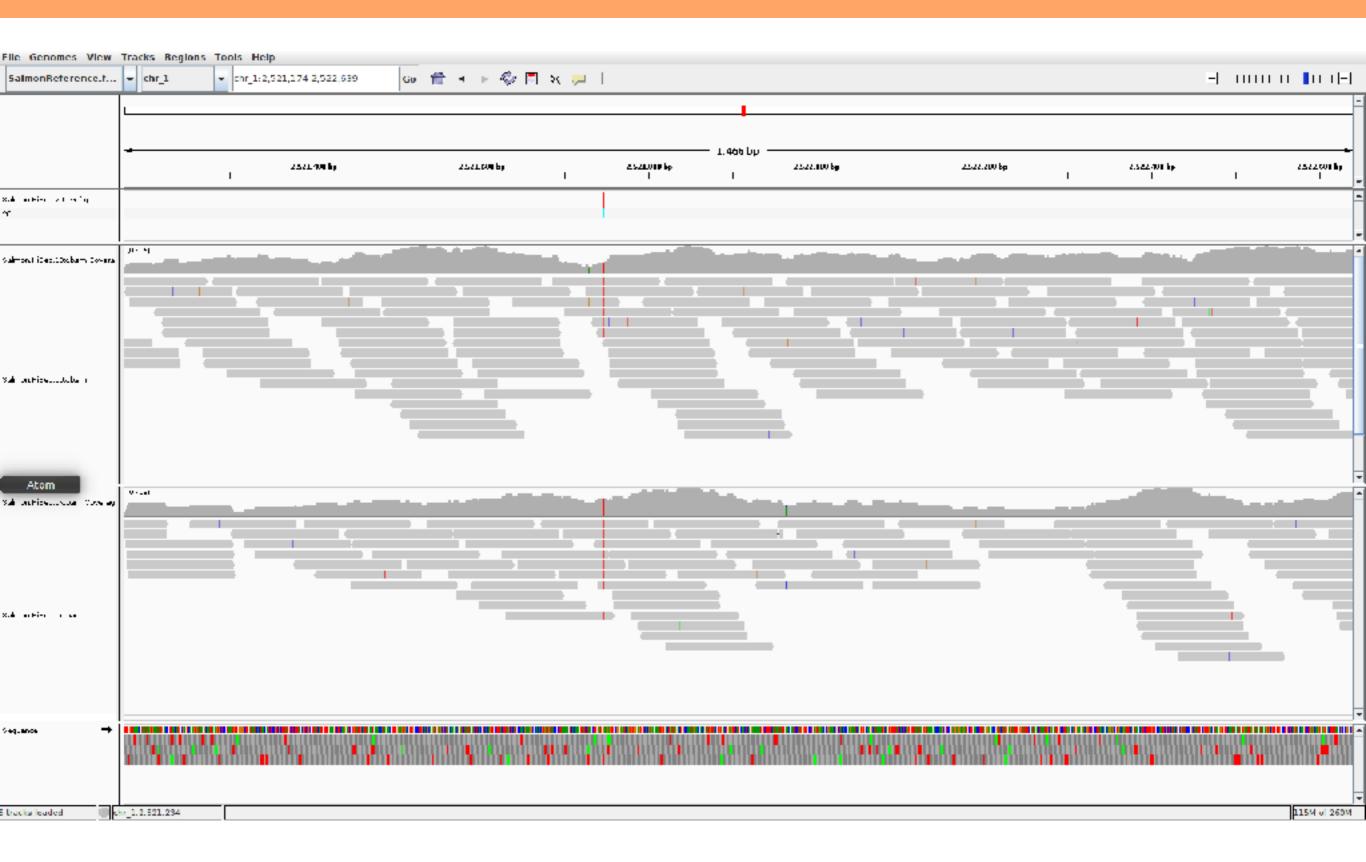
# Topic 8: Variant Calling

# Learning outcomes

- The problem with duplicates and how it's addressed
- Benefits of genotyping with GATK (e.g.the N+1 problem)
- Overview of GATK best practices
- VCF file structure

- 1. Get a reference genome
- 2. Index it
- 3. Map/align reads to it
- 4. Mark duplicates
- 5. Call variants
- 6. Filter variants



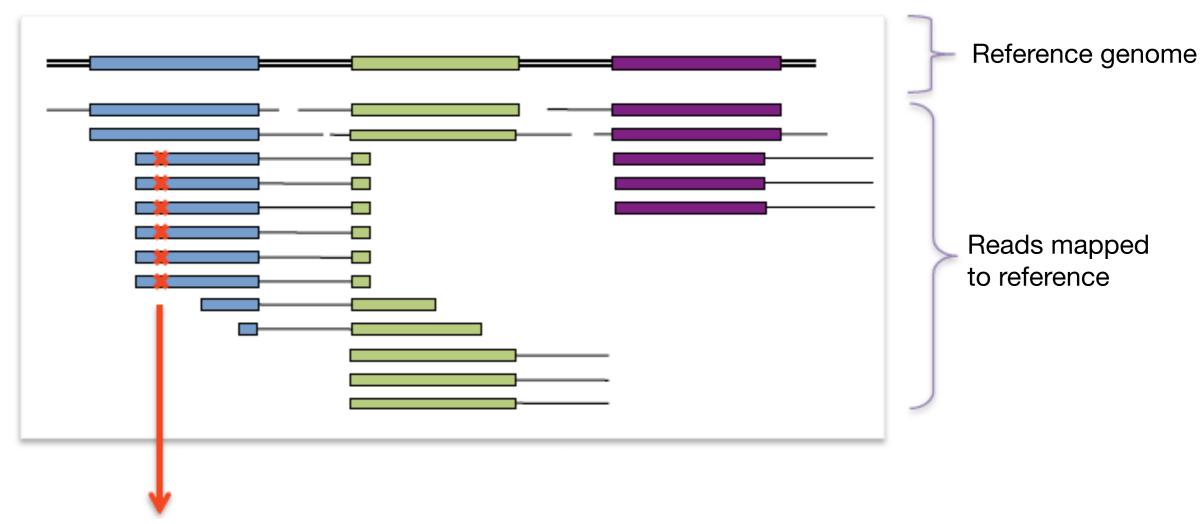
From the first tutorial

- 1. Get a reference genome (we did that already)
- 2. Index it (and this)
- 3. Map/align reads to it (and this)
- 4. Mark duplicates
- 5. Call variants
- 6. Filter variants (next session)



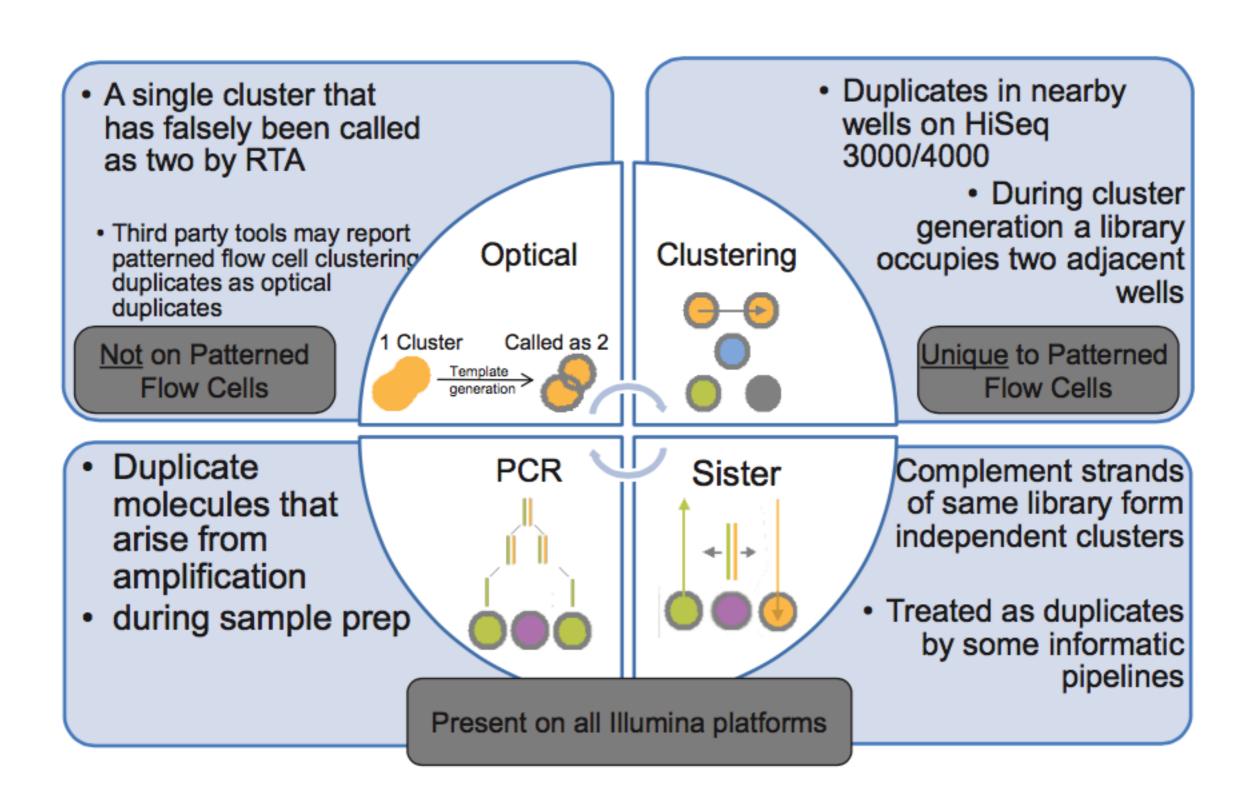
### Why duplicates are bad for variant calling

