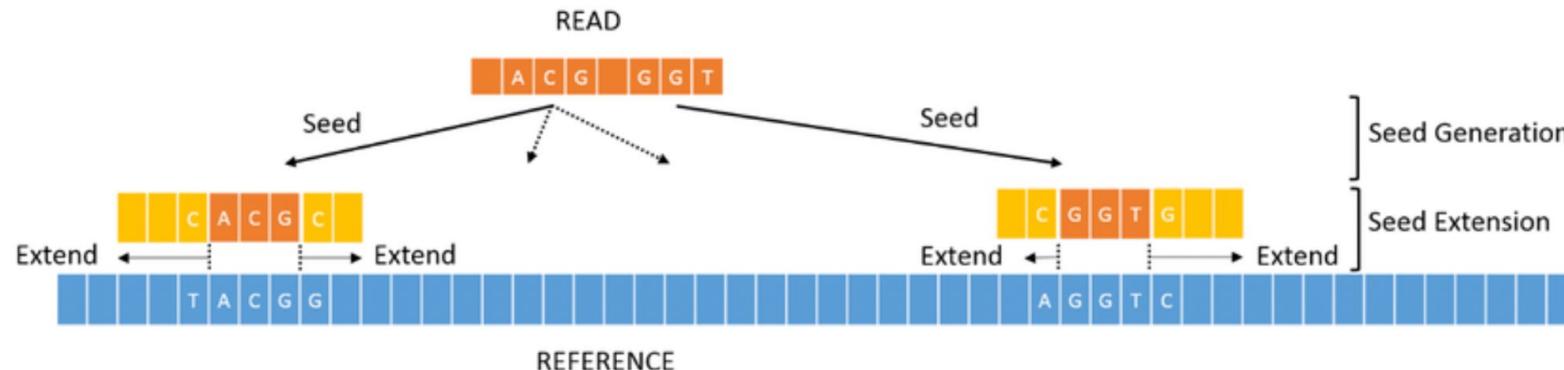


“Seed and extend”

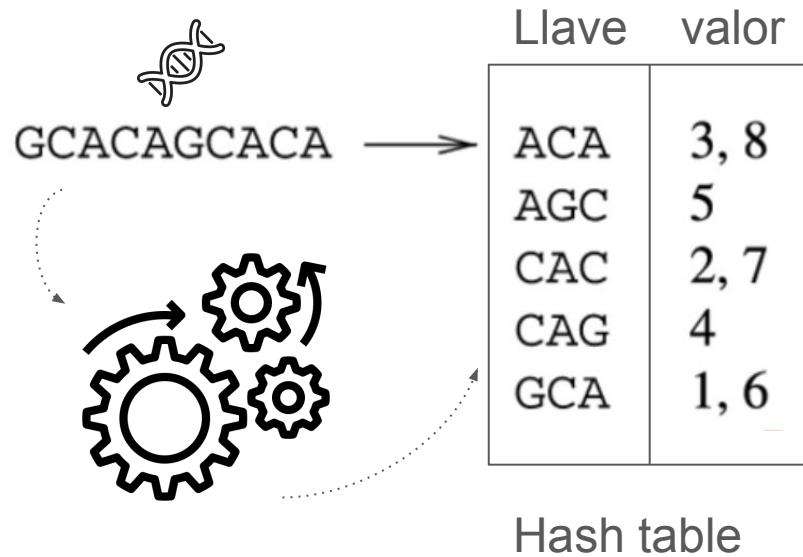
Muchos de estos algoritmos se basan en el mismo principio “seed and extend”.

Seed: mapear una pequena *semilla* al genoma

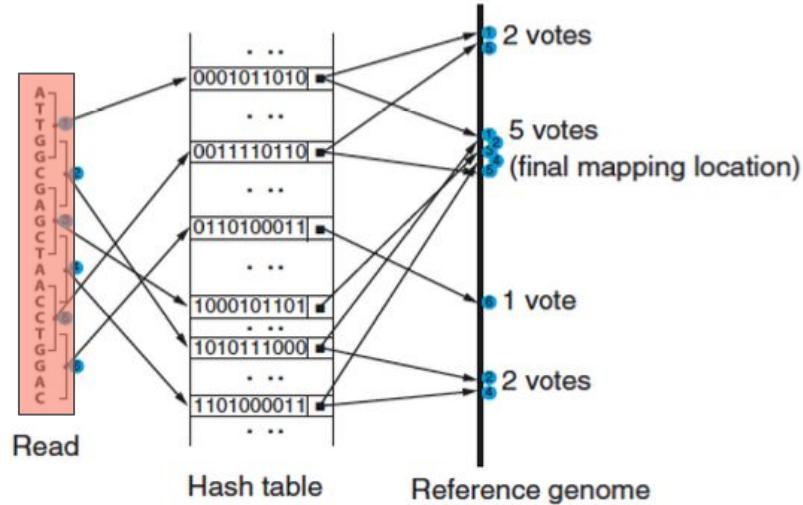
Extension: partiendo de donde mapeo la semilla, mapear el resto del read al genoma



Por qué es eficiente?



“Seed and vote”



SAM/BAM

```
@HD VN:1.3 SO:coordinate
@SQ SN:chr1 LN:249250621
@SQ SN:chr2 LN:243199373
@SQ SN:chr3 LN:198022430
@SQ SN:chr4 LN:191154276
@SQ SN:chr5 LN:180915260
@SQ SN:chr6 LN:171115067
@SQ SN:chr7 LN:159138663
@SQ SN:chr8 LN:146364022
@SQ SN:chr9 LN:141213431
@SQ SN:chr10 LN:135534747
@SQ SN:chr11 LN:135006516
@SQ SN:chr12 LN:133851895
@SQ SN:chr13 LN:115169878
@SQ SN:chr14 LN:107349540
@SQ SN:chr15 LN:102531392
@SQ SN:chr16 LN:90354753
@SQ SN:chr17 LN:81195210
@SQ SN:chr18 LN:78077248
@SQ SN:chr19 LN:59128983
@SQ SN:chr20 LN:63025520
@SQ SN:chr21 LN:48129895
@SQ SN:chr22 LN:51304566
@SQ SN:chrX LN:155270560
@SQ SN:chrY LN:59373566
@PG ID:bwa PN:bwa VN:0.7.12-r1039 CL:bwa sampe ../../bundle/bwa_index/Hs37.2.fa sample_output_file1.sai sample_output_file2.sai
/ /fastq/sample_output_file1.fastq / /fastq/sample_output_file2.fastq
```

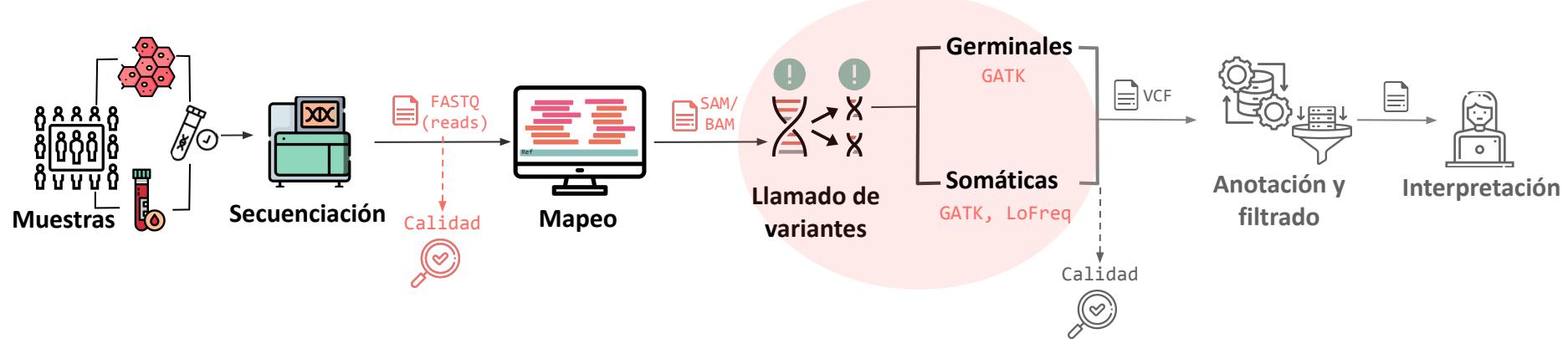
Header

Alineamiento

Pipeline bioinformático



Pipeline bioinformático



Por qué no es tan trivial?

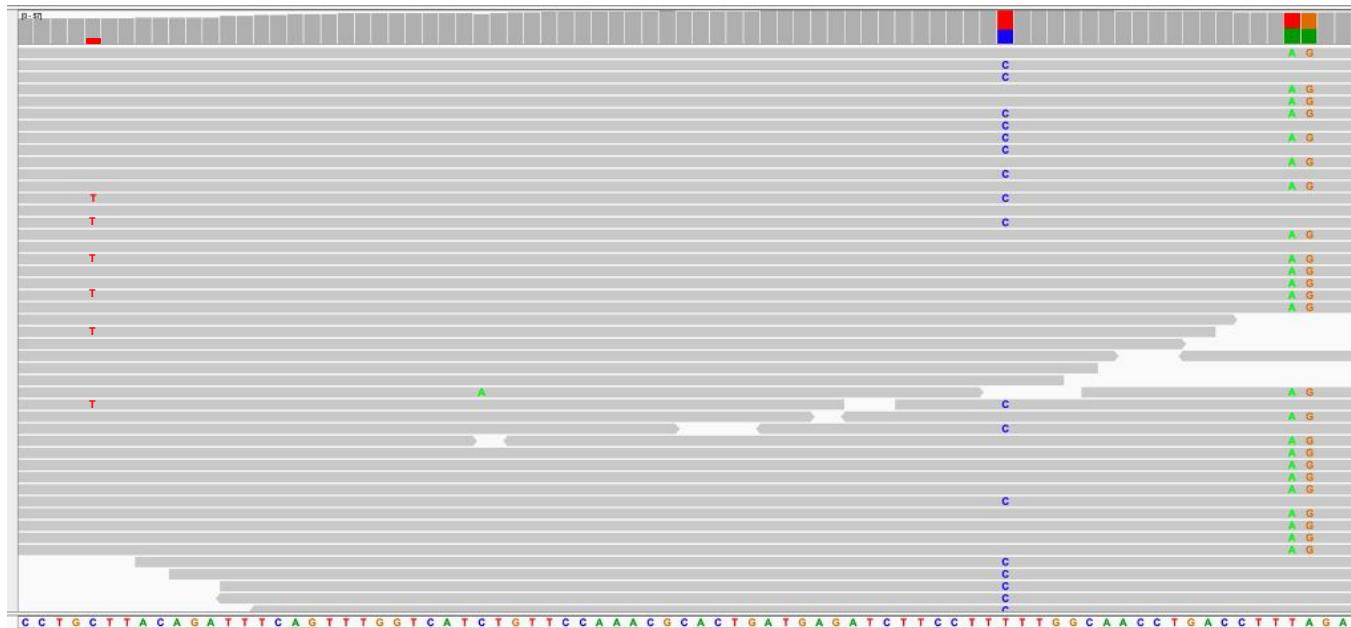


El proceso implica mucho más que simplemente contar los *reads* que caen en una posición determinada.

Por qué no es tan trivial?



El proceso implica mucho más que simplemente contar los *reads* que caen en una posición determinada.



Por qué no es tan trivial?



El proceso implica mucho más que simplemente contar los *reads* que caen en una posición determinada.



Los datos de secuenciación contienen errores inherentes (por ejemplo, tasas de error de la plataforma, artefactos de PCR o errores de alineamiento) que pueden generar señales espurias.

Por qué no es tan trivial?



El proceso implica mucho más que simplemente contar los *reads* que caen en una posición determinada.



Los datos de secuenciación contienen errores inherentes (por ejemplo, tasas de error de la plataforma, artefactos de PCR o errores de alineamiento) que pueden generar señales espurias.



Además, regiones repetitivas en el genoma dificultan el mapeo correcto de los *reads*, lo que incrementa los falsos positivos.

Por qué no es tan trivial?



El proceso implica mucho más que simplemente contar los *reads* que caen en una posición determinada.



Los datos de secuenciación contienen errores inherentes (por ejemplo, tasas de error de la plataforma, artefactos de PCR o errores de alineamiento) que pueden generar señales espurias.



Además, regiones repetitivas en el genoma dificultan el mapeo correcto de los *reads*, lo que incrementa los falsos positivos.

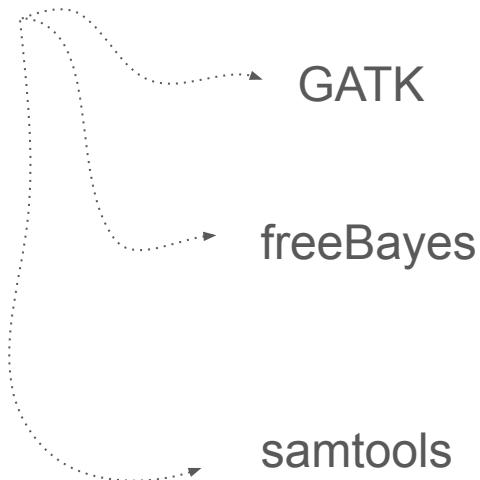


Factores como la cobertura desigual, el sesgo de GC, las inserciones y delecciones pequeñas, y las variantes estructurales complejas también complican la interpretación.

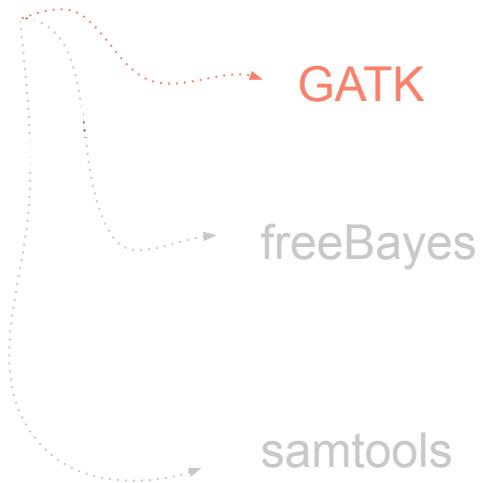


Los algoritmos de VC utilizan modelos estadísticos sofisticados, información de calidad de base, contexto del alineamiento y frecuencias alélicas para distinguir variantes reales de ruido técnico o errores de secuenciación.

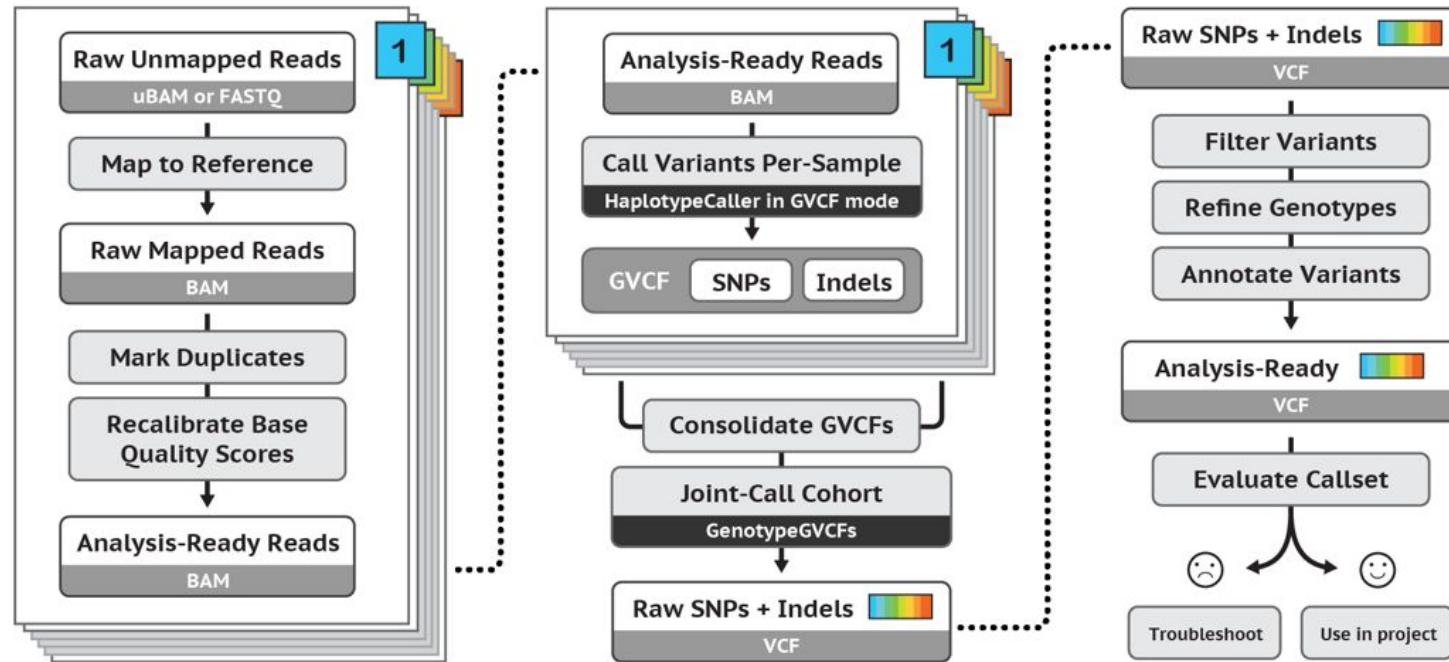
Llamado de variantes



Llamado de variantes



Genome Analysis ToolKit



Llamado de variantes

GATK
(Germinal)

Mark Duplicates

Base Recalibration

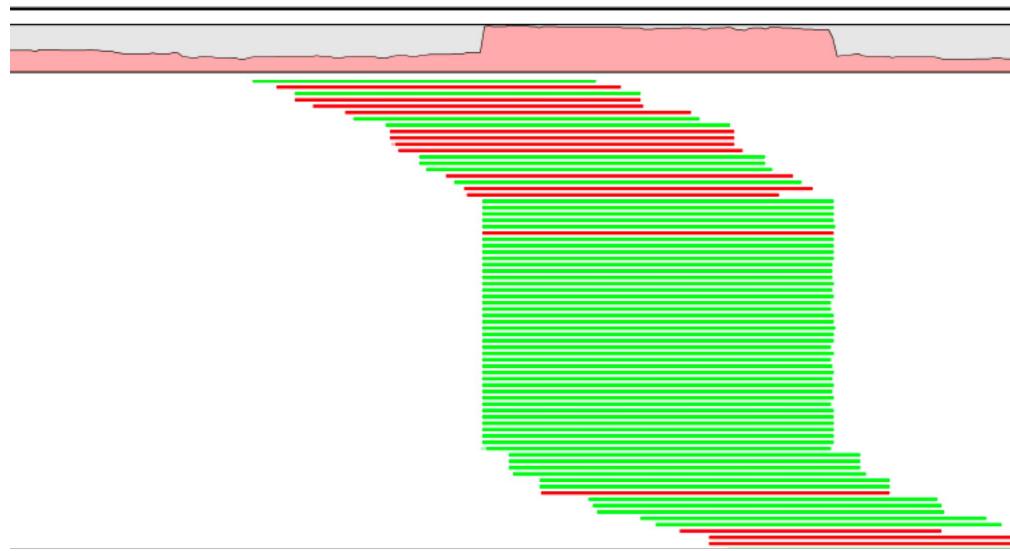
Haplotype Caller

Variant Filtration

Llamado de variantes



Mark Duplicates



Llamado de variantes

GATK
(Germinal)

Mark Duplicates

Base Recalibration

Haplotype Caller

Variant Filtration

Llamado de variantes

GATK
(Germinal)

Base Recalibration

Sesgos en calidad
dependiendo del
secuenciador, ciclo, etc.

Para cada lane, ciclo, phred score, base -> recalibro el PHRED

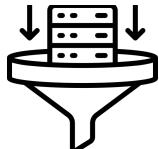
Llamado de variantes

GATK
(Germinal)

Base Recalibration

Sesgos en calidad
dependiendo del
secuenciador, ciclo, etc.

