ICS 07.80
CCS A40

T/LTIA 15—2021

Human genomics' genotype imputation and variation interpretation in lowpass whole genome resequencing

2021 - 11 - 23 2021 - 11 - 30

II

1 1

2 1

3 1

4 1

5 1

6 2

7 2

8 3

A 4

B 5

C 6

7

GB/T 1.1—2020 1

1

2

T/LTIA 14-2021

3

T/LTIA 14-2021

4

bp base pair
dbSNP (the single nucleotide polymorphism database)
NCBI (national center for biotechnology information)
OMIM (online mendelian inheritance in man)
GRCh genome reference consortium human build
hg human reference genome
HGVS human genome variation society
GP (genotype posterior probability)

5

5.1

T/LTIA 14-2021 5

5.2

NCBI

1 NCBI GRCh hg GRCh hg GRCh38 hg38 GRCh37 hg19
2 2017 12 GRCh38[1] hg38[2]
3

5.3

1

1

	PE
	≥ 100bp
Q30	≥ 80%
GC	≤ 45%
	≤ 20%
	≥ 80%
1X 1X	≥ 60%
4X 1X	≥ 90%

6

6.1

T/LTIA 14-2021 7.1

6.2 VCF

VCF T/LTIA 14-2021 7.2

7

7.1

T/LTIA 14-2021 8 A

7.2

7.2.1 1000 Genomes Project

1000 Genomes Project 2015 Phase 3 [3]

7.2.2

EAS

7.3

7.3.1

Phase 3

NA12878 2014 (NIST) [4]

7.3.2

[4]

7.3.3 NA12878

7.3.4 2

2 NA12878 SNP

		GP ≥ 0.6	GP ≥ 0.9
SNP	1X	≥ 93%	≥ 93%
SNP	1X	≥ 80%	≥ 70%
SNP	4X	≥ 97%	≥ 97%
SNP	4X	≥ 95%	≥ 90%
1 Beagle[5] IMPUTE[6] Minimac[7] ( )			
2 GP GP 1 GP GP ≥ 0.6			
3 SNP SNP Indel			
4 Joint Calling Joint Calling			

8

8.1

B

8.2

8.2.1 SNP Indel HGVS

8.2.2 T/LTIA 14-2021 9.2.1

8.2.3 T/LTIA 14-2021 9.2.2

8.3

dbSNP[8] OMIM[9] ClinVar[10] gnomAD[11]

C

A

beagle5 VCF

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	NA12878
chr1	10519	.	G	A	.	PASS	DR2=0.01;AF=0.0067;IMP	GT:DS:GP	0|0:0.01:0.99,0.01,0
chr1	10563	.	C	A	.	PASS	DR2=0;AF=0;IMP	GT:DS:GP	0|0:0:1,0,0
chr1	10575	.	C	G	.	PASS	DR2=0;AF=0;IMP	GT:DS:GP	0|0:0:1,0,0
chr1	10582	.	T	C	.	PASS	DR2=0;AF=0;IMP	GT:DS:GP	0|0:0:1,0,0
chr1	10583	.	G	A	.	PASS	DR2=0.02;AF=0.0209;IMP	GT:DS:GP	0|0:0.04:0.96,0.04,0

B

bcfanno1.4 (https://github.com/shiquan/bcfanno)
(Clinvar OMIM) VCF

HGVS

OMIM

```
#CHROM  POS     ID      REF     ALT     QUAL    FILTER  INFO    FORMAT  NA12878
chr1    20604981        rs60369023      G       A       951.77  .
DP=2;AF=0.5;AN=2;AC=1;ExonIntron=E2;Gene=CDA;Transcript=NM_001785.2;HGVSnom=NM_001785.2:c.208G>A(p.A
la70Thr/p.A70T);AAlength=147;VarType=Missense;RS=60369023;g1000_AF=0.00139776;g1000_EAS_AF=0.005;gno
mad_exomes_AF=0.000252139;gnomad_exomes_AF_EAS=0.00301554;Gene_MIM_id=123920;Phenotypes=.;Possible_I
nheritance=NA;SIFT_pred=D;SIFT_score=0.002;LRT_pred=D;LRT_score=0.000000;MutationTaster_pred=D;Mutat
ionTaster_score=1;FATHMM_pred=T;FATHMM_score=-0.53  GT  0/1
```

C

C.1

C.1.1 dbSNP https://www.ncbi.nlm.nih.gov/projects/SNP

C.1.2 gnomAD https://gnomad.broadinstitute.org/

C.2

C.2.1 OMIM https://www.ncbi.nlm.nih.gov/omim

C.2.2 ClinVar , https://www.ncbi.nlm.nih.gov/clinvar