SIFT Missense Predictions for Genomes and 1000 Genomes Data

March 26, 2017

1 Title: SIFT Missense Predictions for Genomes and 1000 Genomes Data

2 Course: MED263, "Bioinformatics Applications To Human Disease"

3 Preparer: Nathaniel Delos Santos

4 1) Introduction

In this practical, you are going to use linux command line tools, the SIFT4G Variant annotator, and data from the 1000 Genomes Project to predict deleterious missense mutations from human samples. Predicting variant deleteriousness is an important part of analyzing human genome variants in disease, because it provides insight into which genes have been affected by a variant, and how bad the effect might be.
After this tutorial, you should be able to:
- Download variant information from the Ensembl project - Download aligned sequence data from the 1000 Genomes project - Call variants from aligned sequence data - Annotate variants from aligned sequence data with deleteriousness and amino acid change predictions - Prioritize variants by a deleteriousness score - Perform data manipulation with basic command line tools in BASH

4.1 1.1) Download SIFT 4G

SIFT 4G, (Sorting Intolerant From Tolerant, For Genomes) uses variant calls to predict what amino acid substitutions occur, and how deleterious they are. SIFT 4G requires Java and a reference database to run.

We will download SIFT 4G directly from their website at (http://sift.bii.a-star.edu.sg/sift4g/) using wget. Make sure wget is installed on your system.

```
In [1]: %%bash
        wget -q http://sift.bii.a-star.edu.sg/sift4g/SIFT4G_Annotator_v2.4.jar
```

4.2 1.2) Download Homo Sapiens Database (GRCh38.78) for SIFTG

We must download the reference database for GRCh38, the newest version of the human genome reference available from Ensembl. We will download SIFT 4G's version of this database directly from their website. Make sure to choose GRCh38.78. Decompress the

```
In [2]: %%bash
        wget -q http://sift.bii.a-star.edu.sg/sift4g/public/Homo_sapiens/GRCh38.78.
```

```
In [ ]: %%bash
        unzip GRCh38.78.zip
```

4.3 1.3) SAMTools

SAMTools is a general toolkit for use with aligned sequencing data. We will use it here to call variants from sequence alignments, using the 'samtools mpileup' command. We will install version 1.4 here, since the specific version matters for our purposes. Make sure that GCC and your build environment are up to date.

```
In [4]: %%bash
        wget -q https://github.com/samtools/samtools/releases/download/1.4/samtools
```

```
In [ ]: %%bash
        tar -vxjf samtools-1.4.tar.bz2
        cd samtools-1.4
        ./configure
        make
        cd ..
```

4.4 1.4) BCFTools

BCFTools is a general toolkit for use with variant call format (VCF) files. We will use it here to filter and query variants. We install version 1.4 here as we did for SAMTools

```
In [6]: %%bash
        wget -q https://github.com/samtools/bcftools/releases/download/1.4/bcftools
```

```
In [ ]: %%bash
        tar -vxjf bcftools-1.4.tar.bz2
        cd bcftools-1.4
        ./configure
        make
        cd ..
```

5 2) Data

5.1 2.1) Craig Venter Germline Variations

Craig Venter's genome was among the first sequenced. These Variant Call Format (VCF) files summarize the variants observed in his genome from the GRCh38.78 reference.

```
In [8]: %%bash
        wget -q http://ftp.ensembl.org/pub/release-78/variation/vcf/homo_sapiens/Ve
        wget -q http://ftp.ensembl.org/pub/release-78/variation/vcf/homo_sapiens/Ve
```

5.2 Question 1)

How many variants are in the Venter VCF?

5.3 Answer 1)

```
In [9]: %%bash
        zcat Venter.vcf.gz|grep -v '#'|wc -l
```

```
3266109
```

3266109 Variants

5.4 2.2) James Watson Germline Variations

James Watson is famous for discovering the double helix structure of DNA with Francis Crick. He has his own tribute in VCF format here.

```
In [10]: %%bash
         wget -q http://ftp.ensembl.org/pub/release-78/variation/vcf/homo_sapiens/W
         wget -q http://ftp.ensembl.org/pub/release-78/variation/vcf/homo_sapiens/W
```

5.5 2.3) 1000 Genomes human sample exome data

The 1000 Genomes project was an international effort to catalog most variants with more than 1% frequency in the human population. It is a valuable source of human sequencing data. We will not be using the VCFs directly, but instead will be analyzing aligned sequences from a single human sample.

5.5.1 2.3.1) CRAM files

CRAM files are compressed sequence alignment files that use delta compression from a reference to store sequence information, rather than containining the sequence data themselves. Therefore, we must download the CRAM file, CRAM index, and the corresponding reference files to use them.

```
In [11]: %%bash
         wget -q ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/data_collections/1000_ger
         wget -q ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/data_collections/1000_ger
```

The reference files for the CRAM file are downloaded below

```
In [12]: %%bash
         wget -q ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/GRCh3
         wget -q ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/technical/reference/GRCh3
```

5.6 Question 2)

From the README provided by the 1000 Genomes Project (ftp://ftp.1000genomes.ebi.ac.uk/vol1/ftp/data_collections/1000_genomes_project/README.1000genomes.G , what steps have already been performed for these CRAM files to make them ready for analysis?

5.7 Answer 2)

1. Read alignment
2. Local realignment around Indels
3. Recalibration of base quality scores
4. Marking of duplicate reads
5. Merging multiple sequencing libraries into a single sample alignment file
6. Lossless compression using CRAM

6 3) Methods/Results

We will now run SIFT4G to predict the deleteriousness of variants found in the Venter VCF.

6.1 3.1) Analysis of Craig Venter germline variants

First we must decompress the gzipped VCF to an uncompressed VCF using zcat.

SIFT4G is run using java, so we must call it using 'java -jar', passing the SIFT4G program as the '-jar' option. The '-c' option will run SIFT4G in command line mode, and the '-t' option will cause SIFT4G to output additional annotations for each transcript of a gene affected. The '-i' option specifies the input VCF, in this case 'Venter.vcf'. The '-d' option specifies the database we will be using, in this case the GRCh38.78 database. The '-r' option will determine where the results of the SIFT annotation will be located relative to our current directory.

```
In [13]: %%bash
         zcat Venter.vcf.gz > Venter.vcf
         java -jar SIFT4G_Annotator_v2.4.jar -c -t -i Venter.vcf -d GRCh38.78  -r V

Start Time for SIFT4G code: Mon Mar 27 01:34:17 PDT 2017
Updates:
No updates from server!! Please go to http:sift-dna.org for updates.

Started Running ...
Running in Multitranscripts mode

Chromosome        WithSIFT4GAnnotations        WithoutSIFT4GAnnotations        Prog
MT                          1                           0                    Complet
Y                         130                       18578                        Co
22                         2914                       22450
20                         5175                       66630
13                         9175                      116297
21                         7318                       59106
X                          3941                       79652                       C
```

```
10                          10038                          163272
9                          12656                          155721
19                          9448                          57848
18                          11794                          79902
7                          10687                          138041
17                          8831                          79089
16                          13997                          99378
14                          12880                          86113
6                          9155                          147054
11                          15317                          153409
15                          16910                          83760
12                          17570                          132459
8                          18309                          155222
4                          24268                          224475
5                          18220                          144711
3                          20600                          205453
1                          23661                          239132
2                          30376                          244986

Merging temp files...
SIFT4G Annotation completed !
Output directory:Venter.SIFT4G
End Time for parallel code: Mon Mar 27 01:43:35 PDT 2017
```

6.2 Question 3)

On Chromosome 17, how many variants are annotated? How many are unnannotated?

6.3 Answer 3)

8831 annotated, 79089 unnannotated

6.3.1 3.1.1 SIFT 4G Output

The output of SIFT 4G includes a VCF file and an excel (.xls) file that describe the amino acid changes and the predicted deleteriousness of each variant. The excel file is formatted similarly to a tab-separated values file, with the exception of a carriage return (‘^M’) before each new line. We will use this to navigate the SIFT 4G output.

6.4 Question 4)

How many columns Does the SIFT4G output contain? What does each column contain?

6.5 Answer 4)

```
In [14]: %%bash
         cat Venter.SIFT4G/Venter_SIFTannotations.xls|head -n1|tr '\t' '\n'|cat -n
```

```
 1        CHROM
 2        POS
 3        REF_ALLELE
 4        ALT_ALLELE
 5        TRANSCRIPT_ID
 6        GENE_ID
 7        GENE_NAME
 8        REGION
 9        VARIANT_TYPE
10        REF_AMINO
11        ALT_AMINO
12        AMINO_POS
13        SIFT_SCORE
14        SIFT_MEDIAN
15        NUM_SEQS
16        dbSNP
17        SIFT_PREDICTION
```

17 Columns, contents are described above.

6.6 Question 5)

How many deleterious (not 'Low confidence') variants are found from these variants?

6.7 Answer 5)

```
In [15]: %%bash
         cat Venter.SIFT4G/Venter_SIFTannotations.xls|tail -n+2 \
         |grep 'DELETERIOUS'|grep -v 'Low confidence'|cut -f1,2,3,4 \
         |sort|uniq|wc -l
```

```
1561
```

1561 deleterious variants.

6.8 Question 6)

How many genes have deleterious variants? Output the list of genes names into a file. Display the gene names.

6.9 Answer 6)

```
In [16]: %%bash
         cat Venter.SIFT4G/Venter_SIFTannotations.xls|tail -n+2 \
         |grep 'DELETERIOUS'|grep -v 'Low confidence'|cut -f7 \
         |sort|uniq \
         > Venter.SIFT4G.genes_with_deleterious_variants.txt
```

```
wc -l Venter.SIFT4G.genes_with_deleterious_variants.txt
cat Venter.SIFT4G.genes_with_deleterious_variants.txt
```

```
1186 Venter.SIFT4G.genes_with_deleterious_variants.txt
A2ML1
ABCA10
ABCA6
ABCA7
ABCC8
ABCD1
AC008686.1
AC073657.1
ACACB
ACADS
ACAN
ACSM2A
ACTL9
ACTN3
ACTRT2
ADAM19
ADAM7
ADAMDEC1
ADAMTS13
ADAMTS14
ADAMTSL3
ADCK5
ADH1C
ADH4
ADORA3
ADPGK
AGAP10
AGT
AHCY
AHNAK
AIM1L
AKAP13
AKR1C2
ALDH1B1
ALDH5A1
ALOX15
ALPK2
ALPK3
ALPP
ALX4
AMACR
AMPD3
AMY2A
ANAPC1
```

ANKDD1A
ANKLE1
ANKRD1
ANKRD30A
ANKRD33
ANKRD35
ANKRD36
ANKRD36B
ANKRD36C
ANKRD60
ANKRD62
ANXA13
AOAH
AP1G2
AP2S1
APIP
APOA1BP
APOB
APOBEC3H
APOC4
APOC4-APOC2
APOL4
AQP12A
AQP7
ARHGAP9
ARHGEF19
ARHGEF3
ARHGEF37
ARL5C
ARMC4
ARMC9
ARPP21
ARR3
ASB16
ATAD3C
ATF7IP
ATP12A
ATP6V1C2
ATP8B3
ATP8B4
AVPR1B
B3GNT3
BAG3
BANK1
BCL9
BMP2
BMP3
BPIFB2

BRCA1
BTBD16
BTD
BTK
BTN1A1
BTN2A1
C14orf37
C16orf46
C19orf53
C1orf158
C1orf167
C1orf177
C1orf194
C1orf87
C1QTNF9B
C2orf16
C2orf61
C2orf70
C2orf73
C2orf74
C4B
C4orf33
C5orf45
C7
C7orf31
C7orf57
C7orf72
C9orf114
CABIN1
CACNA1S
CACNB2
CACNG6
CALCOCO2
CALR
CAPN5
CAPN8
CAPN9
CAPS2
CAPZA3
CASC1
CASC5
CBLC
CBS
CBWD3
CBWD5
CBWD7
CCDC130
CCDC14

CCDC141
CCDC157
CCDC178
CCDC18
CCDC181
CCDC36
CCDC40
CCDC57
CCDC6
CCDC64B
CCDC74B
CCDC83
CCDC93
CCNL2
CCP110
CCR5
CCSER2
CCT6B
CD101
CD163
CD200R1
CD27
CD300LF
CD3G
CD6
CDAN1
CDC25C
CDC6
CDH11
CDHR2
CDK11A
CDK11B
CDK5RAP2
CECR2
CEL
CELSR2
CENPQ
CEP120
CEP170
CEP89
CES5A
CFAP69
CFAP74
CFHR4
CHFR
CHI3L1
CHIA
CHIT1

CHMP4A
CHRNA3
CHRNB4
CLCN2
CLCNKA
CLEC17A
CLEC4M
CLIP1
CNDP2
CNTNAP3
CNTNAP3B
CNTRL
CNTROB
COASY
COL12A1
COL15A1
COL17A1
COL23A1
COL24A1
COL2A1
COL4A2
COL4A3
COL4A4
COL6A2
COL6A5
COL6A6
COL9A3
COMT
COQ7
COX10
CPAMD8
CPNE6
CPS1
CR1L
CRIM1
CRYBG3
CRYGB
CRYZL1
CSF2RA
CSRP2BP
CTNNAL1
CTSE
CTSH
CUBN
CWH43
CYFIP1
CYLC2
CYP11B1

CYP2A7
CYP2B6
CYP2C18
CYP2F1
CYP3A43
CYP4B1
DAPL1
DCLK3
DCLRE1C
DDIAS
DDRGK1
DDX58
DDX60L
DEAF1
DEFB128
DFNB31
DGKB
DHDH
DHRS4
DHTKD1
DHX34
DHX35
DHX37
DISC1
DLAT
DLEC1
DLGAP2
DLL3
DMD
DNAAF2
DNAH1
DNAH14
DOCK10
DOCK6
DOCK8
DPYD
DPYSL2
DSEL
DSG1
DSP
DUSP27
DYX1C1
ECE2
ECH1
ECHDC3
EDA2R
EDN3
EFCAB8

EFS
EHBP1L1
EIF5AL1
ELAVL1
ELMO3
ELTD1
EML2
EMP2
EMR1
EMR2
EMR3
ENGASE
EPHA10
EPHA8
EPHX1
EPHX2
EPN3
EPPK1
EPS8L1
ERBB3
ERCC4
ERCC5
ERICH6B
ERP27
ESAM
ESPL1
EVC2
EVI5
EXOC3L1
EYA3
FABP6
FADS6
FAM115C
FAM120A
FAM135A
FAM155B
FAM171A2
FAM178A
FAM181B
FAM186A
FAM187A
FAM188B
FAM205A
FAM214A
FAM220A
FAM26F
FAM35A
FAM53C

FAM83E
FAM86B1
FAM86B2
FANCI
FANK1
FASTKD2
FAT2
FAXDC2
FBLIM1
FBN3
FBXW10
FBXW8
FCHSD1
FHDC1
FLJ22184
FLNC
FLVCR1
FOXD4
FPR1
FRAS1
FREM1
FREM2
FRG1
FRG1B
FRG2C
FRMD4B
FSIP1
FSIP2
FUT2
FUT3
FUT9
FYCO1
GABRA4
GAD2
GAGE1
GAGE12J
GAL3ST1
GALNT8
GALNTL5
GALP
GBA
GBGT1
GBP1
GBP3
GBP6
GCAT
GDPD4
GDPGP1

GEMIN4
GFY
GGT2
GGT6
GIMAP6
GIPR
GJB7
GLYATL3
GNA12
GOLGA6C
GOLGA6L2
GOLGA6L4
GOLGA8H
GOLGA8R
GORAB
GPAA1
GPR111
GPR112
GPR137C
GPR144
GPR98
GPRIN2
GPX4
GRAMD2
GRB14
GRIN3A
GRXCR2
GSG2
GSTA5
GTF3C1
GUCA1C
GUCY2F
HADHA
HEATR2
HEATR5A
HELZ2
HHAT
HIBCH
HIGD1B
HIST1H1A
HLA-C
HMCN2
HMGXB4
HNRNPCL1
HNRNPCL2
HOXB1
HPS4
HRNR

HSD17B14
HSDL1
HSPG2
HTR3B
HTR3D
HUS1B
HYDIN
ICAM1
IDH3A
IDO2
IFITM3
IFT81
IFT88
IGF2R
IGSF10
IGSF9
IKBKAP
IL12B
IL12RB1
IL17F
IL18R1
IL1RL2
ILDR2
INMT
INVS
IP6K2
IQCF6
IQGAP3
IQUB
IRAK2
ISM2
ITGA10
ITGA11
ITGA2B
ITGA9
ITGAE
ITGB4
ITIH1
ITIH4
ITIH6
ITPR2
JAG2
JMJD1C
KCNAB2
KCNE1
KCNJ12
KCNK4
KDELR3

KDR
KIAA0226
KIAA0753
KIAA1524
KIAA1549
KIAA1551
KIAA1755
KIF27
KIF2C
KIF9
KL
KLHDC1
KLHDC7A
KLHL38
KLRB1
KLRC3
KRAS
KRI1
KRT13
KRT26
KRT27
KRT3
KRT32
KRT33B
KRT35
KRT40
KRT5
KRT6A
KRT6C
KRT72
KRT74
KRT76
KRT77
KRT83
KRTAP10-1
KRTAP10-11
KRTAP10-3
KRTAP10-4
KRTAP10-5
KRTAP1-1
KRTAP12-3
KRTAP1-5
KRTAP15-1
KRTAP29-1
KRTAP4-1
KRTAP4-3
KRTAP4-4
KRTAP4-6

KRTAP4-8
KRTAP5-9
KRTAP9-2
KRTAP9-4
KRTAP9-6
KRTAP9-8
LAMC3
LARP1B
LBP
LCE3D
LCE5A
LECT1
LETMD1
LHX8
LIG1
LIG3
LIG4
LIPF
LIPT2
LMO7
LPCAT1
LRIT2
LRMP
LRP2
LRRC25
LRRC34
LRRC37A
LRRC6
LRRN4
LTBP3
LYSMD4
MADCAM1
MAF1
MAGEA3
MAGEB16
MAGEC1
MAGEF1
MALRD1
MAP1A
MAP1S
MAP2K3
MAP2K5
MAP7
MAP9
MASP2
MATN2
MBD1
MCEE

MCF2L2
MCM2
MCPH1
MEGF6
MEP1A
MEP1B
MERTK
MGA
MGST1
MKI67
MMS22L
MPHOSPH10
MRC2
MRGPRX4
MROH7
MROH7-TTC4
MRPS7
MS4A14
MS4A6E
MSMB
MSRA
MST1
MST1R
MT1A
MTCH2
MTMR1
MTNR1B
MTR
MTRR
MTUS2
MUC12
MUC15
MUC16
MUC4
MUC5AC
MUS81
MXRA5
MYBPC2
MYCBPAP
MYH15
MYH3
MYH4
MYH6
MYH7B
MYH8
MYO18A
MYOM3
MYPN

NAAA
NAALADL2
NBPF1
NBPF10
NBPF11
NBPF15
NBPF3
NBPF9
NCAPG
NCKAP5
NCR2
NDUFA10
NEK11
NELL1
NEMF
NFATC1
NGRN
NHLRC1
NINL
NIPA1
NIPA2
NIPAL1
NIPSNAP3A
NLRP13
NLRP14
NME6
NOD1
NOP14
NOTCH2NL
NOTCH3
NOX5
NPHP4
NPIPA5
NPIPB11
NPIPB15
NPIPB4
NPIPB6
NPLOC4
NPPC
NPY4R
NRBP1
NRG1
NT5C3B
NTMT1
NTSR1
NUP160
NUPL1
NUTM2B

NUTM2D
NXN
NXPE1
OAS3
OBSCN
OBSL1
OLFM2
OLFML1
OMA1
OR10A2
OR10A6
OR10G3
OR10H1
OR10H3
OR10J1
OR10J5
OR10R2
OR10Z1
OR11G2
OR11H12
OR11H2
OR11H6
OR11L1
OR13D1
OR13G1
OR13J1
OR14C36
OR1A2
OR1D5
OR1E1
OR1I1
OR1L4
OR1L6
OR1N2
OR1Q1
OR1S2
OR2A2
OR2AE1
OR2AG2
OR2AK2
OR2B11
OR2B2
OR2C1
OR2F2
OR2G2
OR2G6
OR2L8
OR2M7

OR2T12
OR2T27
OR2T29
OR2T33
OR2T5
OR2T7
OR2T8
OR3A2
OR3A3
OR4A16
OR4A5
OR4B1
OR4C11
OR4C3
OR4C46
OR4C5
OR4D6
OR4E2
OR4F17
OR4K1
OR4K14
OR4L1
OR4M1
OR4M2
OR4N4
OR4S1
OR51A4
OR51B2
OR51B6
OR51F2
OR51G1
OR51I1
OR51J1
OR51M1
OR51Q1
OR51S1
OR52D1
OR52E2
OR52J3
OR52N1
OR52R1
OR52W1
OR56B1
OR5A1
OR5AK2
OR5AU1
OR5B2
OR5B3

OR5D14
OR5D16
OR5F1
OR5H15
OR5H6
OR5I1
OR5K3
OR5R1
OR6B3
OR6M1
OR6N1
OR7A10
OR7E24
OR7G1
OR8B3
OR8D1
OR8G5
OR8H2
OR8J1
OR8K1
OR8S1
OR9G1
OR9Q2
OTOP2
OTOR
OVCH1
P2RY2
PADI4
PALD1
PAPLN
PAPPA
PARM1
PARP12
PARP14
PARP15
PARS2
PATE1
PCDH15
PCDHA1
PCDHA3
PCDHB12
PCDHB13
PCDHB7
PCDHB8
PCNT
PCNXL3
PCSK4
PDE12