Para cada lane, ciclo, phred score, base -> recalibro el PHRED



Lane 1, 4ta pos. del read, en donde tengo una A y después otra A y donde
la A tiene un PHRED de 35

5 millones

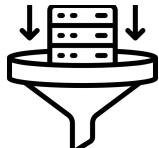
Llamado de variantes

GATK
(Germinal)

Base Recalibration

Sesgos en calidad
dependiendo del
secuenciador, ciclo, etc.

Para cada lane, ciclo, phred score, base -> recalibro el PHRED



Lane 1, 4ta pos. del read, en donde tengo una A y después otra A y donde la A tiene un PHRED de 35

5 millones

$$\frac{\# \text{mismatches} + 1}{\# \text{bases observadas} + 2} \rightarrow \text{Calidad recalibrada (nuevo Phred)}$$

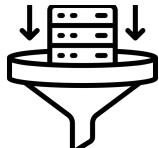
Llamado de variantes

GATK
(Germinal)

Base Recalibration

Sesgos en calidad
dependiendo del
secuenciador, ciclo, etc.

Para cada lane, ciclo, phred score, base -> recalibro el PHRED



Lane 1, 4ta pos. del read, en donde tengo una A y después otra A y donde la A tiene un PHRED de 35



Llamado de variantes

GATK
(Germinal)

Mark Duplicates

Base Recalibration

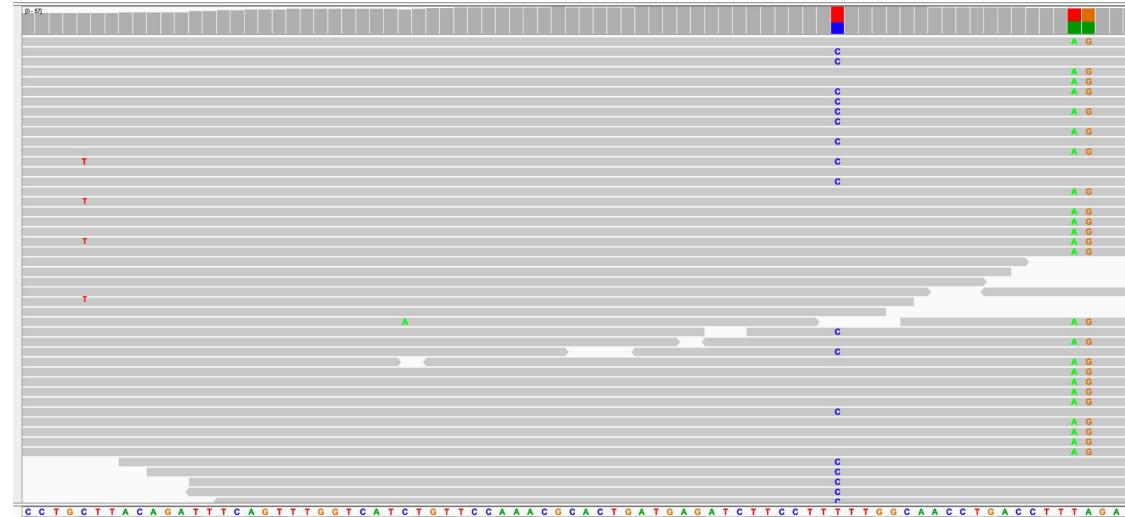
Haplotype Caller

Variant Filtration

Llamado de variantes

GATK
(Germinal)

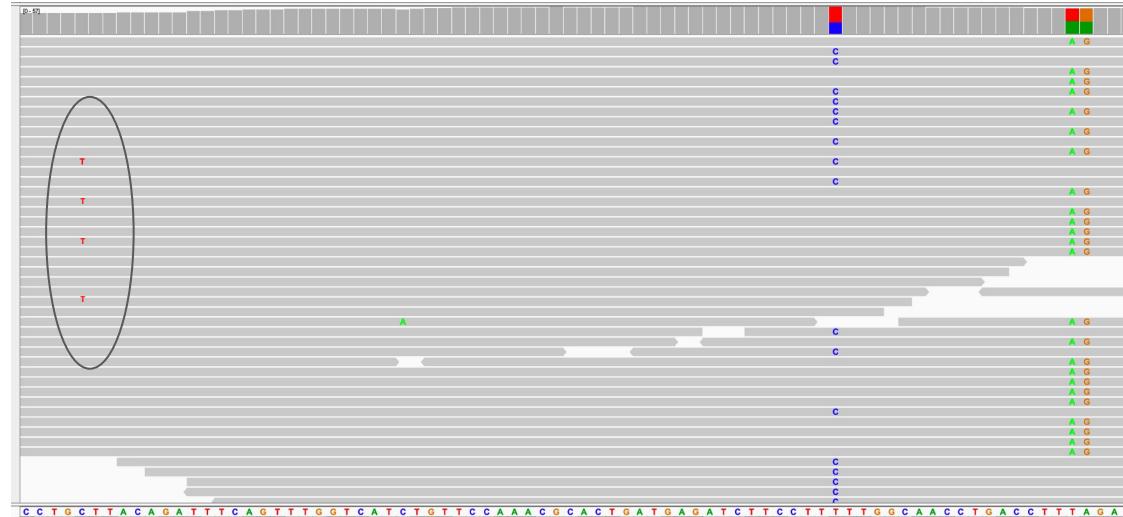
Haplotype Caller



Llamado de variantes

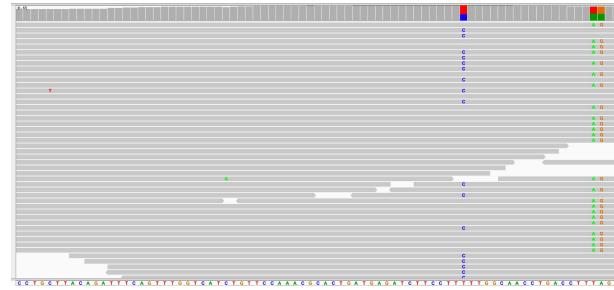
GATK
(Germinal)

Haplotype Caller



Identify ActiveRegions

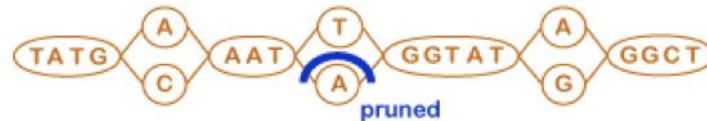
REF	TATGAAATTGGTATAGGCT
READS	A T G A T G A T G A A A C T A C T A C T A C T A T A



Identify ActiveRegions

REF	TATGAAATTGGTATAGGCT
READS	A T A T G A T G A A A C T A C T A C T A C T A T A

Assemble plausible haplotypes



→

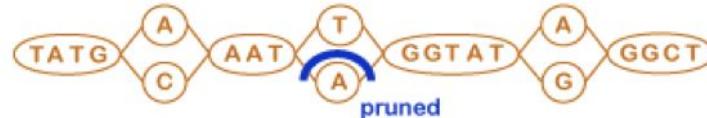
(-bamOut)

TATGAAATTGGTATAGGCT
A T A
A T G
C T A
C T A

Identify ActiveRegions

REF	TATGAAATTGGTATAGGCT	
READS		
A	T	→
A	T	G
A	T	G
A	A	A
C	T	A
C	T	A
C	T	A
C	T	A
T	A	

Assemble plausible haplotypes

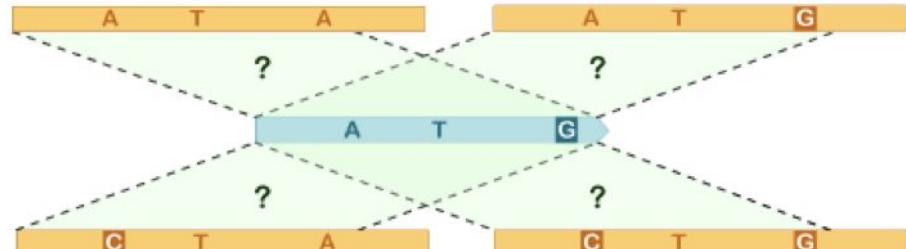


(-bamOut)

TATGAAATTGGTATAGGCT
A T A
A T G
C T A
C T G

Cuántos datos avalan cual haplotipo?

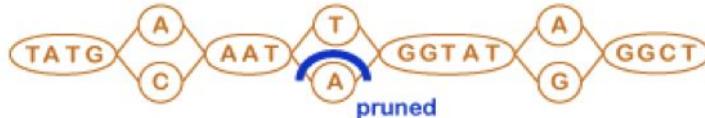
Determine per-read likelihoods (PairHMM)



Identify ActiveRegions

REF	TATGAAATTGGTATAGGCT	
READS	A T	A T G
	A A	G
	A T	G
	A A	A
	C T	A
	C T	A
	C T	A
	T A	

Assemble plausible haplotypes

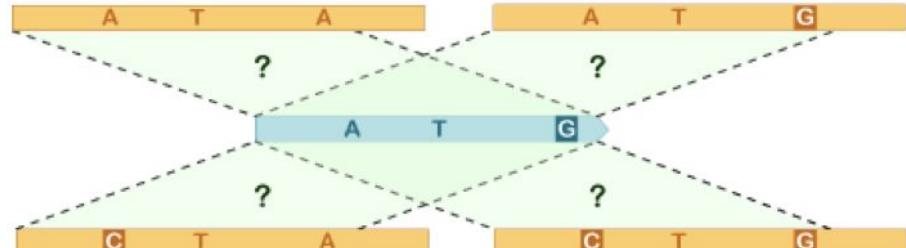


(-bamOut)

TATGAAATTGGTATAGGCT
A T A
A T G
C T A
C T G

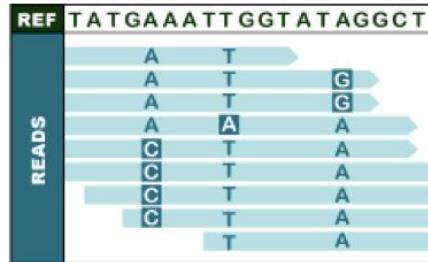
Cuántos datos avalan cual haplotipo?

Determine per-read likelihoods (PairHMM)

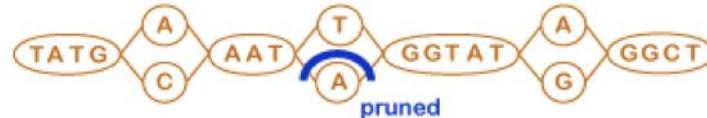


- Cada read se alinea con cada haplotipo
- Incluye calidad
- Se genera un score para cada par

Identify ActiveRegions



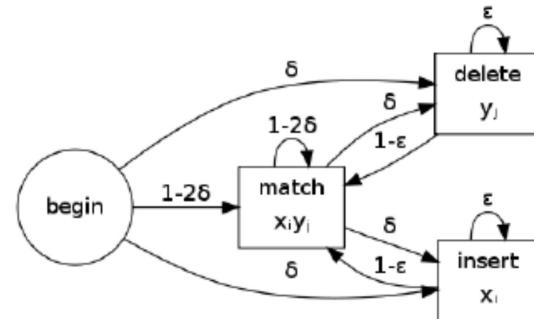
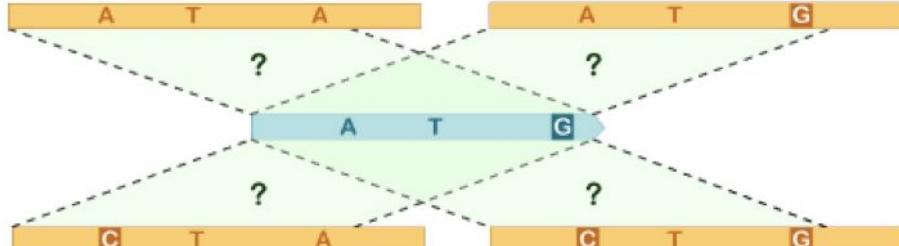
Assemble plausible haplotypes



(-bamOut)

T	A	T
A	T	G
C	T	A
C	T	G

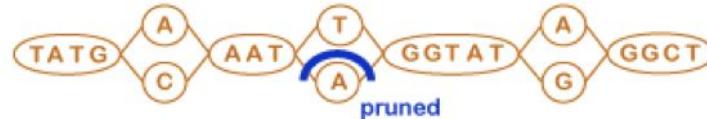
Determine per-read likelihoods (PairHMM)



Identify ActiveRegions

REF	TATGAAATTGGTATAGGCT
READS	A T A T G A T G A A A C T A C T A C T A C T A T A

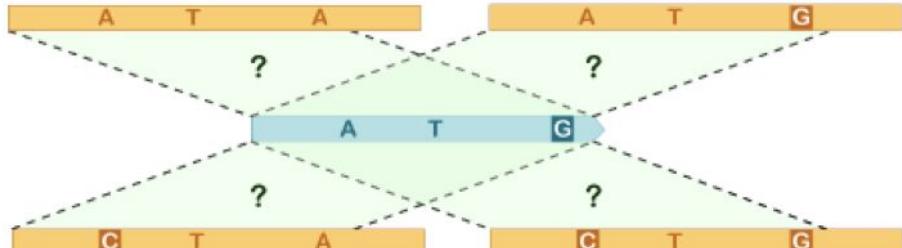
Assemble plausible haplotypes



(-bamOut)

T	A	G	A	T	G	C	T	A	T	A	G	G	C	T	G
---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

Determine per-read likelihoods (PairHMM)

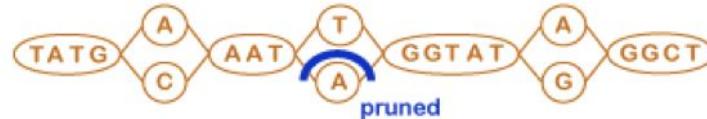


	hap1	hap2	hap3	hap4
read1	23	45	12	94
read2	12	34	21	78

Identify ActiveRegions

REF	TATGAAATTGGTATAGGCT
READS	A T A T G A T G A A A C T A C T A C T A C T A T A

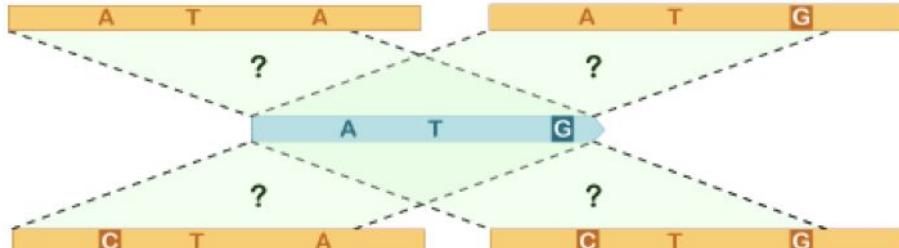
Assemble plausible haplotypes



(-bamOut)

TATGAAATTGGTATAGGCT
A T A
A T G
C T A
C T A

Determine per-read likelihoods (PairHMM)



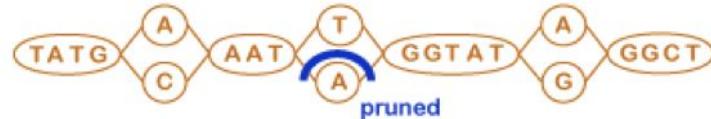
	hap1	hap2	hap3	hap4
read1	23	45	12	94
read2	12	34	21	78

$$P(G | D) = \frac{P(D | G) \times P(G)}{P(D)}$$

Identify ActiveRegions

REF	TATGAAATTGGTATAGGCT
READS	<pre> A T A T G A T G A A A C T A C T A C T A C T A T A </pre>

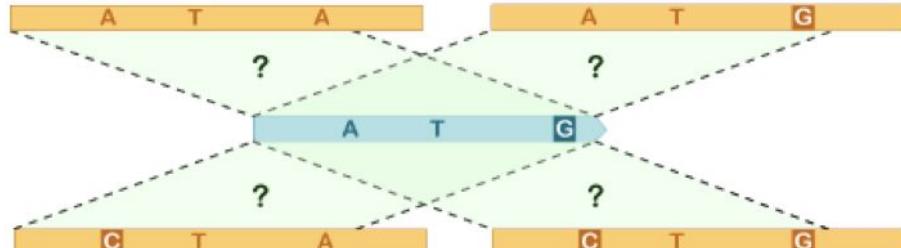
Assemble plausible haplotypes



(-bamOut)

T	A	T
A	T	G
C	T	A
C	T	G

Determine per-read likelihoods (PairHMM)



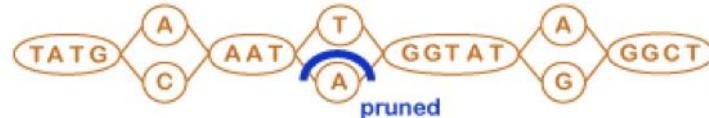
Depende de la frecuencia alélica asumida (0.001 para nuevos alelos, o se usan frecuencias de una base de datos)

$$P(G | D) = \frac{P(D | G) \times P(G)}{P(D)}$$

Identify ActiveRegions

REF	TATGAAATTGGTATAGGCT	
READS	A T	A T G
	A A	G
	A T	G
	A A	A
	C T	A
	C T	A
	C T	A
	T A	

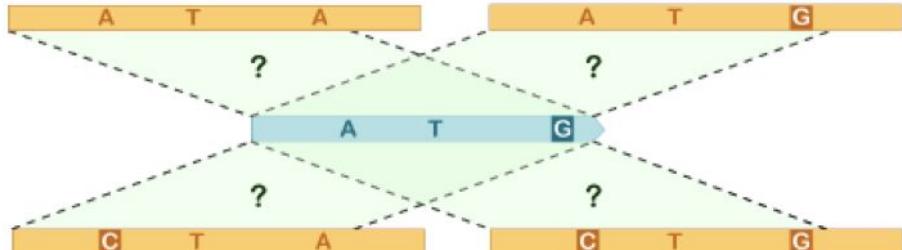
Assemble plausible haplotypes



(-bamOut)

TATGAAATTGGTATAGGCT
A T A
A T G
C T A
C T G

Determine per-read likelihoods (PairHMM)



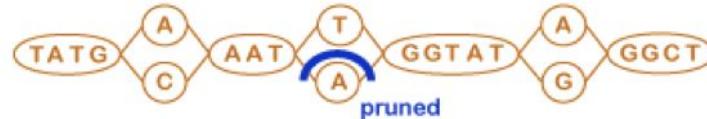
$$P(G | D) = \frac{P(D | G) \times P(G)}{P(D)}$$

$$P(D) = \sum_G P(D | G)P(G)$$

Identify ActiveRegions

REF	TATGAAATTGGTATAGGCT
READS	A T A T G A T G A A A C T A C T A C T A C T A T A

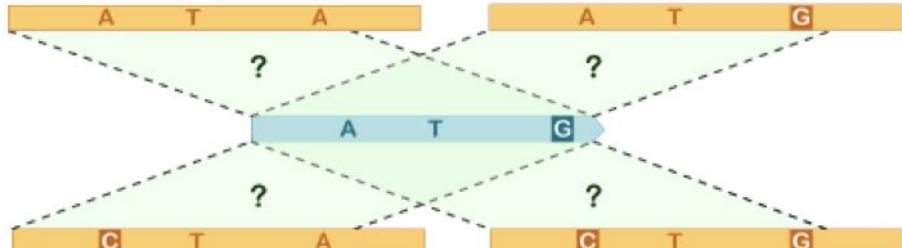
Assemble plausible haplotypes



(-bamOut)

T	A	G
A	T	G
C	T	A
C	T	G

Determine per-read likelihoods (PairHMM)



Genotype sample

	0/0	0/1	1/1
A/C			
A/G			

GLs + annotations

Llamado de variantes

GATK
(Germinal)

Mark Duplicates

Base Recalibration

Haplotype Caller

Variant Filtration

Llamado de variantes

GATK
(Germinal)