Sequencing error propagated in duplicates



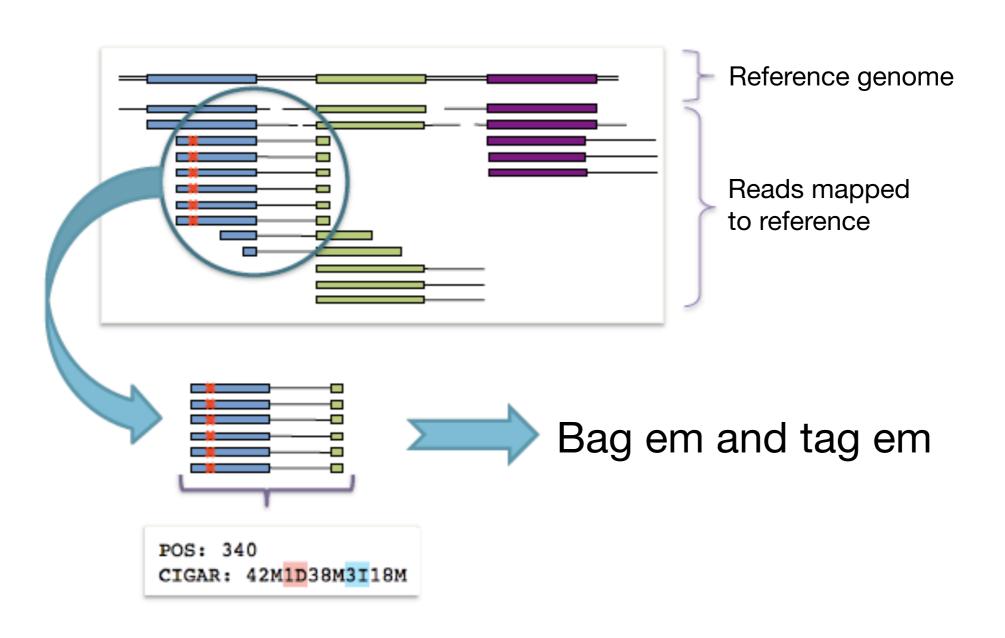
False variant call (bad)

## Where does duplication come from?



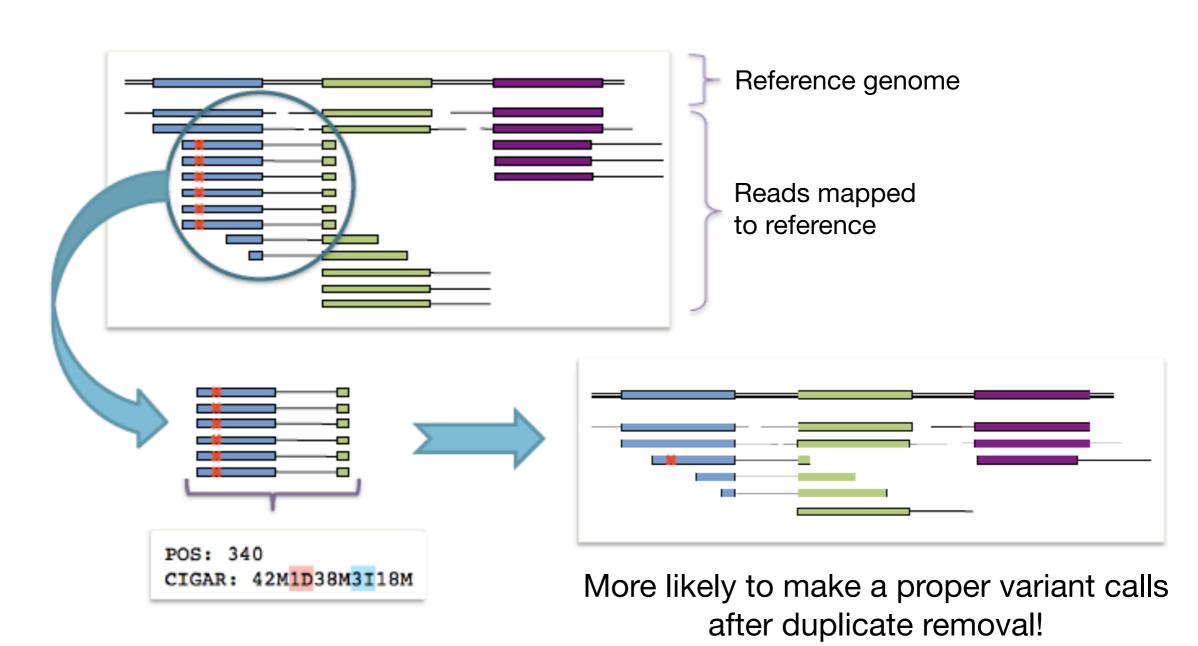
## How to identify duplicates?

Actually pretty easy - duplicated reads would have the same start/end positions and CIGAR strings



## How to identify duplicates?

Actually pretty easy - duplicated reads would have the same start/end positions and CIGAR strings



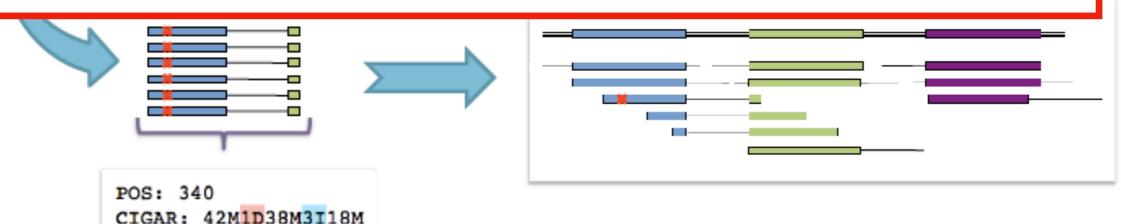
## How to identify duplicates?

Actually pretty easy - duplicated reads would have the same start/end positions and CIGAR strings



There are standard tools to remove duplicates as part of SNP calling pipelines

Are there cases where we wouldn't want to do this?