PDE4DIP
PDLIM5
PDP2
PDZD8
PDZRN4
PER3
PEX11B
PFKFB3
PIAS3
PIEZO1
PIGC
PIK3C2G
PITRM1
PKD1L3
PKHD1L1
PLA1A
PLAUR
PLBD1
PLCB3
PLCL1
PLCZ1
PLEKHG3
PLEKHG4B
PLEKHH1
PLEKHM3
PLET1
PLIN5
PLXNA2
PM20D1
PMS2
POLR2J3
POM121
POM121C
POM121L2
PON2
POTEB3
POTED
POTEE
POTEM
POU5F1B
PPA2
PPARG
PPEF2
PPIAL4B
PPIAL4D
PPIAL4G
PPIP5K1
PPM1F

PPOX
PRAMEF1
PRAMEF11
PRAMEF14
PRAMEF18
PRAMEF26
PRAMEF4
PRAMEF9
PRB4
PRDM15
PRDM7
PRIM2
PRKAG3
PRMT7
PRODH
PRR14
PRRC2C
PRRT4
PRSS48
PRUNE2
PSG5
PSG8
PSMB11
PSMB4
PSMD13
PSMF1
PSMG1
PTGER3
PTGES3L
PTPLA
PTPN20A
PTPRB
PTPRH
PTPRQ
PTX4
QRFPR
RAB11FIP1
RAB2A
RABL6
RAD51C
RASAL1
RBM19
RBMX
RBP3
RD3L
REPIN1
REXO1
RFPL1

RFPL2
RGPD3
RGS12
RGS9
RHBG
RHCE
RHD
RHOT2
RHPN1
RICTOR
RIMBP3B
RIPK2
RNF115
RNF213
RNF43
RP1
RP11-400G3.5
RP11-507M3.1
RP11-545J16.1
RP11-697E2.6
RP1L1
RP4-576H24.4
RPL28
RREB1
RSPH10B2
RSPO1
RTN4
RTP5
RTTN
SACS
SCLT1
SCNN1A
SDHA
SDK2
SEC23B
SEMA4D
SEMA4G
SENP5
SEPN1
SEPT4
SERPINB12
SERPINB8
SERPINF1
SGK223
SH2D4B
SHARPIN
SHFM1
SHMT1

SIGLEC5
SIRPB1
SLC16A8
SLC22A10
SLC22A24
SLC22A4
SLC24A1
SLC25A45
SLC25A47
SLC26A6
SLC39A8
SLC52A1
SLCO1B3
SLCO1B7
SLFN5
SLIT3
SMPDL3B
SMYD4
SNTG2
SON
SOS2
SPATA3
SPATA31A1
SPATA31A6
SPATA31E1
SPATA33
SPEM1
SPIN2A
SPINK5
SPINT2
SPTA1
SRGAP2B
SRP14
SSX5
ST18
STAB2
STEAP1B
STEAP2
STK31
STK36
STON1
STON1-GTF2A1L
STX2
STXBP5L
SULT1C3
SUN1
SUPV3L1
SVOPL

SYCP2
SYNE1
SYNE2
SYPL1
SYT8
TACC2
TACR2
TAF1
TAS2R4
TBC1D28
TBL3
TBX10
TCEB3B
TCEB3C
TCEB3CL
TCF7
TDRD6
TEKT4
TEKT5
TEX13A
TG
TGOLN2
TGS1
THOC1
TICRR
TIMELESS
TIMM23
TLE4
TLR3
TLR5
TMBIM1
TMEM106C
TMEM161A
TMEM171
TMEM185B
TMEM244
TMIGD2
TMPRSS15
TMPRSS2
TMPRSS9
TNK1
TNKS1BP1
TNN
TNP2
TOP1MT
TPRX1
TPSAB1
TPTE

TRAPPC12
TRIM16
TRIM22
TRIM43
TRIM51
TRIM64
TRIOBP
TRNT1
TRPM8
TRPT1
TSEN54
TSKU
TSPAN8
TSPY4
TTC21B
TTC22
TTC24
TTC26
TTC27
TTC30B
TTC6
TTI2
TTLL4
TTN
TUBA3E
TUBB8
TYW1B
UBAP2
UBR1
UCK1
UGT1A6
UGT2A1
UGT2A2
UGT2B28
UGT2B4
UHRF1BP1
UNC5C
UNC93A
URAD
USP17L15
USP17L24
USP17L25
USP17L26
USP17L27
USP17L28
USP17L29
USP17L30
USP17L5

USP36
USP8
UTP20
VCAN
VCX
VCX2
VRK2
VWA5B1
VWDE
WASH4P
WBSCR28
WDR20
WDR49
WDR87
WDR91
WDYHV1
WNK2
WRNIP1
XYLT1
XYLT2
YAF2
ZAN
ZBBX
ZBTB5
ZC3H3
ZDBF2
ZFR2
ZKSCAN7
ZNF114
ZNF117
ZNF131
ZNF155
ZNF177
ZNF180
ZNF19
ZNF208
ZNF211
ZNF214
ZNF221
ZNF229
ZNF239
ZNF28
ZNF30
ZNF33A
ZNF404
ZNF415
ZNF417
ZNF436

ZNF443
ZNF45
ZNF493
ZNF534
ZNF540
ZNF541
ZNF543
ZNF544
ZNF549
ZNF559-ZNF177
ZNF568
ZNF57
ZNF571
ZNF573
ZNF578
ZNF607
ZNF611
ZNF626
ZNF658
ZNF667
ZNF675
ZNF676
ZNF679
ZNF700
ZNF705A
ZNF717
ZNF728
ZNF736
ZNF737
ZNF761
ZNF799
ZNF804B
ZNF880
ZNF98
ZNF99
ZPBP2
ZSCAN5A
ZSCAN5D

1186 genes. Gene names listed above.

6.9.1 3.1.2) SIFT Scores

SIFT scores less than 0.05 are considered deleterious. Anything greater is considered tolerated. Lower SIFT scores are considered more deleterious.

6.10 Question 7)

What is the lowest SIFT score of the deleterious variants?

6.11 Answer 7)

```
In [17]: %%bash
         cat Venter.SIFT4G/Venter_SIFTannotations.xls|tail -n+2 \
         |grep 'DELETERIOUS'|grep -v 'Low confidence' \
         |cut -f1,2,3,4,13 \
         |sort|uniq \
         |sort -k1,1 -k2,2n \
         |sort -k5,5n \
         |head

10        122336645        A        G        0.000
10        125980182        C        T        0.000
10        128113592        C        G        0.000
10        26219214        C        A        0.000
10        46461688        A        C        0.000
10        46549695        C        G        0.000
10        46549695        C        T        0.000
10        48086        G        A        0.000
10        59792934        G        T        0.000
10        6224537        G        T        0.000
```

0.0 is the lowest SIFT score.

6.12 Question 8)

What variants are annotated with the lowest SIFT score? Output the chromosome, coordinate, reference base, alternate base, gene name, reference amino acid, alternate amino acid, amino acid position, and sift score into a file. Display the first 10 lines of this file.

6.13 Answer 8)

```
In [18]: %%bash
         cat Venter.SIFT4G/Venter_SIFTannotations.xls|cut -f1,2,3,4,7,10,11,12,13,1
         |grep '^CHROM\|DELETERIOUS'|grep -v 'Low confidence' \
         |awk '($9==0.0)||$1=="CHROM"' \
         > Venter.SIFT4G.sift_score_0.txt
         head -n10 Venter.SIFT4G.sift_score_0.txt

CHROM        POS        REF_ALLELE        ALT_ALLELE        GENE_NAME        REF_AM
1        1956754        C        A        CFAP74        G        C        628
1        3497541        C        T        MEGF6        G        R        1152
1        11789390        A        G        C1orf167        R        G        810
1        17334004        G        C        PADI4        G        A        112
```

```
1        25321889        G        C        RHD        G        A        385
1        54670856        T        C        MROH7-TTC4        V        A        534
1        54801124        G        C        TTC22        L        V        14
1        54801124        G        C        TTC22        L        V        14
1        120889909        T        G        PPIAL4B        L        R        30
```

6.14 3.2) Analysis of James Watson germline variants

```
In [19]: %%bash
         zcat Watson.vcf.gz > Watson.vcf
         java -jar SIFT4G_Annotator_v2.4.jar -c -t -i Watson.vcf -d GRCh38.78  -r W

Start Time for SIFT4G code: Mon Mar 27 01:43:46 PDT 2017
Updates:
No updates from server!! Please go to http:sift-dna.org for updates.

Started Running ...
Running in Multitranscripts mode

Chromosome        WithSIFT4GAnnotations        WithoutSIFT4GAnnotations        Prog
MT                        0                        1                        Complet
Y                        119                        20889                        Co
22                        3226                        25060
20                        5293                        69584
13                        10529                        120358
21                        6616                        52355
X                        4172                        70126                        C
10                        10105                        166368
9                        9591                        135019
18                        11786                        83503
19                        9681                        67255
17                        8476                        82316
7                        13094                        182383
16                        13877                        101925
14                        13495                        90997
6                        11816                        193121
11                        15692                        165100
15                        16943                        82199
12                        17819                        145339
8                        18021                        156492
4                        25496                        219125
3                        20838                        223465
5                        21886                        184437
1                        21091                        243196
2                        31034                        255084

Merging temp files...
```

```
SIFT4G Annotation completed !
Output directory:Watson.SIFT4G
End Time for parallel code: Mon Mar 27 01:52:23 PDT 2017
```

6.15 Question 9)

On Chromosome 17, how many variants are annotated? How many are unnannotated?

6.16 Answer 9)

8476 annotated, 82316 unnannotated

6.17 Question 10)

How many deleterious (not 'Low confidence') variants are found from these variants?

6.18 Answer 10)

```
In [20]: %%bash
         cat Watson.SIFT4G/Watson_SIFTannotations.xls|tail -n+2 \
         |grep 'DELETERIOUS'|grep -v 'Low confidence'|cut -f1,2,3,4 \
         |sort|uniq|wc -l
```

```
1970
```

1970 deleterious variants.

6.19 Question 11)

How many genes have deleterious variants? Output the list of genes names into a file. Display the gene names.

6.20 Answer 11)

```
In [21]: %%bash
         cat Watson.SIFT4G/Watson_SIFTannotations.xls|tail -n+2 \
         |grep 'DELETERIOUS'|grep -v 'Low confidence'|cut -f7 \
         |sort|uniq \
         > Watson.SIFT4G.genes_with_deleterious_variants.txt
         wc -l Watson.SIFT4G.genes_with_deleterious_variants.txt
         cat Watson.SIFT4G.genes_with_deleterious_variants.txt
```

```
1528 Watson.SIFT4G.genes_with_deleterious_variants.txt
A2ML1
AADACL3
AASDHPPT
```

ABCA5
ABCA9
ABCB5
ABCC10
ABCC11
ABCC8
ABCC9
ABCG8
ABHD17B
ABL1
ABO
AC073657.1
ACACB
ACAN
ACAT1
ACOT4
ACP5
ACSM4
ACSS1
ADA
ADAM15
ADAM21
ADAM30
ADAMTS16
ADAMTS17
ADAMTS18
ADAMTS7
ADAMTSL3
ADD2
ADH1A
ADH1C
ADH5
ADNP
ADRBK1
AGAP2
AGBL1
AGBL2
AGPAT3
AGPAT9
AHCTF1
AHNAK
AHNAK2
AK9
AKAP13
AKAP3
AKR1C2
ALDH1B1
ALDH1L1

ALDH3B2
ALG9
ALOX5AP
ALPK1
ALPK2
ALPK3
ALPP
ALX4
AMACR
AMICA1
AMIGO2
AMPD3
ANGEL1
ANK1
ANKK1
ANKLE1
ANKLE2
ANKRD12
ANKRD30A
ANKRD33
ANKRD35
ANKRD36
ANKRD36C
ANKRD60
ANKS1A
ANKS3
ANO1
ANO10
ANP32E
ANTXRL
ANXA13
AP4S1
APOA1BP
APOA5
APOB
APOBEC2
APOL1
APOL4
APPL1
AQP7
AQR
ARCN1
ARFGEF1
ARHGAP17
ARHGAP19
ARHGAP19-SLIT1
ARHGEF19
ARHGEF28

ARHGEF37
ARID2
ARID4B
ARMC9
ARPP21
ART1
ASAP2
ASB16
ASH1L
ASMTL
ASNA1
ASPG
ASPM
ASPSCR1
ATAD5
ATF7IP
ATG2B
ATG9B
ATP13A4
ATP6V0D1
ATP7B
ATP8B4
ATPAF2
ATXN1
AVEN
AVIL
AVPR1B
BAG3
BARD1
BCAS1
BCAS3
BCL11A
BCL2A1
BCLAF1
BEST4
BICD1
BIRC8
BMP2
BMP3
BMP4
BRCA1
BTC
BTK
BUB1B
C10orf120
C10orf54
C12orf29
C14orf37

C15orf39
C15orf52
C16orf71
C17orf70
C18orf25
C18orf8
C1orf167
C1orf87
C1QTNF6
C1R
C2orf61
C2orf73
C3orf20
C4orf33
C5orf34
C5orf52
C6orf222
C7
C7orf31
C7orf57
C7orf72
C9
C9orf114
C9orf156
C9orf66
CA1
CAAP1
CACNA1B
CACNA2D2
CAGE1
CALR
CAMSAP3
CAND1
CAPN5
CAPN8
CAPN9
CAPZA3
CARD14
CASC5
CBWD7
CC2D1B
CCDC124
CCDC137
CCDC178
CCDC18
CCDC180
CCDC181
CCDC28A

CCDC40
CCDC42
CCDC6
CCDC64B
CCDC93
CCKBR
CCL20
CCNB1
CCNG2
CCNH
CCPG1
CCSER1
CCT6A
CD163
CD164
CD5
CD6
CD80
CDC20B
CDC27
CDC34
CDC40
CDC5L
CDH11
CDH24
CDH3
CDHR2
CDHR3
CDK11A
CDK11B
CDK5RAP2
CDYL2
CECR5
CELSR3
CENPE
CENPQ
CEP120
CEP135
CEP192
CEP290
CERS2
CES1
CES3
CFAP46
CFAP69
CFAP74
CFH
CGREF1

CHAT
CHD1
CHD3
CHD4
CHD6
CHIA
CHIT1
CHM
CHPT1
CHRNA1
CIZ1
CKM
CLASP2
CLCA2
CLCN6
CLCNKB
CLEC4A
CLK1
CLTCL1
CMTR2
CMYA5
CNDP1
CNGB1
CNN2
CNP
CNPPD1
CNR2
CNTN3
CNTNAP5
COASY
COL14A1
COL15A1
COL17A1
COL2A1
COL4A3
COL4A4
COL6A2
COL6A5
COL6A6
COL7A1
COMMD10
COMT
COQ7
COX10
COX11
COX19
CRMP1
CRYBG3

CRYGB
CSGALNACT1
CSRNP3
CTBP2
CTNNAL1
CTNNB1
CTSB
CTSE
CUBN
CUL9
CWH43
CYB5RL
CYBRD1
CYFIP1
CYP2S1
CYP46A1
CYP4F12
CYP4F2
CYP4F8
DACT1
DAK
DAPL1
DCDC2C
DCHS1
DCHS2
DCT
DCTN1
DDB1
DDB2
DDIT3
DDRGK1
DDX10
DDX20
DDX4
DDX43
DDX47
DDX53
DDX56
DDX58
DDX60L
DENND1C
DENND2A
DES
DGCR14
DGKG
DHRS1
DHTKD1
DHX33

DIAPH3
DIS3
DISC1
DLEC1
DLGAP2
DMBT1
DMKN
DMP1
DNAAF3
DNAH1
DNAH14
DNAH17
DNAH7
DOCK6
DOCK8
DPY19L4
DPYD
DPYSL2
DUOX1
DUOX2
DUX4L2
DUX4L4
DUX4L8
DYX1C1
EBNA1BP2
ECHDC3
EEF1G
EHBP1
EID2
EIF3L
EIF4ENIF1
EIF4G1
ELAC2
ELAVL1
ELN
EML6
EMR2
ENPP5
EPHA4
EPHA6
EPHB1
EPPK1
EPS8L1
ERC1
ERCC5
ERICH6B
ERO1LB
ERP27

ESPL1
ESYT2
ESYT3
ETFDH
EVA1A
EVA1C
EVC2
EVL
EXOC8
EXTL1
EYS
F5
FAM124A
FAM136A
FAM160A1
FAM160B2
FAM173A
FAM178A
FAM186A
FAM188B
FAM220A
FAM35A
FAM47C
FAM47E
FAM47E-STBD1
FAM53A
FAM65C
FAM71D
FAM71F1
FAM83G
FAM86C1
FANCA
FANCI
FANK1
FARP2
FAT2
FBLIM1
FBN3
FBXL18
FBXO18
FBXO2
FBXW8
FBXW9
FCGR1B
FCGRT
FCN2
FCRLB
FER

FERMT1
FERMT2
FFAR4
FGF5
FGFR1
FGL1
FHDC1
FHL5
FILIP1
FIP1L1
FJX1
FLJ22184
FLT3
FLVCR1
FMN1
FMO2
FNDC3B
FNIP1
FNIP2
FOPNL
FOXA1
FPR1
FRAS1
FREM2
FRG1B
FRG2C
FRMD4B
FRS2
FSD1L
FTSJ3
FUT2
FUT3
FUT9
FXYD4
GABRG2
GAK
GALC
GALNT12
GALNTL5
GALP
GAMT
GARS
GAS6
GBA
GBP3
GBP6
GCAT
GFM1

GFRA2
GGH
GGT2
GGT6
GHRHR
GIMAP1
GIMAP5
GIMAP6
GIMAP7
GJA4
GJB7
GLS2
GLTSCR1L
GLYATL3
GNL3L
GOLGA6L2
GORAB
GPR114
GPR137C
GPR157
GPR64
GPR98
GPRIN2
GPRIN3
GRIN2C
GRIN2D
GRIN3A
GRIP1
GRM2
GSDMC
GSTA5
GSTZ1
GUCY2F
GXYLT1
GYS1
HADHA
HAL
HAP1
HBS1L
HEATR2
HEATR5A
HEATR5B
HECW1
HERC1
HERC6
HHAT
HHLA1
HIBCH

HMCN1
HMGB4
HMGXB4
HMOX2
HNRNPCL2
HNRNPUL2
HNRNPUL2-BSCL2
HOXB1
HPCA
HPS4
HRNR
HSD17B13
HSD17B4
HSDL1
HSPG2
HTR5A
HUNK
HUS1B
HYDIN
HYOU1
ICAM1
IDH1
IDNK
IDO2
IFI30
IFIT2
IFT172
IFT74
IFT88
IGF2R
IGFN1
IGLL5
IGSF5
IGSF8
IKBKAP
IL6ST
INADL
INMT
INO80
INPP5B
INPP5D
INTU
IPO5
IQGAP3
IRAK2
ITGA10
ITGA11
ITGA9

ITGAE
ITGB4
ITIH1
ITIH3
ITLN1
ITLN2
ITPR2
JMY
JUP
KAT6A
KBTBD13
KCNA2
KCNAB2
KCNE1
KCNJ1
KCNJ12
KCNK1
KCNN3
KCNQ3
KCNRG
KDM7A
KDSR
KIAA0020
KIAA0100
KIAA0368
KIAA0753
KIAA1024L
KIAA1033
KIAA1549
KIAA1715
KIAA1755
KIF23
KIF24
KIF26A
KIF4B
KLHDC1
KLHDC8A
KLHL34
KLHL38
KLK11
KLK3
KLRB1
KLRC3
KMT2C
KPNA1
KRAS
KRI1
KRR1

KRT13
KRT18
KRT28
KRT32
KRT36
KRT37
KRT72
KRT76
KRT83
KRTAP10-1
KRTAP10-11
KRTAP10-12
KRTAP10-3
KRTAP10-4
KRTAP10-5
KRTAP10-7
KRTAP11-1
KRTAP12-3
KRTAP17-1
KRTAP5-6
KRTAP9-1
KRTAP9-4
KSR2
LAMA1
LAMB4
LAMTOR1
LARP1B
LARS
LCE2D
LCE3D
LHX9
LIFR
LIG1
LIPG
LIPT2
LITAF
LMO7
LNP1
LOH12CR1
LPAR3
LPCAT1
LRBA
LRIG2
LRMP
LRP1
LRP12
LRP1B
LRP2

LRP4
LRP8
LRRC41
LRRC48
LRRC74A
LRRD1
LSG1
LSR
LTBP1
LYSMD3
LYSMD4
MACC1
MADCAM1
MAEA
MAF1
MAGEA3
MAK
MALRD1
MAN2C1
MAP1A
MAP2K3
MAP2K5
MAP3K19
MAP4
MAP6D1
MAP7
MAP9
MAPK8IP3
MAPT
MARVELD3
MASP2
MBD1
MBTD1
MCF2L2
MCMDC2
MCPH1
MDN1
MED11
MED12
MEF2A
MEGF6
MEP1A
METTL18
MGAT5B
MIA3
MICAL2
MICALCL
MKI67

MLF1
MMAB
MMP17
MMP21
MMRN2
MMS22L
MOAP1
MOB3B
MOCOS
MOGAT1
MON1A
MOXD1
MPEG1
MPHOSPH10
MPP3
MPPE1
MPRIP
MRC2
MRGPRX2
MRGPRX4
MRI1
MROH7
MROH7-TTC4
MRPL18
MRPS10
MRPS7
MS4A6E
MSL2
MSL3
MSMB
MTA1
MTCH2
MTERF4
MTFP1
MTMR1
MTO1
MTPN
MTR
MTRR
MTUS2
MUC12
MUC16
MUC3A
MUC4
MUC5AC
MUS81
MXRA5
MYCBP2

