



HaplotypeCaller Tutorial

This tutorial will show you how to run the gold-standard GATK variant caller, HaplotypeCaller, which takes your aligned output BAM from the FQ2BAM Tutorial, assembles plausible haplotypes from active regions, and identifies genotype likelihoods according to Bayes' Rule. The result is a Variant Call Format (VCF) file of all the variant calls across the genome, including their position in the genome and the allelic information.

To do this, run the following command:

```
$ docker run \ --gpus all \ --rm \ --volume $(pwd):/workdir \ --volume  
$(pwd):/outputdir \ nvcr.io/nvidia/clara/clara-parabricks:4.3.1-1 \ pbrun  
haplotypcaller \ --ref  
/workdir/parabricks_sample/Ref/Homo_sapiens_assembly38.fasta \ --in-bam  
/workdir/fq2bam_output.bam \ --out-variants /outputdir/variants.vcf [PB Info 2022-  
Sep-02 20:08:13] ----- [PB Info  
2022-Sep-02 20:08:13] || Parabricks accelerated Genomics Pipeline || [PB Info  
2022-Sep-02 20:08:13] || Version 4.0.0-1 || [PB Info 2022-Sep-02 20:08:13] || GPU-  
GATK4 HaplotypeCaller || [PB Info 2022-Sep-02 20:08:13] -----  
----- [PB Info 2022-Sep-02 20:08:48] 0  
/outputdir/fq2bam_output.bam/outputdir/variants.vcf [PB Info 2022-Sep-02  
20:08:48] ProgressMeter - Current-Locus Elapsed-Minutes Regions-Processed  
Regions/Minute [PB Info 2022-Sep-02 20:08:58] chr1:26179201 0.2 143431 860586  
[PB Info 2022-Sep-02 20:09:08] chr1:88982401 0.3 428734 1286202 [PB Info 2022-  
Sep-02 20:09:18] chr1:155342401 0.5 622031 1244062 [PB Info 2022-Sep-02  
20:09:28] chr1:222508801 0.7 921183 1381774 [PB Info 2022-Sep-02 20:09:38]  
chr2:38697601 0.8 1206544 1447852 [PB Info 2022-Sep-02 20:09:48]  
chr2:101697601 1.0 1469587 1469585 [PB Info 2022-Sep-02 20:09:58]  
chr2:176107201 1.2 1802198 1544741 [PB Info 2022-Sep-02 20:10:08] chr3:8990401  
1.3 2124644 1593483 [PB Info 2022-Sep-02 20:10:18] chr3:55473601 1.5 2349628  
1566418 [PB Info 2022-Sep-02 20:10:28] chr3:133680001 1.7 2671435 1602861 [PB  
Info 2022-Sep-02 20:10:38] chr3:192724801 1.8 2922628 1594160 [PB Info 2022-  
Sep-02 20:10:48] chr4:61934401 2.0 3222268 1611134 [PB Info 2022-Sep-02  
20:10:58] chr4:143524801 2.2 3573491 1649303 [PB Info 2022-Sep-02 20:11:08]  
chr5:38947201 2.3 3949926 1692825 [PB Info 2022-Sep-02 20:11:18]  
chr5:116337601 2.5 4255653 1702261 [PB Info 2022-Sep-02 20:11:28]  
chr5:176035201 2.7 4508018 1690506 [PB Info 2022-Sep-02 20:11:38]
```

```
chr6:54768001 2.8 4780747 1687322 [PB Info 2022-Sep-02 20:11:48]
chr6:152140801 3.0 5176363 1725454 [PB Info 2022-Sep-02 20:11:58]
chr7:27105601 3.2 5377117 1698036 [PB Info 2022-Sep-02 20:12:08]
chr7:105105601 3.3 5713863 1714158 [PB Info 2022-Sep-02 20:12:18]
chr8:18086401 3.5 6002114 1714889 [PB Info 2022-Sep-02 20:12:28] chr8:73915201
3.7 6267992 1709451 [PB Info 2022-Sep-02 20:12:38] chr9:5553601 3.8 6588617
1718769 [PB Info 2022-Sep-02 20:12:48] chr9:93672001 4.0 6886981 1721745 [PB
Info 2022-Sep-02 20:12:58] chr10:4094401 4.2 7095584 1702939 [PB Info 2022-Sep-
02 20:13:08] chr10:94593601 4.3 7489282 1728295 [PB Info 2022-Sep-02 20:13:18]
chr11:18398401 4.5 7757420 1723871 [PB Info 2022-Sep-02 20:13:28]
chr11:95976001 4.7 8083942 1732273 [PB Info 2022-Sep-02 20:13:38]
chr12:6652801 4.8 8282324 1713584 [PB Info 2022-Sep-02 20:13:48]
chr12:70632001 5.0 8560282 1712056 [PB Info 2022-Sep-02 20:13:58]
chr13:24446401 5.2 8860588 1714952 [PB Info 2022-Sep-02 20:14:08]
chr13:99038401 5.3 9200934 1725175 [PB Info 2022-Sep-02 20:14:18]
chr14:68467201 5.5 9480546 1723735 [PB Info 2022-Sep-02 20:14:28]
chr15:66432001 5.7 9821985 1733291 [PB Info 2022-Sep-02 20:14:38]
chr16:31128001 5.8 10123582 1735471 [PB Info 2022-Sep-02 20:14:48]
chr17:15782401 6.0 10402640 1733773 [PB Info 2022-Sep-02 20:14:58]
chr17:55262401 6.2 10553414 1711364 [PB Info 2022-Sep-02 20:15:08]
chr18:27960001 6.3 10790487 1703761 [PB Info 2022-Sep-02 20:15:18]
chr19:15883201 6.5 11074759 1703809 [PB Info 2022-Sep-02 20:15:28]
chr20:16108801 6.7 11311965 1696794 [PB Info 2022-Sep-02 20:15:38]
chr21:10468801 6.8 11563626 1692237 [PB Info 2022-Sep-02 20:15:48]
chr22:29289601 7.0 11766272 1680896 [PB Info 2022-Sep-02 20:15:58]
chrX:77452801 7.2 12197330 1701952 [PB Info 2022-Sep-02 20:16:08]
chrUn_JTFH01000876v1_decoy:1 7.3 12604293 1718767 [PB Info 2022-Sep-02
20:16:18] Total time taken: 450.258102 [PB Info 2022-Sep-02 20:16:18] -----
----- [PB Info 2022-Sep-02 20:16:18] ||
Program: GPU-GATK4 HaplotypeCaller || [PB Info 2022-Sep-02 20:16:18] || Version:
4.0.0-1 || [PB Info 2022-Sep-02 20:16:18] || Start Time: Fri Sep 2 20:08:13 2022 ||
[PB Info 2022-Sep-02 20:16:18] || End Time: Fri Sep 2 20:16:18 2022 || [PB Info
2022-Sep-02 20:16:18] || Total Time: 8 minutes 5 seconds || [PB Info 2022-Sep-02
20:16:18] -----
/usr/local/parabricks/binaries/bin/htvc
/workdir/parabricks_sample/Ref/Homo_sapiens_assembly38.fasta
```

```
/outputdir/fq2bam_output.bam 1 -o /outputdir/variants.vcf -nt 5 Please visit  
https://docs.nvidia.com/claras/#parabricks for detailed documentation
```