- 1. Get a reference genome (we did that already)
- 2. Index it (and this)
- 3. Map/align reads to it (and this)
- 4. Mark duplicates
- 5. Call variants
- 6. Filter variants (next session)

#### **Variant Callers**

```
Research article Open access | Published: 22 February 2022
```

Systematic benchmark of state-of-the-art variant calling pipelines identifies major factors affecting accuracy of coding sequence variant discovery

```
Yury A. Barbitoff ☑, Ruslan Abasov, Varvara E. Tvorogova, Andrey S. Glotov & Alexander V. Predeus ☑
```

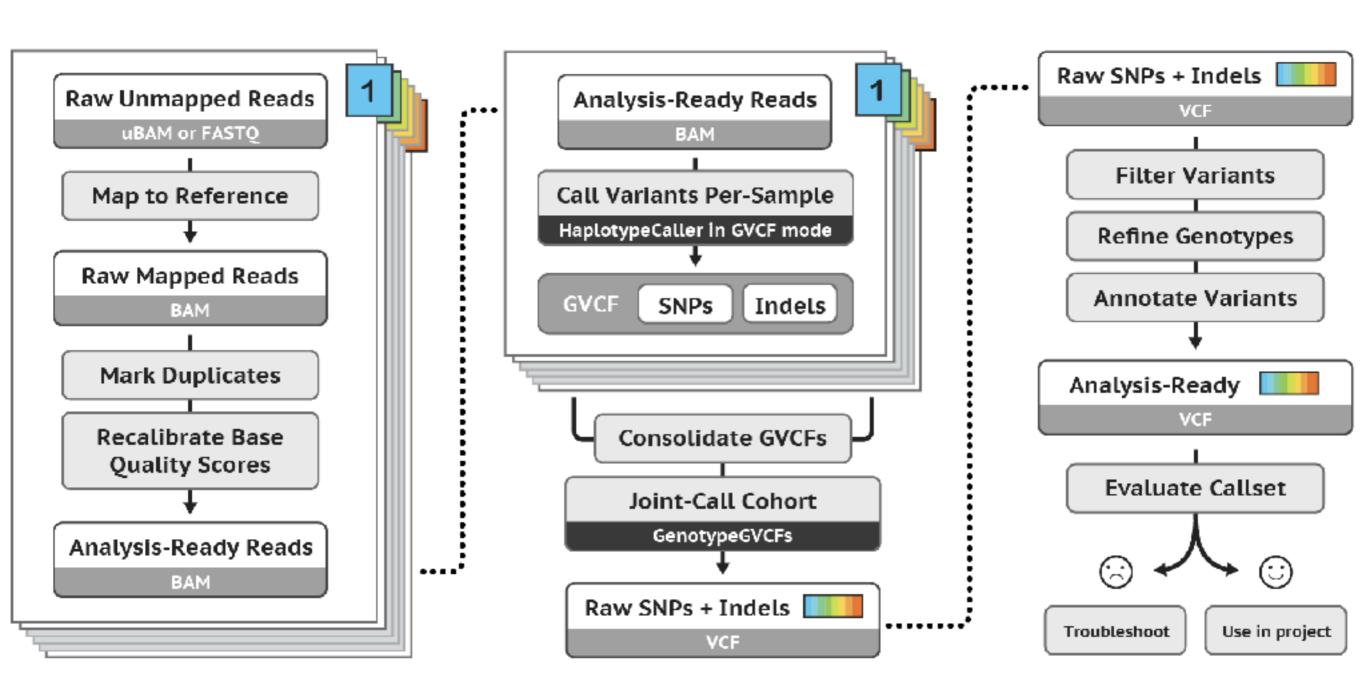
```
BMC Genomics 23, Article number: 155 (2022) Cite this article
```

23k Accesses 34 Citations 4 Altmetric Metrics

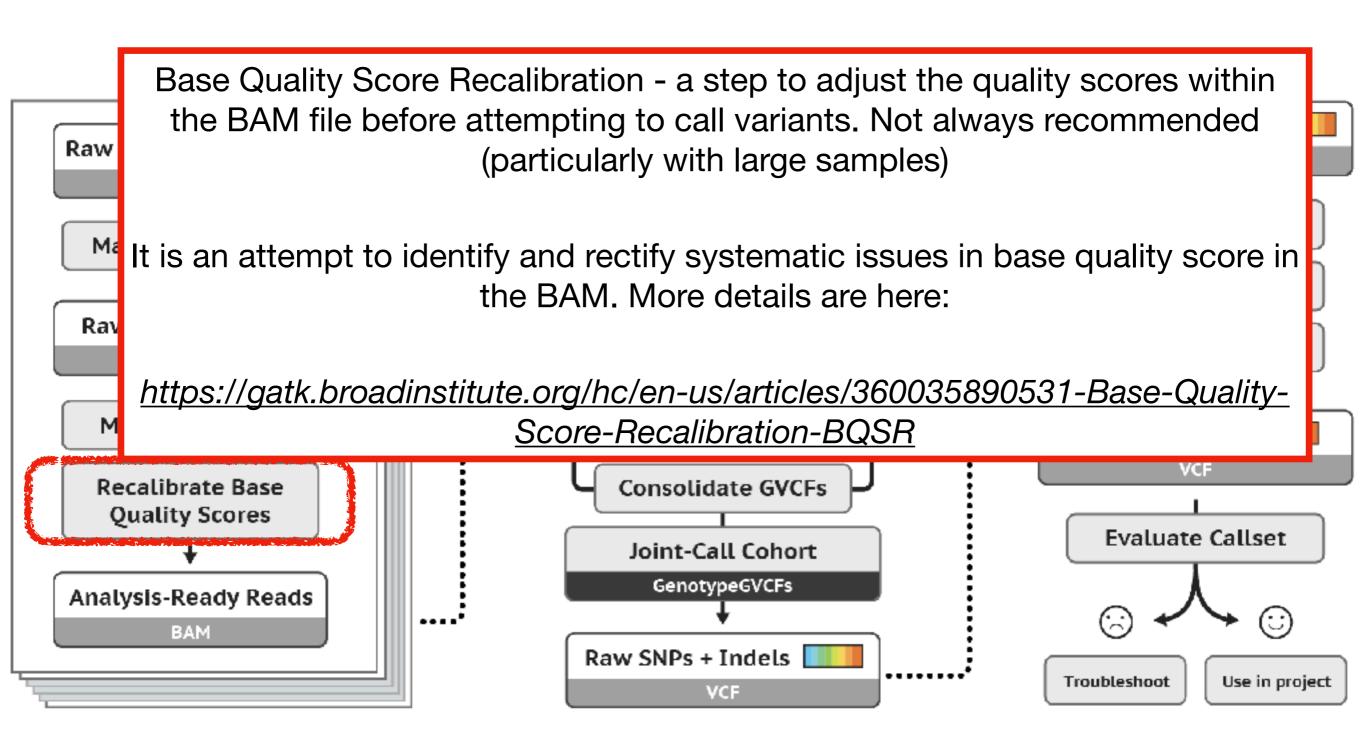
There are lots of variant callers out there, with pros and cons

Probably the most widely used method is GATK, which has similar performance to other methods that are perhaps a little slower

## Overview of GATK Pipeline



## Overview of GATK Pipeline



#### **GATK: Variant callers**

# Unified Genotyper (sunsetted)

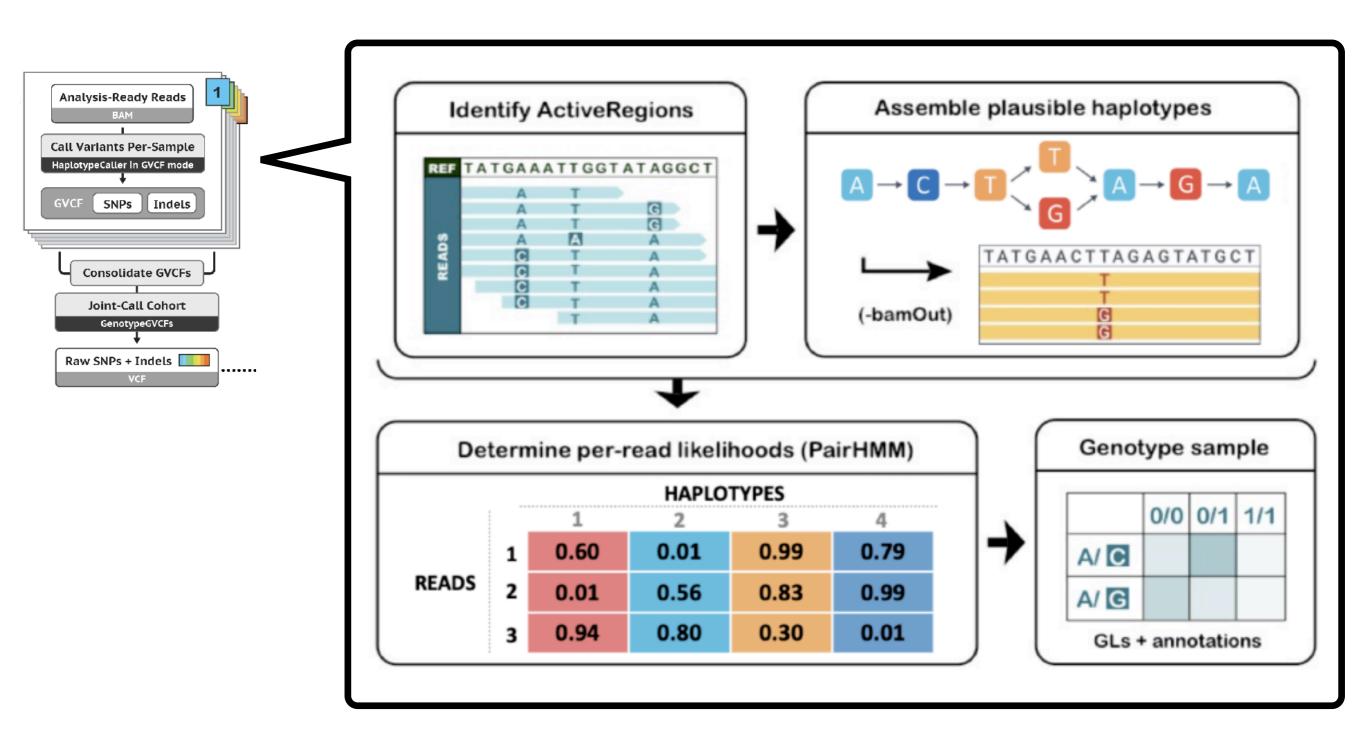
Calls SNPs and in/dels separately

(it did handle multiple ploidy levels and pooled data though)