MYCBPAP
MYH15
MYH4
MYH7B
MYL6
MYO1E
MYO3A
MYO3B
MYO5C
MYO9A
MYOF
MYOM3
MYPN
NAA35
NAA60
NAAA
NAALADL2
NADK2
NADSYN1
NALCN
NARFL
NAV2
NBEA
NBEAL1
NBEAL2
NBPF1
NBPF10
NBPF9
NCAPG
NCDN
NCK1
NCKAP5
NCOA3
NDUFAF1
NDUFS1
NEB
NEDD4
NEDD4L
NEDD8
NEDD8-MDP1
NEIL1
NEK11
NELL1
NEMF
NEO1
NEUROD2
NF1

NFASC
NFATC1
NFX1
NHLRC1
NIPA1
NIPAL1
NIPSNAP3A
NKX3-2
NLN
NLRP1
NLRP12
NLRP13
NLRP3
NME8
NMS
NNT
NOD2
NOL9
NOM1
NOTCH3
NPBWR1
NPHS1
NPIPB15
NPNT
NPR1
NPY4R
NR3C2
NRDE2
NRG1
NRXN1
NRXN3
NT5C3B
NTHL1
NTMT1
NTNG2
NUDT15
NUDT6
NUDT7
NUMA1
NUP107
NUP133
NUP160
OBSCN
OBSL1
OMD
ONECUT1
OPRM1
OPTN

OR10A6
OR10AD1
OR10G2
OR10H1
OR10J1
OR10J5
OR10T2
OR11G2
OR11H1
OR11H6
OR11L1
OR13F1
OR13G1
OR13J1
OR14C36
OR14I1
OR1A1
OR1A2
OR1E1
OR1I1
OR1L1
OR1L4
OR1L6
OR1Q1
OR1S1
OR2C1
OR2D3
OR2M7
OR2T12
OR2T7
OR4A16
OR4A5
OR4B1
OR4C11
OR4C15
OR4C3
OR4C46
OR4C5
OR4C6
OR4D6
OR4K14
OR4K17
OR4L1
OR4M1
OR4M2
OR4N4
OR4P4
OR4S2

OR4X1
OR51A4
OR51A7
OR51G1
OR51I2
OR51J1
OR51M1
OR51Q1
OR51V1
OR52E2
OR52I2
OR52J3
OR52K1
OR52N1
OR52N2
OR52W1
OR56B1
OR5AU1
OR5B2
OR5B3
OR5D16
OR5D18
OR5H15
OR5H6
OR5K3
OR5M1
OR5M3
OR5M8
OR5P3
OR5R1
OR5T2
OR6B2
OR6B3
OR6C68
OR6C70
OR6C74
OR6C76
OR6M1
OR6S1
OR6X1
OR7A10
OR7D2
OR8A1
OR8D1
OR8D4
OR8G5
OR8H1
OR8H2

OR8H3
OR8K1
OR8U1
OR9G1
OR9Q2
ORC3
OSBPL1A
OSBPL2
OTOG
OTOL1
OTOP2
OTOR
P2RX4
P2RX7
P2RY2
P4HA2
PABPC3
PADI4
PALLD
PAN3
PANX3
PAPD7
PAPL
PAPLN
PAPPA
PAPPA2
PARK7
PARP10
PARP14
PATE1
PAXIP1
PCDH15
PCDH17
PCDHA1
PCDHA11
PCDHA3
PCDHA9
PCDHAC1
PCDHB16
PCDHB7
PCDHB8
PCDHGA1
PCDHGA5
PCDHGA8
PCDHGB5
PCDHGC4
PCOLCE2
PDE4C

PDE4D
PDE4DIP
PDGFRA
PDHB
PDLIM2
PDZD9
PDZRN4
PEAR1
PER3
PEX11B
PFKFB3
PGAP1
PGM1
PHACTR4
PHAX
PHC2
PHF7
PIAS3
PICALM
PIEZO1
PIEZO2
PIGC
PIGZ
PIK3CB
PKD1L3
PKHD1
PKHD1L1
PKP1
PKP3
PLB1
PLCD1
PLD3
PLEC
PLEKHG2
PLEKHG4B
PLEKHS1
PLET1
PLIN4
PLIN5
PM20D1
PMS2
PNLIPRP1
PNMAL1
PNPLA6
POC1A
POLI
POLL
POLM

POLQ
POLR1A
POLR3D
POMZP3
PON2
POP1
POU5F1B
PPA2
PPARGC1B
PPEF2
PPIL6
PPME1
PPP1CB
PPP1R15A
PPP1R2
PPP1R9A
PPP2CB
PPP2R3A
PPP2R5B
PPP2R5C
PPP6R1
PPRC1
PRAMEF1
PRAMEF26
PRDM13
PRDM5
PRIM1
PRIM2
PRKAA2
PRKAG3
PRKDC
PRKRA
PRMT6
PRODH2
PROZ
PRPF4B
PRR15
PRR16
PRRC2C
PRSS22
PRSS55
PSD4
PSG1
PSMD13
PSMD6
PSMF1
PSRC1
PTAFR

PTGER3
PTPLA
PTPN13
PTPN23
PTPRB
PTPRQ
PYGO1
QPRT
QRFPR
QSER1
QSOX1
RAB11FIP1
RAD51C
RAET1E
RAI1
RANBP6
RAPGEF2
RASAL1
RBBP4
RBL2
RBM19
RBMX
RBP3
RCN3
RECQL4
RELB
REP15
RFPL1
RFPL2
RFX5
RHAG
RHBDL3
RHBG
RHOBTB2
RHOT1
RHOT2
RIBC1
RICTOR
RIMS2
RIPK2
RLF
RNASEH1
RNASEL
RNF111
RNF115
RNF121
RNF43
ROBO2

RP1
RP11-404P21.8
RP11-683L23.1
RP11-723O4.6
RP1L1
RPAP1
RPGRIP1
RRS1
RSPH14
RTN3
RTN4
S100Z
SAA4
SACM1L
SACS
SAFB2
SALL2
SAMD14
SAMSN1
SART3
SCLT1
SCLY
SCML2
SCN2A
SCN5A
SCYL2
SDCBP2
SDF2
SDK2
SEC14L3
SEC14L4
SEC31A
SECISBP2
SEMA4D
SEMA4G
SEPN1
SERPINB10
SERPINI2
SESN1
SHANK2
SHFM1
SIGLEC12
SIGLEC5
SIGLEC9
SIRPB1
SLC12A1
SLC12A4
SLC15A4

SLC16A1
SLC16A8
SLC1A1
SLC22A10
SLC22A14
SLC22A2
SLC22A24
SLC22A25
SLC22A9
SLC24A1
SLC25A41
SLC25A48
SLC25A5
SLC34A2
SLC35E3
SLC46A3
SLC4A3
SLC6A18
SLC6A3
SLCO1C1
SLFN12
SLFN12L
SLFNL1
SLIT3
SLK
SMCHD1
SMEK1
SMPDL3A
SMPDL3B
SMYD4
SNAP23
SNF8
SNRNP27
SNTG2
SNW1
SNX31
SORBS1
SOS2
SPAG16
SPAG17
SPAG4
SPANXN4
SPATA24
SPATA31A6
SPATA6L
SPEN
SPHK1
SPHKAP

SPINK5
SPNS3
SPRN
SPRYD7
SPTA1
SPTBN2
SRGAP2B
SRP14
SRRM2
SSX5
ST6GAL2
STEAP2
STK3
STK33
STK36
STPG2
STRN4
STXBP2
SUN1
SUN2
SUV39H1
SVEP1
SVOPL
SYNE1
SYNE2
SYNE3
SYT17
SYT8
TAAR5
TAF1
TAOK3
TAPBPL
TAS2R19
TAS2R4
TAS2R42
TBC1D1
TBC1D26
TBC1D9
TBL3
TBX2
TCEB3B
TCF7L2
TCOF1
TDRD5
TEAD4
TEKT4
TENM2
TENM4

TERF2IP
TESK1
TEX14
TGM4
TGOLN2
TGS1
THOC1
THSD7A
TIMM17A
TLDC1
TLN2
TLR3
TLR5
TM4SF19
TM4SF19-TCTEX1D2
TM9SF3
TMBIM1
TMBIM6
TMC6
TMEM120A
TMEM130
TMEM132B
TMEM144
TMEM174
TMEM175
TMEM176A
TMEM185B
TMEM209
TMEM237
TMEM244
TMEM26
TMEM5
TMEM79
TMOD1
TMPRSS15
TMPRSS9
TNC
TNFRSF10A
TNFRSF11A
TNFRSF13B
TNK1
TNN
TNP2
TNS1
TNS3
TOM1
TONSL
TOP1MT

TOX2
TPD52L3
TPO
TPPP2
TPSAB1
TPSG1
TPTE
TRAK2
TRANK1
TREML2
TRIM22
TRIM51
TRIM55
TRIM64C
TRIM66
TRIML2
TRNT1
TROAP
TSEN15
TSEN54
TSNARE1
TSR1
TSSC1
TTC14
TTC21A
TTC24
TTC26
TTC30B
TTC37
TTI2
TTLL4
TTN
TUBB4B
TUBB8
TUBGCP3
TXNDC2
TYR
U2AF2
UAP1
UBN1
UBR2
UBR5
UBXN11
UCHL5
UCK1
UGCG
UGGT2
UGT1A5

UGT2A3
UGT2B28
ULK1
UNC13A
UNC5B
UNC5C
UQCRH
URGCP
USH2A
USP28
USP35
USP45
USP9X
UTP14C
UTP20
VCAN
VDAC3
VIT
VLDLR
VPS13A
VRK2
VWA3B
WBSCR27
WBSCR28
WDR49
WDR6
WDR63
WDR64
WDR72
WDR74
WDR90
WDR91
WDYHV1
WIPF3
WNK2
WNK4
WNT10B
WWC1
XIRP1
XRN2
YAF2
YRDC
ZAN
ZBTB11
ZFHX4
ZGRF1
ZKSCAN2
ZNF101

ZNF131
ZNF134
ZNF160
ZNF174
ZNF177
ZNF180
ZNF184
ZNF19
ZNF214
ZNF215
ZNF221
ZNF225
ZNF236
ZNF239
ZNF273
ZNF28
ZNF281
ZNF283
ZNF285
ZNF286A
ZNF3
ZNF30
ZNF365
ZNF404
ZNF415
ZNF426
ZNF443
ZNF45
ZNF460
ZNF471
ZNF474
ZNF501
ZNF502
ZNF514
ZNF518B
ZNF527
ZNF530
ZNF534
ZNF549
ZNF550
ZNF556
ZNF559-ZNF177
ZNF562
ZNF568
ZNF57
ZNF573
ZNF587B
ZNF594

ZNF595
ZNF598
ZNF607
ZNF611
ZNF658
ZNF667
ZNF680
ZNF681
ZNF683
ZNF705A
ZNF708
ZNF717
ZNF718
ZNF721
ZNF728
ZNF729
ZNF737
ZNF766
ZNF775
ZNF778
ZNF79
ZNF790
ZNF808
ZNF827
ZNF835
ZNF841
ZNF845
ZNF85
ZNF853
ZNF880
ZNF99
ZPBP2
ZSCAN5A
ZSCAN5D
ZSWIM2
ZSWIM4
ZZEF1

1528 genes. Gene names listed above.

6.21 Question 12)

What genes do Craig Venter and James Watson both have deleterious variants in? How many genes is this?

6.22 Answer 12)

```
In [22]: %%bash
         join Venter.SIFT4G.genes_with_deleterious_variants.txt Watson.SIFT4G.genes
         > Venter_and_Watson.SIFT4G.genes_with_deleterious_variants.txt
         wc -l Venter_and_Watson.SIFT4G.genes_with_deleterious_variants.txt
         cat Venter_and_Watson.SIFT4G.genes_with_deleterious_variants.txt
```

```
524 Venter_and_Watson.SIFT4G.genes_with_deleterious_variants.txt
A2ML1
ABCC8
AC073657.1
ACACB
ACAN
ADAMTSL3
ADH1C
AHNAK
AKAP13
AKR1C2
ALDH1B1
ALPK2
ALPK3
ALPP
ALX4
AMACR
AMPD3
ANKLE1
ANKRD30A
ANKRD33
ANKRD35
ANKRD36
ANKRD36C
ANKRD60
ANXA13
APOA1BP
APOB
APOL4
AQP7
ARHGEF19
ARHGEF37
ARMC9
ARPP21
ASB16
ATF7IP
ATP8B4
AVPR1B
BAG3
BMP2
BMP3
```

BRCA1
BTK
C14orf37
C1orf167
C1orf87
C2orf61
C2orf73
C4orf33
C7
C7orf31
C7orf57
C7orf72
C9orf114
CALR
CAPN5
CAPN8
CAPN9
CAPZA3
CASC5
CBWD7
CCDC178
CCDC18
CCDC181
CCDC40
CCDC6
CCDC64B
CCDC93
CD163
CD6
CDH11
CDHR2
CDK11A
CDK11B
CDK5RAP2
CENPQ
CEP120
CFAP69
CFAP74
CHIA
CHIT1
COASY
COL15A1
COL17A1
COL2A1
COL4A3
COL4A4
COL6A2
COL6A5

COL6A6
COMT
COQ7
COX10
CRYBG3
CRYGB
CTNNAL1
CTSE
CUBN
CWH43
CYFIP1
DAPL1
DDRGK1
DDX58
DDX60L
DHTKD1
DISC1
DLEC1
DLGAP2
DNAH1
DNAH14
DOCK6
DOCK8
DPYD
DPYSL2
DYX1C1
ECHDC3
ELAVL1
EMR2
EPPK1
EPS8L1
ERCC5
ERICH6B
ERP27
ESPL1
EVC2
FAM178A
FAM186A
FAM188B
FAM220A
FAM35A
FANCI
FANK1
FAT2
FBLIM1
FBN3
FBXW8
FHDC1

FLJ22184
FLVCR1
FPR1
FRAS1
FREM2
FRG1B
FRG2C
FRMD4B
FUT2
FUT3
FUT9
GALNTL5
GALP
GBA
GBP3
GBP6
GCAT
GGT2
GGT6
GIMAP6
GJB7
GLYATL3
GOLGA6L2
GORAB
GPR137C
GPR98
GPRIN2
GRIN3A
GSTA5
GUCY2F
HADHA
HEATR2
HEATR5A
HHAT
HIBCH
HMGXB4
HNRNPCL2
HOXB1
HPS4
HRNR
HSDL1
HSPG2
HUS1B
HYDIN
ICAM1
IDO2
IFT88
IGF2R

IKBKAP
INMT
IQGAP3
IRAK2
ITGA10
ITGA11
ITGA9
ITGAE
ITGB4
ITIH1
ITPR2
KCNAB2
KCNE1
KCNJ12
KIAA0753
KIAA1549
KIAA1755
KLHDC1
KLHL38
KLRB1
KLRC3
KRAS
KRI1
KRT13
KRT32
KRT72
KRT76
KRT83
KRTAP10-1
KRTAP10-11
KRTAP10-3
KRTAP10-4
KRTAP10-5
KRTAP12-3
KRTAP9-4
LARP1B
LCE3D
LIG1
LIPT2
LMO7
LPCAT1
LRMP
LRP2
LYSMD4
MADCAM1
MAF1
MAGEA3
MALRD1

MAP1A
MAP2K3
MAP2K5
MAP7
MAP9
MASP2
MBD1
MCF2L2
MCPH1
MEGF6
MEP1A
MKI67
MMS22L
MPHOSPH10
MRC2
MRGPRX4
MROH7
MROH7-TTC4
MRPS7
MS4A6E
MSMB
MTCH2
MTMR1
MTR
MTRR
MTUS2
MUC12
MUC16
MUC4
MUC5AC
MUS81
MXRA5
MYCBPAP
MYH15
MYH4
MYH7B
MYOM3
MYPN
NAAA
NAALADL2
NBPF1
NBPF10
NBPF9
NCAPG
NCKAP5
NEK11
NELL1
NEMF

NFATC1
NHLRC1
NIPA1
NIPAL1
NIPSNAP3A
NLRP13
NOTCH3
NPIPB15
NPY4R
NRG1
NT5C3B
NTMT1
NUP160
OBSCN
OBSL1
OR10A6
OR10H1
OR10J1
OR10J5
OR11G2
OR11H6
OR11L1
OR13G1
OR13J1
OR14C36
OR1A2
OR1E1
OR1I1
OR1L4
OR1L6
OR1Q1
OR2C1
OR2M7
OR2T12
OR2T7
OR4A16
OR4A5
OR4B1
OR4C11
OR4C3
OR4C46
OR4C5
OR4D6
OR4K14
OR4L1
OR4M1
OR4M2
OR4N4

OR51A4
OR51G1
OR51J1
OR51M1
OR51Q1
OR52E2
OR52J3
OR52N1
OR52W1
OR56B1
OR5AU1
OR5B2
OR5B3
OR5D16
OR5H15
OR5H6
OR5K3
OR5R1
OR6B3
OR6M1
OR7A10
OR8D1
OR8G5
OR8H2
OR8K1
OR9G1
OR9Q2
OTOP2
OTOR
P2RY2
PADI4
PAPLN
PAPPA
PARP14
PATE1
PCDH15
PCDHA1
PCDHA3
PCDHB7
PCDHB8
PDE4DIP
PDZRN4
PER3
PEX11B
PFKFB3
PIAS3
PIEZO1
PIGC

PKD1L3
PKHD1L1
PLEKHG4B
PLET1
PLIN5
PM20D1
PMS2
PON2
POU5F1B
PPA2
PPEF2
PRAMEF1
PRAMEF26
PRIM2
PRKAG3
PRRC2C
PSMD13
PSMF1
PTGER3
PTPLA
PTPRB
PTPRQ
QRFPR
RAB11FIP1
RAD51C
RASAL1
RBM19
RBMX
RBP3
RFPL1
RFPL2
RHBG
RHOT2
RICTOR
RIPK2
RNF115
RNF43
RP1
RP1L1
RTN4
SACS
SCLT1
SDK2
SEMA4D
SEMA4G
SEPN1
SHFM1
SIGLEC5

SIRPB1
SLC16A8
SLC22A10
SLC22A24
SLC24A1
SLIT3
SMPDL3B
SMYD4
SNTG2
SOS2
SPATA31A6
SPINK5
SPTA1
SRGAP2B
SRP14
SSX5
STEAP2
STK36
SUN1
SVOPL
SYNE1
SYNE2
SYT8
TAF1
TAS2R4
TBL3
TCEB3B
TEKT4
TGOLN2
TGS1
THOC1
TLR3
TLR5
TMBIM1
TMEM185B
TMEM244
TMPRSS15
TMPRSS9
TNK1
TNN
TNP2
TOP1MT
TPSAB1
TPTE
TRIM22
TRIM51
TRNT1
TSEN54

TTC24
TTC26
TTC30B
TTI2
TTLL4
TTN
TUBB8
UCK1
UGT2B28
UNC5C
UTP20
VCAN
VRK2
WBSCR28
WDR49
WDR91
WDYHV1
WNK2
YAF2
ZAN
ZNF131
ZNF177
ZNF180
ZNF19
ZNF214
ZNF221
ZNF239
ZNF28
ZNF30
ZNF404
ZNF415
ZNF443
ZNF45
ZNF534
ZNF549
ZNF559-ZNF177
ZNF568
ZNF57
ZNF573
ZNF607
ZNF611
ZNF658
ZNF667
ZNF705A
ZNF717
ZNF728
ZNF737
ZNF880

```
ZNF99
ZPBP2
ZSCAN5A
ZSCAN5D
```

Gene names provided above. 524 genes in common.

6.23 Question 13)

What is the lowest SIFT score of the deleterious variants?

6.24 Answer 13)

```
In [23]: %%bash
         cat Watson.SIFT4G/Watson_SIFTannotations.xls|tail -n+2 \
         |grep 'DELETERIOUS'|grep -v 'Low confidence' \
         |cut -f1,2,3,4,13 \
         |sort|uniq \
         |sort -k1,1 -k2,2n \
         |sort -k5,5n \
         |head

10        113766634        T        C        0.000
10        19387657        A        G        0.000
10        26157364        C        A        0.000
10        46461688        A        C        0.000
10        48086        G        A        0.000
10        59792934        G        T        0.000
10        62376867        C        T        0.000
10        86936837        C        G        0.000
10        89307233        A        T        0.000
10        95339252        C        A        0.000
```

0.0 is the lowest SIFT score.

6.25 Question 14)

What variants are annotated with the lowest SIFT score? Output the chromosome, coordinate, reference base, alternate base, gene name, reference amino acid, alternate amino acid, amino acid position, and sift score into a file. Display the first 10 lines of this file.