If you get an out-of-memory error make sure your computer has enough RAM, and that large amounts of memory aren't being used by other programs.

You should now have the following files in your directory:

```
$ ls -lrt -rw-r--r-- 1 root root 4819386804 Sep 2 15:58 fq2bam_output.bam -rw-r--r-- 1  
root root 6882792 Sep 2 15:59 fq2bam_output.bam.bai -rw-r--r-- 1 root root 87690  
Sep 2 15:59 fq2bam_output_chrs.txt -rw-r--r-- 1 root root 23643404 Sep 2 16:16  
variants.vcf (sample data not shown)
```

The first ten non-header lines of `variants.vcf` should be as follows:

```
chr1 16378 . T C 45.28 .  
AC=2;AF=1.00;AN=2;DP=2;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.500;MQ=23  
GT:AD:DP:GQ:PL 1/1:0,2:2:6:57,6,0 chr1 63268 . T C 43.28 .  
AC=2;AF=1.00;AN=2;DP=2;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.500;MQ=23  
GT:AD:DP:GQ:PL 1/1:0,2:2:6:55,6,0 chr1 63516 . A G 1202.03 .  
AC=2;AF=1.00;AN=2;DP=40;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=26  
GT:AD:DP:GQ:PL 1/1:0,40:40:99:1216,120,0 chr1 63527 . T C 1002.03 .  
AC=2;AF=1.00;AN=2;DP=33;ExcessHet=3.0103;FS=0.000;MLEAC=2;MLEAF=1.00;MQ=26  
GT:AD:DP:GQ:PL 1/1:0,33:33:99:1016,99,0 chr1 131609 . C A 83.60 .  
AC=1;AF=0.500;AN=2;BaseQRankSum=-0.275;DP=11;ExcessHet=3.0103;FS=0.000;MLEAF=0.500;MQ=23  
GT:AD:DP:GQ:PL 0/1:6,5:11:91:91,0,115 chr1 133483 . G T 70.60 .  
AC=1;AF=0.500;AN=2;BaseQRankSum=-1.534;DP=8;ExcessHet=3.0103;FS=3.332;MLEAC=1;MLEAF=0.500;MQ=23  
GT:AD:DP:GQ:PL 0/1:5,3:8:78:78,0,100 chr1 264627 . A G 37.28 .  
AC=2;AF=1.00;AN=2;DP=2;ExcessHet=3.0103;FS=0.000;MLEAC=1;MLEAF=0.500;MQ=49  
GT:AD:DP:GQ:PL 1/1:0,2:2:6:49,6,0 chr1 268012 . G A 330.60 .  
AC=1;AF=0.500;AN=2;BaseQRankSum=0.636;DP=116;ExcessHet=3.0103;FS=1.639;MLEAC=1;MLEAF=0.500;MQ=23  
GT:AD:DP:GQ:PL 0/1:93,23:116:99:338,0,3307 chr1 268130 . G T 2188.60 .  
AC=1;AF=0.500;AN=2;BaseQRankSum=1.543;DP=228;ExcessHet=3.0103;FS=8.632;MLEAC=1;MLEAF=0.500;MQ=23  
GT:AD:DP:GQ:PL 0/1:128,100:228:99:2196,0,4482 chr1 268516 . C T 59.60 .
```

AC=1;AF=0.500;AN=2;BaseQRankSum=1.465;DP=7;ExcessHet=3.0103;FS=3.680;MLEAC
GT:AD:DP:GQ:PL 0/1:5,2:7:67:67,0,138

© Copyright 2024, Nvidia.. PDF Generated on 06/05/2024