#### Haplotype Caller

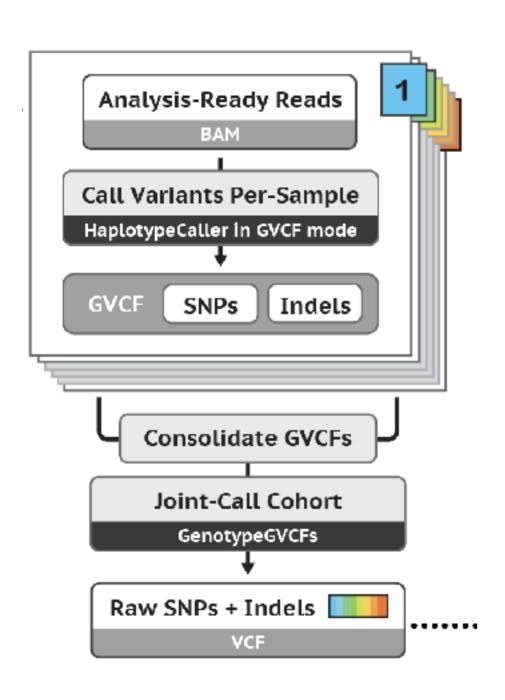
Calls SNPs, in/dels and small structural variants by doing local re-assembly and considering haplotypes

## **GATK:** Haplotype caller



\*Is capable of phasing data as well!

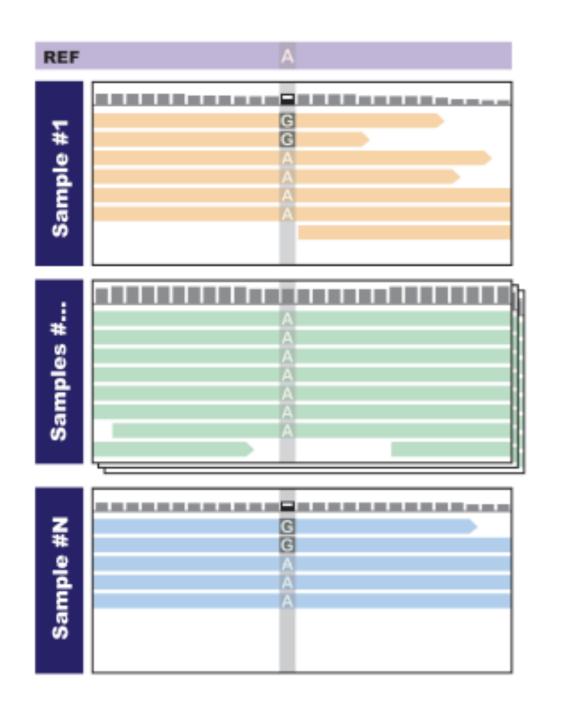
## **Joint Discovery**



Why analyse the data separately, and then together?

Why would we compare information across samples?

## **Joint Discovery**

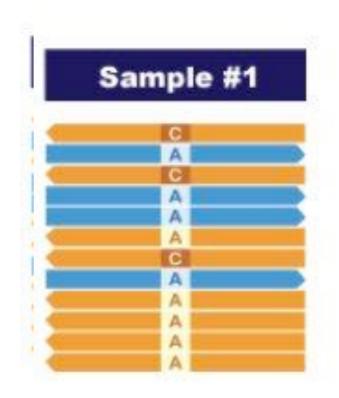


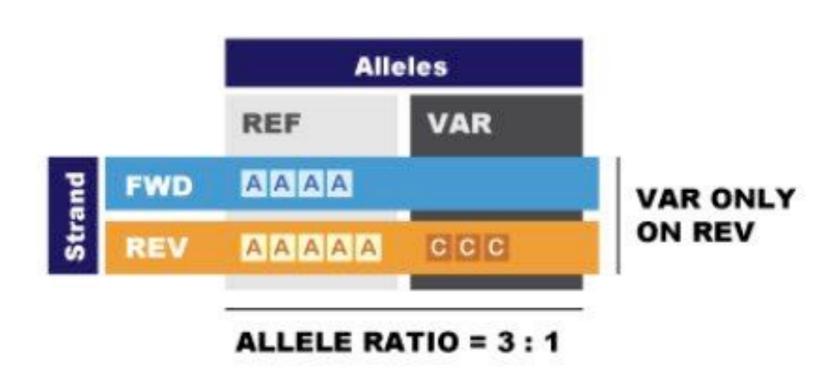
If we analyse Sample #1 or Sample # N alone we may not be confident that the variant is legitimate

If we see the same variant in multiple individuals we are more confident that it is real and not a sequencing artefact

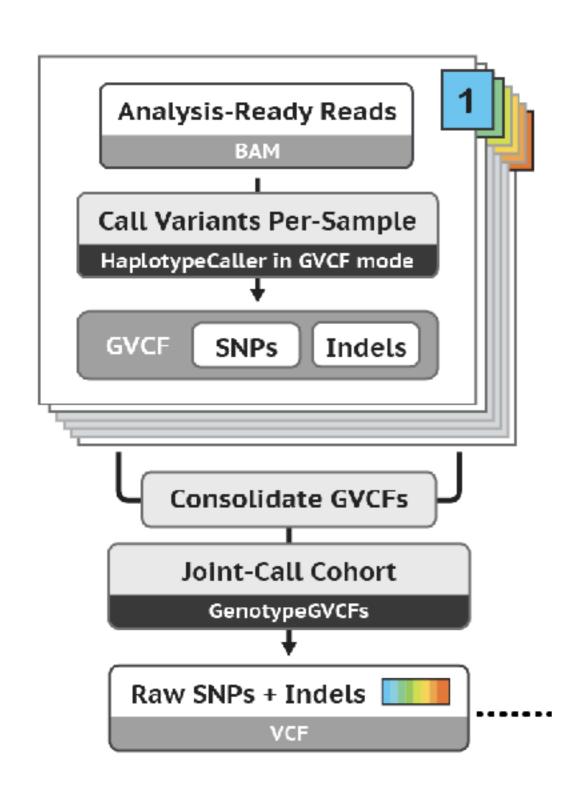
## **Joint Discovery**

For example, in the case of strand bias





#### The N+1 Problem



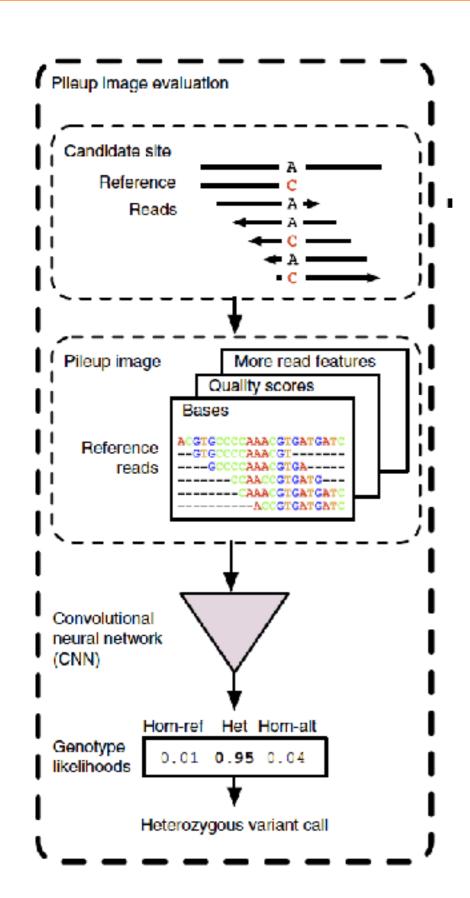
Often, we may receive our genomic data in batches as it is generated by the sequencing facility