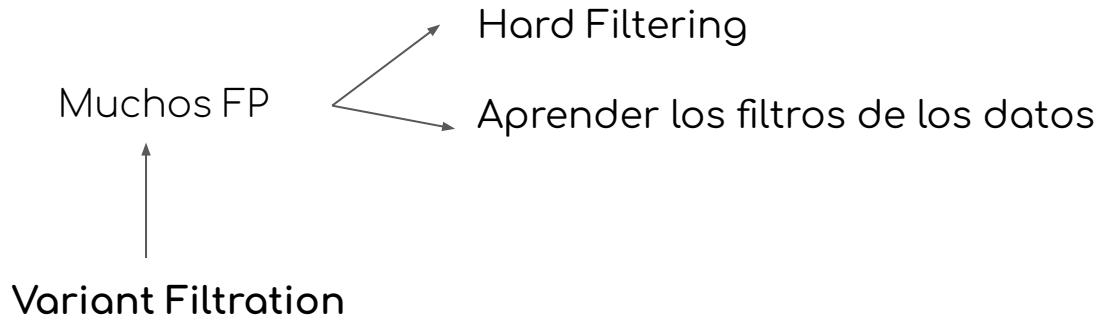
Variant Filtration

Llamado de variantes



Muchos FP
↑
Variant Filtration

Llamado de variantes

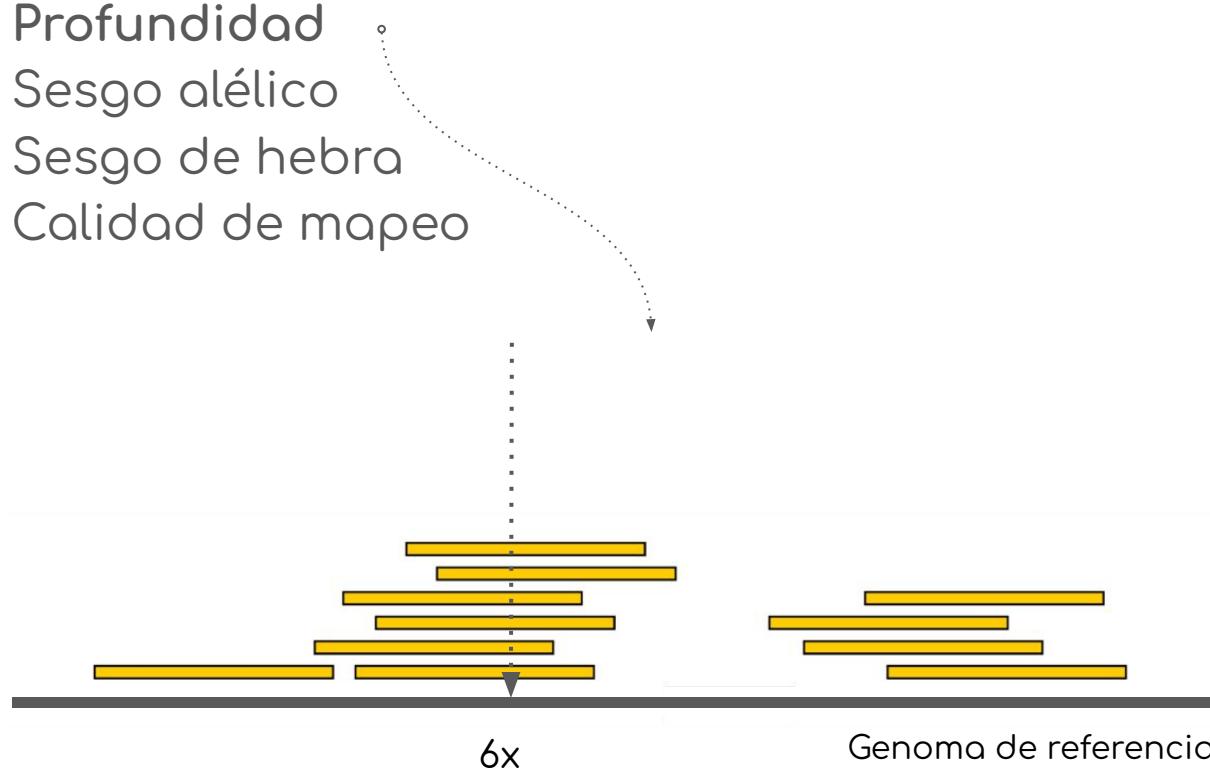


“Hard filtering”

- Profundidad
- Sesgo alélico
- Sesgo de hebra
- Calidad de mapeo

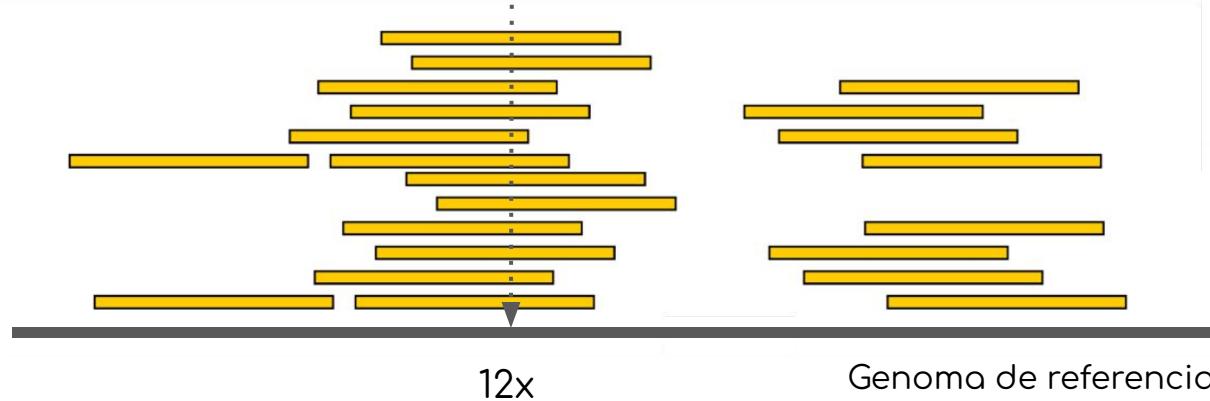
“Hard filtering”

- Profundidad
- Sesgo alélico
- Sesgo de hebra
- Calidad de mapeo



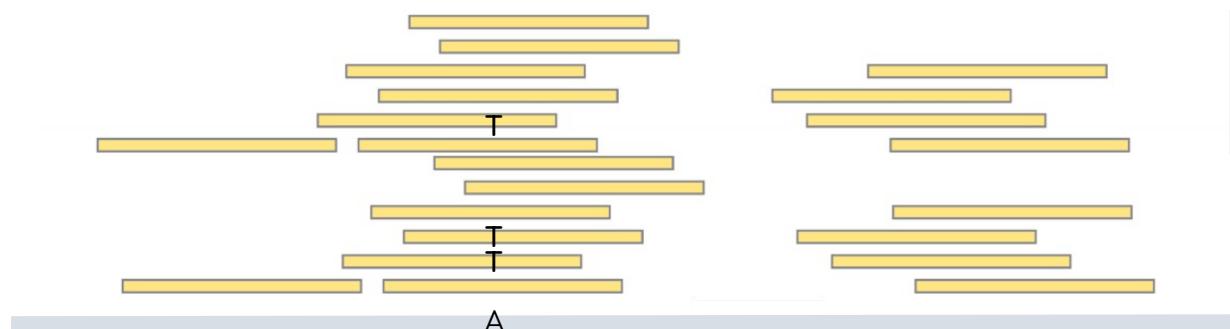
“Hard filtering”

- Profundidad
- Sesgo alélico
- Sesgo de hebra
- Calidad de mapeo



“Hard filtering”

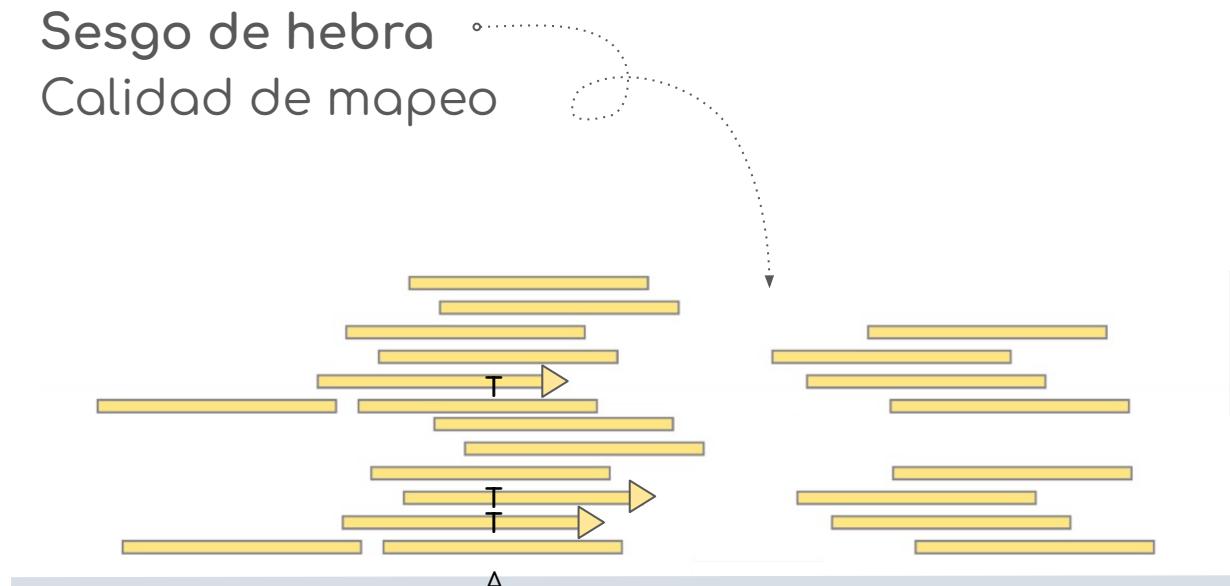
- Profundidad
- Sesgo alélico
- Sesgo de hebra
- Calidad de mapeo



3T y 9A

“Hard filtering”

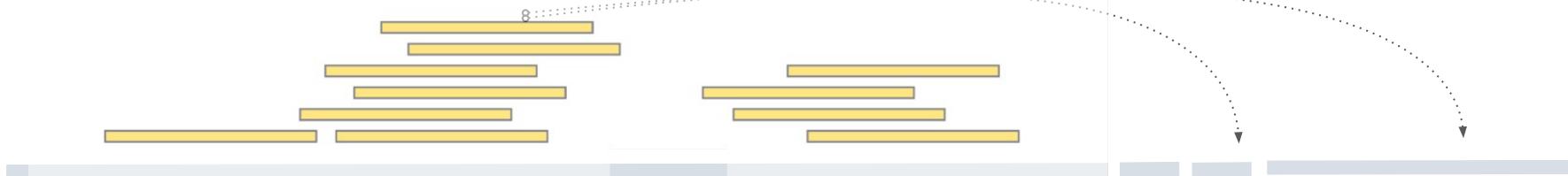
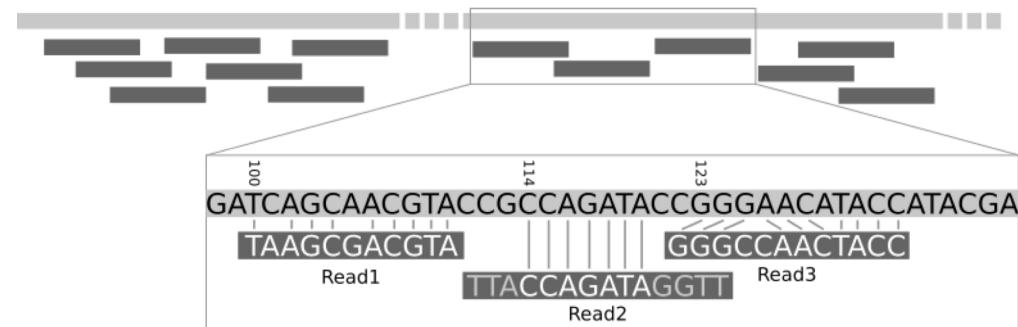
- Profundidad
- Sesgo alélico
- **Sesgo de hebra**
- Calidad de mapeo



3T y 9A

“Hard filtering”

- Profundidad
- Sesgo alélico
- Sesgo de hebra
- Calidad de mapeo



Aprender los patrones de error de los datos

Variant Quality Score Recalibration

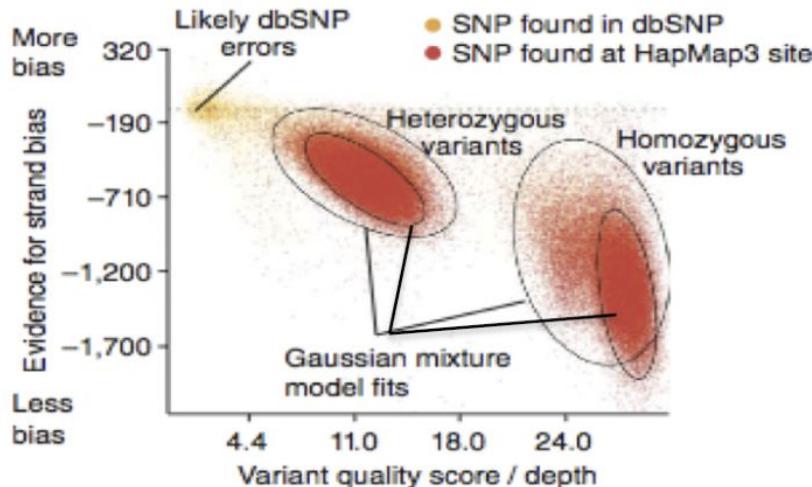
El VQSR en GATK es un algoritmo basado en modelos mixtos gausianos que buscan clasificar varintesthat aims to classify variants based on how their annotation values cluster given a training set of high-confidence variants([2.3](#)).

Aprender los patrones de error de los datos

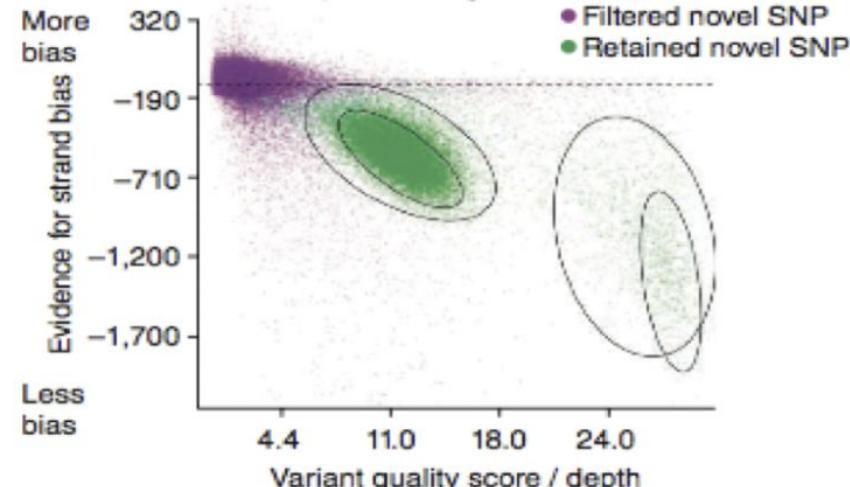
Variant Quality Score Recalibration

El VQSR en GATK es un algoritmo basado en modelos mixtos gausianos que buscan clasificar variantes que aims to classify variants based on how their annotation values cluster given a training set of high-confidence variants(2.3).

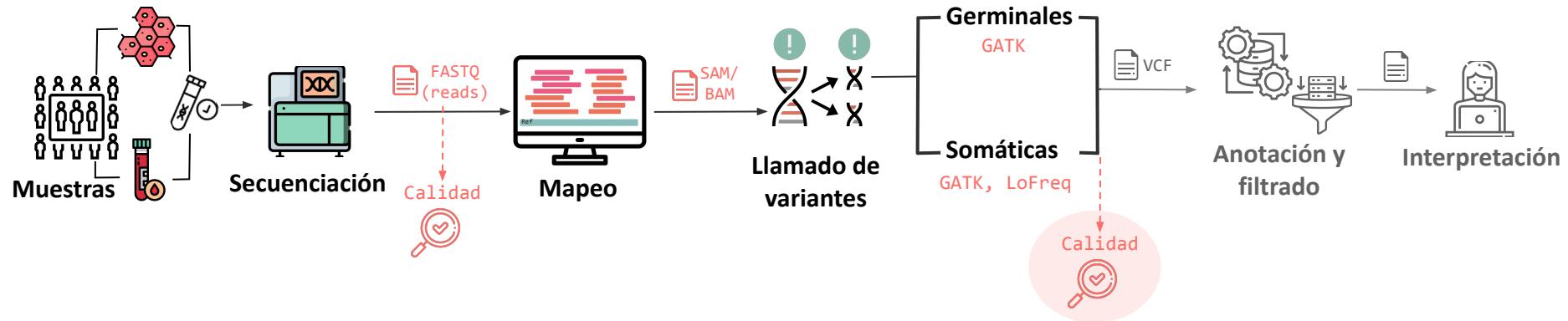
Model trained on validated VCF



Model applied to new SNPs

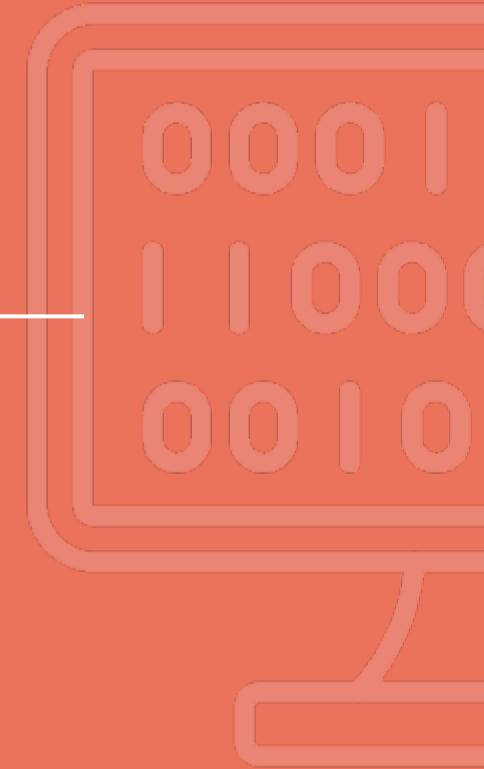


Pipeline bioinformático





Calidad Concepto #2



GACCGATGACCCGGTTCAGGCTTCACCACAGTGTGGAACGCCGCGTCCTCGAACTTAACGGCGACC
TAAGGTTGACGACGGACCCAGCAGTGTCAATCTCAGCGCTCCGCTGACCCCTCAGCAAAGGGCTTG
GCTCAATCTCGTCCAGCCATTGACCATCGTGAGGGGTTGCTCTGTATCCGTGCCGAGCAGCTTG
AAAAACGAAATCGAGGCCATCTGGGGCCCCGATTACCGACGCTCTAGCCGCCACTCGGACATCAGA
CCAACCTGGGGTCCGCATCGCTCCGCCGGCAGCAGAAGCCGACGACACTACCGTCCGCCCTCGA
ATCCTGCTACCACATGCCAGACACCACAACGACAACGAGGAGATTGATGACAGCGCTGCCAACGG
CGATAACCAGCACAGTGGCAAGTTACTTCACCGAGCGCCGACAATACCGATTCCGTAACCGCTG
GCGTAACCAGCTTAACCGTCGCTACACCTTGATACGTTGTTACGGCCCTCCAACCGTTCGCGCA
CGCCGCCGCTTGGCGATCGCAGAAGCACCCGCCGCGTTACAACCCCCCTGTTCATCTGGGGCGAGTCC
GGTCTCGGAAGACACACCTGCTACACGCCAGGCAACTATGCCAACGGTTCCGGGAATGCCGG
AAAATATGCTCCACCGAGGAATTACCAACGACTTCATTAACTCGCTCCGCCATGACCGCAAGGTCGC
CTCAACCGCAGTACCGCAGTAGACGTTGCTGCGACCGACATCCAATTCAAGGCAAAGAG
GTATTCAAGAGGGAGTTCTCCACACCTTCAACACCGTCAACCGTCCACCGGAGGAGGAGGAGGAG
CTGACCGCCACCCAAAGCAGCTGCCACCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
ACTGACGTACAACCAACCCGAGCTGGAGACCCGATCGGCAAGGAGGAGGAGGAGGAGGAGGAGG
TCGCGGTCCCCGACGATGTCCTCGAACTCATCGCCAGCAGTATCGAACGCAATATCCGTGAACTCGAGG
CGCGCTGATCCGGTCAACCGCGTTGCCGATCGGCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
TGCTTCGGATCTGATCGCCACGCCATCGGCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
GCCGAATACTCGACACTACCGTCGAAGAGCTCGGGGGCCGCAAGGCCGAGCAGTGGCCAGTCAC
GACAGATTGCGA
CCGTGATCACACAACCGTCACTGTACGCCAACGCAAGATCCTGCGAGATGGCCGAGGCCGTCAGGTC
TTGATCACGTCAAAGAACTCACCACTCGCATCCGTCAGCGCTCCAAGCGTAGCACGGCGTTCTCC
CAACGTTCTAAAAAAACTCTCTCCAGGTACACCCAGTCAACAGAGATTGGCTGTGAGTGTGCT
TGCACAAACCGCGCACAGACTCATACAGTCCCAGGTTCCGTTACAACCCACGCCCTCATCCCCACCG
ACCCAAACACACACCCACAGTCATGCCACCGTCATCCACAACCTCGACCGACGTCGACCTGCACCAAGA
CCAGACTGCCCCAAACTGCACACCCCTCTAATACTGTTACCGAGATTCTCGTCGTTGTTCTGGAAA
ACAGCGCTGGGATCGTCGCTGGATACCAACCCGCTAACACTGGCTCGCGGTGGTCAGAGGTCAAT
CTCAACTTCAAGTTGACGTGAGAACGCTACCGTTGTTGACTGCTGTTGCCGCCGTCGGCG
TCACCGCGTCATGGGCATTGTCGTTGGCAGTCCCCACGCTAGCGGGGCGTAGCCACGGGATCGAACTC
ATCGTGAGGTGAAAGGGCGCAATGGACGCCGCTACGACAAGAGATTGGCCTCACCGACTTGACGTTG
TGCTACGAGAGTCTTCGCCATGCCGTTGCGTGGCTAAAAATCTGCCAGCCAGGCCGCCGCGTGC
GGTCTCCGGCGTGTGTTGACCGGCTCGGACAACGGTCTGACGATTCCGGATTGACTACCGAGGTT
CGCCGAGGCCAGGTTGGCGCTGAAATTGTTCTCCTGGAAGCGTTAGTTCTGGCCGATTGTTGT

Calidad: Concepto #2

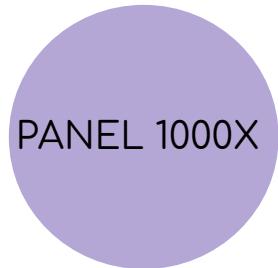
1. Profundidad
2. On Target
3. Cobertura

1.