6.26 Answer 14)

```
In [24]: %%bash
         cat Watson.SIFT4G/Watson_SIFTannotations.xls|cut -f1,2,3,4,7,10,11,12,13,1
         |grep '^CHROM\|DELETERIOUS'|grep -v 'Low confidence' \
         |awk '($9==0.0)||$1=="CHROM"' \
```

```
> Watson.SIFT4G.sift_score_0.txt
head -n10 Watson.SIFT4G.sift_score_0.txt
```

```
CHROM       POS        REF_ALLELE        ALT_ALLELE        GENE_NAME        REF_AN
1        1956754        C        A        CFAP74        G        C        628
1        3497541        C        T        MEGF6        G        R        1152
1        11789390        A        G        C1orf167        R        G        810
1        12725782        C        T        AADACL3        P        L        280
1        17334004        G        C        PADI4        G        A        112
1        26367769        T        C        ZNF683        D        G        48
1        26367769        T        C        ZNF683        D        G        48
1        26367769        T        C        ZNF683        D        G        48
1        28490968        C        T        PHACTR4        R        C        622
```

6.27 3.3) Analysis of 1000 Genomes Sample Human Data

6.27.1 3.3.1) Calling variants from aligned sequencing data

The 1000 Genomes exome sequencing data for this sample is not yet in VCF format. We must use samtools mpileup and bcftools call to convert it.

For samtools mpileup, we use the following options: - '-u' generate uncompressed VCF/BCF output. This saves time on compression and decompression, since we pipe to bcftools. - '-g' generate output in BCF format. This is a more compact binary format, ideal for transferring between programs. - '-f' the FASTA file used as reference for the CRAM file. Required to determine if something varies from the reference, and to decompress the CRAM data.

For bcftools call, we use the following options to call variants: - '-f GQ,GP' output genotype quality and genotype probability. We care about GQ for filtering. - '-v' output variant sites only. We don't care about sites that match the reference. - '-m' we use the multiallelic caller, upon recommendation by the samtools website. - '-O v' output VCF formatted file. - '-o' output variants to the specified file

We connect the output of samtools mpileup to the input of bcftools using a pipe '|'.

In [25]:
```
%%bash
date
samtools-1.4/samtools mpileup \
-ugf GRCh38_full_analysis_set_plus_decoy_hla.fa \
NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.cram \
| bcftools-1.4/bcftools call \
-f GQ,GP \
-vmO v \
-o NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.gq.gp.vcf
date
```

```
Sun Mar 26 01:52:24 PDT 2017
Sun Mar 26 04:28:51 PDT 2017


Note: none of --samples-file, --ploidy or --ploidy-file given, assuming all sites a
[mpileup] 1 samples in 1 input files
```

```
<mpileup> Set max per-file depth to 8000
```

6.27.2 3.3.2) Filtering variants by read depth, quality, and genotype quality

Not all variant calls are made equal. We want to avoid predicting the deleteriousness of variants that may not be real. So we use filtering to filter for the depth of sequencing at each variant coordinate, and the confidence the variant caller has in the variant. This is encapsulated in the DP, QUAL, and GQ fields.

The command bcftools filter is used to implement these filters. - '-i' specifies an expression for variants to include. - 'INFO/DP>10': We want raw read depth to be greater than 10 - 'QUAL>20': We want the quality of any variant called here to be greater than 20 - 'FMT/GQ>20': We want the genotype to be called with a confidence greater than 20.

We then combine these criteria using logical AND ('&&') to yield the final filter inclusion statement, '(QUAL>20)&&(INFO/DP>10)&&(FMT/GQ>20)'.

For more details on DP, QUAL, and GQ, see the guide from GATK (http://gatkforums.broadinstitute.org/gatk/discussion/1268/what-is-a-vcf-and-how-should-i-interpret-it).

```
In [26]: %%bash
         date
         bcftools-1.4/bcftools filter -i '(QUAL>20)&&(INFO/DP>10)&&(FMT/GQ>20)' \
         -O v \
         -o NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_g
         NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.gq.gp.vcf
         date

Sun Mar 26 04:28:51 PDT 2017
Sun Mar 26 04:28:56 PDT 2017
```

6.28 Question 15)

How many variants are in the VCF before filtering? How many after filtering?

6.29 Answer 15)

```
In [27]: %%bash
         cat NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.gq.gp.vcf|grep -v '^#'|
         cat NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_

2254572
93617
```

2254572 variants before filtering. 93617 variants after filtering.