Or, extra money is made available so additional samples can be sequenced

Variant calling is computationally intensive so having to re-run from scratch each time additional samples are added would be a waste of resources

This is the N+1 problem

#### Other Variant Callers

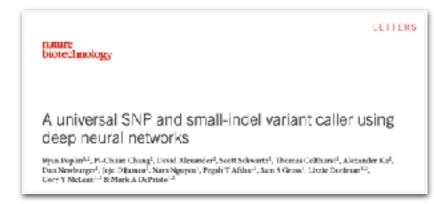




A Machine Learning based variant caller developed by Google

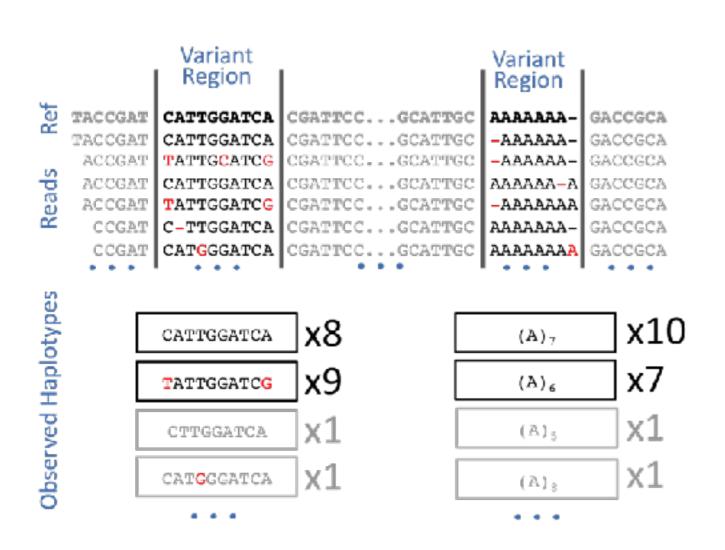
Uses CNNs trained on images of alignments to identify and call variants and extract statistics

Limited to individual samples or trios



#### Other Variant Callers

## freebayes



Uses the reads themselves, rather than the alignment of the reads

Generally faster than GATK (more RAM intensive)

No solution to N+1

Provides lots of summary statistics

Garrison and Marth 2012 arXiv

#### Other Variant Callers

#### **ANGSD**



Calls SNPs based on BAM - no realignment

Outputs genotype likelihoods comes with a suite of analyses that link with it

Recommended for low coverage data (e.g. ancient DNA)

Korneliussen et al 2015 BMC Bioinformatics

#### Structural Variation

For structural variation, the choice of software is dependant on the data type and the kind of variation you're looking for (e.g. CNVs, inversions or deletions)

Article Open access | Published: 14 March 2024

Benchmarking long-read aligners and SV callers for structural variation detection in Oxford nanopore sequencing data

Asmaa A. Helal, Bishoy T. Saad ™, Mina T. Saad, Gamal S. Mosaad & Khaled M. Aboshanab ™

Scientific Reports 14, Article number: 6160 (2024) | Cite this article

4215 Accesses | 1 Citations | 1 Altmetric | Metrics

#### Comparison of structural variant callers for massive wholegenome sequence data

Soobok Joe, Jong-Lyul Park, Jun Kim, Sangok Kim, Ji-Hwan Park, Min-Kyung Yeo, Dongyoon Lee, Jin Ok Yang ≅ & Seon-Young Kim ≅

BMC Genomics 25, Article number: 318 (2024) | Cite this article

3237 Accesses 2 Citations 2 Altmetric Metrics

# The Variant Call Format

```
Milleronmoterblyd.
Millet et killerssyld desigle de Feel liller de mode Po
FALT-viD-tCh_AEF.Descript on- Represente any possible alternative allele not already represented at this location by AEF and ALT'Y
"FRID ER KID Lo-December of ion "Look cooling"
FFFCFF NT=11E=ND.Hurber=N.Type=IntegeryEestript on=1Alletic cepths for the refigholds, alletes in the order Histedii
MRUFFHINVILTUM,Hurbonni, yponintegon,Leconipt enniHpproximete hode copth (roadwill in Hyrkop wrill in bed mated and fillonde). N
  €ECELOT=11E=5),Hurbasel,Typa=Integas,Feacsips.co="Genotypa Quality"≥
MANAPHAIRKILTSI,Murberti, yeerString,Description Senetype /
MECEL AT - 410 - 1100 CE, Momber - "Type- otherer, Dexor of time-"divinus DP of Servence of this tractivation for Electrical
MRAFFHITALITYSI. Number 1, goetetring Lessarips, and Physical phabling habitages internation, describing the attended alleases in relation to one crothers will always be necessarily and is not intended to describe equal alleases in
WECFLAT-ALC-ALD, housewell, Time-Strong, Description-Alphasing. Or information, nonce and not pied for it for a given some a Character second 
WESSIBLE-15-25, Musland, To exhibit year, Testrip to the Plastag set (tag multiplier and to the first variable to eset)'?
MESSIBLE-18-18-190, Automotive generating the Unconditional reference genotype confidence, encoded as a prived condity 19*log18 ofgenotype cold to unong)"?
MENTER OF St. Multiplier E. Grand Compared to the order of the Unong of the Uniong of th
MONTKODIBLOCK December 2008 Control of the Control 
Fib Communities Colling Grand for Signify (Miles Fred at the 20 incl. 1885) in Lea Lyand // 86% make about the Life Community and A. / 86 to be about the contract of the cont
FOFTNOWMERCLINGS OF DELIGIOUS AND COMMENDATION OF THE CONTROL OF T
MINE (ID Ft. Marker Hylige Integerydeser of ich Millete boart in geretgoest i om ecchiebliet beleit in die oanvierdomies violeem.
 FINTCAMIDAGE, Mumberan, Typea-Took, Jason ptionan, Itale Prequency, for each fLT allele, in the wore order as lieted
MinHumkiBhHk, Numberh],iyperinteger,Jeser oblernflote, humber of al.e.eo in egilee genotypeo'∠
Filtfor(In-Fase)NorkSum_burners[] peef cat, Near vintion="2 acces (non dilocked bank sum test of 6 t Ma. Def bose qual tiex")
MINFUrviDHLF, fumborni,luporintecer,Jeser otierniHopres mate read cepths sere reces may hove been filtered'∠
FiltEC-/ID-ENDyMonter-Lifypa-In égaryNextricio-'S optiosit ou of thatioter-of t
MinkurviDrbaccookst, Humberri.igper-lost, Jecor pliantiknhad-semiod p-value foh excel tost af excesa hetehazyges ty 🗸
MileBC-/ID-BS, "bodier-li, Type-Floot, Dascolytion-"Erwaltscalad year becausing Elector's as no it action in latential son or biox".
MititariDa norsecting/oath, lumberal, lypea loat, Lescriptional norsecting coefficient as estimated from the genotype like. Inocce per schole when compared against the Hardy Weinberg expectation??
With First India First production of the same of the same of the same of the first production of the same of the same of the first production of the same of 
"THECK IN IC. Howler I, Type Flows. Ocean plane "FIS Mapping Country".
MikfCeviDeDCCankCon, Humbere1.Type=Floot, Description="Ziscore From Alicoxon rank sun test of Alt vs. Fef read modoling qualities" -
"MBH ՀՄԱՄԻ համից 1,1թբ. - համե համայինամ Գամայի Համայի գոխանից հայ հայ հայ հա
```

MINTCHNIDED TO BE A CONSTRUCTION - 10 CONSTRUCTI

Filebook Hitera Problem Switz aloo 1. Type book place in the case of a colline or make a contract of the School problem. For a second or a colline book is the colline book in the colline book in the colline book is the colline book in the colline

FiltCeriDeCCN,Humberel.Type=Floot,Dewor ation="Symmetric Occe Patic of 252 cantingency tople to detect strand bids">