Profundidad esperada



→ 60MB tamaño de exoma



→ 5MB tamaño de panel

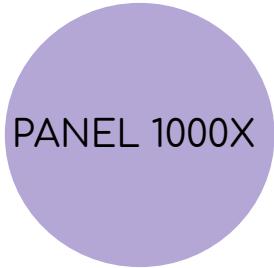


1.

Profundidad esperada



→ 60MB tamaño de exoma



→ 5MB tamaño de panel

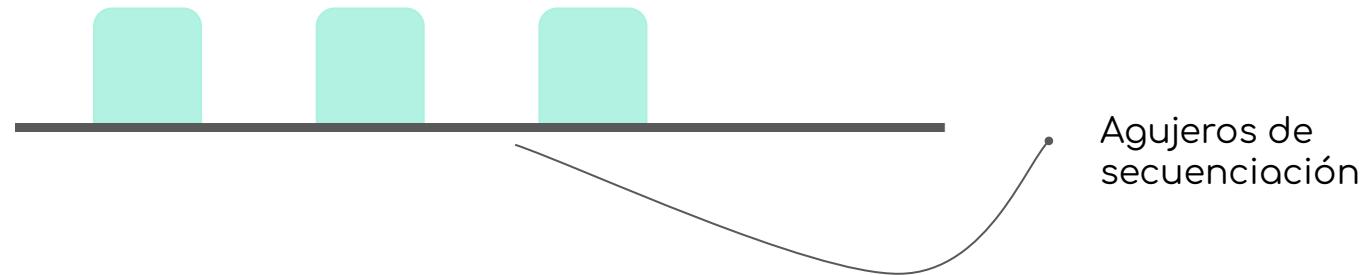


1.

Profundidad
esperada



Profundidad
obtenida

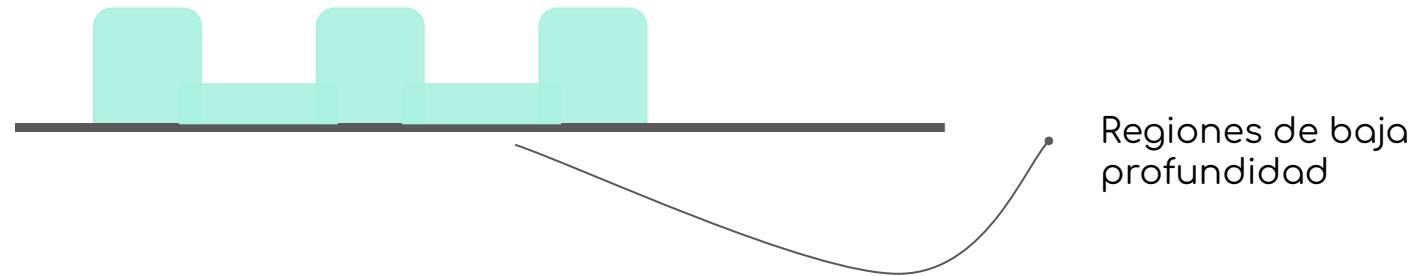


1.

Profundidad
esperada



Profundidad
obtenida



2.

On Target

¿Cuánto cubro de mis regiones de interés?



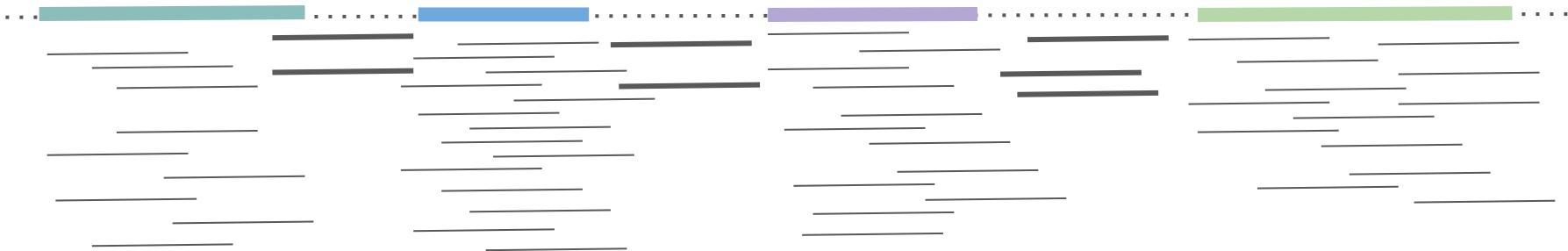
On Target

¿Cuánto cubro de mis regiones de interés?



On Target

¿Cuánto cubro de mis regiones de interés?



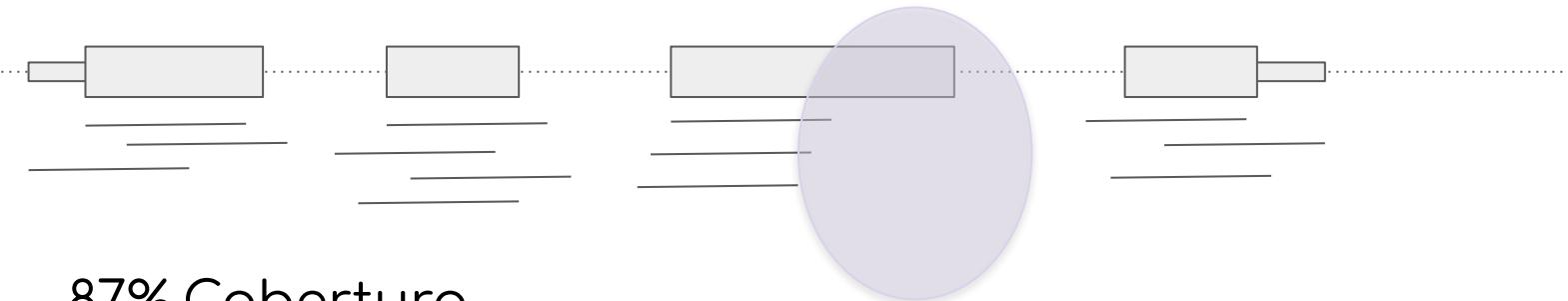
% de reads que cubren mi región de interés

31/38=0.82 -> 82% on target

3.

Cobertura

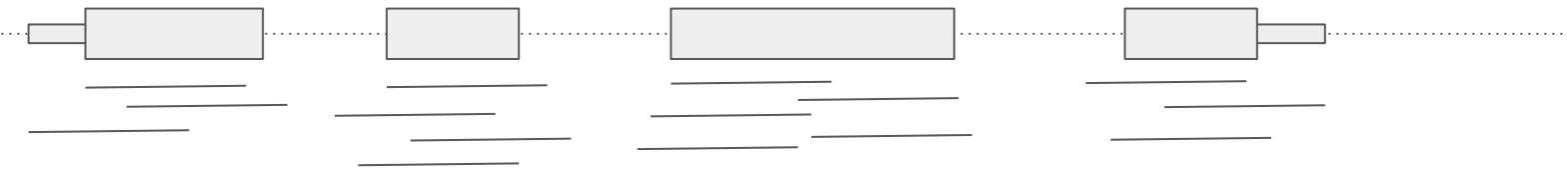
BRCA1



3.

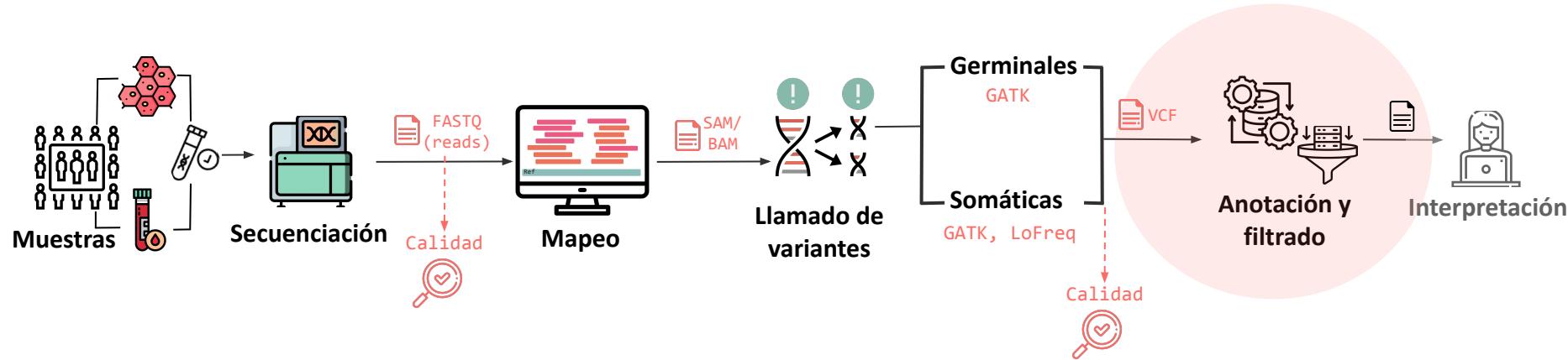
Cobertura

BRCA1

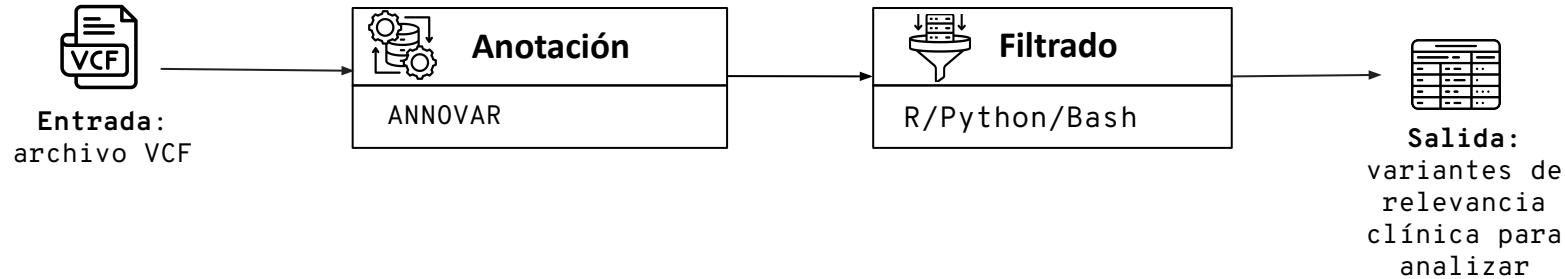


100% Cobertura

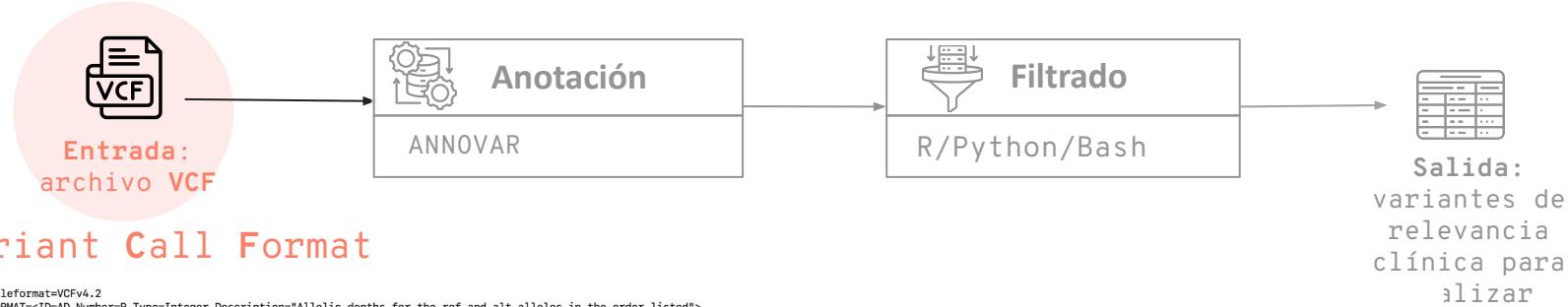
Pipeline bioinformático



Anotación y filtrado



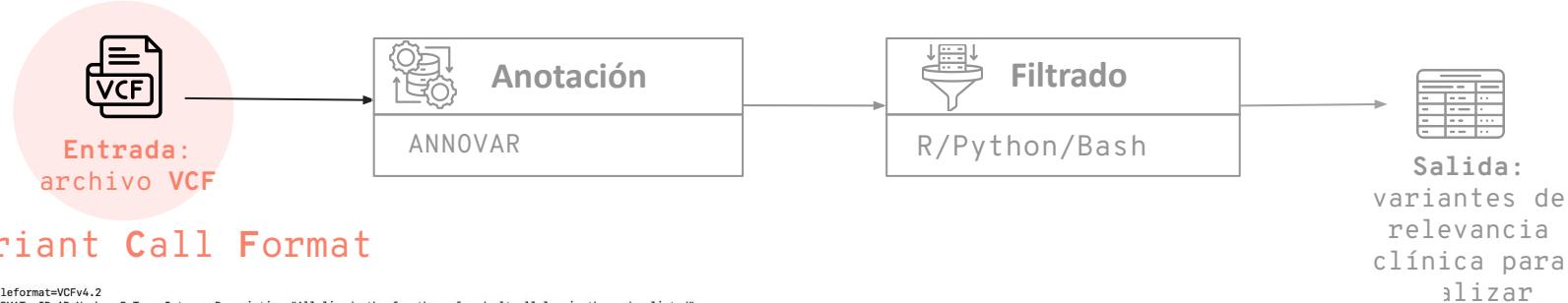
Anotación y filtrado



Variant Call Format

```
#fileformat=VCFv4.2
##FORMAT=Type=Integer,Description="Allelic depths for the ref and alt alleles in the order listed">
##FORMAT=Type=Integer,Description="Read depth">
##FORMAT=Type=Integer,Description="Genotype quality">
##FORMAT=Type=String,Description="Genotype">
##FORMAT=Type=String,Description="Physical phasing haplotype information, describing how the alternate alleles are phased in relation to one another">
##FORMAT=Type=String,Description="Physical phasing ID information, where each unique ID within a given sample (but not across samples) connects records within a phasing group">
##FORMAT=Type=Integer,Description="The phred-scaled genotype likelihoods rounded to the closest integer">
##INFO=Type=Integer,Description="Allele count in genotypes, for each ALT allele, in the same order as listed">
##INFO=Type=Float,Description="Allele frequency, for each ALT allele, in the same order as listed">
##INFO=Type=Integer,Description="Total number of alleles in called genotypes">
##INFO=Type=Float,Description="Z-score from Wilcoxon rank sum test of Alt Vs. Ref base qualities">
##INFO=Type=Float,Description="Z-score From Wilcoxon rank sum test of Alt vs. Ref number of hard clipped bases">
##INFO=Type=Flag,Description="dbSNP membership">
##INFO=Type=Flag,Description="Combined depth across samples">
##INFO=Type=Float,Description="Phred-scaled p-value for exact test of excess heterozygosity">
##INFO=Type=Float,Description="Phred-scaled p-value using Fisher's exact test to detect strand bias">
##INFO=Type=Float,Description="Inbreeding coefficient as estimated from the genotype likelihoods per-sample when compared against the Hardy-Weinberg expectation">
##INFO=Type=Integer,Description="Maximum likelihood expectation (MLE) for the allele counts, for each ALT allele, in the same order as listed">
##INFO=Type=Float,Description="Maximum likelihood expectation (MLE) for the allele frequency, for each ALT allele, in the same order as listed">
##INFO=Type=Float,Description="MQ mapping quality">
##INFO=Type=Float,Description="Z-score From Wilcoxon rank sum test of Alt vs. Ref read mapping qualities">
##INFO=Type=Float,Description="Variant Confidence/Quality by Depth">
##INFO=Type=Float,Description="Z-score from Wilcoxon rank sum test of Alt vs. Ref read position bias">
##INFO=Type=Float,Description="Symmetric Odds Ratio of 2x2 contingency table to detect strand bias">
##SentieonCommandLine.GVCFtyper,Version="sentieon-genomics-202112.07",Date="2025-01-27T20:26:17Z",CommandLine="/opt/sentieon/sentieon-genomics-202112.07/libexec/driver -r bundle/b37decoy_FASTA/human_g1k_v37_decoy.fasta -t -3 -interval bundle/b37decoy_FASTA/human_g1k_v37_decoy.nodecoy.bed -algo GVCFtyper -d bundle/b37decoy_VCF/db.snp_138.b37.vcf.gz -v 189DV-output.g.vcf.gz 189DV-output.vcf.gz"
##SentieonCommandLine.Haplotype,Version="sentieon-genomics-202112.07",Date="2025-01-27T20:23:54Z",CommandLine="/opt/sentieon/sentieon-genomics-202112.07/libexec/driver -r bundle/b37decoy_FASTA/human_g1k_v37_decoy.fasta -t 36 -i 189DV-deduped.bam -o 189DV-recal_data.table --interval bundle/b37decoy_FASTA/human_g1k_v37_decoy.nodecoy.bed --algo Haplotype -d bundle/b37decoy_VCF/db.snp_138.b37.vcf.gz --emit_mode gvcf --emit_conf=30 -call_conf=30 189DV-output.g.vcf.gz"
##contig=<ID=1,length=249250621,assembly=b37>
##contig=<ID=2,length=243199373,assembly=b37>
##contig=<ID=3,length=198022439,assembly=b37>
##contig=<ID=4,length=191154276,assembly=b37>
##contig=<ID=5,length=188915268,assembly=b37>
##contig=<ID=6,length=171115067,assembly=b37>
##contig=<ID=7,length=159138663,assembly=b37>
##contig=<ID=8,length=146364022,assembly=b37>
##contig=<ID=9,length=141213431,assembly=b37>
##contig=<ID=10,length=135534747,assembly=b37>
##contig=<ID=11,length=135086516,assembly=b37>
##contig=<ID=12,length=133851895,assembly=b37>
##contig=<ID=13,length=115169278,assembly=b37>
##contig=<ID=14,length=107349540,assembly=b37>
##contig=<ID=15,length=102531392,assembly=b37>
##contig=<ID=16,length=98354753,assembly=b37>
```