6.29.1 3.3.3) Annotating variants with SIFT4G

In [34]:
```
%%bash
java -jar SIFT4G_Annotator_v2.4.jar -c -t \
-i NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_g
-d GRCh38.78  \
-r NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_g
```

```
Start Time for SIFT4G code: Mon Mar 27 11:51:12 PDT 2017
Updates:
No updates from server!! Please go to http:sift-dna.org for updates.

Started Running ...
Running in Multitranscripts mode

Chromosome        WithSIFT4GAnnotations        WithoutSIFT4GAnnotations        Prog
The following chromosomes (or scaffolds/contigs) are not found in the SIFT 4G datab
HLA-B*08:33, Un_KN707967v1_decoy, HLA-DQB1*06:03:01, 19_KI270882v1_alt, Un_JTFH0100
```

```
HLA-A*02:68                         0                         9
GRCh38.78/1_KI270764v1_alt.regions does not exist
GRCh38.78/Un_JTFH01001889v1_decoy.regions does not exist
Un_JTFH01001889v1_decoy                         0                         1
GRCh38.78/10_KI270824v1_alt.regions does not exist
GRCh38.78/18_KI270863v1_alt.regions does not exist
GRCh38.78/HLA-DRB1*15:02:01.regions does not exist
HLA-DRB1*15:02:01                         0                         2
GRCh38.78/HLA-A*02:77.regions does not exist
8_KI270822v1_alt                         0                        22
HLA-A*02:77                         0                         9
GRCh38.78/HLA-B*67:01:01.regions does not exist
HLA-B*67:01:01                         0                         5
GRCh38.78/HLA-A*26:50.regions does not exist
1_KI270764v1_alt                         0                         4
HLA-A*26:50                         0                         2
GRCh38.78/Un_KI270746v1.regions does not exist
GRCh38.78/HLA-A*03:02:01.regions does not exist
GRCh38.78/HLA-B*55:12.regions does not exist
HLA-B*55:12                         0                         2
GRCh38.78/M.regions does not exist
18_KI270863v1_alt                         0                        72
GRCh38.78/HLA-DQB1*05:03:01:01.regions does not exist
HLA-A*03:02:01                         0                         2
HLA-DQB1*05:03:01:01                         0                         3
GRCh38.78/19_KI270933v1_alt.regions does not exist
Un_KI270746v1                         0                        65
10_KI270824v1_alt                         0                        11
GRCh38.78/Un_JTFH01000544v1_decoy.regions does not exist
GRCh38.78/HLA-B*39:01:01:02L.regions does not exist
GRCh38.78/HLA-A*01:16N.regions does not exist
Un_JTFH01000544v1_decoy                         0                        10
HLA-B*39:01:01:02L                         0                         2
HLA-A*01:16N                         0                         6
GRCh38.78/HLA-DQA1*05:11.regions does not exist
HLA-DQA1*05:11                         0                         1
GRCh38.78/19_GL949749v2_alt.regions does not exist
GRCh38.78/HLA-B*55:24.regions does not exist
HLA-B*55:24                         0                         2
19_GL949749v2_alt                         0                       149
GRCh38.78/HLA-A*68:02:01:01.regions does not exist
HLA-A*68:02:01:01                         0                         8
GRCh38.78/HLA-C*07:02:01:04.regions does not exist
HLA-C*07:02:01:04                         0                         2
GRCh38.78/Un_KN707964v1_decoy.regions does not exist
GRCh38.78/HLA-DQA1*05:03.regions does not exist
HLA-DQA1*05:03                         0                         1
Un_KN707964v1_decoy                         0                         4
```

```
GRCh38.78/HLA-A*02:57.regions does not exist
HLA-A*02:57                        0                          11
GRCh38.78/2_KI270770v1_alt.regions does not exist
GRCh38.78/HLA-C*07:02:01:03.regions does not exist
HLA-C*07:02:01:03                  0                           2
GRCh38.78/14_GL000225v1_random.regions does not exist
Y                        0                      1594                        Compl
GRCh38.78/19_KI270887v1_alt.regions does not exist
GRCh38.78/HLA-B*41:01:01.regions does not exist
HLA-B*41:01:01                     0                           4
GRCh38.78/Un_JTFH01001545v1_decoy.regions does not exist
14_GL000225v1_random                 0                             1296
M                        0                      31                        Complet
Un_JTFH01001545v1_decoy                0                           3
GRCh38.78/HLA-B*45:04.regions does not exist
GRCh38.78/HLA-A*02:51.regions does not exist
HLA-B*45:04                        0                           5
HLA-A*02:51                        0                           9
GRCh38.78/Un_JTFH01001117v1_decoy.regions does not exist
GRCh38.78/HLA-A*68:02:01:03.regions does not exist
19_KI270887v1_alt                  0                          96
Un_JTFH01001117v1_decoy                0                           3
HLA-A*68:02:01:03                  0                           7
19_KI270933v1_alt                  0                          99
GRCh38.78/HLA-B*67:01:02.regions does not exist
HLA-B*67:01:02                     0                           5
2_KI270770v1_alt                     0                           2
GRCh38.78/HLA-B*18:17N.regions does not exist
GRCh38.78/HLA-C*05:01:01:02.regions does not exist
HLA-B*18:17N                       0                           2
HLA-C*05:01:01:02                  0                           2
GRCh38.78/Un_JTFH01000972v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01001223v1_decoy.regions does not exist
Un_JTFH01001223v1_decoy                0                           4
GRCh38.78/HLA-A*68:08:01.regions does not exist
HLA-A*68:08:01                     0                           7
GRCh38.78/Un_JTFH01000144v1_decoy.regions does not exist
Un_JTFH01000144v1_decoy                0                           2
GRCh38.78/HLA-A*68:02:01:02.regions does not exist
Un_JTFH01000972v1_decoy                0                          18
HLA-A*68:02:01:02                  0                           7
GRCh38.78/Un_JTFH01001008v1_decoy.regions does not exist
Un_JTFH01001008v1_decoy                0                           1
GRCh38.78/Un_KI270757v1.regions does not exist
GRCh38.78/Un_JTFH01000329v1_decoy.regions does not exist
GRCh38.78/HLA-C*07:02:01:05.regions does not exist
HLA-C*07:02:01:05                  0                           2
GRCh38.78/Un_JTFH01001058v1_decoy.regions does not exist
```

```
GRCh38.78/Un_JTFH01000667v1_decoy.regions does not exist
Un_JTFH01000329v1_decoy                         0                       1
Un_JTFH01000667v1_decoy                         0                       1
Un_JTFH01001058v1_decoy                         0                       7
GRCh38.78/Un_JTFH01001212v1_decoy.regions does not exist
Un_JTFH01001212v1_decoy                         0                       2
GRCh38.78/Un_JTFH01000650v1_decoy.regions does not exist
Un_JTFH01000650v1_decoy                         0                       1
GRCh38.78/22_KI270735v1_random.regions does not exist
22_KI270735v1_random                            0                      35
GRCh38.78/HLA-A*02:48.regions does not exist
HLA-A*02:48                                     0                       9
GRCh38.78/15_KI270852v1_alt.regions does not exist
GRCh38.78/HLA-B*46:01:05.regions does not exist
GRCh38.78/HLA-B*08:01:01.regions does not exist
HLA-B*46:01:05                                  0                       6
HLA-B*08:01:01                                  0                       3
GRCh38.78/HLA-A*02:65.regions does not exist
GRCh38.78/Un_JTFH01001377v1_decoy.regions does not exist
GRCh38.78/Un_KN707906v1_decoy.regions does not exist
HLA-A*02:65                                     0                       2
Un_JTFH01001377v1_decoy                         0                       4
GRCh38.78/Un_KI270515v1.regions does not exist
GRCh38.78/Un_JTFH01000997v1_decoy.regions does not exist
GRCh38.78/HLA-B*55:48.regions does not exist
HLA-B*55:48                                     0                       2
15_KI270852v1_alt                               0                      32
Un_JTFH01000997v1_decoy                         0                       2
GRCh38.78/8_KI270819v1_alt.regions does not exist
GRCh38.78/HLA-B*08:20.regions does not exist
GRCh38.78/Un_JTFH01000715v1_decoy.regions does not exist
HLA-B*08:20                                     0                       4
Un_JTFH01000715v1_decoy                         0                      17
GRCh38.78/19_KI270866v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000225v1_decoy.regions does not exist
Un_JTFH01000225v1_decoy                         0                       6
GRCh38.78/Un_KI270590v1.regions does not exist
Un_KN707906v1_decoy                             0                       5
GRCh38.78/Un_JTFH01001934v1_decoy.regions does not exist
GRCh38.78/Un_KI270330v1.regions does not exist
Un_KI270515v1                                   0                      15
Un_JTFH01001934v1_decoy                         0                       1
Un_KI270330v1                                   0                       3
GRCh38.78/Un_JTFH01000096v1_decoy.regions does not exist
Un_JTFH01000096v1_decoy                         0                       8
Un_KI270757v1                                   0                     134
GRCh38.78/Un_KN707740v1_decoy.regions does not exist
Un_KI270590v1                                   0                      39
```

```
Un_KN707740v1_decoy                          0                          1
GRCh38.78/3_KI270895v1_alt.regions does not exist
GRCh38.78/HLA-B*44:02:27.regions does not exist
HLA-B*44:02:27                               0                          3
GRCh38.78/11_KI270902v1_alt.regions does not exist
GRCh38.78/Un_JTFH01001724v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000090v1_decoy.regions does not exist
Un_JTFH01000090v1_decoy                      0                          1
Un_JTFH01001724v1_decoy                      0                          2
GRCh38.78/Un_JTFH01000206v1_decoy.regions does not exist
3_KI270895v1_alt                             0                         57
11_KI270902v1_alt                            0                        113
19_KI270866v1_alt                            0                          5
8_KI270819v1_alt                             0                         12
GRCh38.78/Un_KN707959v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01001142v1_decoy.regions does not exist
GRCh38.78/HLA-B*46:01:01.regions does not exist
GRCh38.78/Un_JTFH01000762v1_decoy.regions does not exist
Un_JTFH01001142v1_decoy                      0                          1
HLA-B*46:01:01                               0                          6
GRCh38.78/3_KI270777v1_alt.regions does not exist
GRCh38.78/Un_JTFH01001506v1_decoy.regions does not exist
Un_JTFH01001506v1_decoy                      0                          8
GRCh38.78/Un_KN707866v1_decoy.regions does not exist
Un_KN707866v1_decoy                          0                         11
GRCh38.78/17_KI270860v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000885v1_decoy.regions does not exist
Un_JTFH01000885v1_decoy                      0                          1
3_KI270777v1_alt                             0                          4
17_KI270860v1_alt                            0                         16
GRCh38.78/Un_JTFH01000264v1_decoy.regions does not exist
GRCh38.78/17_KI270908v1_alt.regions does not exist
Un_JTFH01000264v1_decoy                      0                          6
Un_JTFH01000206v1_decoy                      0                          6
GRCh38.78/HLA-B*54:01:01.regions does not exist
Un_JTFH01000762v1_decoy                      0                          4
HLA-B*54:01:01                               0                          2
GRCh38.78/6_KI270797v1_alt.regions does not exist
Un_KN707959v1_decoy                          0                          2
17_KI270908v1_alt                            0                          8
GRCh38.78/Un_JTFH01001251v1_decoy.regions does not exist
Un_JTFH01001251v1_decoy                      0                          1
GRCh38.78/Un_KI270744v1.regions does not exist
GRCh38.78/17_KI270857v1_alt.regions does not exist
GRCh38.78/HLA-B*44:02:17.regions does not exist
HLA-B*44:02:17                               0                          5
GRCh38.78/HLA-A*01:01:01:02N.regions does not exist
GRCh38.78/Un_JTFH01000116v1_decoy.regions does not exist
```

```
GRCh38.78/Un_KN707972v1_decoy.regions does not exist
GRCh38.78/HLA-A*11:02:01.regions does not exist
HLA-A*11:02:01                          0                          9
Un_KN707972v1_decoy                          0                         12
Un_KI270744v1                         0                        140
GRCh38.78/Un_JTFH01000653v1_decoy.regions does not exist
Un_JTFH01000116v1_decoy                          0                          6
GRCh38.78/HLA-A*02:376.regions does not exist
HLA-A*01:01:01:02N                          0                          7
GRCh38.78/HLA-A*24:10:01.regions does not exist
Un_JTFH01000653v1_decoy                          0                          1
HLA-A*02:376                         0                         9
GRCh38.78/2_GL383522v1_alt.regions does not exist
6_KI270797v1_alt                          0                          3
17_KI270857v1_alt                          0                        119
GRCh38.78/HLA-B*18:26.regions does not exist
HLA-B*18:26                          0                         2
GRCh38.78/HLA-A*30:02:01:02.regions does not exist
GRCh38.78/12_GL383550v2_alt.regions does not exist
HLA-A*24:10:01                          0                         14
GRCh38.78/Un_JTFH01000493v1_decoy.regions does not exist
HLA-A*30:02:01:02                          0                          6
GRCh38.78/HLA-C*05:01:01:01.regions does not exist
HLA-C*05:01:01:01                          0                          2
GRCh38.78/Un_JTFH01001956v1_decoy.regions does not exist
GRCh38.78/3_KI270781v1_alt.regions does not exist
GRCh38.78/17_KI270861v1_alt.regions does not exist
Un_JTFH01000493v1_decoy                          0                          3
GRCh38.78/HLA-B*14:02:01.regions does not exist
HLA-B*14:02:01                          0                          3
GRCh38.78/HLA-C*01:21.regions does not exist
HLA-C*01:21                          0                         1
12_GL383550v2_alt                          0                          4
17_KI270861v1_alt                          0                         37
GRCh38.78/Un_JTFH01000180v1_decoy.regions does not exist
3_KI270781v1_alt                          0                          4
GRCh38.78/19_GL383575v2_alt.regions does not exist
GRCh38.78/HLA-A*30:02:01:01.regions does not exist
Un_JTFH01001956v1_decoy                          0                          1
Un_JTFH01000180v1_decoy                          0                         11
HLA-A*30:02:01:01                          0                          5
2_GL383522v1_alt                          0                          8
GRCh38.78/HLA-B*35:14:02.regions does not exist
GRCh38.78/Un_JTFH01001875v1_decoy.regions does not exist
GRCh38.78/HLA-B*15:18:01.regions does not exist
Un_JTFH01001875v1_decoy                          0                          1
HLA-B*35:14:02                          0                          2
GRCh38.78/Un_JTFH01000004v1_decoy.regions does not exist
```

```
Un_JTFH01000004v1_decoy                          0                          1
GRCh38.78/Un_JTFH01001234v1_decoy.regions does not exist
Un_JTFH01001234v1_decoy                          0                          1
HLA-B*15:18:01                             0                            4
GRCh38.78/Un_KN707986v1_decoy.regions does not exist
GRCh38.78/Un_KI270742v1.regions does not exist
Un_KN707986v1_decoy                              0                            1
19_GL383575v2_alt                            0                           14
GRCh38.78/HLA-B*37:01:05.regions does not exist
GRCh38.78/HLA-B*18:03.regions does not exist
HLA-B*18:03                            0                           2
GRCh38.78/Un_JTFH01001998v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01001615v1_decoy.regions does not exist
Un_JTFH01001615v1_decoy                          0                          1
Un_JTFH01001998v1_decoy                          0                          5
GRCh38.78/HLA-B*18:02.regions does not exist
HLA-B*18:02                            0                           1
GRCh38.78/HLA-A*02:10.regions does not exist
HLA-A*02:10                            0                           8
GRCh38.78/HLA-C*05:08.regions does not exist
HLA-C*05:08                            0                           1
GRCh38.78/3_KI270780v1_alt.regions does not exist
GRCh38.78/HLA-B*51:07:01.regions does not exist
HLA-B*51:07:01                             0                            1
HLA-B*37:01:05                             0                            1
GRCh38.78/Un_JTFH01000348v1_decoy.regions does not exist
Un_JTFH01000348v1_decoy                          0                         15
GRCh38.78/Un_JTFH01000127v1_decoy.regions does not exist
Un_JTFH01000127v1_decoy                          0                          7
GRCh38.78/16_KI270854v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000896v1_decoy.regions does not exist
Un_JTFH01000896v1_decoy                          0                          5
3_KI270780v1_alt                           0                            6
GRCh38.78/9_KI270718v1_random.regions does not exist
GRCh38.78/22_KI270738v1_random.regions does not exist
16_KI270854v1_alt                            0                           10
GRCh38.78/HLA-B*27:04:01.regions does not exist
HLA-B*27:04:01                             0                            5
GRCh38.78/Un_JTFH01001929v1_decoy.regions does not exist
Un_JTFH01001929v1_decoy                          0                         22
GRCh38.78/HLA-C*17:03.regions does not exist
HLA-C*17:03                            0                           2
Un_KI270742v1                              0                             16
GRCh38.78/2_KI270767v1_alt.regions does not exist
GRCh38.78/15_KI270848v1_alt.regions does not exist
GRCh38.78/1_KI270712v1_random.regions does not exist
GRCh38.78/Un_JTFH01000704v1_decoy.regions does not exist
GRCh38.78/6_KI270801v1_alt.regions does not exist
```

```
GRCh38.78/19_GL949752v1_alt.regions does not exist
GRCh38.78/17_JH159148v1_alt.regions does not exist
Un_JTFH01000704v1_decoy                         0                         1
GRCh38.78/HLA-B*47:01:01:01.regions does not exist
HLA-B*47:01:01:01                         0                         4
GRCh38.78/Un_JTFH01001332v1_decoy.regions does not exist
19_GL949752v1_alt                         0                       293
9_KI270718v1_random                         0                        15
Un_JTFH01001332v1_decoy                         0                         2
GRCh38.78/HLA-B*15:108.regions does not exist
HLA-B*15:108                         0                         4
17_JH159148v1_alt                         0                        19
GRCh38.78/14_KI270724v1_random.regions does not exist
22_KI270738v1_random                         0                         9
15_KI270848v1_alt                         0                        22
GRCh38.78/19_KI270929v1_alt.regions does not exist
2_KI270767v1_alt                         0                         2
6_KI270801v1_alt                         0                         9
19_KI270929v1_alt                         0                       113
1_KI270712v1_random                         0                        15
GRCh38.78/16_KI270855v1_alt.regions does not exist
GRCh38.78/19_GL383576v1_alt.regions does not exist
GRCh38.78/Un_GL000219v1.regions does not exist
GRCh38.78/HLA-B*44:02:01:03.regions does not exist
GRCh38.78/3_GL383526v1_alt.regions does not exist
GRCh38.78/HLA-B*15:17:01:02.regions does not exist
HLA-B*44:02:01:03                         0                         5
HLA-B*15:17:01:02                         0                         4
GRCh38.78/HLA-B*47:01:01:02.regions does not exist
HLA-B*47:01:01:02                         0                         4
GRCh38.78/HLA-B*15:17:01:01.regions does not exist
19_GL383576v1_alt                         0                         6
GRCh38.78/Un_KN707863v1_decoy.regions does not exist
HLA-B*15:17:01:01                         0                         4
GRCh38.78/HLA-B*38:01:01.regions does not exist
HLA-B*38:01:01                         0                         2
16_KI270855v1_alt                         0                        18
GRCh38.78/14_GL000009v2_random.regions does not exist
Un_KN707863v1_decoy                         0                         1
GRCh38.78/Un_JTFH01001111v1_decoy.regions does not exist
GRCh38.78/12_KI270837v1_alt.regions does not exist
Un_GL000219v1                         0                        34
GRCh38.78/HLA-A*68:71.regions does not exist
Un_JTFH01001111v1_decoy                         0                         1
14_KI270724v1_random                         0                        15
HLA-A*68:71                         0                         8
3_GL383526v1_alt                         0                         2
GRCh38.78/22_KI270733v1_random.regions does not exist
```

```
12_KI270837v1_alt                          0                          28
GRCh38.78/HLA-B*37:01:01.regions does not exist
HLA-B*37:01:01                             0                           2
GRCh38.78/Un_KN707668v1_decoy.regions does not exist
Un_KN707668v1_decoy                        0                           1
GRCh38.78/Un_JTFH01000402v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000351v1_decoy.regions does not exist
Un_JTFH01000351v1_decoy                    0                          23
GRCh38.78/HLA-B*35:02:01.regions does not exist
14_GL000009v2_random                       0                          10
HLA-B*35:02:01                             0                           2
GRCh38.78/HLA-A*24:61.regions does not exist
22_KI270733v1_random                       0                          88
GRCh38.78/HLA-B*35:41.regions does not exist
GRCh38.78/Un_JTFH01001066v1_decoy.regions does not exist
GRCh38.78/HLA-A*24:09N.regions does not exist
GRCh38.78/4_GL000257v2_alt.regions does not exist
Un_JTFH01001066v1_decoy                    0                           1
HLA-A*24:61                            0                           7
HLA-B*35:41                            0                           2
Un_JTFH01000402v1_decoy                    0                           8
GRCh38.78/HLA-B*44:02:01:01.regions does not exist
HLA-B*44:02:01:01                          0                           5
HLA-A*24:09N                           0                          14
GRCh38.78/Un_JTFH01001878v1_decoy.regions does not exist
Un_JTFH01001878v1_decoy                    0                           2
GRCh38.78/Un_JTFH01000343v1_decoy.regions does not exist
Un_JTFH01000343v1_decoy                    0                           1
4_GL000257v2_alt                       0                          37
GRCh38.78/Un_JTFH01000672v1_decoy.regions does not exist
Un_JTFH01000672v1_decoy                    0                          17
GRCh38.78/HLA-B*52:01:01:03.regions does not exist
HLA-B*52:01:01:03                          0                           1
GRCh38.78/Un_JTFH01001011v1_decoy.regions does not exist
Un_JTFH01001011v1_decoy                    0                           1
GRCh38.78/Un_KI270435v1.regions does not exist
GRCh38.78/Un_JTFH01000600v1_decoy.regions does not exist
GRCh38.78/HLA-B*52:01:01:01.regions does not exist
HLA-B*52:01:01:01                          0                           1
Un_JTFH01000600v1_decoy                    0                           2
GRCh38.78/18_GL383572v1_alt.regions does not exist
18_GL383572v1_alt                          0                          11
GRCh38.78/HLA-B*52:01:01:02.regions does not exist
HLA-B*52:01:01:02                          0                           1
GRCh38.78/HLA-C*15:17.regions does not exist
HLA-C*15:17                            0                           3
GRCh38.78/19_KI270891v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000418v1_decoy.regions does not exist
```

```
Un_JTFH01000418v1_decoy                          0                          4
Un_KI270435v1                          0                         48
GRCh38.78/HLA-B*40:01:02.regions does not exist
HLA-B*40:01:02                          0                          4
GRCh38.78/HLA-C*15:16.regions does not exist
HLA-C*15:16                          0                         3
GRCh38.78/Un_JTFH01000040v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01001109v1_decoy.regions does not exist
Un_JTFH01000040v1_decoy                          0                          1
Un_JTFH01001109v1_decoy                          0                          7
19_KI270891v1_alt                           0                        102
GRCh38.78/Un_JTFH01000194v1_decoy.regions does not exist
Un_JTFH01000194v1_decoy                          0                          2
GRCh38.78/8_KI270813v1_alt.regions does not exist
GRCh38.78/HLA-B*82:02:01.regions does not exist
HLA-B*82:02:01                          0                          5
GRCh38.78/HLA-B*15:16:01.regions does not exist
HLA-B*15:16:01                          0                          3
GRCh38.78/Un_JTFH01001961v1_decoy.regions does not exist
GRCh38.78/19_KI270886v1_alt.regions does not exist
GRCh38.78/1_GL383520v2_alt.regions does not exist
Un_JTFH01001961v1_decoy                          0                         12
GRCh38.78/1_KI270760v1_alt.regions does not exist
GRCh38.78/4_GL383527v1_alt.regions does not exist
1_GL383520v2_alt                           0                          1
8_KI270813v1_alt                           0                         72
GRCh38.78/19_KI270883v1_alt.regions does not exist
GRCh38.78/14_KI270845v1_alt.regions does not exist
GRCh38.78/HLA-B*15:04:01.regions does not exist
HLA-B*15:04:01                          0                          4
19_KI270886v1_alt                            0                        100
GRCh38.78/HLA-B*40:01:01.regions does not exist
1_KI270760v1_alt                           0                          1
HLA-B*40:01:01                          0                          4
GRCh38.78/16_GL383556v1_alt.regions does not exist
19_KI270883v1_alt                            0                         99
4_GL383527v1_alt                           0                          4
GRCh38.78/13_KI270839v1_alt.regions does not exist
14_KI270845v1_alt                            0                          4
GRCh38.78/6_KI270800v1_alt.regions does not exist
16_GL383556v1_alt                            0                         36
GRCh38.78/19_KI270890v1_alt.regions does not exist
GRCh38.78/HLA-A*01:11N.regions does not exist
GRCh38.78/HLA-A*26:15.regions does not exist
GRCh38.78/11_KI270831v1_alt.regions does not exist
13_KI270839v1_alt                            0                          2
HLA-A*01:11N                         0                          7
HLA-A*26:15                          0                         2
```

```
GRCh38.78/HLA-C*04:128.regions does not exist
HLA-C*04:128                        0                        1
GRCh38.78/Un_KN707645v1_decoy.regions does not exist
Un_KN707645v1_decoy                         0                         1
GRCh38.78/12_KI270833v1_alt.regions does not exist
6_KI270800v1_alt                       0                       3
11_KI270831v1_alt                       0                       16
19_KI270890v1_alt                       0                       66
12_KI270833v1_alt                       0                       6
GRCh38.78/Un_JTFH01001973v1_decoy.regions does not exist
Un_JTFH01001973v1_decoy                         0                         2
GRCh38.78/Un_KN707626v1_decoy.regions does not exist
GRCh38.78/14_KI270847v1_alt.regions does not exist
Un_KN707626v1_decoy                         0                         4
14_KI270847v1_alt                       0                       66