Medity KIL omiltiong, PS3998601

₹8000683=1tgre.,2\_nc=31'=p trace? Moduree-UdranieoLt hoort

ēžkourseeberot (pa£907s). Moduree-Hop.otypolo.len

Min Trads ar am Serkinn- "2+h s. a-1.2.

```
Mbertoploweenpatternandreenpat Ht. a wer/Chinochus/Hs.enhwi.veruga wer/ChinophwG/Hs.enhwz.vot.ga
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               He China scalaid
                                                                                                                                                                                                                                         CIA FLITER MEC FORLET HE Colomba dails — Ha Chica Maple
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               He Cl. code, p. 12
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                             ta Olimbia in the
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    He Brimak, daif
                                                                                                                                                                                                                                                                                                                                               10-4; FT-9. 626; TH-156; BaseQFmx9cr-3.524; ET-209; Elicesel et =9. 6869; T9-6. 399; Inbreeding Doe FF-8. 2702; hubfe-8. 332; hQ-59. 66; dQTarkCur-9. 66; QD-14. 17; Necc Tos Nork Cur-6. 69; D0F-8. 598.
                                                                                                                                                                                                                                              :73.69
                                                                                                                                                                                                                                                                                                                                             97-27: 97-8.27 (98-40) The Angelog in State - 1.066 a - 0.066 a - 
                                                                                                                                                                                                                                             3919.39
                                         16.7
 hr_1
                                    196
                                                                                                                                                                                                                                              19739.54
                                                                                                                                                                                                                                                                                                                                                                                                * 1/1.14
shr_1
                                      379
                                                                                                                                                                                                                                            513.88 .
                                                                                                                                                                                                                                                                                                                                                10-7;61-9;646;14-164;89xeQFan:Sur=3:90;EF=258;Expession=9:0661;10-4;631;intreedingCoeff=6:3271;thETO=6:fLETNF=6:319;thQ=50:06:fQTankCun=9:66;QD=81:96;Toe6CoeffosTankCun=1:64;00f=8:125
                                                                                                                                                                                                                                          : i.14 .
                                                                                                                                                                                                                                                                                                                                               45 See Louis See Copen Margin of Books of Early 18 Mary Section 65 Contracted Contracted
                                       726
                                                                                                                                                                                                                                             .5758.77
                                                                                                                                                                                                                                                                                                                                                                                                10-.33;17-..33;18-.33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18-33;18
hr__
                                                                                                                                                                                                                                              15009.40
                                                                                                                                                                                                                                                                                                                                                                                                 10-4;61-9.623;14-176;07-886;Excess lat-6.6392;60-3.006; horsed hqDosff-9.8499; 1266-3; 1266-3; 10-66.63;00-83.12;006-3.912 67;60;07;00;00;00;10 673;5.6;6;13;6,13;206 370;2;3;2;6;6;5,66 4273;6-73;60;410;6-73;60;410;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;6-73;60;
                                 804
                                                                                                                                                                                                                                             200.52 .
 hr_-
                                                                                                                                                                                                                                           13.52 ...
hr_1
                                      1866

    61: 43:00-36-30:24.
    6/3:88,88:38-24.
    7: 3/8:3. 3:38:38.

                                         104.3
                                                                                                                                                                                                                                            G7.35
                                                                                                                                                                                                                                                                                                                                                                                                40-15/24/F-1. 334/HF-10/24/F-43/54/P/COORD-40/T-2. 31/85/F-3-8/EBC: InbroodingCont-W. 46/214/ILEF/71/96/FLEAF-1. 334/ILEF/31/F-3. 38/34/F-1. 27/5 - 61/4/BED/4/01/F-3/6/F-3/4/F-1. 27/1/E. 3/1/5/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4/F-3/4
shr_1
                                        1852
                                                                                                                                                                                                                                             17515.61
                                                                                                                                                                                                                                                                                                                                               30-1 (67-5), Cla-30(66-170) Fulsa) Buck5 (a-1), 67(3) F-47(3) F-47(3) F-48(4), -20(3) 3(6) F-30(66) - (100-6) (104) (107-3) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-30) (104-
                                                                                                                                                                                                                                                   jalar .
                                                                                                                                                                                                                                                                                                                                                                                                HUMO 1954 HAD. 335 HAM. 1950 HAM 125 Excoponios TO. 30 MOS HISTS. BED. Inbrooding Local THU. 38 FILEH (TO MOS FILEH HAL. 335 HILLES). BES. SURTE. 6.13 | BIO HISTORY PL | 1/15 M. C. 5: 18: 25 d., 18: 8 | 1/15 D. C. 5: 18: 25 d., 18: 8 | 1/15 D. C. 5: 18: 25 d., 18: 8 | 1/15 D. C. 5: 18: 25 d., 18: 8 | 1/15 D. C. 5: 18: 25 d., 18: 8 | 1/15 D. C. 5: 18: 25 d., 18:
                                        1456
                                                                                                                                                                                                                                             23573,26
                                          1-56
                                                                                                                                                                                                                                                                                                                                               30-1;65-7.200-490;66-196;60-490;65-0-1.090-6-;08-470;55-6-8-44-7.6-30;65-9-306; | 109-6-1-1006 | 156-7-200-490;6(-3-06);1 | 30-05-1-3-6-; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470;55-9-306; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; | 19-470; |
                                                                                                                                                                                                                                                   79.29 .
  hr_:
                                                                                                                                                                                                                                                                                                                                                12-by: -8.00b; 14-19b; Saset; and Sure 1.00be 30:th = 132:th cessilet-30:30b; 3-0.0bb; intreeding to e14-0.00b; 1-5:th Life 5:th Life 5:
                                     11.78
                                                                                                                                                                                                                                             lod.2c .
                                                                                                                                                                                                                                                                                                                                             30- Jagafest 70; fb- 70; fb-40; fb-30; fb-30
                                          1-53
                                                                                                                                                                                                                                                  777A.60
                                         1886
                                                                                                                                                                                                                                              7835,43%
                                                                                                                                                                                                                                                  2770. C
                                                                                                                                                                                                                                                                                                                                               10-5; 17-9: 611; 14-162; BaseQFankOur=8: 90: EF =121: Et desidet =3: 0514; T0=5: 115; IntrredlingCoef (=6: 1895; hLETC=6: 815; hCTC=6: 815; hQ=50: 06; hQTankCur=9: 66; QD=12: 83; headFosTankCur= | 1: 053e (66; DDT=6: 635; 07: f D: DI
                                                                                                                                                                                                                                            231,72 ...
```

# The Variant Call Format

Millietormoterskynds Millietormoter Pass, biologij biolik i libiografijos de Pa

"FRID ER KID Lo-December of ion "Look cooling"

104.5

1852

1-56

11 rk

1886

G7.35

17515.61

jalar . 23573,26

79.29

155.2c .

777A.60

7835,43% 2570. C 231,72 ...

FALT-viD-tCh\_AEF.Descript on- Represente any possible alternative allele not already represented at this location by AEF and ALT'Y