Anotación y filtrado



Variant Call Format

```
#fileformat=VCFv4.2
##FORMAT=Description="Allelic depths for the ref and alt alleles in the order listed"
##FORMAT=Description="Read depth"
##FORMAT=Description="Genotype quality"
##FORMAT=Description="Genotype"
##FORMAT=Description="Physical phasing haplotype information, describing how the alternate alleles are phased in relation to one another"
##FORMAT=Description="Physical phasing ID information, where each unique ID within a given sample (but not across samples) connects records within a phasing group"
##FORMAT=Description="The phred-scaled genotype likelihoods rounded to the closest integer"
##INFO=Description="Allele count in genotypes, for each ALT allele, in the same order as listed"
##INFO=Description="Allele frequency, for each ALT allele, in the same order as listed"
##INFO=Description="Z-score from Wilcoxon rank sum test of Alt Vs. Ref base qualities"
##INFO=Description="Z-score From Wilcoxon rank sum test of Alt vs. Ref number of hard clipped bases"
##INFO=Description="dbSNP membership"
##INFO=Description="Combined depth across samples"
##INFO=Description="Phred-scaled p-value for exact test of excess heterozygosity"
##INFO=Description="Phred-scaled p-value using Fisher's exact test to detect strand bias"
##INFO=Description="Inbreeding coefficient as estimated from the genotype likelihoods per-sample when compared against the Hardy-Weinberg expectation"
##INFO=Description="Maximum likelihood expectation (MLE) for the allele counts, for each ALT allele, in the same order as listed"
##INFO=Description="Maximum likelihood expectation (MLE) for the allele frequency, for each ALT allele, in the same order as listed"
##INFO=Description="MQ mapping quality"
##INFO=Description="Z-score From Wilcoxon rank sum test of Alt vs. Ref read mapping qualities"
##INFO=Description="Variant Confidence/Quality by Depth"
##INFO=Description="Z-score from Wilcoxon rank sum test of Alt vs. Ref read position bias"
##INFO=Description="Symmetric Odds Ratio of 2x2 contingency table to detect strand bias"
##SentieonCommandLine.GVCFtyper,Version="sentieon-genomics-202112.07",Date="2025-01-27T20:26:17Z",CommandLine="/opt/sentieon/sentieon-genomics-202112.07/libexec/driver -r bundle/b37decoy_FASTA/human_g1k_v37_decoy.fasta -t -3 -interval bundle/g1k_v37_decoy.fasta -algo GVCFtyper -d bundle/b37decoy_VCF/db.snp.138.b37.vcf.gz -v 189DV-output.g.vcf.gz -o 189DV-output.vcf.gz"
##SentieonCommandLine.Haplotype,Version="sentieon-genomics-202112.07",Date="2025-01-27T20:23:54Z",CommandLine="/opt/sentieon/sentieon-genomics-202112.07/libexec/driver -r bundle/b37decoy_FASTA/human_g1k_v37_decoy.nodecoy.bed --algo Haplotype -d bundle/b37decoy_VCF/db.snp.138.b37.vcf.gz --emit_mode gvft --emit_conf=30 -call_conf=30 189DV-output.g.vcf.gz"
##contig=<ID=1,length=2475b621,assembly=b37>
##contig=<ID=2,length=243199373,assembly=b37>
##contig=<ID=3,length=198022439,assembly=b37>
##contig=<ID=4,length=191154276,assembly=b37>
##contig=<ID=5,length=188915268,assembly=b37>
##contig=<ID=6,length=171115067,assembly=b37>
##contig=<ID=7,length=159138663,assembly=b37>
##contig=<ID=8,length=146364022,assembly=b37>
##contig=<ID=9,length=141213431,assembly=b37>
##contig=<ID=10,length=135534747,assembly=b37>
##contig=<ID=11,length=135086516,assembly=b37>
##contig=<ID=12,length=133851895,assembly=b37>
##contig=<ID=13,length=115169278,assembly=b37>
##contig=<ID=14,length=107349540,assembly=b37>
##contig=<ID=15,length=102531392,assembly=b37>
##contig=<ID=16,length=98354753,assembly=b37>
```

metainformación

Header

Anotación y filtrado



Variant Call Format

```
#fileformat=VCFv4.2
##FORMAT=Type=Integer,Description="Allelic depths for the ref and alt alleles in the order listed"
##FORMAT=Type=Integer,Description="Read depth"
##FORMAT=Type=Integer,Description="Genotype quality"
##FORMAT=Type=String,Description="Genotype"
##FORMAT=Type=String,Description="Physical phasing haplotype information, describing how the alternate alleles are phased in relation to one another"
##FORMAT=Type=String,Description="Physical phasing ID information, where each unique ID within a given sample (but not across samples) connects records within a phasing group"
##FORMAT=Type=Integer,Description="The phred-scaled genotype likelihoods rounded to the closest integer"
##INFO=Type=Integer,Description="Allele count in genotypes, for each ALT allele, in the same order as listed"
##INFO=Type=Float,Description="Allele frequency, for each ALT allele, in the same order as listed"
##INFO=Type=Integer,Description="Total number of alleles in called genotypes"
##INFO=Type=Float,Description="Z-score from Wilcoxon rank sum test of Alt Vs. Ref base qualities"
##INFO=Type=Float,Description="Z-score From Wilcoxon rank sum test of Alt vs. Ref number of hard clipped bases"
##INFO=Type=Flag,Description="dbSNP membership"
##INFO=Type=Integer,Description="Combined depth across samples"
##INFO=Type=Float,Description="Phred-scaled p-value for exact test of excess heterozygosity"
##INFO=Type=Float,Description="Phred-scaled p-value using Fisher's exact test to detect strand bias"
##INFO=Type=Float,Description="Inbreeding coefficient as estimated from the genotype likelihoods per-sample when compared against the Hardy-Weinberg expectation"
##INFO=Type=Integer,Description="Maximum likelihood expectation (MLE) for the allele counts, for each ALT allele, in the same order as listed"
##INFO=Type=Float,Description="Maximum likelihood expectation (MLE) for the allele frequency, for each ALT allele, in the same order as listed"
##INFO=Type=Float,Description="MQ mapping quality"
##INFO=Type=Float,Description="Z-score From Wilcoxon rank sum test of Alt vs. Ref read mapping qualities"
##INFO=Type=Float,Description="Variant Confidence/Quality by Depth"
##INFO=Type=Float,Description="Z-score from Wilcoxon rank sum test of Alt vs. Ref read position bias"
##INFO=Type=Float,Description="Symmetric Odds Ratio of 2x2 contingency table to detect strand bias"
##SentieonCommandLine.GVCFtyper,Version="sentieon-genomics-202112.07",CommandLine="/opt/sentieon/sentieon-genomics-202112.07/libexec/driver -r bundle/b37decoy_FASTA/human_g1k_v37_decoy.fasta -t 36 -interval bundle/b37decoy_FASTA/human_g1k_v37_decoy.nodecoy.bed --algo GVCFtyper -bundle/b37decoy_VCF/dbnsnp_138.b37.vcf.gz -v 189DV-output.g.vcf.gz 189DV-output.vcf.gz"
##SentieonCommandLine.Haplotype,Version="sentieon-genomics-202112.07",CommandLine="/opt/sentieon/sentieon-genomics-202112.07/libexec/driver -r bundle/b37decoy_FASTA/human_g1k_v37_decoy.fasta -t 36 -i 189DV-deduced.bam -q 189DV-recal_data.table --interval bundle/b37decoy_FASTA/human_g1k_v37_decoy.nodecoy.bed --algo Haplotype -d bundle/b37decoy_VCF/dbnsnp_138.b37.vcf.gz --emit_mode gvcf --emit_conf=30 -call_conf=30 189DV-output.g.vcf.gz"
##contig<ID=1,length=249250621,assembly=b37>.....
##contig<ID=2,length=243199373,assembly=b37>
##contig<ID=3,length=198022439,assembly=b37>
##contig<ID=4,length=19154276,assembly=b37>
##contig<ID=5,length=188915260,assembly=b37>
##contig<ID=6,length=171115067,assembly=b37>
##contig<ID=7,length=159138663,assembly=b37>
##contig<ID=8,length=146364022,assembly=b37>
##contig<ID=9,length=141213431,assembly=b37>
##contig<ID=10,length=135534747,assembly=b37>
##contig<ID=11,length=135086516,assembly=b37>
##contig<ID=12,length=133851895,assembly=b37>
##contig<ID=13,length=115169278,assembly=b37>
##contig<ID=14,length=107349540,assembly=b37>
##contig<ID=15,length=102531392,assembly=b37>
##contig<ID=16,length=98354753,assembly=b37>
```

cromosomas

Header

Anotación y filtrado

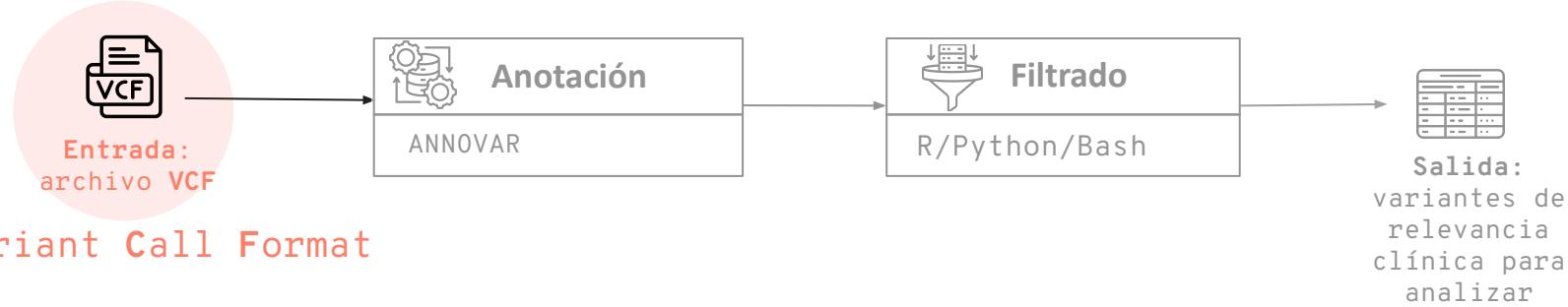


Variant Call Format

body

```
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT 189DV
1 12783 . G A 168.88 . AC=2;AF=1;AN=2;DP=7;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1;MQ=26.59;QD=24.11;SOR=4.174 GT:AD:DP:GQ:PL 1/1:0,7:7:21:197,21,0
1 13116 rs201725126 T G 875.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=-3.183;ClippingRankSum=0.000;DB;DP=41;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=30.01;MQRankSum=-5.408;QD=21.36;ReadPosRankSum=1.215;
SOR=0.837 GT:AD:DP:GQ:PGT:PID:PL 0/1:18,23:41:99:0|1:13116_T_G:984,0,992
1 13118 rs200579949 A G 875.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=-3.811;ClippingRankSum=0.000;DB;DP=40;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=29.58;MQRankSum=-5.320;QD=21.89;ReadPosRankSum=0.806;
SOR=0.931 GT:AD:DP:GQ:PGT:PID:PL 0/1:17,23:40:99:0|1:13116_T_G:984,0,992
1 13302 rs108734498 C T 188.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=5.029;ClippingRankSum=0.000;DB;DP=77;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=36.93;MQRankSum=-6.427;QD=2.45;ReadPosRankSum=-5.290;S
OR=0.662 GT:AD:DP:GQ:PL 0/1:61,16:77:99:217,0,1583
1 13621 . A G 134.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=4.197;ClippingRankSum=0.000;DP=100;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=28.21;MQRankSum=-0.513;QD=1.35;ReadPosRankSum=1.294;SOR=0.559
GT:AD:DP:GQ:PL 0/1:85,18:100:99:163,0,2226
1 13656 . CAG C 1428.73 . AC=1;AF=0.5;AN=2;BaseQRankSum=-6.852;ClippingRankSum=0.000;DP=65;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=24.12;MQRankSum=-5.028;QD=22.32;ReadPosRankSum=-1.665;SOR=1.160
GT:AD:DP:GQ:PL 0/1:26,38:64:99:1466,0,1955
1 14673 rs369473859 G C 53.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=3.034;ClippingRankSum=0.000;DB;DP=15;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.81;MQRankSum=-1.540;QD=3.58;ReadPosRankSum=0.620;S
R=1.179 GT:AD:DP:GQ:PL 0/1:11,4:15:82:82,0,335
1 14677 rs201327123 G A 34.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=2.834;ClippingRankSum=0.000;DB;DP=16;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.10;MQRankSum=-1.169;QD=2.17;ReadPosRankSum=-0.889;S
OR=0.693 GT:AD:DP:GQ:PL 0/1:13,3:16:63:63,0,485
1 14699 rs372910670 C G 44.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=2.076;ClippingRankSum=0.000;DB;DP=18;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.08;MQRankSum=-0.871;QD=2.49;ReadPosRankSum=0.127;S
R=1.329 GT:AD:DP:GQ:PL 0/1:14,4:18:73:73,0,413
1 14907 rs79585140 A G 2100.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=-0.052;ClippingRankSum=0.000;DB;DP=139;ExcessHet=3.0103;FS=0.647;MLEAC=1;MLEAF=0.5;MQ=47.89;MQRankSum=0.136;QD=15.11;ReadPosRankSum=2.890;
SOR=0.606 GT:AD:DP:GQ:PL 0/1:53,86:139:99:2129,0,1105
```

Anotación y filtrado



Variant Call Format

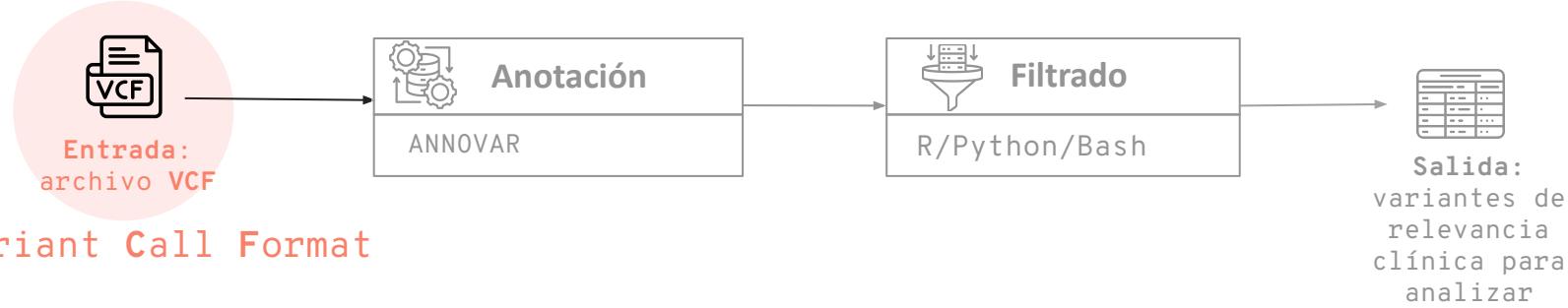
cromosoma

1

body

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	189DV
1	12783	.	G	A	168.80	.	AC=2;AF=1;AN=2;DP=7;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1;MQ=26.59;QD=24.11;SOR=4.174	GT:AD:DP:GQ:PL	1/1:0,7:7:21:197,21,0
SOR=0.837	13116	rs201725126	T	G	875.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-3.183;ClippingRankSum=0.008;DB=41;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=30.01;MQRankSum=-5.408;QD=21.36;ReadPosRankSum=1.215;		
SOR=0.931	13118	rs200579949	A	G	875.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-3.811;ClippingRankSum=0.000;DB=40;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=29.58;MQRankSum=-5.320;QD=21.89;ReadPosRankSum=0.806;		
SOR=0.662	13302	rs180734498	C	T	188.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=5.029;ClippingRankSum=0.000;DB=77;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=36.93;MQRankSum=-6.427;QD=2.45;ReadPosRankSum=-5.290;SOR=0.559	GT:AD:DP:GQ:PL	0/1:61,167:77:99:217,0,1583
GT:AD:DP:GQ:PL	13621	.	A	G	134.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=4.197;ClippingRankSum=0.000;DP=100;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=28.21;MQRankSum=-0.513;QD=1.35;ReadPosRankSum=1.294;SOR=0.559		
GT:AD:DP:GQ:PL	13656	.	CAG	C	1428.73	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-6.852;ClippingRankSum=0.000;DP=65;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=24.12;MQRankSum=-5.028;QD=22.32;ReadPosRankSum=-1.665;SOR=1.160		
R=1.179	14673	rs369473859	G	C	53.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=3.034;ClippingRankSum=0.000;DB=15;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.81;MQRankSum=-1.540;QD=3.58;ReadPosRankSum=0.620;SOR=0.653		
GT:AD:DP:GQ:PL	14677	.	G	A	0/1:11,45:18:82:0,335	.	AC=1;AF=0.5;AN=2;BaseQRankSum=2.834;ClippingRankSum=0.000;DB=16;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.10;MQRankSum=-1.169;QD=2.17;ReadPosRankSum=-0.889;SOR=0.653		
R=1.329	14699	rs372918670	C	G	44.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=2.076;ClippingRankSum=0.000;DB=18;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.08;MQRankSum=-0.871;QD=2.49;ReadPosRankSum=0.127;SOR=0.06		
GT:AD:DP:GQ:PL	14907	rs79585140	A	G	2100.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-0.052;ClippingRankSum=0.000;DB=139;ExcessHet=3.0103;FS=0.647;MLEAC=1;MLEAF=0.5;MQ=47.89;MQRankSum=0.136;QD=15.11;ReadPosRankSum=2.890;SOR=0.06	GT:AD:DP:GQ:PL	0/1:53,86:139:99:2129,0,1105