GRCh38.78/7_KI270803v1_alt.regions does not exist
GRCh38.78/15_GL383554v1_alt.regions does not exist
GRCh38.78/HLA-A*24:07:01.regions does not exist
HLA-A*24:07:01                      0                      14
GRCh38.78/Un_JTFH01000133v1_decoy.regions does not exist
Un_JTFH01000133v1_decoy                         0                         1
15_GL383554v1_alt                       0                       1
GRCh38.78/14_KI270846v1_alt.regions does not exist
GRCh38.78/19_GL000209v2_alt.regions does not exist
GRCh38.78/16_GL383557v1_alt.regions does not exist
16_GL383557v1_alt                       0                       1
7_KI270803v1_alt                       0                       179
GRCh38.78/Un_JTFH01000561v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000796v1_decoy.regions does not exist
GRCh38.78/HLA-B*39:10:01.regions does not exist
Un_JTFH01000561v1_decoy                         0                         2
19_GL000209v2_alt                       0                       52
HLA-B*39:10:01                      0                      2
GRCh38.78/5_GL339449v2_alt.regions does not exist
Un_JTFH01000796v1_decoy                         0                         5
GRCh38.78/3_KI270934v1_alt.regions does not exist
GRCh38.78/11_KI270830v1_alt.regions does not exist
3_KI270934v1_alt                       0                       61
GRCh38.78/HLA-A*02:95.regions does not exist
14_KI270846v1_alt                       0                       160
HLA-A*02:95                      0                      9
GRCh38.78/19_KI270938v1_alt.regions does not exist
GRCh38.78/Un_KN707925v1_decoy.regions does not exist
Un_KN707925v1_decoy                         0                         3
GRCh38.78/15_KI270849v1_alt.regions does not exist
GRCh38.78/Un_KI270751v1.regions does not exist
5_GL339449v2_alt                       0                       25
GRCh38.78/Un_KI270467v1.regions does not exist
```

```
11_KI270830v1_alt                          0                          26
GRCh38.78/19_GL949753v2_alt.regions does not exist
GRCh38.78/22_KI270877v1_alt.regions does not exist
19_KI270938v1_alt                          0                         274
GRCh38.78/Un_JTFH01001394v1_decoy.regions does not exist
Un_JTFH01001394v1_decoy                    0                           1
Un_KI270751v1                          0                          41
19_GL949753v2_alt                          0                         164
22_KI270877v1_alt                          0                           8
Un_KI270467v1                          0                          76
GRCh38.78/HLA-DQA1*03:02.regions does not exist
HLA-DQA1*03:02                          0                          18
GRCh38.78/HLA-A*03:01:01:01.regions does not exist
HLA-A*03:01:01:01                          0                           3
GRCh38.78/HLA-B*44:56N.regions does not exist
GRCh38.78/HLA-C*12:02:02.regions does not exist
HLA-C*12:02:02                          0                           3
GRCh38.78/2_KI270776v1_alt.regions does not exist
GRCh38.78/19_GL949748v2_alt.regions does not exist
HLA-B*44:56N                           0                           5
19_GL949748v2_alt                          0                          95
GRCh38.78/Un_KN707896v1_decoy.regions does not exist
Un_KN707896v1_decoy                         0                          39
GRCh38.78/HLA-A*11:69N.regions does not exist
GRCh38.78/HLA-B*49:32.regions does not exist
HLA-B*49:32                           0                          5
HLA-A*11:69N                           0                          10
GRCh38.78/HLA-A*31:01:02.regions does not exist
HLA-A*31:01:02                          0                           3
GRCh38.78/HLA-A*02:89.regions does not exist
HLA-A*02:89                           0                          8
GRCh38.78/Un_KN707687v1_decoy.regions does not exist
GRCh38.78/17_KI270907v1_alt.regions does not exist
GRCh38.78/9_KI270720v1_random.regions does not exist
17_KI270907v1_alt                          0                           3
Un_KN707687v1_decoy                         0                           3
GRCh38.78/Un_JTFH01001957v1_decoy.regions does not exist
Un_JTFH01001957v1_decoy                    0                           1
GRCh38.78/HLA-A*02:533.regions does not exist
15_KI270849v1_alt                          0                          12
2_KI270776v1_alt                         0                          13
HLA-A*02:533                           0                          9
GRCh38.78/Un_JTFH01001099v1_decoy.regions does not exist
Un_JTFH01001099v1_decoy                    0                           1
GRCh38.78/22_KI270732v1_random.regions does not exist
GRCh38.78/Un_KN707661v1_decoy.regions does not exist
Un_KN707661v1_decoy                         0                           7
GRCh38.78/HLA-A*02:01:01:02L.regions does not exist
```

```
HLA-A*02:01:01:02L                          0                          8
GRCh38.78/HLA-B*13:02:09.regions does not exist
HLA-B*13:02:09                              0                          2
GRCh38.78/HLA-C*05:93.regions does not exist
HLA-C*05:93                                 0                          1
GRCh38.78/Un_KI270750v1.regions does not exist
9_KI270720v1_random                          0                         17
22_KI270732v1_random                         0                         18
GRCh38.78/HLA-C*07:49.regions does not exist
GRCh38.78/HLA-DQB1*02:01:01.regions does not exist
HLA-DQB1*02:01:01                           0                          4
HLA-C*07:49                                 0                          2
GRCh38.78/5_KI270793v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000851v1_decoy.regions does not exist
GRCh38.78/HLA-A*68:22.regions does not exist
HLA-A*68:22                                 0                          7
Un_KI270750v1                                0                         5
GRCh38.78/Un_JTFH01001271v1_decoy.regions does not exist
Un_JTFH01001271v1_decoy                                 0                  1
Un_JTFH01000851v1_decoy                                 0                  1
GRCh38.78/Un_JTFH01000258v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000212v1_decoy.regions does not exist
Un_JTFH01000212v1_decoy                                 0                  1
Un_JTFH01000258v1_decoy                                 0                  1
GRCh38.78/HLA-B*55:01:01.regions does not exist
HLA-B*55:01:01                              0                          2
GRCh38.78/Un_JTFH01000396v1_decoy.regions does not exist
Un_JTFH01000396v1_decoy                                 0                  8
GRCh38.78/HLA-B*13:02:03.regions does not exist
HLA-B*13:02:03                              0                          2
5_KI270793v1_alt                             0                         5
GRCh38.78/Un_KN707970v1_decoy.regions does not exist
GRCh38.78/HLA-A*68:03:01.regions does not exist
GRCh38.78/HLA-DRB1*12:17.regions does not exist
HLA-DRB1*12:17                              0                          4
HLA-A*68:03:01                              0                          7
Un_KN707970v1_decoy                          0                         6
GRCh38.78/HLA-A*02:02:01.regions does not exist
GRCh38.78/HLA-B*13:02:01.regions does not exist
HLA-A*02:02:01                              0                          8
GRCh38.78/HLA-A*68:17.regions does not exist
HLA-B*13:02:01                              0                          2
HLA-A*68:17                                 0                          7
GRCh38.78/HLA-B*35:05:01.regions does not exist
HLA-B*35:05:01                              0                          1
GRCh38.78/HLA-A*24:02:03Q.regions does not exist
GRCh38.78/Un_JTFH01000269v1_decoy.regions does not exist
HLA-A*24:02:03Q                              0                         14
```

```
GRCh38.78/HLA-B*55:01:03.regions does not exist
Un_JTFH01000269v1_decoy                         0                                1
HLA-B*55:01:03                          0                          2
GRCh38.78/HLA-A*31:01:23.regions does not exist
HLA-A*31:01:23                          0                          1
GRCh38.78/HLA-DRB1*15:03:01:01.regions does not exist
HLA-DRB1*15:03:01:01                          0                          2
GRCh38.78/Un_JTFH01001021v1_decoy.regions does not exist
GRCh38.78/HLA-DRB1*15:03:01:02.regions does not exist
HLA-DRB1*15:03:01:02                          0                          2
GRCh38.78/HLA-DRB1*16:02:01.regions does not exist
HLA-DRB1*16:02:01                         0                          4
GRCh38.78/HLA-C*16:02:01.regions does not exist
HLA-C*16:02:01                          0                          3
Un_JTFH01001021v1_decoy                         0                                1
GRCh38.78/Un_JTFH01000981v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01001613v1_decoy.regions does not exist
GRCh38.78/HLA-A*03:01:01:03.regions does not exist
Un_JTFH01001613v1_decoy                         0                                1
HLA-A*03:01:01:03                         0                          1
Un_JTFH01000981v1_decoy                         0                                2
GRCh38.78/HLA-B*40:02:01.regions does not exist
HLA-B*40:02:01                          0                          5
GRCh38.78/Un_JTFH01000423v1_decoy.regions does not exist
GRCh38.78/HLA-C*08:03:01.regions does not exist
GRCh38.78/HLA-B*08:08N.regions does not exist
HLA-C*08:03:01                          0                          3
Un_JTFH01000423v1_decoy                         0                                4
GRCh38.78/19_GL383574v1_alt.regions does not exist
HLA-B*08:08N                         0                         3
GRCh38.78/HLA-DRB1*01:02:01.regions does not exist
GRCh38.78/Un_KI270337v1.regions does not exist
GRCh38.78/Un_JTFH01000317v1_decoy.regions does not exist
HLA-DRB1*01:02:01                         0                          4
GRCh38.78/HLA-A*02:53N.regions does not exist
GRCh38.78/3_GL000221v1_random.regions does not exist
GRCh38.78/HLA-B*39:14.regions does not exist
HLA-A*02:53N                          0                         10
Un_JTFH01000317v1_decoy                         0                                2
HLA-B*39:14                        0                         2
19_GL383574v1_alt                         0                          7
GRCh38.78/HLA-A*02:32N.regions does not exist
GRCh38.78/6_GL383533v1_alt.regions does not exist
HLA-A*02:32N                          0                          9
GRCh38.78/7_KI270807v1_alt.regions does not exist
GRCh38.78/Un_KN707969v1_decoy.regions does not exist
GRCh38.78/8_KI270901v1_alt.regions does not exist
Un_KN707969v1_decoy                           0                           13
```

```
GRCh38.78/HLA-B*53:01:01.regions does not exist
GRCh38.78/17_KI270862v1_alt.regions does not exist
7_KI270807v1_alt                        0                          4
GRCh38.78/HLA-B*40:03.regions does not exist
HLA-B*40:03                        0                          4
HLA-B*53:01:01                        0                          1
6_GL383533v1_alt                        0                          7
GRCh38.78/HLA-DQB1*03:03:02:03.regions does not exist
HLA-DQB1*03:03:02:03                          0                          1
GRCh38.78/Un_JTFH01000280v1_decoy.regions does not exist
Un_JTFH01000280v1_decoy                          0                          8
GRCh38.78/HLA-DQB1*03:03:02:01.regions does not exist
HLA-DQB1*03:03:02:01                          0                          1
Un_KI270337v1                      0                         21
GRCh38.78/Un_JTFH01001946v1_decoy.regions does not exist
Un_JTFH01001946v1_decoy                          0                          6
3_GL000221v1_random                         0                          4
GRCh38.78/Un_JTFH01001184v1_decoy.regions does not exist
Un_JTFH01001184v1_decoy                          0                          1
GRCh38.78/HLA-C*16:04:01.regions does not exist
HLA-C*16:04:01                        0                          3
8_KI270901v1_alt                        0                          1
GRCh38.78/HLA-A*68:01:01:02.regions does not exist
17_KI270862v1_alt                         0                         23
HLA-A*68:01:01:02                         0                          8
GRCh38.78/Un_KI270465v1.regions does not exist
GRCh38.78/HLA-DQB1*03:03:02:02.regions does not exist
HLA-DQB1*03:03:02:02                          0                          1
Un_KI270465v1                      0                          5
GRCh38.78/HLA-DQB1*03:05:01.regions does not exist
HLA-DQB1*03:05:01                         0                          2
GRCh38.78/8_KI270821v1_alt.regions does not exist
GRCh38.78/HLA-DQA1*06:01:01.regions does not exist
HLA-DQA1*06:01:01                         0                          6
GRCh38.78/Un_JTFH01000342v1_decoy.regions does not exist
GRCh38.78/HLA-A*68:01:01:01.regions does not exist
GRCh38.78/HLA-DRB1*15:01:01:03.regions does not exist
HLA-DRB1*15:01:01:03                          0                          2
GRCh38.78/HLA-DRB1*15:01:01:01.regions does not exist
HLA-A*68:01:01:01                         0                          7
HLA-DRB1*15:01:01:01                          0                          2
8_KI270821v1_alt                        0                         50
Un_JTFH01000342v1_decoy                          0                          2
GRCh38.78/3_KI270779v1_alt.regions does not exist
GRCh38.78/HLA-DRB1*15:01:01:04.regions does not exist
HLA-DRB1*15:01:01:04                          0                          2
GRCh38.78/7_KI270809v1_alt.regions does not exist
GRCh38.78/HLA-B*50:01:01.regions does not exist
```

```
HLA-B*50:01:01                          0                          5
GRCh38.78/Un_KN707885v1_decoy.regions does not exist
Un_KN707885v1_decoy                     0                          1
GRCh38.78/22_KI270736v1_random.regions does not exist
GRCh38.78/HLA-DRB1*15:01:01:02.regions does not exist
HLA-DRB1*15:01:01:02                    0                          2
GRCh38.78/Un_JTFH01000150v1_decoy.regions does not exist
Un_JTFH01000150v1_decoy                 0                          5
3_KI270779v1_alt                        0                         64
GRCh38.78/Un_JTFH01001884v1_decoy.regions does not exist
7_KI270809v1_alt                        0                         14
GRCh38.78/6_KI270798v1_alt.regions does not exist
GRCh38.78/HLA-B*35:01:01:02.regions does not exist
GRCh38.78/19_KI270867v1_alt.regions does not exist
GRCh38.78/HLA-A*24:08.regions does not exist
22_KI270736v1_random                    0                        474
HLA-A*24:08                             0                         14
GRCh38.78/HLA-A*24:20.regions does not exist
HLA-A*24:20                             0                         14
GRCh38.78/HLA-B*44:150.regions does not exist
HLA-B*44:150                            0                          5
GRCh38.78/HLA-B*35:01:01:01.regions does not exist
HLA-B*35:01:01:02                       0                          2
HLA-B*35:01:01:01                       0                          2
GRCh38.78/2_KI270773v1_alt.regions does not exist
19_KI270867v1_alt                       0                         12
GRCh38.78/Un_JTFH01001040v1_decoy.regions does not exist
Un_JTFH01001040v1_decoy                 0                          1
Un_JTFH01001884v1_decoy                 0                          1
GRCh38.78/Un_KN707966v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000153v1_decoy.regions does not exist
GRCh38.78/HLA-A*03:36N.regions does not exist
HLA-A*03:36N                            0                          2
6_KI270798v1_alt                        0                         14
2_KI270773v1_alt                        0                          2
GRCh38.78/2_KI270774v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000517v1_decoy.regions does not exist
Un_JTFH01000517v1_decoy                 0                          5
GRCh38.78/HLA-A*03:11N.regions does not exist
HLA-A*03:11N                            0                          3
2_KI270774v1_alt                        0                         13
GRCh38.78/Un_JTFH01001039v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000843v1_decoy.regions does not exist
Un_JTFH01000843v1_decoy                 0                         19
Un_KN707966v1_decoy                     0                         11
Un_JTFH01001039v1_decoy                 0                          4
GRCh38.78/HLA-C*07:384.regions does not exist
Un_JTFH01000153v1_decoy                 0                          1
```

```
GRCh38.78/HLA-B*39:01:03.regions does not exist
HLA-B*39:01:03                          0                          2
GRCh38.78/HLA-B*39:34.regions does not exist
HLA-B*39:34                           0                        2
GRCh38.78/Un_JTFH01000098v1_decoy.regions does not exist
HLA-C*07:384                           0                         2
Un_JTFH01000098v1_decoy                           0                          1
GRCh38.78/HLA-DQA1*04:01:02:01.regions does not exist
HLA-DQA1*04:01:02:01                            0                           5
GRCh38.78/HLA-C*07:392.regions does not exist
HLA-C*07:392                           0                         2
GRCh38.78/Un_KN707649v1_decoy.regions does not exist
Un_KN707649v1_decoy                            0                           1
GRCh38.78/1_KI270759v1_alt.regions does not exist
GRCh38.78/HLA-B*39:13:02.regions does not exist
HLA-B*39:13:02                          0                          2
GRCh38.78/HLA-B*39:01:16.regions does not exist
HLA-B*39:01:16                          0                          2
GRCh38.78/17_KI270909v1_alt.regions does not exist
GRCh38.78/HLA-C*07:149.regions does not exist
HLA-C*07:149                           0                         2
17_KI270909v1_alt                            0                          31
GRCh38.78/HLA-A*29:46.regions does not exist
1_KI270759v1_alt                           0                         5
HLA-A*29:46                           0                        9
GRCh38.78/Un_JTFH01001478v1_decoy.regions does not exist
Un_JTFH01001478v1_decoy                           0                          3
GRCh38.78/HLA-B*27:131.regions does not exist
HLA-B*27:131                           0                         5
GRCh38.78/HLA-B*15:02:01.regions does not exist
GRCh38.78/Un_JTFH01000136v1_decoy.regions does not exist
HLA-B*15:02:01                          0                          4
Un_JTFH01000136v1_decoy                           0                          1
GRCh38.78/5_KI270795v1_alt.regions does not exist
GRCh38.78/HLA-A*34:01:01.regions does not exist
HLA-A*34:01:01                          0                          5
GRCh38.78/HLA-A*36:01.regions does not exist
GRCh38.78/Un_JTFH01001390v1_decoy.regions does not exist
GRCh38.78/HLA-B*39:01:21.regions does not exist
Un_JTFH01001390v1_decoy                           0                          1
HLA-A*36:01                           0                        5
HLA-B*39:01:21                          0                          2
GRCh38.78/HLA-DQA1*04:01:02:02.regions does not exist
HLA-DQA1*04:01:02:02                            0                           5
GRCh38.78/Un_JTFH01000628v1_decoy.regions does not exist
GRCh38.78/Un_KN707883v1_decoy.regions does not exist
GRCh38.78/HLA-C*05:09:01.regions does not exist
5_KI270795v1_alt                           0                         7
```

```
HLA-C*05:09:01                          0                          2
Un_JTFH01000628v1_decoy                         0                          22
GRCh38.78/HLA-B*40:40.regions does not exist
HLA-B*40:40                        0                        4
GRCh38.78/HLA-B*35:241.regions does not exist
HLA-B*35:241                       0                        2
Un_KN707883v1_decoy                          0                        1
GRCh38.78/Un_KI270508v1.regions does not exist
GRCh38.78/Un_JTFH01000999v1_decoy.regions does not exist
GRCh38.78/HLA-DQA1*03:03:01.regions does not exist
Un_JTFH01000999v1_decoy                         0                          8
HLA-DQA1*03:03:01                       0                         17
GRCh38.78/HLA-A*29:02:01:01.regions does not exist
HLA-A*29:02:01:01                       0                         9
Un_KI270508v1                        0                         16
GRCh38.78/Un_JTFH01000340v1_decoy.regions does not exist
GRCh38.78/20_KI270870v1_alt.regions does not exist
GRCh38.78/HLA-A*29:02:01:02.regions does not exist
GRCh38.78/Un_JTFH01000870v1_decoy.regions does not exist
Un_JTFH01000870v1_decoy                         0                          4
HLA-A*29:02:01:02                       0                         7
GRCh38.78/HLA-B*39:05:01.regions does not exist
HLA-B*39:05:01                          0                          2
20_KI270870v1_alt                       0                         19
GRCh38.78/HLA-B*40:150.regions does not exist
GRCh38.78/Un_JTFH01001086v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01001405v1_decoy.regions does not exist
Un_JTFH01000340v1_decoy                         0                          1
HLA-B*40:150                       0                        4
Un_JTFH01001405v1_decoy                         0                          3
GRCh38.78/HLA-B*07:33:01.regions does not exist
Un_JTFH01001086v1_decoy                         0                          3
GRCh38.78/Un_JTFH01000277v1_decoy.regions does not exist
GRCh38.78/20_KI270871v1_alt.regions does not exist
HLA-B*07:33:01                          0                          2
Un_JTFH01000277v1_decoy                         0                          4
GRCh38.78/HLA-C*17:01:01:01.regions does not exist
HLA-C*17:01:01:01                       0                         2
GRCh38.78/Un_JTFH01001045v1_decoy.regions does not exist
Un_JTFH01001045v1_decoy                         0                          2
20_KI270871v1_alt                       0                         3
GRCh38.78/10_GL383545v1_alt.regions does not exist
GRCh38.78/HLA-A*24:215.regions does not exist
HLA-A*24:215                       0                        12
GRCh38.78/21_GL383581v2_alt.regions does not exist
GRCh38.78/HLA-B*52:01:02.regions does not exist
GRCh38.78/HLA-C*17:01:01:02.regions does not exist
HLA-C*17:01:01:02                       0                         2
```

```
HLA-B*52:01:02                          0                          1
10_GL383545v1_alt                          0                          12
GRCh38.78/Un_KI270519v1.regions does not exist
GRCh38.78/Un_JTFH01001102v1_decoy.regions does not exist
GRCh38.78/HLA-B*15:32:01.regions does not exist
HLA-B*15:32:01                          0                          4
21_GL383581v2_alt                          0                          6
Un_JTFH01001102v1_decoy                          0                          1
GRCh38.78/Un_KN707647v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000899v1_decoy.regions does not exist
GRCh38.78/HLA-B*42:08.regions does not exist
HLA-B*42:08                          0                          4
Un_JTFH01000899v1_decoy                          0                          2
GRCh38.78/19_KI270922v1_alt.regions does not exist
Un_KI270519v1                          0                          117
GRCh38.78/Un_JTFH01001305v1_decoy.regions does not exist
Un_JTFH01001305v1_decoy                          0                          2
GRCh38.78/EBV.regions does not exist
GRCh38.78/Un_JTFH01000366v1_decoy.regions does not exist
EBV                          0                          9                          Comple
GRCh38.78/Un_JTFH01001337v1_decoy.regions does not exist
Un_JTFH01000366v1_decoy                          0                          2
Un_KN707647v1_decoy                          0                          1
Un_JTFH01001337v1_decoy                          0                          3
GRCh38.78/22_KI270879v1_alt.regions does not exist
19_KI270922v1_alt                          0                          107
GRCh38.78/HLA-B*42:02.regions does not exist
HLA-B*42:02                          0                          4
GRCh38.78/19_KI270923v1_alt.regions does not exist
22_KI270879v1_alt                          0                          30
GRCh38.78/Un_JTFH01000732v1_decoy.regions does not exist
Un_JTFH01000732v1_decoy                          0                          22
GRCh38.78/14_GL000194v1_random.regions does not exist
19_KI270923v1_alt                          0                          68
GRCh38.78/Un_JTFH01000242v1_decoy.regions does not exist
Un_JTFH01000242v1_decoy                          0                          2
GRCh38.78/HLA-DQB1*06:01:01.regions does not exist
HLA-DQB1*06:01:01                          0                          6
14_GL000194v1_random                          0                          38
GRCh38.78/Un_KI270583v1.regions does not exist
GRCh38.78/Un_JTFH01000528v1_decoy.regions does not exist
Un_JTFH01000528v1_decoy                          0                          1
Un_KI270583v1                          0                          1
GRCh38.78/19_KI270865v1_alt.regions does not exist
GRCh38.78/Un_GL000214v1.regions does not exist
19_KI270865v1_alt                          0                          19
GRCh38.78/HLA-A*33:01:01.regions does not exist
HLA-A*33:01:01                          0                          3
```

```
GRCh38.78/Un_JTFH01000477v1_decoy.regions does not exist
GRCh38.78/HLA-B*40:79.regions does not exist
HLA-B*40:79                         0                         4
GRCh38.78/10_GL383546v1_alt.regions does not exist
GRCh38.78/Un_JTFH01001056v1_decoy.regions does not exist
Un_JTFH01000477v1_decoy                         0                         1
GRCh38.78/Un_KN707828v1_decoy.regions does not exist
Un_JTFH01001056v1_decoy                         0                         1
Un_KN707828v1_decoy                           0                        34
Un_GL000214v1                          0                        15
GRCh38.78/21_GL383580v2_alt.regions does not exist
GRCh38.78/HLA-DRB1*09:21.regions does not exist
HLA-DRB1*09:21                          0                         3
GRCh38.78/22_GL383582v2_alt.regions does not exist
21_GL383580v2_alt                           0                         1
10_GL383546v1_alt                           0                         6
GRCh38.78/HLA-A*24:02:10.regions does not exist
HLA-A*24:02:10                          0                        14
GRCh38.78/HLA-A*32:06.regions does not exist
GRCh38.78/Un_JTFH01000645v1_decoy.regions does not exist
HLA-A*32:06                         0                        10
22_GL383582v2_alt                           0                        20
GRCh38.78/Un_JTFH01000799v1_decoy.regions does not exist
Un_JTFH01000645v1_decoy                         0                         9
Un_JTFH01000799v1_decoy                         0                         5
GRCh38.78/HLA-B*56:01:01.regions does not exist
HLA-B*56:01:01                          0                         3
GRCh38.78/HLA-B*15:01:01:01.regions does not exist
HLA-B*15:01:01:01                           0                         4
GRCh38.78/19_KI270921v1_alt.regions does not exist
GRCh38.78/HLA-DQA1*01:01:02.regions does not exist
HLA-DQA1*01:01:02                           0                        10
GRCh38.78/Un_JTFH01000191v1_decoy.regions does not exist
Un_JTFH01000191v1_decoy                         0                         6
GRCh38.78/22_KI270878v1_alt.regions does not exist
22_KI270878v1_alt                           0                         8
GRCh38.78/HLA-C*07:02:01:02.regions does not exist
HLA-C*07:02:01:02                           0                         2
19_KI270921v1_alt                           0                        78
GRCh38.78/HLA-DQA1*03:01:01.regions does not exist
HLA-DQA1*03:01:01                           0                        16
GRCh38.78/HLA-C*07:02:01:01.regions does not exist
HLA-C*07:02:01:01                           0                         2