FFFCFF NT=11E=ND.Hurber=N.Type=IntegeryEestript on=1Alletic cepths for the refigholds, alletes in the order Histedii

```
MRUFFHINVILTUM,Hurbonni, yponintegon,Leconipt enniHpproximete hode copth (roadwill in Hyrkop wrill in bed mated and fillonde). N
 €ECELOT=11E=5),Hurbasel,Typa=Integas,Feacsips.co="Genotypa Quality"≥
MRUHFHIRKILHSI,Humberni, yeerString,DescriptionerGenetype /
MECELAT-200-100 CE, Nomber-1, Type-1 obeyer, Dexor Intim-" Divisor Development of this the GMCE Into k"z
MREAFFAITS LETS I. No spently goethering. Losses Ipt and Physical phabling habitotype Internation, describing how the externete alletes are phable to relation to the enterexygous and is not intended to describe equipment.
WECFLAT-ALC-ALD, but now-1, Time-Str. of Cestading to the Ministration, whose endowing in the original shape shape and the contract shape shape and the Ministration of the Ministration o
WESSIBLE-15-25, Musland, To medite jer, Sescription of Plus by set (typically the bost or of the Cost variable the set)'?
MESSIBLE-18-18-190, Automotive jer, Sescription = Unconditional reference genotype confidence, encoded as a prived condity 19*log18 of genotype cold to unong)
MENTER OF St. Muslander August Orbitation of Marketing Companies and other conditions of the Companies of the Compan
_CFFO.samp.elmap i reader threads 3 intervals shrill genom sact segment size 1918
File Communation Colling Lyndiw England of the Complete Health and Colombia 1985 and Lyndia Lyndia / Ho/Sacombia example of Leerch or to 2.7.
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 Kalamana kalama/kalamka lamana, luaru ee naharasan eesamin erahar luarus ee rangeel aar
#667%Commond_ine=v Dellap.clypeCaller.CornandLine='llap.otypeCaller | native pair from threads 6 | emit mef borf denoe CPCT | cutout grof/Chinopl..pt. iB.cecub.g.vcf
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                   Milita (ID-Fallamber Ajigge Integenjdeser olien Maliate boart in gerelgbest om esen Abi ettelet in til bane erdom es tibleem
 FINTCAMIDAGE, Mumberan, Typea-Took, Jason ptionan, Itale Prequency, for each fLT allele, in the wore order as lieted
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  Header
MinHumkiUmHk, dumbermi,iypeminteger,Jeder oblichmilote, humber of alleleo in edilee genotypeo'∠
Filtfor(In-Fase)NorkSum_burners[] peef cat, Near vintion="2 acces (non dilocked bank sum test of 6 t Ma. Def bose qual tiex")
MINFUrviDHLF, fumborni,luporintecer,Jeser otierniHopres mate read cepths sere reces may hove been filtered'∠
FiltEC-/ID-ENDyMonter-Lifypa-In égaryNextricio-'S optiosit ou of thatioter-of t
MinkurviDrbaccookst, Humberri.igper-lost, Jecor pliantiknhad-semiod p-value foh excel tost af excesa hetehazyges ty 🗸
MileBC-/ID-BS, "bodier-li, Type-Floot, Dascolytion-"Erwaltscalad year becausing Elector's as no it action in latential son or biox".
Mititarila norsecting/oeth, lumberal, lypea loat, Leschipt onal norsecting coeth clerk as estimated from the genotype like. Inocce per schole when compared against the large deliberage expectation?

With First In-Li Fac, has non-first product the state of the scholar action of the scho
"FIREC SIN IC. Smaller I, Type Floot. Ocean plaine "FIS Mapping Case High.
MikfCeviDeDCCankCon, Humbere1.Type=Floot, Description="Ziscore From Alicoxon rank sun test of Alt vs. Fef read modoling qualities" -
FileD : HEIL Ander 1, byr - hall tens y binn "Swind Certific reγtag iby by tgrt">.
MINITER OF THE PROPERTY OF A STATE OF THE PROPERTY OF A STATE OF A
MIDEL AND France Foods and National Author Englished Assistance and the Community of March 1995 and the Advanced France Community of the Commu
∺IFFC=/ID=CCN,Humber=I.Type=Floot,Dewor ation="Symmetric Occe Patic of 252 contingency toble to detect strand bigs")
<sup>M</sup>eer tog KIE om<u>met Long</u>t (19989€E)
€S000663=ftgre.,2_nrc=11*=p trace?
Moduree-UdranieoLt hoort
ēžkourseeberot (pa£907s).
Mocumee-Hop Jotypolo . Ten
Min Trads ar am Serkinn- "2+h s. a-1.2.
Mbertoploweenpatternandreenpat -t. ლ ver/Jhlinech_AHS.enh_1.verugz ver/Jhlinep.ლწ.ტნ.enh_2.ve1.gz
                                                                                                                                                                                                       NATIONAL SET OF CONTROL OF CONTRO
                                                                                                                                             ÇIA ELITER MEC FOELET -- Crimok, duif
                                                                                                                                                :73.69
                                                                                                                                               3919.39
hr_1
                     196
                                                                                                                                                19739.54
                                                                                                                                                 * 171.14
                                                                                                                                              513.88 .
                                                                                                                                              1 6.14
                                                                                                                                                                                                                                       10-135;17-1.30;18-135;DF-395;Excess Not-3.3103;FS-9.666; InbreedingCoeff-0.5.32; MLEFD-195;FLEFT-1.7; MC-63.30;CE-33.57; SON-1.4.3 CT; NOsDF;CQ; PL 1/1:0.6; 5:18;247, 18.6 1/1:0,5:5:15; 290
                                                                                                                                               .5758.77
hr_.
                                                                                                                                               15009.40
                                                                                                                                                                                                                                        40 (1/254) 1.0359) 1.0250) 1.0550/ccos40 1.3190; 63 3.660; 1.60 ccd; gloc 1.8.4 4/580; 62 1.0270; 4.1.070; 63 38; 64 27.54; 65 4050) 1.02 65 4050) 1.02 7.54; 63 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050; 65 4050
                                                                                                                                                                                                        80:1
                                                                                                                                               200.52 .
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  3/0:2,3:2:6:6,5,66
                                                                                                                                              03.52 ...
hr_1
                       THEC
                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                24,311 3/8:0.de3:0:d.J
```

/339||U\_C3J39:UL\_25.88;3J4-8.6J3 | 61:4J8DF:6J;7L | 1/189.6:5:18:253,18.8 | 1/1:6,282:6:54,1 70; 1 FEC- ; 1 FEE-7,2004-90; tig-19,46; tig8 mbSc (-8,66; g0-25,0); 846, E1,8m bSc (-1,40; 508-6, 220

21: PEL 12-3: PEL 1 -6. 615; III,-63. 30: PL Innkton-3. 06: IJ-23. 16; II-adi os Innkton-11. 65 I-306: IJ-11. 162 1 - 4.0070: 1 FEC+ 14;11 F9F+4.059: 10+06.04;10F±48.0±40.46;00+79.50: 8±41F+8±4541+7.0164+41;508+4

30-1;65-300 Ca-30;66-170;Cuka)8uuk5 u-1067;38-459;50 exikta 1-0,0180;65-3066; iliumen ilijoje 17-3,0140; tl 660

90-1;65-5,200-4-30;66-196;60k4)80k5 (4-1),0334-6;03-470;506-844-206,30;65-9,366; (109-6),1;00-60-3

90- 19;95-6, 701;65- 70;66-408m k5 w-8.66;09-51 (Facesardet -6.6099;65-9.986) (Indeed Ing

12-5;1 =0.016; 14-156; Saseti andSure (1.516e 01:11 = 152:15 dessileted, 0596; 0=0.06d; intreeding (certeb)

HUTO 15gHFT0, 33gHFT0 15gUFT4 12gExocopHetT0, 3195eFBT3, 886g Inbrood Ingloon1T8,5643gHLEFCT196eFLEFF

# VCF: Header

```
##FILTER=<ID=PASS,Description="All filters passed">
##ALT=<ID=NON_REF,Description="Represents any possible alternative allele not already represented at this
##FILTER=<ID=LowQual, Description="Low quality">
##FORMAT=<ID=AD, Number=R, Type=Integer, Description="Allelic depths for the ref and alt alleles in the order
##FORMAT=<ID=DP, Number=1, Type=Integer, Description="Approximate read depth (reads with MQ=255 or with bad make the contraction of the contractio
##FORMAT=<ID=GQ, Number=1, Type=Integer, Description="Genotype Quality">
##FORMAT=<ID=GT, Number=1, Type=String, Description="Genotype">
##FORMAT=<ID=MIN_DP, Number=1, Type=Integer, Description="Minimum DP observed within the GVCF block">
##FORMAT=<ID=PGT, Number=1, Type=String, Description="Physical phasing haplotype information, describing how
phased in relation to one another; will always be heterozygous and is not intended to describe called alle
##FORMAT=<ID=PID, Number=1, Type=String, Description="Physical phasing ID information, where each unique ID v
not across samples) connects records within a phasing group">
##FORMAT=<ID=PL, Number=G, Type=Integer, Description="Normalized, Phred-scaled likelihoods for genotypes as a
specification">
##FORMAT=<ID=PS, Number=1, Type=Integer, Description="Phasing set (typically the position of the first variar
##FORMAT=<ID=RGQ, Number=1, Type=Integer, Description="Unconditional reference genotype confidence, encoded of
p(genotype call is wrong)">
##FORMAT=<ID=SB, Number=4, Type=Integer, Description="Per-sample component statistics which comprise the Fish
```

Contains detailed information on what each column contains, the file version, commands used to generate file etc.