Anotación y filtrado



Variant Call Format

posición

10

body

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	189DV
1	12783	.	G	A	168.80	.	AC=2;AF=1;AN=2;DP=7;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1;MQ=26.59;QD=24.11;SOR=4.174	GT:AD:DP:GQ:PL	1/1:0,7:7:21:197,21,0
SOR=0.837	13116	rs201725126	T	G	875.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-3.183;ClippingRankSum=-0.006;DB=41;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=30.01;MQRankSum=-5.408;QD=21.36;ReadPosRankSum=1.215;	GT:AD:DP:GQ:PL	1/1:0,7:7:21:197,21,0
SOR=0.931	13118	rs200579949	A	G	875.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-3.811;ClippingRankSum=0.006;DB=40;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=29.58;MQRankSum=-5.320;QD=21.89;ReadPosRankSum=0.806;	GT:AD:DP:GQ:PL:ID:PL	0/1:17,23:40:99:08:1:13116_T_G:984,0,992
OR=0.662	13382	rs180734948	C	T	188.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=5.029;ClippingRankSum=0.000;DB=77;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=36.93;MQRankSum=-6.427;QD=2.45;ReadPosRankSum=-5.290;SOR=0.559	GT:AD:DP:GQ:PL:ID:PL	0/1:61,16:77:99:21,0,1583
GT:AD:DP:GQ:PL	13621	.	A	G	134.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=4.197;ClippingRankSum=0.000;DP=100;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=28.21;MQRankSum=-0.513;QD=1.35;ReadPosRankSum=1.294;SOR=0.559	GT:AD:DP:GQ:PL	0/1:85,15:100:99:163,0,2226
GT:AD:DP:GQ:PL	13656	.	CAG	C	1428.73	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-6.852;ClippingRankSum=0.000;DP=65;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=24.12;MQRankSum=-5.028;QD=22.32;ReadPosRankSum=-1.665;SOR=1.160	GT:AD:DP:GQ:PL	0/1:26,38:64:99:1466,0,1055
R=1.179	14673	rs369473859	G	C	53.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=3.034;ClippingRankSum=0.000;DB=DP=15;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.81;MQRankSum=-1.540;QD=3.58;ReadPosRankSum=0.620;SOR=0.663	GT:AD:DP:GQ:PL	0/1:11,45:15:82:82,0,335
OR=0.693	14677	rs201327213	G	A	34.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=2.834;ClippingRankSum=0.000;DB=DP=16;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.10;MQRankSum=-1.169;QD=2.17;ReadPosRankSum=-0.889;SOR=0.663	GT:AD:DP:GQ:PL	0/1:13,31:63:63,0,405
R=1.329	14699	rs372919670	C	G	44.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=2.076;ClippingRankSum=0.000;DB=DP=18;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.08;MQRankSum=-0.871;QD=2.49;ReadPosRankSum=0.127;SOR=0.663	GT:AD:DP:GQ:PL	0/1:14,4:18:73:73,0,413
GT:AD:DP:GQ:PL	14987	rs79585140	A	G	2100.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-0.052;ClippingRankSum=0.000;DB=DP=139;ExcessHet=3.0103;FS=0.647;MLEAC=1;MLEAF=0.5;MQ=47.89;MQRankSum=0.136;QD=15.11;ReadPosRankSum=2.890;SOR=0.466	GT:AD:DP:GQ:PL	0/1:53,86:139:99:2129,0,1105

Anotación y filtrado



Variant Call Format

ref

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	189DV	body
1	12783	.	G	A	168.88	.	AC=2;AF=1;AN=2;DP=7;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1;MQ=26.59;QD=24.11;SOR=4.174	GT:AD:DP:GQ:PL	1/1:0,7:7:21:197,21,0	
1	13116	rs201725126	T	G	875.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-3.183;ClippingRankSum=0.000;DB:DP=41;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=30.01;MQRankSum=-5.408;QD=21.36;ReadPosRankSum=1.215;SOR=0.837	GT:AD:DP:GQ:PL	0/1:18,23:41:99:0 1:13116_T_G:984,0,992	
1	13118	rs200579949	A	G	875.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-3.811;ClippingRankSum=0.000;DB:DP=40;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=29.58;MQRankSum=-5.320;QD=21.89;ReadPosRankSum=0.806;SOR=0.931	GT:AD:DP:GQ:PL	0/1:17,23:40:99:0 1:13118_T_G:984,0,992	
1	13302	rs100734498	C	T	188.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=5.029;ClippingRankSum=0.000;DB:DP=77;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=36.93;MQRankSum=-6.427;QD=2.45;ReadPosRankSum=-5.290;SOR=0.662	GT:AD:DP:GQ:PL	0/1:61,16:77:99:217,0,1583	
1	13621	.	A	G	134.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=4.197;ClippingRankSum=0.000;DP=100;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=28.21;MQRankSum=-0.513;QD=1.35;ReadPosRankSum=1.294;SOR=0.559	GT:AD:DP:GQ:PL	0/1:85,15:100:99:163,0,2226	
1	13656	.	CAG	C	1428.73	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-6.852;ClippingRankSum=0.000;DP=65;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=24.12;MQRankSum=-5.028;QD=22.32;ReadPosRankSum=-1.665;SOR=1.160	GT:AD:DP:GQ:PL	0/1:26,38:64:99:1466,0,1955	
1	14673	rs369473859	G	C	53.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=3.034;ClippingRankSum=0.000;DB:DP=15;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.81;MQRankSum=-1.540;QD=3.58;ReadPosRankSum=0.620;SOR=1.179	GT:AD:DP:GQ:PL	0/1:11,4:15:82:82,0,335	
1	14677	rs201327123	G	A	34.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=2.834;ClippingRankSum=0.000;DB:DP=16;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.10;MQRankSum=-1.169;QD=2.17;ReadPosRankSum=-0.889;SOR=0.693	GT:AD:DP:GQ:PL	0/1:13,3:16:63,0,485	
1	14699	rs372910670	C	G	44.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=2.076;ClippingRankSum=0.000;DB:DP=18;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.08;MQRankSum=-0.871;QD=2.49;ReadPosRankSum=0.127;SOR=1.329	GT:AD:DP:GQ:PL	0/1:14,4:18:73:73,0,413	
1	14907	rs79585140	A	G	2100.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-0.052;ClippingRankSum=0.000;DB:DP=139;ExcessHet=3.0103;FS=0.647;MLEAC=1;MLEAF=0.5;MQ=47.89;MQRankSum=0.136;QD=15.11;ReadPosRankSum=2.890;SOR=0.606	GT:AD:DP:GQ:PL	0/1:53,86:139:99:2129,0,1105	

Anotación y filtrado



Variant Call Format

alt

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	189DV	body
1	12783	.	G	A	168.88	.	AC=2;AF=1;AN=2;DP=7;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1;MQ=26.59;QD=24.11;SOR=4.174	GT:AD:DP:GQ:PL	1/1:0,7:7:21:197,21,0	
1	13116	rs201725126	T	G	875.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-3.183;ClippingRankSum=0.000;DB:DP=41;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=30.01;MQRankSum=-5.408;QD=21.36;ReadPosRankSum=1.215;SOR=0.837	GT:AD:DP:GQ:PL	0/1:18,23:41:99:0 1:13116_T_G:984,0,992	
1	13118	rs200579949	A	G	875.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-3.811;ClippingRankSum=0.000;DB:DP=40;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=29.58;MQRankSum=-5.320;QD=21.89;ReadPosRankSum=0.806;SOR=0.931	GT:AD:DP:GQ:PL:PID:PL	0/1:17,23:40:99:0 1:13118_T_G:984,0,992	
1	13302	rs108734498	C	T	188.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=5.029;ClippingRankSum=0.000;DB:DP=77;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=36.93;MQRankSum=-6.427;QD=2.45;ReadPosRankSum=-5.290;SOR=0.662	GT:AD:DP:GQ:PL	0/1:61,16:77:99:217,0,1583	
1	13621	.	A	G	134.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=4.197;ClippingRankSum=0.000;DP=100;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=28.21;MQRankSum=-0.513;QD=1.35;ReadPosRankSum=1.294;SOR=0.559	GT:AD:DP:GQ:PL	0/1:85,18:100:99:163,0,2226	
1	13656	.	CAG	C	1428.73	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-6.852;ClippingRankSum=0.000;DP=65;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=24.12;MQRankSum=-5.028;QD=22.32;ReadPosRankSum=-1.665;SOR=1.160	GT:AD:DP:GQ:PL	0/1:26,38:64:9:1466,0,1955	
1	14673	rs369473859	G	C	53.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=3.034;ClippingRankSum=0.000;DB:DP=15;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.81;MQRankSum=-1.540;QD=3.58;ReadPosRankSum=0.620;SOR=1.179	GT:AD:DP:GQ:PL	0/1:11,4:15:82:82,0,335	
1	14677	rs201327123	G	A	34.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=2.834;ClippingRankSum=0.000;DB:DP=16;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.10;MQRankSum=-1.169;QD=2.17;ReadPosRankSum=-0.889;SOR=0.693	GT:AD:DP:GQ:PL	0/1:13,3:16:63,0,485	
1	14699	rs372910670	C	G	44.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=2.076;ClippingRankSum=0.000;DB:DP=18;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.08;MQRankSum=-0.871;QD=2.49;ReadPosRankSum=0.127;SOR=1.329	GT:AD:DP:GQ:PL	0/1:14,4:18:73:73,0,413	
1	14907	rs79585140	A	G	2100.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-0.052;ClippingRankSum=0.000;DB:DP=139;ExcessHet=3.0103;FS=0.647;MLEAC=1;MLEAF=0.5;MQ=47.89;MQRankSum=0.136;QD=15.11;ReadPosRankSum=2.890;SOR=0.606	GT:AD:DP:GQ:PL	0/1:53,86:139:99:2129,0,1105	

Anotación y filtrado



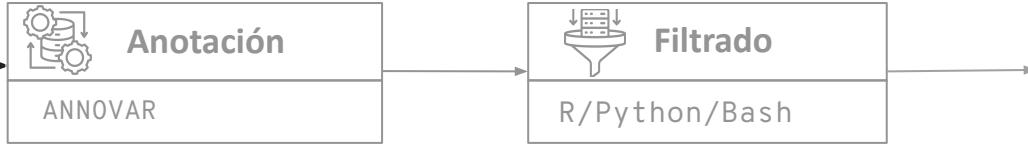
Variant Call Format

QUAL del GATK

body

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	189DV	GT:AD:DP:GQ:PL	1/1:0,7:7:21:197,21,0
1	12783	.	G	A	168.88	.	AC=2;AF=0.5;AN=2;BaseQRankSum=-3.183;ClippingRankSum=0.000;DB=41;ExcessHet=3.0103;FS=0.000;MLEAC=1;MQ=26.59;QD=24.11;SOR=4.174	GT:AD:DP:GQ:PL	1/1:0,7:7:21:197,21,0		
1	13116	rs201725126	T	G	875.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-3.183;ClippingRankSum=0.000;DB=41;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=30.01;MQRankSum=-5.408;QD=21.36;ReadPosRankSum=1.215;SOR=0.837	GT:AD:DP:GQ:PL	1/1:0,7:7:21:197,21,0		
1	13118	rs200579949	A	G	875.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-3.811;ClippingRankSum=0.000;DB=40;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=29.58;MQRankSum=-5.320;QD=21.89;ReadPosRankSum=0.806;SOR=0.931	GT:AD:DP:GQ:PL	1/1:0,7:7:21:197,21,0		
1	13302	rs108734498	C	T	188.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=5.029;ClippingRankSum=0.000;DB=77;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=36.93;MQRankSum=-6.427;QD=2.45;ReadPosRankSum=-5.290;SOR=0.662	GT:AD:DP:GQ:PL	0/1:61,16:77:99:217,0,1583		
1	13621	.	A	G	134.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=4.197;ClippingRankSum=0.000;DP=100;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=28.21;MQRankSum=-0.513;QD=1.35;ReadPosRankSum=1.294;SOR=0.559	GT:AD:DP:GQ:PL	0/1:85,15:100:99:163,0,2226		
1	13656	.	CAG	C	1428.73	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-6.852;ClippingRankSum=0.000;DP=65;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=24.12;MQRankSum=-5.028;QD=22.32;ReadPosRankSum=-1.665;SOR=1.160	GT:AD:DP:GQ:PL	0/1:26,38:64:99:1466,0,1955		
1	14673	rs369473859	G	C	53.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=3.034;ClippingRankSum=0.000;DB=15;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.81;MQRankSum=-1.540;QD=3.58;ReadPosRankSum=0.620;SOR=1.179	GT:AD:DP:GQ:PL	0/1:11,4:15:82:82,0,335		
1	14677	rs201327123	G	A	34.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=2.834;ClippingRankSum=0.000;DB=16;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.10;MQRankSum=-1.169;QD=2.17;ReadPosRankSum=-0.889;SOR=0.693	GT:AD:DP:GQ:PL	0/1:13,3:16:63:53,0,485		
1	14699	rs372910670	C	G	44.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=2.076;ClippingRankSum=0.000;DB=18;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.08;MQRankSum=-0.871;QD=2.49;ReadPosRankSum=0.127;SOR=1.329	GT:AD:DP:GQ:PL	0/1:14,4:18:73:73,0,413		
1	14907	rs79585140	A	G	2100.77	.	AC=1;AF=0.5;AN=2;BaseQRankSum=-0.052;ClippingRankSum=0.000;DB=139;ExcessHet=3.0103;FS=0.647;MLEAC=1;MLEAF=0.5;MQ=47.89;MQRankSum=0.136;QD=15.11;ReadPosRankSum=2.890;SOR=0.606	GT:AD:DP:GQ:PL	0/1:53,86:139:99:2129,0,1105		

Anotación y filtrado



Variant Call Format

fromato

body

```
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT 189DV
1 12783 . G A 168.88 . AC=2;AF=1;AN=2;DP=7;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1;MQ=26.59;QD=24.11;SOR=4.174
1 13116 rs201725126 T G 875.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=-3.183;ClippingRankSum=0.000;DB;DP=41;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=30.01;QD=21.36;ReadPosRankSum=1.215;
SOR=0.837 GT:AD:DP:GQ:PGT:PID:PL 0/1:18,23:41:99:0|1:13116_T_G:904,0,992
1 13118 rs200579949 A G 875.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=-3.811;ClippingRankSum=0.000;DB;DP=40;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=29.58;QRankSum=-5.320;QD=21.89;ReadPosRankSum=0.806;
SOR=0.931 GT:AD:DP:GQ:PGT:PID:PL 0/1:17,23:40:99:0|1:13116_T_G:904,0,992
1 13302 rs108734498 C T 188.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=5.029;ClippingRankSum=0.000;DB;DP=77;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=36.93;QRankSum=-6.427;QD=2.45;ReadPosRankSum=-5.290;S
OR=0.662 GT:AD:DP:GQ:PL 0/1:61,16:77:99:217,0,1583
1 13621 . A G 134.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=4.197;ClippingRankSum=0.000;DP=100;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=28.21;QRankSum=-0.513;QD=1.35;ReadPosRankSum=1.294;SOR=0.559
GT:AD:DP:GQ:PL 0/1:85,18:100:99:163,0,2226
1 13656 . CAG C 1428.73 . AC=1;AF=0.5;AN=2;BaseQRankSum=-6.852;ClippingRankSum=0.000;DP=65;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=24.12;QRankSum=-5.028;QD=22.32;ReadPosRankSum=-1.665;SOR=1.160
GT:AD:DP:GQ:PL 0/1:26,38:64:99:1466,0,1955
1 14673 rs369473859 G C 53.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=3.034;ClippingRankSum=0.000;DB;DP=15;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.81;QRankSum=-1.540;QD=3.58;ReadPosRankSum=0.620;S
R=1.179 GT:AD:DP:GQ:PL 0/1:11,4:15:82:82,0,335
1 14677 rs201327123 G A 34.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=2.834;ClippingRankSum=0.000;DB;DP=16;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.10;QRankSum=-1.169;QD=2.17;ReadPosRankSum=-0.889;S
OR=0.693 GT:AD:DP:GQ:PL 0/1:13,3:16:63:63,0,485
1 14699 rs372910670 C G 44.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=2.076;ClippingRankSum=0.000;DB;DP=18;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.08;QRankSum=-0.871;QD=2.49;ReadPosRankSum=0.127;S
R=1.329 GT:AD:DP:GQ:PL 0/1:14,4:18:73:73,0,413
1 14907 rs79585140 A G 2100.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=-0.052;ClippingRankSum=0.000;DB;DP=139;ExcessHet=3.0103;FS=0.647;MLEAC=1;MLEAF=0.5;MQ=47.89;QRankSum=0.136;QD=15.11;ReadPosRankSum=2.890;S
OR=0.606 GT:AD:DP:GQ:PL 0/1:53,86:139:99:2129,0,1105
```

Anotación y filtrado



Variant Call Format

Entrada:
archivo VCF

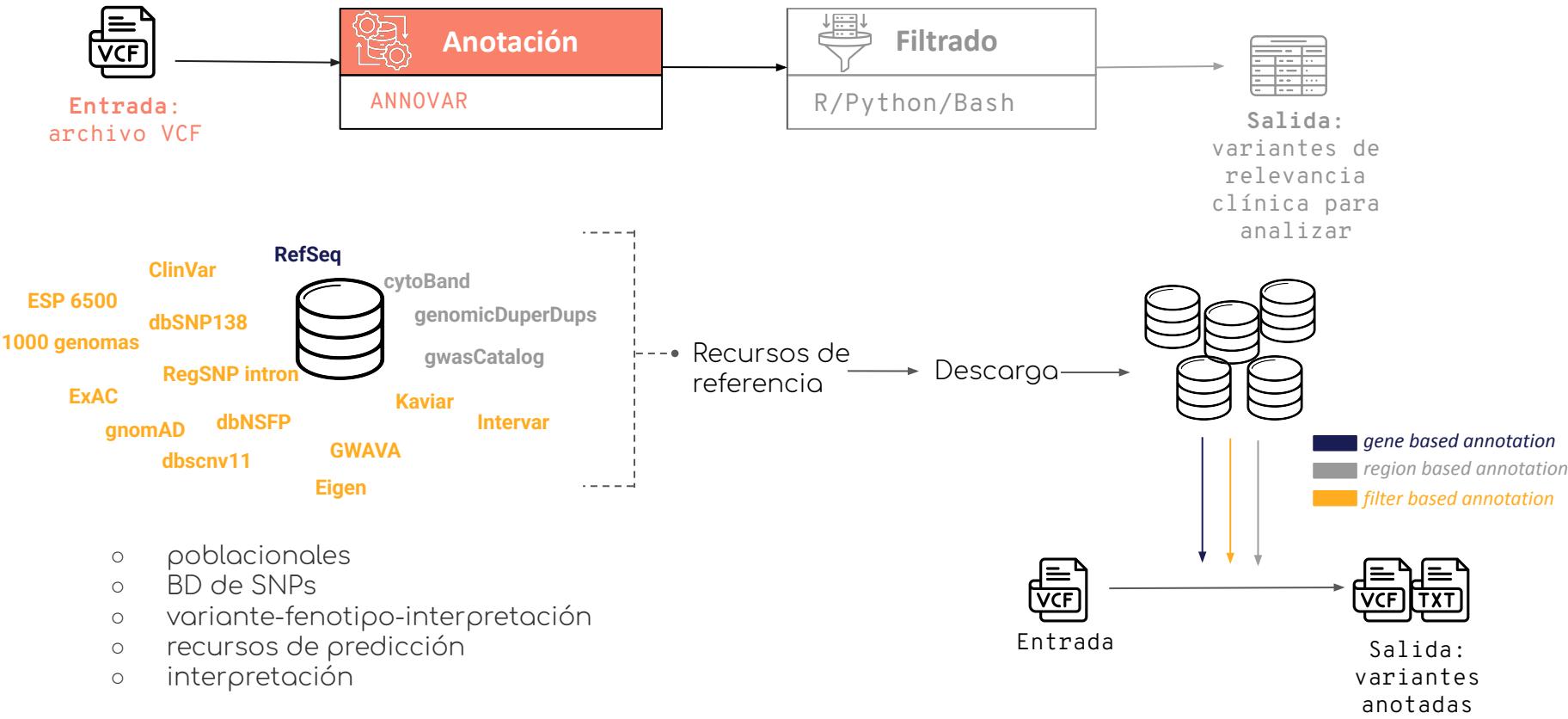
Salida:
variantes de
relevancia
clínica para
analizar