GRCh38.78/Un_JTFH01000510v1_decoy.regions does not exist
Un_JTFH01000510v1_decoy                         0                         2
GRCh38.78/HLA-A*32:01:01.regions does not exist
GRCh38.78/Un_JTFH01001237v1_decoy.regions does not exist
HLA-A*32:01:01                          0                        10
```

```
Un_JTFH01001237v1_decoy                          0                          3
GRCh38.78/Un_JTFH01000112v1_decoy.regions does not exist
Un_JTFH01000112v1_decoy                          0                          4
GRCh38.78/Un_JTFH01000323v1_decoy.regions does not exist
Un_JTFH01000323v1_decoy                          0                          5
GRCh38.78/7_GL383534v2_alt.regions does not exist
GRCh38.78/Un_JTFH01001002v1_decoy.regions does not exist
Un_JTFH01001002v1_decoy                          0                          1
GRCh38.78/22_GL383583v2_alt.regions does not exist
GRCh38.78/22_KI270875v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000383v1_decoy.regions does not exist
22_KI270875v1_alt                              0                            22
GRCh38.78/Un_JTFH01000802v1_decoy.regions does not exist
GRCh38.78/Un_KN707904v1_decoy.regions does not exist
Un_JTFH01000802v1_decoy                          0                          2
Un_JTFH01000383v1_decoy                          0                          6
GRCh38.78/HLA-B*58:01:01.regions does not exist
GRCh38.78/HLA-C*07:32N.regions does not exist
HLA-B*58:01:01                             0                            2
HLA-C*07:32N                             0                            2
22_GL383583v2_alt                              0                            8
GRCh38.78/Un_JTFH01000458v1_decoy.regions does not exist
Un_JTFH01000458v1_decoy                          0                          1
Un_KN707904v1_decoy                              0                            2
GRCh38.78/Un_KI270517v1.regions does not exist
7_GL383534v2_alt                               0                            7
GRCh38.78/HLA-B*54:18.regions does not exist
GRCh38.78/HLA-C*17:01:01:03.regions does not exist
HLA-B*54:18                              0                            2
HLA-C*17:01:01:03                              0                            3
GRCh38.78/HLA-A*29:01:01:01.regions does not exist
HLA-A*29:01:01:01                              0                            7
Un_KI270517v1                                0                              9
GRCh38.78/14_KI270723v1_random.regions does not exist
GRCh38.78/14_KI270726v1_random.regions does not exist
GRCh38.78/17_GL000205v2_random.regions does not exist
GRCh38.78/HLA-A*02:06:01.regions does not exist
GRCh38.78/HLA-A*30:89.regions does not exist
HLA-A*30:89                              0                            5
GRCh38.78/Un_JTFH01000509v1_decoy.regions does not exist
Un_JTFH01000509v1_decoy                          0                          1
GRCh38.78/22_KI270876v1_alt.regions does not exist
HLA-A*02:06:01                               0                              8
GRCh38.78/HLA-A*01:01:38L.regions does not exist
HLA-A*01:01:38L                                0                              8
22_KI270876v1_alt                              0                            6
14_KI270726v1_random                             0                              4
GRCh38.78/HLA-C*14:21N.regions does not exist
```

```
HLA-C*14:21N                          0                          2
GRCh38.78/21_KI270872v1_alt.regions does not exist
17_GL000205v2_random                          0                          220
GRCh38.78/12_KI270835v1_alt.regions does not exist
GRCh38.78/Un_JTFH01001982v1_decoy.regions does not exist
Un_JTFH01001982v1_decoy                       0                            1
21_KI270872v1_alt                         0                          15
GRCh38.78/Un_JTFH01001430v1_decoy.regions does not exist
GRCh38.78/HLA-C*07:66.regions does not exist
Un_JTFH01001430v1_decoy                       0                            7
HLA-C*07:66                           0                          2
14_KI270723v1_random                          0                          31
GRCh38.78/Un_KN707642v1_decoy.regions does not exist
GRCh38.78/HLA-C*08:112.regions does not exist
HLA-C*08:112                          0                          2
Un_KN707642v1_decoy                           0                            1
GRCh38.78/HLA-C*07:67.regions does not exist
HLA-C*07:67                           0                          2
GRCh38.78/HLA-B*15:10:01.regions does not exist
GRCh38.78/19_KI270920v1_alt.regions does not exist
HLA-B*15:10:01                            0                            3
GRCh38.78/HLA-B*15:01:01:03.regions does not exist
12_KI270835v1_alt                         0                          21
GRCh38.78/14_KI270725v1_random.regions does not exist
GRCh38.78/Un_JTFH01000986v1_decoy.regions does not exist
Un_JTFH01000986v1_decoy                       0                            9
HLA-B*15:01:01:03                         0                          4
GRCh38.78/HLA-A*30:01:01.regions does not exist
HLA-A*30:01:01                            0                            6
19_KI270920v1_alt                         0                          98
GRCh38.78/Un_JTFH01001243v1_decoy.regions does not exist
Un_JTFH01001243v1_decoy                       0                            1
GRCh38.78/17_JH159146v1_alt.regions does not exist
GRCh38.78/HLA-A*24:86N.regions does not exist
HLA-A*24:86N                          0                          9
GRCh38.78/HLA-C*14:03.regions does not exist
HLA-C*14:03                           0                          3
GRCh38.78/HLA-A*01:04N.regions does not exist
HLA-A*01:04N                          0                          7
GRCh38.78/HLA-A*11:01:01.regions does not exist
HLA-A*11:01:01                            0                            9
GRCh38.78/Un_GL000216v2.regions does not exist
17_JH159146v1_alt                         0                          57
14_KI270725v1_random                            0                            5
GRCh38.78/Un_JTFH01000249v1_decoy.regions does not exist
Un_JTFH01000249v1_decoy                       0                            15
GRCh38.78/HLA-C*02:02:02:01.regions does not exist
GRCh38.78/HLA-A*74:01.regions does not exist
```

```
GRCh38.78/Un_JTFH01001465v1_decoy.regions does not exist
HLA-A*74:01                         0                         9
Un_JTFH01001465v1_decoy                       0                         9
HLA-C*02:02:02:01                         0                         1
Un_GL000216v2                       0                       527
GRCh38.78/HLA-C*02:02:02:02.regions does not exist
HLA-C*02:02:02:02                         0                         1
GRCh38.78/Un_KI270516v1.regions does not exist
GRCh38.78/Un_KI270429v1.regions does not exist
GRCh38.78/HLA-B*44:09.regions does not exist
HLA-B*44:09                         0                         5
GRCh38.78/HLA-B*18:01:01:02.regions does not exist
HLA-B*18:01:01:02                         0                         2
GRCh38.78/HLA-B*18:01:01:01.regions does not exist
HLA-B*18:01:01:01                         0                         2
GRCh38.78/Un_JTFH01000968v1_decoy.regions does not exist
Un_JTFH01000968v1_decoy                       0                         4
Un_KI270516v1                       0                         9
GRCh38.78/HLA-B*44:04.regions does not exist
HLA-B*44:04                         0                         5
Un_KI270429v1                       0                         3
GRCh38.78/HLA-C*02:11.regions does not exist
HLA-C*02:11                         0                         1
GRCh38.78/HLA-C*07:02:64.regions does not exist
HLA-C*07:02:64                       0                       2
GRCh38.78/Un_KN707887v1_decoy.regions does not exist
GRCh38.78/HLA-C*02:10.regions does not exist
GRCh38.78/Un_JTFH01000515v1_decoy.regions does not exist
HLA-C*02:10                         0                         1
Un_JTFH01000515v1_decoy                       0                         4
Un_KN707887v1_decoy                       0                       2
GRCh38.78/8_KI270900v1_alt.regions does not exist
GRCh38.78/19_GL383573v1_alt.regions does not exist
GRCh38.78/5_KI270898v1_alt.regions does not exist
8_KI270900v1_alt                        0                       4
GRCh38.78/HLA-A*11:01:18.regions does not exist
GRCh38.78/Un_GL000195v1.regions does not exist
HLA-A*11:01:18                       0                       9
19_GL383573v1_alt                       0                       46
GRCh38.78/18_GL383567v1_alt.regions does not exist
5_KI270898v1_alt                        0                       6
GRCh38.78/8_KI270926v1_alt.regions does not exist
18_GL383567v1_alt                       0                       11
GRCh38.78/Un_JTFH01000064v1_decoy.regions does not exist
Un_JTFH01000064v1_decoy                       0                       13
GRCh38.78/HLA-B*41:02:01.regions does not exist
HLA-B*41:02:01                       0                       4
8_KI270926v1_alt                        0                       5
```

```
GRCh38.78/Un_KN707876v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000017v1_decoy.regions does not exist
Un_JTFH01000017v1_decoy                    0                    1
Un_KN707876v1_decoy                     0                    1
GRCh38.78/Un_JTFH01001680v1_decoy.regions does not exist
Un_GL000195v1                      0                   51
GRCh38.78/Un_KI270538v1.regions does not exist
Un_KI270538v1                      0                   33
GRCh38.78/19_KI270919v1_alt.regions does not exist
GRCh38.78/HLA-B*48:01:01.regions does not exist
HLA-B*48:01:01                      0                    7
GRCh38.78/Un_JTFH01001087v1_decoy.regions does not exist
GRCh38.78/9_KI270719v1_random.regions does not exist
GRCh38.78/HLA-A*03:01:01:02N.regions does not exist
Un_JTFH01001680v1_decoy                    0                    1
GRCh38.78/Un_JTFH01001748v1_decoy.regions does not exist
Un_JTFH01001087v1_decoy                    0                    5
GRCh38.78/1_GL383519v1_alt.regions does not exist
GRCh38.78/3_KI270936v1_alt.regions does not exist
3_KI270936v1_alt                     0                   64
Un_JTFH01001748v1_decoy                    0                    1
GRCh38.78/Un_KI270591v1.regions does not exist
1_GL383519v1_alt                     0                   30
GRCh38.78/19_GL949750v2_alt.regions does not exist
19_GL949750v2_alt                      0                   93
Un_KI270591v1                      0                   38
HLA-A*03:01:01:02N                       0                    3
19_KI270919v1_alt                     0                  100
GRCh38.78/Un_JTFH01001132v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000126v1_decoy.regions does not exist
Un_JTFH01000126v1_decoy                    0                   15
Un_JTFH01001132v1_decoy                    0                    3
GRCh38.78/HLA-B*07:44.regions does not exist
HLA-B*07:44                        0                    3
9_KI270719v1_random                         0                   13
GRCh38.78/HLA-B*44:26.regions does not exist
HLA-B*44:26                        0                    6
GRCh38.78/HLA-B*56:04.regions does not exist
GRCh38.78/Un_JTFH01000845v1_decoy.regions does not exist
Un_JTFH01000845v1_decoy                    0                   14
GRCh38.78/17_KI270729v1_random.regions does not exist
HLA-B*56:04                        0                    2
GRCh38.78/HLA-B*56:03.regions does not exist
HLA-B*56:03                        0                    6
GRCh38.78/HLA-B*07:41.regions does not exist
HLA-B*07:41                        0                    2
GRCh38.78/HLA-B*44:03:02.regions does not exist
GRCh38.78/Un_JTFH01001233v1_decoy.regions does not exist
```

```
Un_JTFH01001233v1_decoy                          0                          2
HLA-B*44:03:02                          0                          5
17_KI270729v1_random                          0                        238
GRCh38.78/HLA-A*24:152.regions does not exist
HLA-A*24:152                          0                         12
GRCh38.78/6_GL000251v2_alt.regions does not exist
GRCh38.78/Un_JTFH01001960v1_decoy.regions does not exist
Un_JTFH01001960v1_decoy                          0                          2
GRCh38.78/Un_KN707884v1_decoy.regions does not exist
Un_KN707884v1_decoy                          0                          4
6_GL000251v2_alt                          0                        805
GRCh38.78/HLA-B*14:01:01.regions does not exist
HLA-B*14:01:01                          0                          3
GRCh38.78/1_KI270766v1_alt.regions does not exist
GRCh38.78/11_KI270829v1_alt.regions does not exist
GRCh38.78/HLA-B*07:50.regions does not exist
GRCh38.78/Un_GL000218v1.regions does not exist
GRCh38.78/HLA-B*44:46.regions does not exist
HLA-B*07:50                          0                          2
GRCh38.78/Un_JTFH01000660v1_decoy.regions does not exist
Un_JTFH01000660v1_decoy                          0                         19
GRCh38.78/4_KI270786v1_alt.regions does not exist
1_KI270766v1_alt                          0                         23
HLA-B*44:46                          0                          6
GRCh38.78/HLA-B*44:03:01.regions does not exist
HLA-B*44:03:01                          0                          5
GRCh38.78/19_KI270889v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000134v1_decoy.regions does not exist
Un_JTFH01000134v1_decoy                          0                          1
GRCh38.78/Un_JTFH01001893v1_decoy.regions does not exist
Un_GL000218v1                          0                          6
GRCh38.78/Un_JTFH01000459v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01001193v1_decoy.regions does not exist
Un_JTFH01000459v1_decoy                          0                          3
GRCh38.78/Un_JTFH01001317v1_decoy.regions does not exist
Un_JTFH01001317v1_decoy                          0                          6
Un_JTFH01001193v1_decoy                          0                          1
GRCh38.78/HLA-C*12:22.regions does not exist
HLA-C*12:22                          0                          3
GRCh38.78/HLA-B*39:01:01:03.regions does not exist
19_KI270889v1_alt                          0                        121
HLA-B*39:01:01:03                          0                          2
GRCh38.78/HLA-B*15:07:01.regions does not exist
GRCh38.78/17_GL383564v2_alt.regions does not exist
HLA-B*15:07:01                          0                          4
GRCh38.78/HLA-A*31:14N.regions does not exist
HLA-A*31:14N                          0                          1
GRCh38.78/HLA-B*39:01:01:01.regions does not exist
```

```
HLA-B*39:01:01:01                          0                          2
GRCh38.78/HLA-B*51:01:01.regions does not exist
HLA-B*51:01:01                             0                          1
Un_JTFH01001893v1_decoy                    0                          3
GRCh38.78/Un_JTFH01000123v1_decoy.regions does not exist
17_GL383564v2_alt                          0                          20
Un_JTFH01000123v1_decoy                    0                          12
GRCh38.78/12_GL383553v2_alt.regions does not exist
GRCh38.78/Un_JTFH01001974v1_decoy.regions does not exist
Un_JTFH01001974v1_decoy                    0                          3
11_KI270829v1_alt                          0                          13
12_GL383553v2_alt                          0                          2
GRCh38.78/HLA-A*24:11N.regions does not exist
HLA-A*24:11N                               0                          13
GRCh38.78/HLA-B*44:49.regions does not exist
HLA-B*44:49                                0                          6
4_KI270786v1_alt                           0                          4
GRCh38.78/4_KI270925v1_alt.regions does not exist
GRCh38.78/Un_JTFH01001941v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01001208v1_decoy.regions does not exist
Un_JTFH01001941v1_decoy                    0                          4
Un_JTFH01001208v1_decoy                    0                          7
4_KI270925v1_alt                           0                          167
GRCh38.78/HLA-B*81:01.regions does not exist
HLA-B*81:01                                0                          7
GRCh38.78/Un_JTFH01001418v1_decoy.regions does not exist
Un_JTFH01001418v1_decoy                    0                          2
GRCh38.78/6_KB021644v2_alt.regions does not exist
GRCh38.78/HLA-B*39:06:02.regions does not exist
HLA-B*39:06:02                             0                          2
GRCh38.78/HLA-C*12:03:01:02.regions does not exist
HLA-C*12:03:01:02                          0                          3
GRCh38.78/6_KI270799v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000050v1_decoy.regions does not exist
Un_JTFH01000050v1_decoy                    0                          9
6_KB021644v2_alt                           0                          43
GRCh38.78/HLA-DQA1*01:04:01:02.regions does not exist
HLA-DQA1*01:04:01:02                       0                          12
GRCh38.78/Un_JTFH01000179v1_decoy.regions does not exist
GRCh38.78/2_KI270772v1_alt.regions does not exist
Un_JTFH01000179v1_decoy                    0                          1
GRCh38.78/HLA-A*66:01:01.regions does not exist
GRCh38.78/8_KI270816v1_alt.regions does not exist
HLA-A*66:01:01                             0                          1
GRCh38.78/HLA-C*12:03:01:01.regions does not exist
GRCh38.78/Un_KN707641v1_decoy.regions does not exist
GRCh38.78/HLA-B*40:10:01.regions does not exist
HLA-B*40:10:01                             0                          4
```

```
GRCh38.78/Un_KI270756v1.regions does not exist
6_KI270799v1_alt                          0                          1
HLA-C*12:03:01:01                          0                          3
Un_KN707641v1_decoy                          0                          2
2_KI270772v1_alt                          0                        102
GRCh38.78/HLA-B*51:01:02.regions does not exist
GRCh38.78/Un_KI270743v1.regions does not exist
GRCh38.78/HLA-DQB1*06:02:01.regions does not exist
HLA-DQB1*06:02:01                          0                          7
HLA-B*51:01:02                        0                        1
8_KI270816v1_alt                          0                         34
GRCh38.78/Un_JTFH01001070v1_decoy.regions does not exist
Un_JTFH01001070v1_decoy                          0                          1
Un_KI270756v1                        0                       35
GRCh38.78/Un_JTFH01001018v1_decoy.regions does not exist
Un_JTFH01001018v1_decoy                          0                          4
GRCh38.78/HLA-A*33:03:01.regions does not exist
GRCh38.78/2_KI270769v1_alt.regions does not exist
HLA-A*33:03:01                        0                        3
2_KI270769v1_alt                          0                         2
GRCh38.78/HLA-C*12:13.regions does not exist
GRCh38.78/HLA-A*30:04:01.regions does not exist
HLA-A*30:04:01                        0                        6
GRCh38.78/Un_KI270438v1.regions does not exist
HLA-C*12:13                       0                      3
GRCh38.78/HLA-C*12:19.regions does not exist
HLA-C*12:19                       0                      3
GRCh38.78/1_KI270711v1_random.regions does not exist
GRCh38.78/5_KI270791v1_alt.regions does not exist
GRCh38.78/19_KI270868v1_alt.regions does not exist
Un_KI270743v1                        0                       2
GRCh38.78/Un_JTFH01000526v1_decoy.regions does not exist
Un_JTFH01000526v1_decoy                          0                          4
GRCh38.78/19_KI270931v1_alt.regions does not exist
GRCh38.78/8_KI270812v1_alt.regions does not exist
GRCh38.78/HLA-B*38:02:01.regions does not exist
19_KI270868v1_alt                          0                         13
HLA-B*38:02:01                        0                        2
5_KI270791v1_alt                          0                         17
19_KI270931v1_alt                          0                        108
1_KI270711v1_random                          0                         19
Un_KI270438v1                        0                      313
GRCh38.78/Un_JTFH01000619v1_decoy.regions does not exist
Un_JTFH01000619v1_decoy                          0                          3
8_KI270812v1_alt                          0                         12
GRCh38.78/HLA-A*11:77.regions does not exist
HLA-A*11:77                       0                      10
GRCh38.78/HLA-DQA1*01:04:01:01.regions does not exist
```

```
HLA-DQA1*01:04:01:01                          0                          12
GRCh38.78/15_KI270906v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000806v1_decoy.regions does not exist
GRCh38.78/Un_KI270320v1.regions does not exist
GRCh38.78/HLA-A*02:259.regions does not exist
Un_JTFH01000806v1_decoy                       0                           1
15_KI270906v1_alt                           0                           3
GRCh38.78/4_KI270896v1_alt.regions does not exist
GRCh38.78/HLA-A*24:02:01:01.regions does not exist
HLA-A*02:259                              0                           8
HLA-A*24:02:01:01                           0                          14
Un_KI270320v1                             0                           1
GRCh38.78/Un_JTFH01000543v1_decoy.regions does not exist
4_KI270896v1_alt                            0                          15
Un_JTFH01000543v1_decoy                       0                           1
GRCh38.78/HLA-B*07:06.regions does not exist
GRCh38.78/Un_KN707892v1_decoy.regions does not exist
Un_KN707892v1_decoy                           0                           1
GRCh38.78/Un_KI270512v1.regions does not exist
GRCh38.78/19_GL949747v2_alt.regions does not exist
HLA-B*07:06                             0                           2
GRCh38.78/HLA-B*39:38Q.regions does not exist
HLA-B*39:38Q                              0                           2
GRCh38.78/HLA-B*35:03:01.regions does not exist
HLA-B*35:03:01                            0                           1
19_GL949747v2_alt                           0                         137
GRCh38.78/HLA-B*15:01:01:02N.regions does not exist
HLA-B*15:01:01:02N                            0                           1
Un_KI270512v1                             0                           2
GRCh38.78/HLA-A*02:251.regions does not exist
HLA-A*02:251                              0                           8
GRCh38.78/HLA-A*24:02:01:03.regions does not exist
GRCh38.78/HLA-B*15:03:01.regions does not exist
HLA-B*15:03:01                            0                           3
HLA-A*24:02:01:03                           0                          11
GRCh38.78/X_KI270880v1_alt.regions does not exist
GRCh38.78/12_GL877876v1_alt.regions does not exist
GRCh38.78/HLA-DQA1*01:05:01.regions does not exist
HLA-DQA1*01:05:01                           0                          12
GRCh38.78/HLA-C*14:02:01.regions does not exist
HLA-C*14:02:01                            0                           2
X_KI270880v1_alt                            0                          72
GRCh38.78/15_KI270850v1_alt.regions does not exist
GRCh38.78/HLA-C*08:01:01.regions does not exist
HLA-C*08:01:01                            0                           3
GRCh38.78/Un_JTFH01001216v1_decoy.regions does not exist
12_GL877876v1_alt                           0                          31
15_KI270850v1_alt                           0                          27
```

```
GRCh38.78/Un_JTFH01001446v1_decoy.regions does not exist
Un_JTFH01001446v1_decoy                          0                          17
Un_JTFH01001216v1_decoy                          0                          2
GRCh38.78/HLA-B*15:220.regions does not exist
HLA-B*15:220                          0                          4
GRCh38.78/Un_JTFH01000227v1_decoy.regions does not exist
Un_JTFH01000227v1_decoy                          0                          1
GRCh38.78/Un_JTFH01000263v1_decoy.regions does not exist
Un_JTFH01000263v1_decoy                          0                          13
GRCh38.78/7_KI270899v1_alt.regions does not exist
GRCh38.78/Un_KI270754v1.regions does not exist
GRCh38.78/HLA-C*02:87.regions does not exist
HLA-C*02:87                          0                          1
GRCh38.78/15_GL383555v2_alt.regions does not exist
7_KI270899v1_alt                          0                          52
15_GL383555v2_alt                          0                          3
Un_KI270754v1                          0                          9
GRCh38.78/1_KI270761v1_alt.regions does not exist
GRCh38.78/HLA-B*15:83.regions does not exist
GRCh38.78/HLA-A*02:279.regions does not exist
HLA-B*15:83                          0                          3
GRCh38.78/Un_KN707968v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000274v1_decoy.regions does not exist
HLA-A*02:279                          0                          9
Un_KN707968v1_decoy                          0                          24
1_KI270761v1_alt                          0                          7
GRCh38.78/7_KI270805v1_alt.regions does not exist
7_KI270805v1_alt                          0                          11
Un_JTFH01000274v1_decoy                          0                          12
GRCh38.78/Un_JTFH01000929v1_decoy.regions does not exist
Un_JTFH01000929v1_decoy                          0                          3
GRCh38.78/Un_JTFH01001862v1_decoy.regions does not exist
GRCh38.78/HLA-B*44:02:01:02S.regions does not exist
HLA-B*44:02:01:02S                          0                          3
GRCh38.78/HLA-B*78:01:01.regions does not exist
Un_JTFH01001862v1_decoy                          0                          13
HLA-B*78:01:01                          0                          2
GRCh38.78/Un_JTFH01000420v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01001387v1_decoy.regions does not exist
Un_JTFH01001387v1_decoy                          0                          3
GRCh38.78/14_KI270844v1_alt.regions does not exist
Un_JTFH01000420v1_decoy                          0                          4
14_KI270844v1_alt                          0                          17
GRCh38.78/HLA-DQA1*05:05:01:03.regions does not exist
HLA-DQA1*05:05:01:03                          0                          4
GRCh38.78/5_KI270792v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000984v1_decoy.regions does not exist
Un_JTFH01000984v1_decoy                          0                          2
```

```
5_KI270792v1_alt                          0                          13
GRCh38.78/HLA-DQA1*05:05:01:01.regions does not exist
HLA-DQA1*05:05:01:01                      0                           4
GRCh38.78/17_KI270859v1_alt.regions does not exist
GRCh38.78/HLA-DQA1*02:01.regions does not exist
HLA-DQA1*02:01                            0                          13
GRCh38.78/HLA-B*15:77.regions does not exist
HLA-B*15:77                               0                           3
17_KI270859v1_alt                          0                          2
GRCh38.78/HLA-A*02:266.regions does not exist
HLA-A*02:266                              0                           8
GRCh38.78/Un_JTFH01001980v1_decoy.regions does not exist
Un_JTFH01001980v1_decoy                   0                           3
GRCh38.78/HLA-B*58:31N.regions does not exist
GRCh38.78/HLA-B*15:58.regions does not exist
HLA-B*58:31N                              0                           1
HLA-B*15:58                               0                           4
GRCh38.78/6_KI270802v1_alt.regions does not exist
GRCh38.78/1_KI270762v1_alt.regions does not exist
GRCh38.78/HLA-B*15:66.regions does not exist
HLA-B*15:66                               0                           3
GRCh38.78/Un_KI270333v1.regions does not exist
GRCh38.78/HLA-A*02:269.regions does not exist
HLA-A*02:269                              0                           8
6_KI270802v1_alt                          0                           2
GRCh38.78/HLA-A*02:264.regions does not exist