Lines starting with ##

##fileformat=VCFv4.2

strand bias.">

# VCF: Records

```
##ALT=<ID=NON_REF,Description="Represents any possible alternative allele not already represented at this location by REF and ALT">
##FILTER=<ID=LowQual, Description="Low quality">
##FORMAT=<ID=AD, Number=R, Type=Integer, Description="Allelic depths for the ref and alt alleles in the order listed">
##FORMAT=<ID=DP, Number=1, Type=Integer, Description="Approximate read depth (reads with MQ=255 or with bad mates are filtered)">
##FORMAT=<ID=GQ, Number=1, Type=Integer, Description="Genotype Quality">
##FORMAT=<ID=GT, Number=1, Type=String, Description="Genotype">
##FORMAT=<ID=MIN_DP, Number=1, Type=Integer, Description="Minimum DP observed within the GVCF block">
##FORMAT=<ID=PGT, Number=1, Type=String, Description="Physical phasing haplotype information, describing how the alternate alleles are phased in
be heterozygous and is not intended to describe called alleles">
##FORMAT=<ID=PID, Number=1, Type=String, Description="Physical phasing ID information, where each unique ID within a given sample (but not acros
phasing group">
##FORMAT=<ID=PL, Number=G, Type=Integer, Description="Normalized, Phred-scaled likelihoods for genotypes as defined in the VCF specification">
##FORMAT=<ID=PS, Number=1, Type=Integer, Description="Phasing set (typically the position of the first variant in the set)">
##FORMAT=<ID=RGQ, Number=1, Type=Integer, Description="Unconditional reference genotype confidence, encoded as a phred quality -10*log10 p(genotype)
##FORMAT=<ID=SB, Number=4, Type=Integer, Description="Per-sample component statistics which comprise the Fisher's Exact Test to detect strand bit
                                                            QUAL
                                    REF
#CHROM
           POS
                                                ALT
                                                                        FILTER INFO
                                                                                                FORMAT -e Chinook.p1.i0
```

##fileformat=VCFv4.2

##FILTER=<ID=PASS,Description="All filters passed">

Contains detailed information on what each column contains, the file version, commands used to generate file etc.

# VCF: Records

FORMAT -e Chinook.p1.i0

-e Chinook.p1.i1

FILTER INFO

QUAL

#CHROM POS

ID

REF

#### Records for two SNPs

How do we know that they are SNPs?

# VCF: Records - INFO

```
#CHROM POS
            ID
                  REF
                               QUAL
                                     FILTER INFO FORMAT -e Chinook.p1.i0
                                                                           -e Chinook.p1.i1
                             173.69 .
chr_1 102
AC=4;AF=0.026;AN=156;BaseQRankSum=0.524;DP=209;ExcessHet=0.0860;FS=0.000;InbreedingCoeff=0.2
702;MLEAC=5;MLEAF=0.032;MQ=60.00;MQRankSum=0.00;QD=14.47;ReadPosRankSum=0.00;SOR=0.693
GT:AD:DP:GQ:PL 0/0:3,0:3:9:0,9,102
                            0/0:2,0:2:6
                                       AC=4
                                       AF=0.026
                                       AN=156
                                       BaseQRankSum=0.524
                                      DP=209
                                       ExcessHet=0.0860
                                       FS=0.000
                                       InbreedingCoeff=0.2702
                                      MLEAC=5
                                      MLEAF=0.032
                                      MQ = 60.00
                                      MQRankSum=0.00
```

QD=14.47

SOR = 0.693

Semi-colon separated data held the INFO field

ReadPosRankSum=0.00

# VCF: Records - INFO

FORMAT -e Chinook.p1.i0

FILTER INFO

```
AC=4;AF=0.026;AN=156;BaseQRankSum=0.524;DP=209;ExcessHet=0.0860;FS=0.000;InbreedingCoeff=0.2
702;MLEAC=5;MLEAF=0.032;MQ=60.00;MQRankSum=0.00;QD=14.47;ReadPosRankSum=0.00;SOR=0.693
GT:AD:DP:GQ:PL 0/0:3,0:3:9:0,9,102
                           0/0:2,0:2:6
            AC=4
            AF=0.026
            AN=156
            BaseQRankSum=0.524
            DP = 209
            ExcessHet=0.0860
            FS=0.000
            InbreedingCoeff=0.2702
            MLEAC=5
            MLEAF=0.032
            MQ = 60.00
            MQRankSum=0.00
            0D=14.47
```

#CHROM POS

102

chr\_1

ID

REF

 $\mathsf{C}$ 

ALT

Т

QUAL

173.69

##INFO=<ID=AC, Number=A, Type=Integer, Description="Allele count in a ##INFO=<ID=AF, Number=A, Type=Float, Description="Allele Frequency, f ##INFO=<ID=AN, Number=1, Type=Integer, Description="Total number of a ##INFO=<ID=BaseQRankSum, Number=1, Type=Float, Description="Z-score f ##INFO=<ID=DP, Number=1, Type=Integer, Description="Approximate read ##INFO=<ID=END, Number=1, Type=Integer, Description="Stop position of ##INFO=<ID=ExcessHet, Number=1, Type=Float, Description="Phred-scaled ##INFO=<ID=FS, Number=1, Type=Float, Description="Phred-scaled p-value" ##INFO=<ID=InbreedingCoeff, Number=1, Type=Float, Description="InbreedingCoeff", Numb per-sample when compared against the Hardy-Weinberg expectation"> ##INFO=<ID=MLEAC, Number=A, Type=Integer, Description="Maximum likeli the same as the AC), for each ALT allele, in the same order as lis ##INFO=<ID=MLEAF, Number=A, Type=Float, Description="Maximum likeliho the same as the AF), for each ALT allele, in the same order as lis ##INFO=<ID=MQ, Number=1, Type=Float, Description="RMS Mapping Quality ##INFO=<ID=MQRankSum, Number=1, Type=Float, Description="Z-score From ##INFO=<ID=QD, Number=1, Type=Float, Description="Variant Confidence/ ##INFO=<ID=RAW\_MQandDP, Number=2, Type=Integer, Description="Raw data Quality calculation. Incompatible with deprecated RAW\_MQ formulati ##INFO=<ID=ReadPosRankSum, Number=1, Type=Float, Description="Z-score bias"> ##INFO=<ID=SOR, Number=1, Type=Float, Description="Symmetric Odds Rat

-e Chinook.p1.i1

The Key to the INFO Field is in the header

ReadPosRankSum=0.00

SOR=0.693

# VCF: Records - FORMAT

#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT -e Chinook.p1.i0 -e Chinook.p1.i1 chr\_1 102 . C T 173.69 .

AC=4;AF=0.026;AN=156;BaseQRankSum=0.524;DP=209;ExcessHet=0.0860;FS=0.000;InbreedingCoeff=0.2702;MLEAC=5;MLEAF=0.032;MQ=60.00;MQ RankSum=0.00;QD=14.47;ReadPosRankSum=0.00;SOR=0.693 GT:AD:DP:GQ:PL 0/0:3,0:3:9:0,9,102 0/0:2,0:2:6

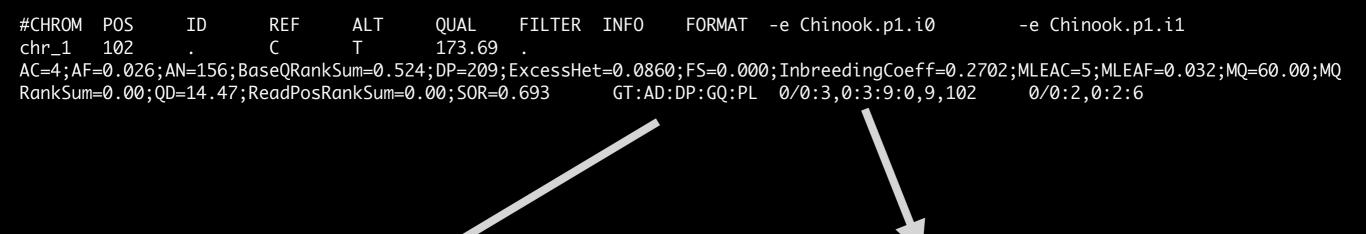
GT:AD:DP:GQ:PL

0/0:3,0:3:9:0,9,102

Colon separated key to the data in the column for each sample

Colon separated data for sample "Chinook.p1.i0"

# VCF: Records - FORMAT



GT:AD:DP:GQ:PL

0/0:3,0:3:9:0,9,102

Colon separated key to the data in the column for each sample

Colon separated data for sample "Chinook.p1.i0"

The Key to abbreviations in the FORMAT field is in the header