los valores

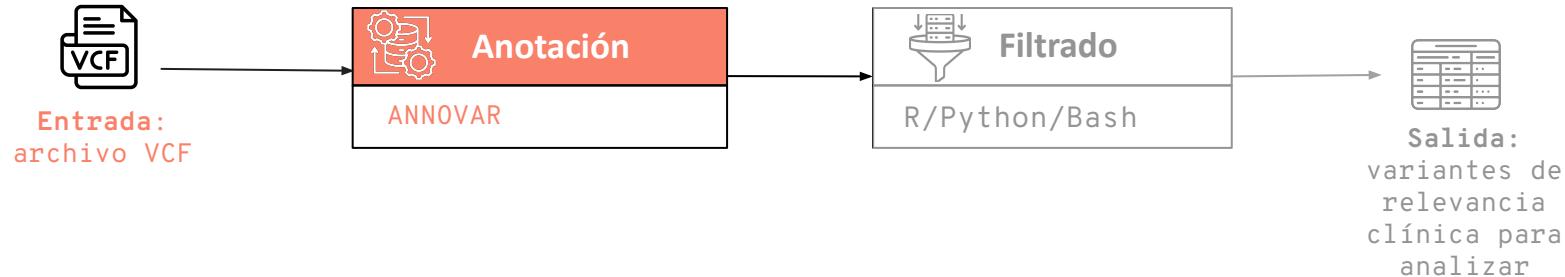
body

```
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT 189DV
1 12783 . G A 168.88 . AC=2;AF=1;AN=2;DP=7;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1;MQ=26.59;QD=24.11;SOR=4.174
1 13116 rs201725126 T G 875.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=-3.183;ClippingRankSum=0.000;DB;DP=41;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=30.01;MQRankSum=-5.408;QD=21.36;ReadPosRankSum=1.215;
SOR=0.837 GT:AD:DP:GQ:PGT:PID:PL 0/1:18,23:41:99:0|1:13116_T_G:984,0,992
1 13118 rs200579949 A G 875.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=-3.811;ClippingRankSum=0.000;DB;DP=40;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=29.58;MQRankSum=-5.320;QD=21.89;ReadPosRankSum=0.806;
SOR=0.931 GT:AD:DP:GQ:PGT:PID:PL 0/1:17,23:40:99:0|1:13116_T_G:984,0,992
1 13302 rs108734498 C T 188.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=5.029;ClippingRankSum=0.000;DB;DP=77;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=36.93;MQRankSum=-6.427;QD=2.45;ReadPosRankSum=-5.290;S
OR=0.662 GT:AD:DP:GQ:PL 0/1:61,16:77:99:217,0,1583
1 13621 . A G 134.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=4.197;ClippingRankSum=0.000;DP=100;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=28.21;MQRankSum=-0.513;QD=1.35;ReadPosRankSum=1.294;SOR=0.559
GT:AD:DP:GQ:PL 0/1:85,18:100:99:163,0,2226
1 13656 . CAG C 1428.73 . AC=1;AF=0.5;AN=2;BaseQRankSum=-6.852;ClippingRankSum=0.000;DP=65;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=24.12;MQRankSum=-5.028;QD=22.32;ReadPosRankSum=-1.665;SOR=1.160
GT:AD:DP:GQ:PL 0/1:26,38:64:99:1466,0,1955
1 14673 rs369473859 G C 53.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=3.034;ClippingRankSum=0.000;DB;DP=15;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.81;MQRankSum=-1.540;QD=3.58;ReadPosRankSum=0.620;S
R=1.179 GT:AD:DP:GQ:PL 0/1:11,4:15:82:82,0,335
1 14677 rs201327123 G A 34.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=2.834;ClippingRankSum=0.000;DB;DP=16;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.10;MQRankSum=-1.169;QD=2.17;ReadPosRankSum=-0.889;S
OR=0.693 GT:AD:DP:GQ:PL 0/1:13,3:16:63:63,0,485
1 14699 rs372910670 C G 44.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=2.076;ClippingRankSum=0.000;DB;DP=18;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.5;MQ=37.08;MQRankSum=-0.871;QD=2.49;ReadPosRankSum=0.127;S
R=1.329 GT:AD:DP:GQ:PL 0/1:14,4:18:73:73,0,413
1 14907 rs79585140 A G 2100.77 . AC=1;AF=0.5;AN=2;BaseQRankSum=-0.052;ClippingRankSum=0.000;DB;DP=139;ExcessHet=3.0103;FS=0.647;MLEAC=1;MLEAF=0.5;MQ=47.89;MQRankSum=0.136;QD=15.11;ReadPosRankSum=2.890;
SOR=0.606 GT:AD:DP:GQ:PL 0/1:53,86:139:99:2129,0,1105
```

Anotación y filtrado

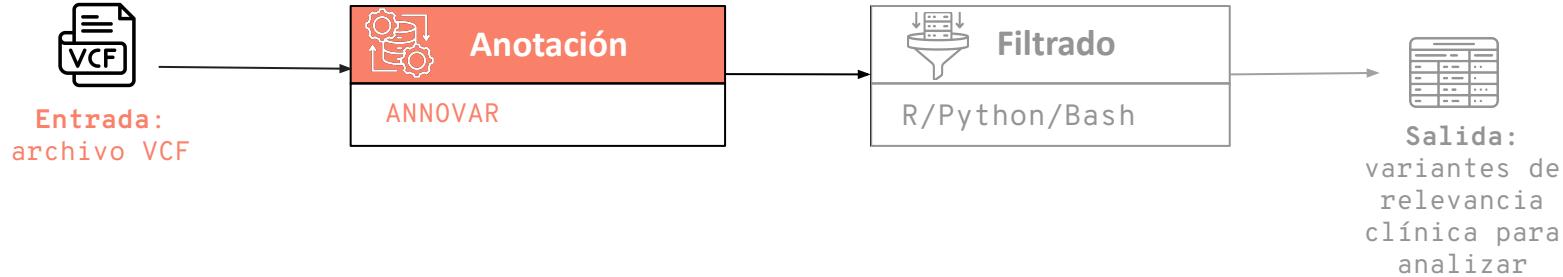


Anotación y filtrado



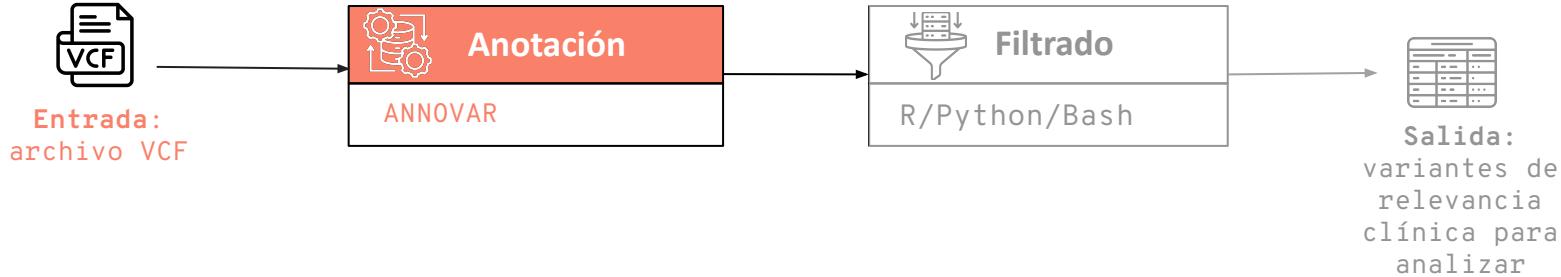
```
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT Sample
chr11 69462910 . G A 103,14 DPFilter;MQFilter
AC=2;AF=1,00000e+00;AN=2;DP=5;ExcessHet=3,01030e+00;FS=0,00000e+00;MLEAC=2;MLEAF=1,00000e+00;MQ=2,90000e+01;QD=
2,57900e+01;SOR=6,93000e-01 GT:AD:DP:GQ:PL 1/1:0,4:4:12:117,12,0
```

Anotación y filtrado



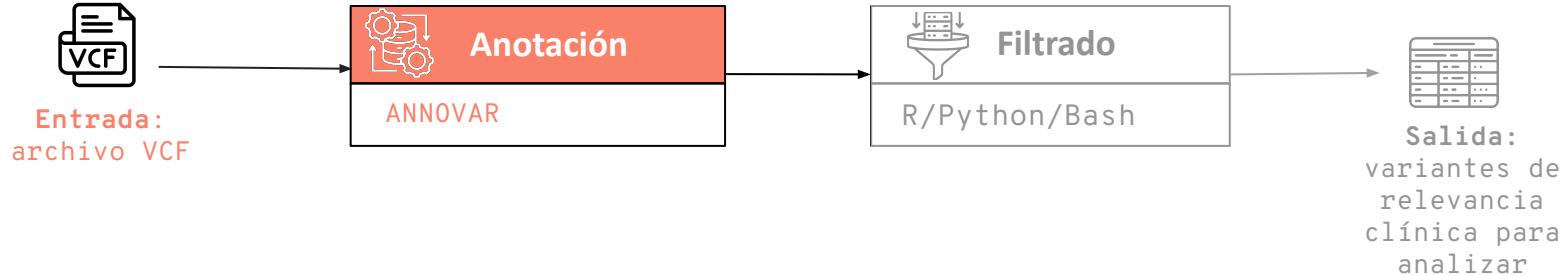
#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	Sample
chr11	69462910	.	G	A	103,14	DPFilter;MQFilter			
AC=2;AF=1,00000e+00;AN=2;DP=5;ExcessHet=3,01030e+00;FS=0,00000e+00;MLEAC=2;MLEAF=1,00000e+00;MQ=2,90000e+01;QD=2,57900e+01;SOR=6,93000e-01;ANNOVAR_DATE=2018-04-16;Func.refGene=exonic;Gene.refGene=CCND1;GeneDetail.refGene=.;ExonicFunc.refGene=synonymous_SNV;AAChange.refGene=CCND1:NM_053056:exon4:c.G723A:p.P241P;cytoBand=11q13.3;genomicSuperDups=.;gwasCatalog=Name\x3dImmunoglobulin_light_chain_(AL)_amyloidosis,Immunoglobulin_light_chain_(AL)_amyloidosis_(serum_Ig_light_chain_only_lambda/Kappa_based_profile),Immunoglobulin_light_chain_(AL)_amyloidosis_(heart_involvement),Immunoglobulin_light_chain_(AL)_amyloidosis_(kidney_involvement),Immunoglobulin_light_chain_(AL)_amyloidosis_(serum_Ig_lambda_profile),Immunoglobulin_light_chain_(AL)_amyloidosis_(serum_Ig_light_chain_only_lambda_based_profile);CLNALLELEID=28794;CLNDN=VON_HIPPEL-LINDAU_SYNDROME,_MODIFIER_OF Colorectal_cancer,_susceptibility_to Multiple_myeloma,_translocation_11,14_type;CLNDISDB=. MedGen:C1858438 MedGen:CN186214;CLNREVSTAT=no_assertion_criteria_provided;CLNSIG=risk_factor;esp6500siv2_all=0.3867;1000g2015aug_a11=0.413538;1000g2015aug_afr=0.1876;1000g2015aug_eas=0.5714;1000g2015aug_eur=0.497;snp138=rs9344;ExAC_ALL=0.4551;ExAC_AFR=0.2187;ExAC_AMR=0.3772;ExAC_EAS=0.5695;ExAC_FIN=0.4830;ExAC_NFE=0.4702;ExAC_OTH=0.4889;ExAC_SAS=0.5245;avsnp150=rs9344;SIFT_score=.;SIFT_converted_rankscore=.;SIFT_pred=.;Polyphen2_HDIV_score=.;Polyphen2_HDIV_rankscore=.;Polyphen2_HVAR_pred=.;Polyphen2_HVAR_rankscore=.;Polyphen2_HVAR_pred=.;LRT_score=.;LRT_converted_rankscore=.;LRT_pred=.;MutationTaster_score=.;MutationTaster_converted_rankscore=.;MutationAssessor_score=.;MutationAssessor_score_rankscore=.;MutationAssessor_pred=.;FATHMM_score=.;FATHMM_converted_rankscore=.;FATHMM_pred=.;PROVEAN_score=.;PROVEAN_converted_rankscore=.;PROVEAN_pred=.;VEST3_score=.;VEST3_rankscore=.;MetaSVM_score=.;MetaSVM_rankscore=.;MetaSVM_pred=.;MetaLR_score=.;MetaLR_rankscore=.;MetaLR_pred=.;M-CAP_score=.;M-CAP_rankscore=.;M-CAP_pred=.;REVEL_score=.;REVEL_rankscore=.;MutPred_score=.;MutPred_rankscore=.;CADD_raw=.;CADD_raw_rankscore=.;DANN_score=.;DANN_rankscore=.;fathmm-MKL_coding_score=.;fathmm-MKL_coding_rankscore=.;Eigen_coding_or_noncoding=.;Eigen_raw=.;Eigen-PC=.;GenoCanyon_score=.;GenoCanyon_score_rankscore=.;integrated_fitCons_score=.;integrated_fitCons_score_rankscore=.;integrated_confidence_value=.;GERP++_RS=.;GERP++_RS_rankscore=.;phyloP100way_vertebrate=.;phyloP100way_vertebrate_rankscore=.;phyloP20way_mammalian=.;phyloP20way_mammalian_rankscore=.;phastCons100way_vertebrate=.;phastCons100way_vertebrate_rankscore=.;phastCons20way_mammalian=.;phastCons20way_mammalian_rankscore=.;SiPhy_29way_logOdds=.;SiPhy_29way_logOdds_rankscore=.;Interpro_domain=.;GTEX_V6p_gene=.;GTEX_V6p_tissue=.;regsnp_fpr=.;regsnp_disease=.;regsnp_splicing_site=.;GWAVA_region_score=0.39;GWAVA_tss_score=0.45;GWAVA_unmatched_score=182;AF=0.3969;AF_popmax=0.5450;AF_male=0.3968;AF_female=0.3971;AF_raw=0.3971;AF_afr=0.2138;AF_sas=.;AF_amr=0.4100;AF_eas=0.5450;AF_nfe=0.4584;AF_fin=0.4856;AF_asj=0.4759;AF_oth=0.4658;non_topmed_AF_popmax=0.5449;non_neuro_AF_popmax=0.5450;non_cancer_AF_popmax=.;controls_AF_popmax=0.5449;InterVar_automated=Benign;PVS1=0;PS1=0;PS2=0;PS3=0;PS4=0;PM1=0;PM2=0;PM3=0;PM4=0;PM5=0;PM6=0;PP1=0;PP2=0;PP3=0;PP4=0;PP5=0;BA1=1;BS1=1;BS2=1;BS3=0;BS4=0;BP1=0;BP2=0;BP3=0;BP4=1;BP5=0;BP6=0;BP7=0;Kaviar_AF=0.442051;Kaviar_AC=68342;Kaviar_AN=154602;Eigen=1.5733;ALLELE_END	GT:AD:DP:GQ:PL	1/1:0,4:4:12:117,12,0							

Anotación y filtrado



```
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT Sample
chr11 69462910 . G A 103,14 DPFilter;MQFilter
AC=2;AF=1,00000e+00;AN=2;DP=5;ExcessHet=3,01030e+00;FS=0,00000e+00;MLEAC=2;MLEAF=1,00000e+00;MQ=2,90000e+01;QD=2,57900e+01;SOR=6,93000e-01;ANNOVAR_DATE=201
8-04-16;Func.refGene=exonic;Gene.refGene=CCND1;GeneDetail.refGene=.;ExonicFunc.refGene=synonymous_SNV;AAChange.refGene=CCND1:NM_053056:exon4:c.G723A:p.P241
P;cytoBand=11q13.3;genomicSuperDups=.;gwasCatalog=Name\x3dImmunoglobulin_light_chain_(AL)_amyloidosis,Immunoglobulin_light_chain_(AL)_amyloidosis_(serum_Ig
_light_chain_only_lambda/kappa_based_profile),Immunoglobulin_light_chain_(AL)_amyloidosis_(heart_involvement),Immunoglobulin_light_chain_(AL)_amyloidosis_(k
idney_involvement),Immunoglobulin_light_chain_(AL)_amyloidosis_(serum_Ig_lambda_profile),Immunoglobulin_light_chain_(AL)_amyloidosis_(serum_Ig_light_chain_o
nly_lambda_based_profile);CLNALLELEID=28794;CLNDN=VON_HIPPEL-LINDAU_SYNDROME,_MODIFIER_OF[Colorectal_cancer,_susceptibility_to|Multiple_myeloma,_translocati
on_11_14_type;CLNDISDB=.|MedGen:C1858438|MedGen:CN186214;CLNREVSTAT=no_assertion_criteria_provided;CLNSIG=risk_factor;esp6500siv2_all=0.3867;1000g2015aug_a
11=0.413538;1000g2015aug_afr=0.1876;1000g2015aug_eas=0.5714;1000g2015aug_eur=0.497;snp138=rs9344;ExAC_ALL=0.4551;ExAC_AFR=0.2187;ExAC_AMR=0.3772;ExAC_EAS=0
.5695;ExAC_FIN=0.4830;ExAC_NFE=0.4702;ExAC_OTH=0.4889;ExAC_SAS=0.5245;avsnp150=rs9344;SIFT_score=.;SIFT_converted_rankscore=.;SIFT_pred=.;Polyphen2_HDIV_sc
ore=.;Polyphen2_HDIV_rankscore=.;Polyphen2_HVAR_pred=.;Polyphen2_HVAR_score=.;Polyphen2_HVAR_rankscore=.;LRT_score=.;LRT_converted_ra
nkscore=.;LRT_pred=.;MutationTaster_score=.;MutationTaster_converted_rankscore=.;MutationAssessor_score=.;MutationAssessor_score_rank
score=.;MutationAssessor_pred=.;FATHMM_score=.;FATHMM_converted_rankscore=.;FATHMM_pred=.;PROVEAN_score=.;PROVEAN_converted_rankscore=.;PROVEAN_pred=.;VEST
3_score=.;VEST3_rankscore=.;MetaSVM_score=.;MetaSVM_rankscore=.;MetaSVM_pred=.;MetaLR_score=.;MetaLR_rankscores=.;MetaLR_pred=.;M-CAP_score=.;M-CAP_ranksco
re=.;M-CAP_pred=.;REVEL_score=.;REVEL_rankscore=.;MutPred_score=.;MutPred_rankscore=.;CADD_raw=.;CADD_raw_rankscore=.;DANN_score=.;DANN_ranksco
re=.;fathmm-MKL_coding_score=.;fathmm-MKL_coding_rankscore=.;Eigen_coding_or_noncoding=.;Eigen_raw=.;Eigen-PC-raw=.;GenoCanyon_sco
re=.;GenoCanyon_score_rankscore=.;integrated_fitCons_score=.;integrated_fitCons_score_rankscore=.;integrated_confidence_value=.;GERP++_RS=.;GERP++_RS_ranksco
re=.;phyloP100way_vertebrate=.;phyloP100way_vertebrate_rankscore=.;phyloP20way_mammalian=.;phyloP20way_mammalian_rankscore=.;phastCons100way_vertebrate=.;ph
astCons100way_vertebrate_rankscore=.;phastCons20way_mammalian=.;phastCons20way_mammalian_rankscore=.;SiPhy_29way_logOdds=.;SiPhy_29way_logOdds_rankscore=.;I
nterpro_domain=.;GTEX_V6p_gene=.;GTEX_V6p_tissue=.;regsnp_fpr=.;regsnp_disease=.;regsnp_splicing_site=.;GWAVA_region_score=0.39;GWAVA_tss_score=0.45;GWAVA
_unmatched_score=182;AF=0.3969;AF_popmax=0.5450;AF_male=0.3968;AF_female=0.3971;AF_raw=0.3971;AF_afr=0.2138;AF_sas=.;AF_amr=0.4100;AF_eas=0.5450;AF_nfe=0.4
584;AF_fin=0.4856;AF_asj=0.4759;AF_oth=0.4658;non_topmed_AF_popmax=0.5449;non_neuro_AF_popmax=0.5450;non_cancer_AF_popmax=.;controls_AF_popmax=0.5449;InterV
ar_automated=Benign;PVS1=0;PS1=0;PS2=0;PS3=0;PS4=0;PM1=0;PM2=0;PM3=0;PM4=0;PM5=0;PM6=0;PP1=0;PP2=0;PP3=0;PP4=0;PP5=0;BA1=1;BS1=1;BS2=1;BS3=0;BS4=0;BP1=0;BP
2=0;BP3=0;BP4=1;BP5=0;BP6=0;BP7=0;Kaviar_AF=0.442051;Kaviar_AC=68342;Kaviar_AN=154602;Eigen=1.5733;ALLELE_END GT:AD:DP:GQ:PL 1/1:0,4:4:12:117,12,0
```

Anotación y filtrado

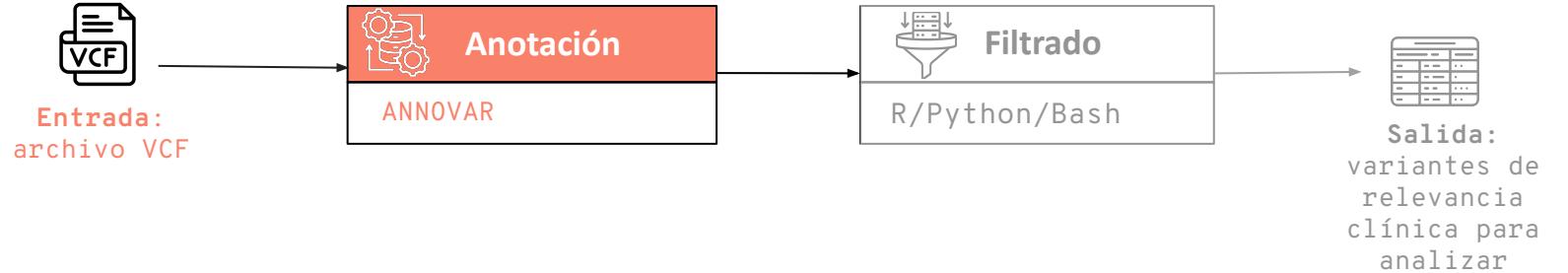


#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT Sample

chr11 69462910 . G A 103,14 DPFilter;MQFilter

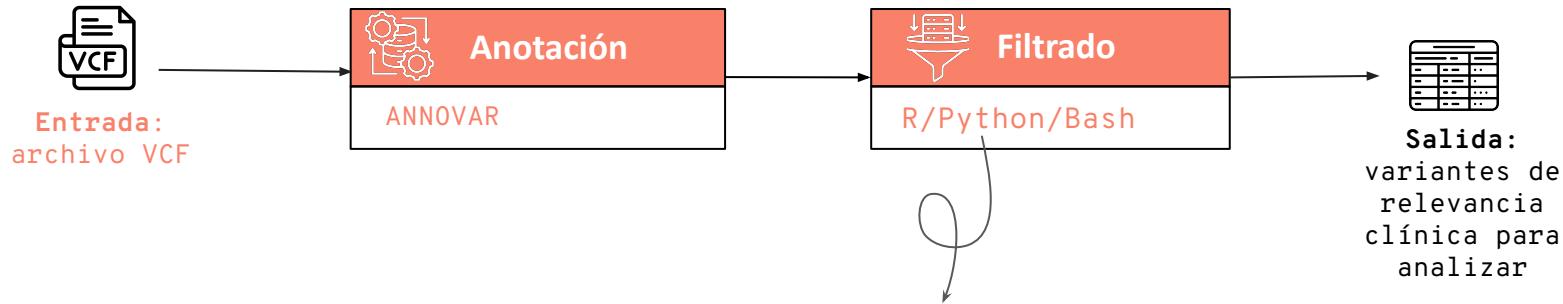
AC=2;AF=1,00000e+00;AN=2;DP=5;ExcessHet=3,01030e+00;FS=0,000000+00.MI_FAF=2.MI_FAF=1 AA000000+00.MN=2 QAA00000+01.0D=2 5790000+01.SOR=6,93000e-01;ANNOVAR_DATE=2018-04-16;Func.refGene=exonic;Gene.refGene=CND1;GeneDetail.: ExonicFunc.refGene=synonymous_SNV;NM_053056:exon4:c.G723A:p.P241P;cytoBand=11q13.3;genomicSuperDups=.;gwasCatalog=Name\x3d\x22CND1\x22\x2c\x22rs9344\x22\x2c\x22rs138\x22\x2c\x22ExAC_ALL\x22\x2c\x22ExAC_AFR\x22\x2c\x22ExAC_AMR\x22\x2c\x22ExAC_EAS\x22\x2c\x22ExAC_FIN\x22\x2c\x22ExAC_NFE\x22\x2c\x22ExAC_OTH\x22\x2c\x22ExAC_SAS\x22\x2c\x22SIFT_score\x22\x2c\x22SIFT_converted_rankscore\x22\x2c\x22Polyphen2_HDIV_score\x22\x2c\x22Polyphen2_HDIV_rankscore\x22\x2c\x22Polyphen2_HVAR_score\x22\x2c\x22Polyphen2_HVAR_rankscore\x22\x2c\x22LRT_score\x22\x2c\x22LRT_converted_rankscore\x22\x2c\x22LRT_pred\x22\x2c\x22MutationTaster_score\x22\x2c\x22MutationTaster_converted_rankscore\x22\x2c\x22MutationAssessor_score\x22\x2c\x22MutationAssessor_score_rank\x22\x2c\x22MutationAssessor_pred\x22\x2c\x22FATHMM_score\x22\x2c\x22FATHMM_converted_rankscore\x22\x2c\x22PROVEAN_score\x22\x2c\x22PROVEAN_converted_rankscore\x22\x2c\x22VEST3_score\x22\x2c\x22VEST3_rankscore\x22\x2c\x22MetaSVM_score\x22\x2c\x22MetaSVM_rankscore\x22\x2c\x22MetaSVM_pred\x22\x2c\x22MetaLR_score\x22\x2c\x22MetaLR_rankscores\x22\x2c\x22MetaLR_pred\x22\x2c\x22M-CAP_score\x22\x2c\x22M-CAP_rankscore\x22\x2c\x22fathmm-MKL_coding_score\x22\x2c\x22fathmm-MKL_coding_rankscore\x22\x2c\x22fathmm-MKL_coding_pred\x22\x2c\x22Eigen_coding_or_noncoding\x22\x2c\x22Eigen_raw\x22\x2c\x22Eigen-PC-raw\x22\x2c\x22GenoCanyon_score\x22\x2c\x22GenoCanyon_score_rankscore\x22\x2c\x22integrated_fitCons_score\x22\x2c\x22integrated_fitCons_score_rankscore\x22\x2c\x22integrated_confidence_value\x22\x2c\x22GERP++_RS\x22\x2c\x22GERP++_RS_rankscore\x22\x2c\x22phyloP100way_vertebrate\x22\x2c\x22phyloP100way_vertebrate_rankscore\x22\x2c\x22phyloP20way_mammalian\x22\x2c\x22phyloP20way_mammalian_rankscore\x22\x2c\x22phastCons100way_vertebrate\x22\x2c\x22phastCons100way_vertebrate_rankscore\x22\x2c\x22phastCons20way_mammalian\x22\x2c\x22phastCons20way_mammalian_rankscore\x22\x2c\x22SiPhy_29way_logOdds\x22\x2c\x22SiPhy_29way_logOdds_rankscore\x22\x2c\x22Interpro_domain\x22\x2c\x22GTEX_V6p_gene\x22\x2c\x22GTEX_V6p_tissue\x22\x2c\x22regsnp_fpr\x22\x2c\x22regsnp_disease\x22\x2c\x22regsnp_splicing_site\x22\x2c\x22GWAVA_region_score\x22\x2c\x22GWAVA_tss_score\x22\x2c\x22GWAVA_unmatched_score\x22\x2c\x22AF\x22\x2c\x22AF.popmax\x22\x2c\x22AF_male\x22\x2c\x22AF_female\x22\x2c\x22AF_raw\x22\x2c\x22AF_afr\x22\x2c\x22AF_sas\x22\x2c\x22AF_amr\x22\x2c\x22AF_eas\x22\x2c\x22AF_nfe\x22\x2c\x22AF_fin\x22\x2c\x22AF_asj\x22\x2c\x22AF_oth\x22\x2c\x22non_topmed_AF_popmax\x22\x2c\x22non_neuro_AF_popmax\x22\x2c\x22non_cancer_AF_popmax\x22\x2c\x22controls_AF_popmax\x22\x2c\x22BP2\x22\x2c\x22BP3\x22\x2c\x22BP4\x22\x2c\x22BP5\x22\x2c\x22BP6\x22\x2c\x22BP7\x22\x2c\x22Kaviar_AF\x22\x2c\x22Kaviar_AC\x22\x2c\x22Kaviar_AN\x22\x2c\x22Eigen\x22\x2c\x22ALLELE_END\x22\x2c\x22GT\x22\x2c\x22AD\x22\x2c\x22DP\x22\x2c\x22GQ\x22\x2c\x22PL\x22

Anotación y filtrado



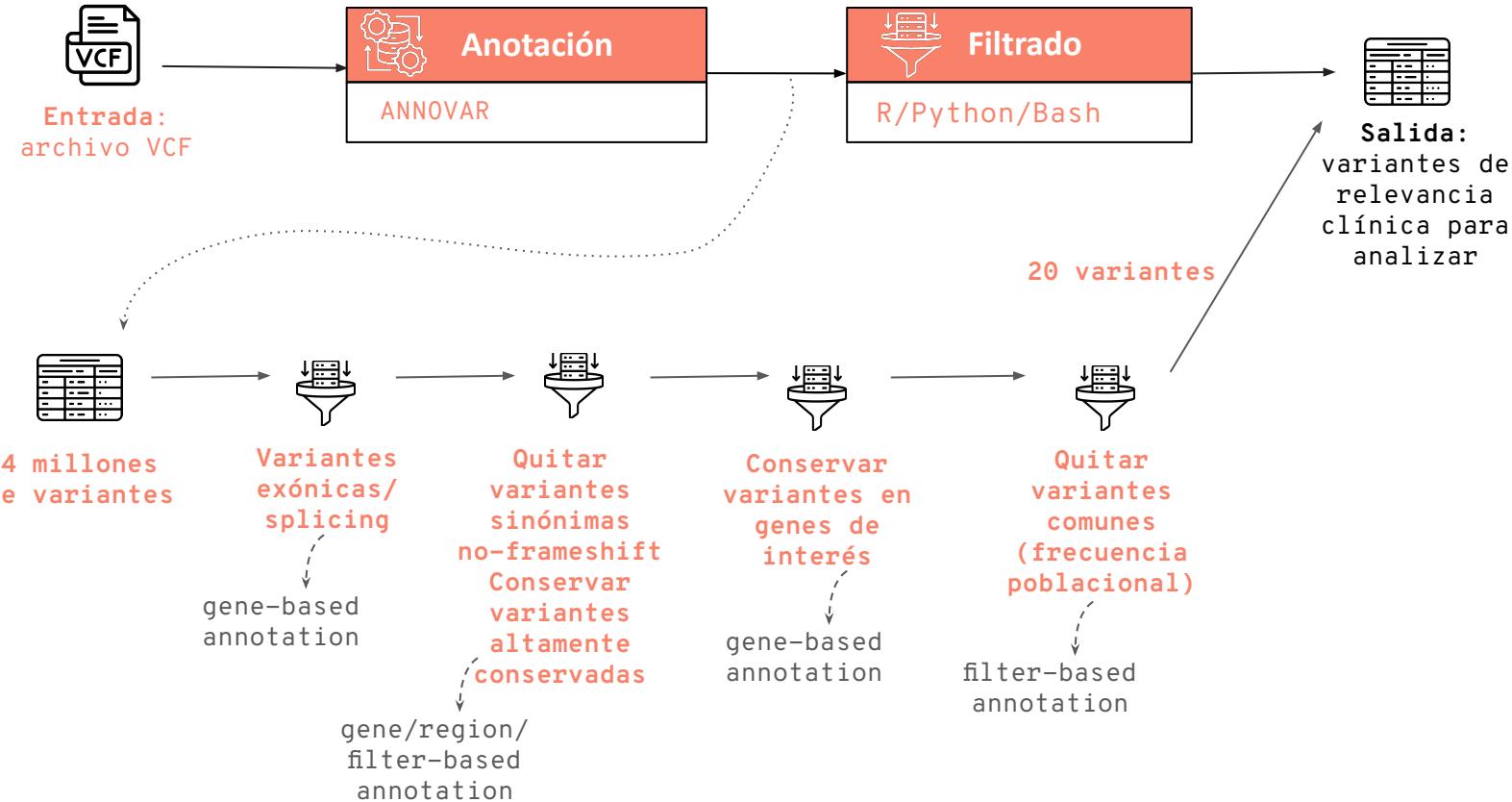
	A	B	C	D	E	F	G	H	I	J	K	L
1	Chr	Start	End	Ref	Alt	Func.refGene	Gene.refGene	GeneDetail.refGene	ExonicFunc.refGene	AAChange.refGene	Xref.refGene	ExAC_Freq
2	1	948921	948921	T	C	UTRS	ISG15	NM_005101:c.-33T>C	.	.	Immunodeficienc	0.941
3	1	1404001	1404001	G	T	UTR3	ATAD3C	NM_001039211:c.*91G>T	.	.	.	0.054
4	1	5935162	5935162	A	T	splicing	NPHP4	NM_001291594:exon17:c.1282-2T>A	.	.	Nephronophthisi	0.825
5	1	162736463	162736463	C	T	intronic	DDR2	.	.	.	Spondylometaepi	.
6	1	84875173	84875173	C	T	intronic	DNASE2B
7	1	13211293	13211294	TC	-	intergenic	PRAMEF36P	F dist=11566;dist=116902
8	1	11403596	11403596	-	AT	intergenic	UBIAD1;PTCH	dist=55105;dist=135699
9	1	105492231	105492231	A	ATAAA	intergenic	LOC10012911	dist=872538;dist=640085
10	1	67705958	67705958	G	A	exonic	IL23R	.	nonsynonymous SNV	IL23R:NM_144701:exon9:c.G1142A:p.R381Q	.	0.041
11	2	234183368	234183368	A	G	exonic	ATG16L1	.	nonsynonymous SNV	ATG16L1:NM_198890:exon5:c.A409G:p.T137A;ATG16L1:NM_0	.	0.457
12	16	50745926	50745926	C	T	exonic	NOD2	.	nonsynonymous SNV	NOD2:NM_001293557:exon3:c.C2023T:p.R675W;NOD2:NM_0 Blau syndrome, A	0.023	.
13	16	50756540	50756540	G	C	exonic	NOD2	.	nonsynonymous SNV	NOD2:NM_001293557:exon7:c.G2641C:p.G881R;NOD2:NM_0 Blau syndrome, A	0.009917	.
14	16	50763778	50763778	-	C	exonic	NOD2	.	frameshift insertion	NOD2:NM_001293557:exon10:c.2936dupC:p.L980Pfs*2;NOD2 Blau syndrome, A	0.013	.
15	13	20763686	20763686	G	-	exonic	GJB2	.	frameshift deletion	GJB2:NM_004004:exon2:c.35delG:p.G12Vfs*2	Bart-Pumpfrey si	0.006038
16	13	20797176	21105944	O	-	exonic	CRYL1;GJB6	.	frameshift deletion	GJB6:NM_001110220:wholegene;GJB6:NM_001110221:whole	.	.
17	8	8887543	8887543	A	T	exonic	ERI1	.	stoploss	ERI1:NM_153332:exon7:c.A1049T:p.X350L	.	.
18	8	8887539	8887539	A	T	exonic	ERI1	.	stopgain	ERI1:NM_153332:exon7:c.A1045T:p.K349X	.	.
19	8	8887536	8887537	AG	GATT	exonic	ERI1	.	frameshift substitutio	ERI1:NM_153332:exon7:c.1042_1043GATT:p.R348Dfs*2	.	.
20	8	8887540	8887540	G	GGAA	exonic	ERI1	.	nonframeshift substit	ERI1:NM_153332:exon7:c.1046delinsGGAA:p.R348_K349insR	.	.
21	5	1295288	1295288	G	A	upstream	TERT	dist=126
22	chr14	95602958	95602958	A	C	splicing	DICER1	NM_001271282:exon1:UTRS	.	.	Goiter, multinodi	.

Anotación y filtrado

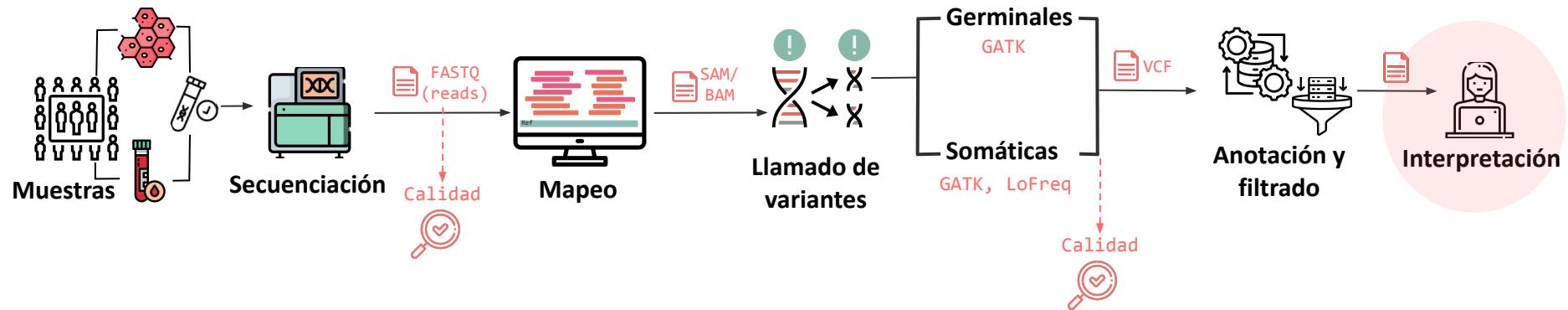


Objetivo: obtener un conjunto reducido
de variantes de interés para su análisis

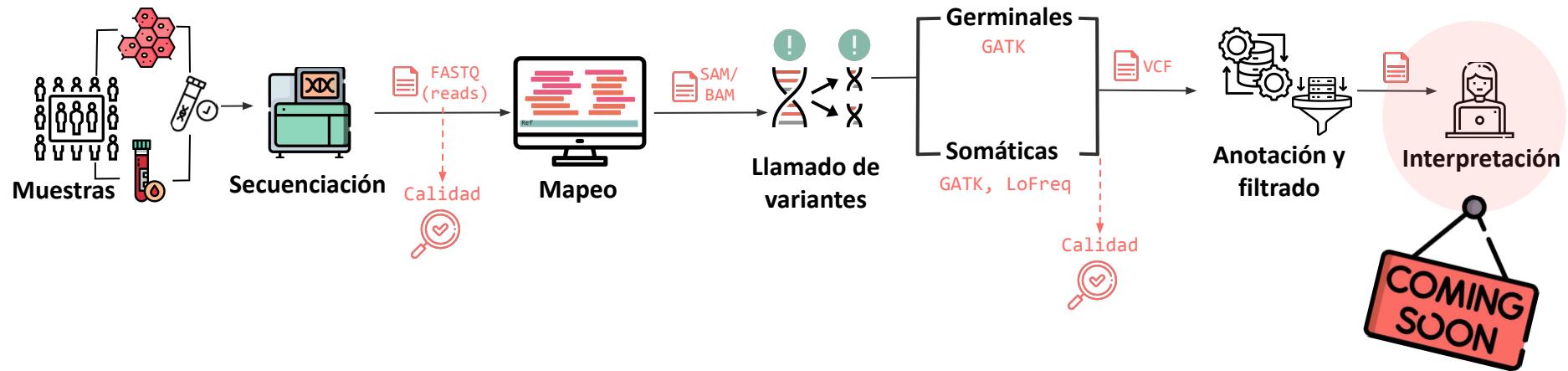
Anotación y filtrado

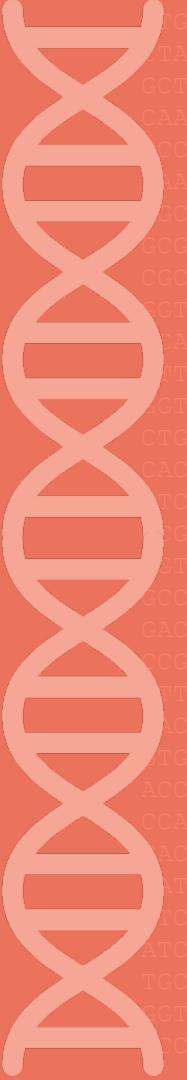


Pipeline bioinformático



Pipeline bioinformático





¡Gracias!



lucia@pasteur.edu.uy
csimoes@pasteur.edu.uy