GRCh38.78/7_KI270806v1_alt.regions does not exist
1_KI270762v1_alt                          0                          43
HLA-A*02:264                              0                           8
GRCh38.78/12_GL877875v1_alt.regions does not exist
GRCh38.78/13_KI270840v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000181v1_decoy.regions does not exist
Un_JTFH01000181v1_decoy                   0                           1
GRCh38.78/HLA-DQA1*05:05:01:02.regions does not exist
HLA-DQA1*05:05:01:02                      0                           4
Un_KI270333v1                             0                          46
GRCh38.78/HLA-A*02:265.regions does not exist
13_KI270840v1_alt                          0                          1
HLA-A*02:265                              0                           7
7_KI270806v1_alt                          0                          17
12_GL877875v1_alt                          0                          10
GRCh38.78/HLA-B*27:32.regions does not exist
HLA-B*27:32                               0                           5
GRCh38.78/9_KI270823v1_alt.regions does not exist
GRCh38.78/HLA-B*35:01:22.regions does not exist
HLA-B*35:01:22                            0                           2
GRCh38.78/HLA-DQB1*05:01:01:01.regions does not exist
HLA-DQB1*05:01:01:01                      0                           1
```

```
GRCh38.78/15_KI270905v1_alt.regions does not exist
9_KI270823v1_alt                        0                        3
GRCh38.78/Un_KN707798v1_decoy.regions does not exist
Un_KN707798v1_decoy                       0                       1
15_KI270905v1_alt                       0                       91
GRCh38.78/HLA-C*03:04:04.regions does not exist
HLA-C*03:04:04                        0                       1
GRCh38.78/HLA-B*48:04.regions does not exist
HLA-B*48:04                       0                       7
GRCh38.78/Un_KN707862v1_decoy.regions does not exist
Un_KN707862v1_decoy                       0                       17
GRCh38.78/Un_JTFH01001345v1_decoy.regions does not exist
Un_JTFH01001345v1_decoy                         0                       1
GRCh38.78/HLA-A*23:38N.regions does not exist
GRCh38.78/Un_KN707638v1_decoy.regions does not exist
HLA-A*23:38N                      0                       12
Un_KN707638v1_decoy                       0                       2
GRCh38.78/HLA-DQB1*05:01:01:02.regions does not exist
HLA-DQB1*05:01:01:02                       0                       1
GRCh38.78/16_KI270856v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000730v1_decoy.regions does not exist
GRCh38.78/HLA-C*04:01:62.regions does not exist
Un_JTFH01000730v1_decoy                         0                       3
HLA-C*04:01:62                        0                       1
GRCh38.78/HLA-A*01:01:01:01.regions does not exist
HLA-A*01:01:01:01                       0                       7
GRCh38.78/16_KI270728v1_random.regions does not exist
GRCh38.78/HLA-C*08:02:01:02.regions does not exist
HLA-C*08:02:01:02                       0                       3
GRCh38.78/Un_KN707879v1_decoy.regions does not exist
GRCh38.78/HLA-B*27:06.regions does not exist
Un_KN707879v1_decoy                       0                       2
GRCh38.78/4_GL000008v2_random.regions does not exist
GRCh38.78/Un_JTFH01001084v1_decoy.regions does not exist
Un_JTFH01001084v1_decoy                         0                       5
GRCh38.78/HLA-B*27:25.regions does not exist
HLA-B*27:06                       0                       5
HLA-B*27:25                       0                       5
16_KI270856v1_alt                       0                       3
GRCh38.78/HLA-C*08:02:01:01.regions does not exist
HLA-C*08:02:01:01                       0                       3
GRCh38.78/Un_KN707881v1_decoy.regions does not exist
GRCh38.78/HLA-C*16:01:01.regions does not exist
HLA-C*16:01:01                       0                       3
GRCh38.78/11_KI270832v1_alt.regions does not exist
GRCh38.78/19_GL949746v1_alt.regions does not exist
Un_KN707881v1_decoy                       0                       1
GRCh38.78/Un_KI270510v1.regions does not exist
```

```
GRCh38.78/HLA-A*29:01:01:02N.regions does not exist
HLA-A*29:01:01:02N                          0                          6
11_KI270832v1_alt                          0                         74
GRCh38.78/HLA-DQA1*04:02.regions does not exist
HLA-DQA1*04:02                             0                          4
16_KI270728v1_random                         0                       21
GRCh38.78/Un_JTFH01000995v1_decoy.regions does not exist
GRCh38.78/HLA-C*02:16:02.regions does not exist
HLA-C*02:16:02                             0                          1
Un_JTFH01000995v1_decoy                            0                          10
GRCh38.78/HLA-C*04:03:01.regions does not exist
4_GL000008v2_random                          0                        168
HLA-C*04:03:01                             0                          2
GRCh38.78/Un_JTFH01000717v1_decoy.regions does not exist
GRCh38.78/HLA-B*44:138Q.regions does not exist
19_GL949746v1_alt                          0                        209
HLA-B*44:138Q                            0                          3
Un_KI270510v1                            0                          5
GRCh38.78/19_KI270884v1_alt.regions does not exist
Un_JTFH01000717v1_decoy                            0                          2
GRCh38.78/Un_JTFH01000302v1_decoy.regions does not exist
GRCh38.78/HLA-C*02:85.regions does not exist
HLA-C*02:85                              0                          1
GRCh38.78/Un_KI270589v1.regions does not exist
GRCh38.78/HLA-B*27:24.regions does not exist
HLA-B*27:24                              0                          5
GRCh38.78/HLA-B*15:42.regions does not exist
GRCh38.78/HLA-B*48:08.regions does not exist
Un_JTFH01000302v1_decoy                            0                          9
HLA-B*15:42                              0                          2
HLA-B*48:08                              0                          7
GRCh38.78/HLA-C*02:86.regions does not exist
HLA-C*02:86                              0                          1
GRCh38.78/Un_JTFH01000700v1_decoy.regions does not exist
GRCh38.78/8_KI270811v1_alt.regions does not exist
Un_JTFH01000700v1_decoy                            0                          1
GRCh38.78/HLA-C*08:01:03.regions does not exist
HLA-C*08:01:03                             0                          2
19_KI270884v1_alt                          0                         92
GRCh38.78/Un_JTFH01001991v1_decoy.regions does not exist
Un_JTFH01001991v1_decoy                            0                          3
8_KI270811v1_alt                           0                          4
GRCh38.78/Un_JTFH01001997v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000480v1_decoy.regions does not exist
Un_JTFH01001997v1_decoy                            0                          4
GRCh38.78/HLA-B*15:27:01.regions does not exist
HLA-B*15:27:01                             0                          3
Un_KI270589v1                            0                         63
```

```
GRCh38.78/HLA-C*07:02:05.regions does not exist
Un_JTFH01000480v1_decoy                         0                         12
GRCh38.78/HLA-C*14:23.regions does not exist
GRCh38.78/13_KI270838v1_alt.regions does not exist
HLA-C*07:02:05                          0                          2
GRCh38.78/HLA-DQA1*05:01:01:01.regions does not exist
HLA-DQA1*05:01:01:01                          0                          1
GRCh38.78/3_KI270935v1_alt.regions does not exist
GRCh38.78/HLA-DQA1*05:01:01:02.regions does not exist
GRCh38.78/HLA-B*07:05:01.regions does not exist
HLA-DQA1*05:01:01:02                          0                          1
HLA-C*14:23                         0                          2
GRCh38.78/HLA-C*12:99.regions does not exist
HLA-C*12:99                         0                          3
3_KI270935v1_alt                            0                         68
HLA-B*07:05:01                           0                          1
13_KI270838v1_alt                             0                          2
GRCh38.78/Un_JTFH01000319v1_decoy.regions does not exist
Un_JTFH01000319v1_decoy                         0                          5
GRCh38.78/HLA-C*07:02:06.regions does not exist
HLA-C*07:02:06                          0                          2
GRCh38.78/Un_KN707901v1_decoy.regions does not exist
Un_KN707901v1_decoy                           0                          4
GRCh38.78/1_KI270713v1_random.regions does not exist
GRCh38.78/Un_JTFH01000226v1_decoy.regions does not exist
Un_JTFH01000226v1_decoy                         0                          4
GRCh38.78/12_KI270834v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000241v1_decoy.regions does not exist
Un_JTFH01000241v1_decoy                         0                         23
GRCh38.78/19_KI270885v1_alt.regions does not exist
19_KI270885v1_alt                           0                         111
GRCh38.78/HLA-C*02:69.regions does not exist
HLA-C*02:69                         0                          1
12_KI270834v1_alt                           0                         12
GRCh38.78/HLA-B*38:14.regions does not exist
1_KI270713v1_random                           0                          16
HLA-B*38:14                         0                          2
GRCh38.78/17_KI270730v1_random.regions does not exist
GRCh38.78/HLA-B*40:72:01.regions does not exist
GRCh38.78/HLA-C*04:06.regions does not exist
HLA-C*04:06                         0                          2
HLA-B*40:72:01                          0                          4
GRCh38.78/Un_JTFH01000273v1_decoy.regions does not exist
Un_JTFH01000273v1_decoy                         0                          1
GRCh38.78/2_KI270894v1_alt.regions does not exist
GRCh38.78/19_GL949751v2_alt.regions does not exist
GRCh38.78/Un_JTFH01000977v1_decoy.regions does not exist
GRCh38.78/HLA-A*02:03:03.regions does not exist
```

```
HLA-A*02:03:03                          0                           8
Un_JTFH01000977v1_decoy                         0                           6
2_KI270894v1_alt                        0                         109
17_KI270730v1_random                        0                         230
19_GL949751v2_alt                       0                         150
GRCh38.78/22_KI270928v1_alt.regions does not exist
GRCh38.78/HLA-A*02:03:01.regions does not exist
HLA-A*02:03:01                          0                           8
22_KI270928v1_alt                       0                          19
GRCh38.78/Un_KI270588v1.regions does not exist
GRCh38.78/1_GL383518v1_alt.regions does not exist
GRCh38.78/HLA-A*03:21N.regions does not exist
HLA-A*03:21N                        0                           1
GRCh38.78/Un_KI270336v1.regions does not exist
GRCh38.78/6_GL000252v2_alt.regions does not exist
GRCh38.78/Un_JTFH01001206v1_decoy.regions does not exist
Un_KI270588v1                       0                           2
Un_JTFH01001206v1_decoy                         0                           2
Un_KI270336v1                       0                          13
GRCh38.78/17_GL000258v2_alt.regions does not exist
6_GL000252v2_alt                        0                         574
1_GL383518v1_alt                        0                           7
GRCh38.78/HLA-C*08:20.regions does not exist
HLA-C*08:20                         0                           2
GRCh38.78/HLA-C*08:22.regions does not exist
GRCh38.78/HLA-DQB1*06:09:01.regions does not exist
HLA-DQB1*06:09:01                       0                           6
17_GL000258v2_alt                       0                          13
GRCh38.78/HLA-A*24:02:01:02L.regions does not exist
HLA-A*24:02:01:02L                      0                          14
GRCh38.78/HLA-C*08:24.regions does not exist
HLA-C*08:24                         0                           2
GRCh38.78/3_KI270924v1_alt.regions does not exist
GRCh38.78/3_KI270937v1_alt.regions does not exist
GRCh38.78/HLA-C*08:21.regions does not exist
HLA-C*08:21                         0                           2
3_KI270937v1_alt                        0                          75
3_KI270924v1_alt                        0                          67
HLA-C*08:22                         0                           2
GRCh38.78/Un_JTFH01001147v1_decoy.regions does not exist
Un_JTFH01001147v1_decoy                         0                           3
GRCh38.78/HLA-DQB1*03:01:01:02.regions does not exist
HLA-DQB1*03:01:01:02                        0                          11
GRCh38.78/15_KI270851v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000111v1_decoy.regions does not exist
15_KI270851v1_alt                       0                          27
Un_JTFH01000111v1_decoy                         0                           5
GRCh38.78/4_KI270789v1_alt.regions does not exist
```

```
GRCh38.78/Un_JTFH01001972v1_decoy.regions does not exist
Un_JTFH01001972v1_decoy                         0                         3
GRCh38.78/19_KI270888v1_alt.regions does not exist
4_KI270789v1_alt                          0                          1
GRCh38.78/HLA-B*42:01:01.regions does not exist
HLA-B*42:01:01                            0                          4
GRCh38.78/22_KB663609v1_alt.regions does not exist
GRCh38.78/HLA-DQA1*01:07.regions does not exist
HLA-DQA1*01:07                            0                          12
GRCh38.78/12_GL383552v1_alt.regions does not exist
GRCh38.78/HLA-DQA1*01:02:01:02.regions does not exist
HLA-DQA1*01:02:01:02                            0                          14
GRCh38.78/Un_JTFH01001133v1_decoy.regions does not exist
Un_JTFH01001133v1_decoy                         0                         1
GRCh38.78/X_KI270913v1_alt.regions does not exist
X_KI270913v1_alt                          0                          29
GRCh38.78/HLA-DQA1*01:02:01:03.regions does not exist
HLA-DQA1*01:02:01:03                            0                          14
GRCh38.78/Un_KN707867v1_decoy.regions does not exist
GRCh38.78/HLA-A*68:01:02:02.regions does not exist
GRCh38.78/Un_JTFH01000346v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000629v1_decoy.regions does not exist
Un_JTFH01000629v1_decoy                         0                         2
GRCh38.78/Un_JTFH01001046v1_decoy.regions does not exist
Un_JTFH01001046v1_decoy                         0                         2
GRCh38.78/14_KI270722v1_random.regions does not exist
GRCh38.78/HLA-DQB1*03:01:01:01.regions does not exist
HLA-DQB1*03:01:01:01                            0                          11
GRCh38.78/HLA-B*44:23N.regions does not exist
HLA-B*44:23N                           0                           6
14_KI270722v1_random                            0                          28
GRCh38.78/HLA-B*13:25.regions does not exist
HLA-B*13:25                           0                           2
Un_JTFH01000346v1_decoy                         0                         1
Un_KN707867v1_decoy                           0                           2
19_KI270888v1_alt                           0                          70
HLA-A*68:01:02:02                           0                          8
GRCh38.78/HLA-A*23:09.regions does not exist
HLA-A*23:09                           0                           11
GRCh38.78/19_KI270918v1_alt.regions does not exist
GRCh38.78/16_KI270853v1_alt.regions does not exist
GRCh38.78/Un_KN707984v1_decoy.regions does not exist
GRCh38.78/17_KI270858v1_alt.regions does not exist
GRCh38.78/HLA-A*74:02:01:02.regions does not exist
17_KI270858v1_alt                           0                          1
HLA-A*74:02:01:02                           0                          10
19_KI270918v1_alt                           0                          84
GRCh38.78/Un_KI270466v1.regions does not exist
```

```
Un_KN707984v1_decoy                          0                          1
GRCh38.78/HLA-A*02:455.regions does not exist
HLA-A*02:455                             0                          8
GRCh38.78/11_KI270927v1_alt.regions does not exist
GRCh38.78/HLA-A*74:02:01:01.regions does not exist
GRCh38.78/HLA-B*13:15.regions does not exist
HLA-A*74:02:01:01                            0                          7
HLA-B*13:15                              0                          2
GRCh38.78/HLA-A*80:01:01:02.regions does not exist
22_KB663609v1_alt                            0                          1
HLA-A*80:01:01:02                            0                          5
12_GL383552v1_alt                            0                          1
GRCh38.78/Un_JTFH01000378v1_decoy.regions does not exist
GRCh38.78/17_GL383563v3_alt.regions does not exist
GRCh38.78/HLA-B*07:02:01.regions does not exist
Un_JTFH01000378v1_decoy                          0                          1
11_KI270927v1_alt                            0                        116
GRCh38.78/17_JH159147v1_alt.regions does not exist
17_JH159147v1_alt                            0                         11
GRCh38.78/Un_JTFH01001539v1_decoy.regions does not exist
Un_JTFH01001539v1_decoy                          0                          1
Un_KI270466v1                            0                         30
GRCh38.78/2_KI270768v1_alt.regions does not exist
GRCh38.78/8_KI270815v1_alt.regions does not exist
GRCh38.78/5_KI270897v1_alt.regions does not exist
GRCh38.78/Un_KI270442v1.regions does not exist
GRCh38.78/Un_KN707905v1_decoy.regions does not exist
16_KI270853v1_alt                            0                         71
GRCh38.78/Un_JTFH01000324v1_decoy.regions does not exist
Un_JTFH01000324v1_decoy                          0                          9
GRCh38.78/HLA-DQA1*01:02:01:01.regions does not exist
GRCh38.78/Un_JTFH01000906v1_decoy.regions does not exist
HLA-DQA1*01:02:01:01                             0                         14
Un_JTFH01000906v1_decoy                          0                          2
5_KI270897v1_alt                             0                         18
GRCh38.78/HLA-B*27:05:18.regions does not exist
GRCh38.78/Un_JTFH01001512v1_decoy.regions does not exist
GRCh38.78/HLA-DQA1*01:11.regions does not exist
HLA-DQA1*01:11                           0                         13
17_GL383563v3_alt                            0                         49
HLA-B*27:05:18                           0                          5
GRCh38.78/HLA-DQB1*03:01:01:03.regions does not exist
HLA-DQB1*03:01:01:03                             0                         11
Un_JTFH01001512v1_decoy                          0                         16
HLA-B*07:02:01                           0                          2
GRCh38.78/HLA-DQA1*01:10.regions does not exist
HLA-DQA1*01:10                           0                         10
GRCh38.78/HLA-A*68:18N.regions does not exist
```

```
HLA-A*68:18N                          0                          8
8_KI270815v1_alt                          0                          2
Un_KN707905v1_decoy                           0                          6
2_KI270768v1_alt                          0                         11
Un_KI270442v1                           0                        425
GRCh38.78/HLA-A*80:01:01:01.regions does not exist
GRCh38.78/HLA-A*68:01:02:01.regions does not exist
GRCh38.78/1_KI270763v1_alt.regions does not exist
HLA-A*68:01:02:01                           0                          8
HLA-A*80:01:01:01                           0                          6
GRCh38.78/HLA-B*15:25:01.regions does not exist
HLA-B*15:25:01                          0                          4
GRCh38.78/Un_KN707686v1_decoy.regions does not exist
Un_KN707686v1_decoy                           0                          1
GRCh38.78/6_GL000250v2_alt.regions does not exist
GRCh38.78/HLA-A*02:43N.regions does not exist
HLA-A*02:43N                          0                          8
1_KI270763v1_alt                          0                          3
GRCh38.78/HLA-DQA1*01:02:01:04.regions does not exist
HLA-DQA1*01:02:01:04                           0                         12
GRCh38.78/3_KI270782v1_alt.regions does not exist
GRCh38.78/Un_JTFH01001669v1_decoy.regions does not exist
GRCh38.78/Un_KN707607v1_decoy.regions does not exist
Un_KN707607v1_decoy                           0                          2
GRCh38.78/HLA-A*02:60:01.regions does not exist
HLA-A*02:60:01                          0                          9
Un_JTFH01001669v1_decoy                               0                          7
GRCh38.78/HLA-B*45:01:01.regions does not exist
GRCh38.78/Un_JTFH01000433v1_decoy.regions does not exist
Un_JTFH01000433v1_decoy                               0                          1
6_GL000250v2_alt                          0                        342
GRCh38.78/HLA-A*02:01:01:03.regions does not exist
GRCh38.78/Un_JTFH01000750v1_decoy.regions does not exist
HLA-A*02:01:01:03                           0                          8
Un_JTFH01000750v1_decoy                               0                          1
GRCh38.78/HLA-B*27:05:02.regions does not exist
HLA-B*45:01:01                          0                          4
HLA-B*27:05:02                          0                          5
GRCh38.78/3_KI270784v1_alt.regions does not exist
GRCh38.78/HLA-A*02:01:01:01.regions does not exist
HLA-A*02:01:01:01                           0                          8
GRCh38.78/15_KI270727v1_random.regions does not exist
GRCh38.78/HLA-A*11:60.regions does not exist
HLA-A*11:60                         0                        8
3_KI270782v1_alt                          0                          7
GRCh38.78/Un_JTFH01001057v1_decoy.regions does not exist
3_KI270784v1_alt                          0                         10
Un_JTFH01001057v1_decoy                               0                          1
```

```
GRCh38.78/HLA-A*02:01:01:04.regions does not exist
HLA-A*02:01:01:04                         0                         8
15_KI270727v1_random                        0                         3
GRCh38.78/19_KI270932v1_alt.regions does not exist
GRCh38.78/8_KI270817v1_alt.regions does not exist
GRCh38.78/HLA-B*13:08.regions does not exist
GRCh38.78/Un_KI270509v1.regions does not exist
HLA-B*13:08                         0                         2
Un_KI270509v1                         0                         2
19_KI270932v1_alt                         0                         102
GRCh38.78/HLA-A*01:20.regions does not exist
GRCh38.78/Un_JTFH01000559v1_decoy.regions does not exist
HLA-A*01:20                         0                         6
GRCh38.78/HLA-A*69:01.regions does not exist
GRCh38.78/Un_JTFH01000589v1_decoy.regions does not exist
HLA-A*69:01                         0                         7
GRCh38.78/Un_JTFH01001217v1_decoy.regions does not exist
Un_JTFH01000559v1_decoy                         0                         1
GRCh38.78/20_KI270869v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000876v1_decoy.regions does not exist
Un_JTFH01001217v1_decoy                         0                         5
Un_JTFH01000589v1_decoy                         0                         5
GRCh38.78/8_KI270818v1_alt.regions does not exist
8_KI270818v1_alt                        0                         43
GRCh38.78/1_KI270765v1_alt.regions does not exist
Un_JTFH01000876v1_decoy                         0                         15
20_KI270869v1_alt                         0                         11
GRCh38.78/19_KI270930v1_alt.regions does not exist
GRCh38.78/HLA-A*11:74.regions does not exist
HLA-A*11:74                         0                         9
1_KI270765v1_alt                        0                         1
GRCh38.78/HLA-B*15:13:01.regions does not exist
GRCh38.78/HLA-B*40:06:01:02.regions does not exist
HLA-B*15:13:01                        0                         3
HLA-B*40:06:01:02                         0                         5
GRCh38.78/Un_JTFH01001214v1_decoy.regions does not exist
GRCh38.78/11_KI270903v1_alt.regions does not exist
Un_JTFH01001214v1_decoy                         0                         1
8_KI270817v1_alt                        0                         8
GRCh38.78/HLA-C*08:04:01.regions does not exist
HLA-C*08:04:01                        0                         2
GRCh38.78/Un_JTFH01000430v1_decoy.regions does not exist
19_KI270930v1_alt                         0                         112
Un_JTFH01000430v1_decoy                         0                         14
11_KI270903v1_alt                         0                         27
GRCh38.78/Un_JTFH01000820v1_decoy.regions does not exist
Un_JTFH01000820v1_decoy                         0                         3
GRCh38.78/HLA-A*11:75.regions does not exist
```

```
HLA-A*11:75                         0                         10
GRCh38.78/HLA-B*48:03:01.regions does not exist
HLA-B*48:03:01                       0                        7
GRCh38.78/12_KI270904v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000159v1_decoy.regions does not exist
Un_JTFH01000159v1_decoy                      0                        1
12_KI270904v1_alt                      0                       52
GRCh38.78/HLA-A*01:14.regions does not exist
HLA-A*01:14                         0                         6
GRCh38.78/Un_JTFH01000411v1_decoy.regions does not exist
GRCh38.78/Un_KN707860v1_decoy.regions does not exist
GRCh38.78/HLA-A*43:01.regions does not exist
Un_KN707860v1_decoy                       0                       3
GRCh38.78/Un_JTFH01000297v1_decoy.regions does not exist
Un_JTFH01000411v1_decoy                      0                        3
GRCh38.78/HLA-B*35:08:01.regions does not exist
HLA-B*35:08:01                       0                        1
Un_JTFH01000297v1_decoy                      0                        1
HLA-A*43:01                         0                         1
GRCh38.78/Un_JTFH01000395v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000987v1_decoy.regions does not exist
GRCh38.78/Un_KI270749v1.regions does not exist
Un_JTFH01000987v1_decoy                      0                        2
GRCh38.78/Un_JTFH01001166v1_decoy.regions does not exist
Un_JTFH01001166v1_decoy                      0                        2
Un_JTFH01000395v1_decoy                      0                        2
GRCh38.78/Un_JTFH01001320v1_decoy.regions does not exist
GRCh38.78/Un_GL000220v1.regions does not exist
Un_JTFH01001320v1_decoy                      0                        1
GRCh38.78/HLA-A*11:50Q.regions does not exist
GRCh38.78/HLA-B*51:42.regions does not exist
HLA-A*11:50Q                         0                        8
GRCh38.78/HLA-A*34:02:01.regions does not exist
HLA-A*34:02:01                       0                        7
GRCh38.78/Un_KN707659v1_decoy.regions does not exist
HLA-B*51:42                         0                         7
Un_KI270749v1                         0                        1
Un_KN707659v1_decoy                       0                       19
GRCh38.78/Un_KN707963v1_decoy.regions does not exist
Un_KN707963v1_decoy                       0                       222
GRCh38.78/HLA-A*01:03.regions does not exist
GRCh38.78/HLA-A*31:04.regions does not exist
HLA-A*31:04                         0                         1
HLA-A*01:03                         0                         8
GRCh38.78/1_KI270709v1_random.regions does not exist
Un_GL000220v1                         0                        1
GRCh38.78/HLA-DQA1*01:03:01:02.regions does not exist
HLA-DQA1*01:03:01:02                       0                       10
```

```
GRCh38.78/Un_KI270507v1.regions does not exist
GRCh38.78/1_KI270706v1_random.regions does not exist
GRCh38.78/Un_JTFH01000557v1_decoy.regions does not exist
GRCh38.78/HLA-A*68:113.regions does not exist
Un_JTFH01000557v1_decoy                         0                               1
HLA-A*68:113                        0                       7
GRCh38.78/HLA-B*08:132.regions does not exist
HLA-B*08:132                        0                       3
GRCh38.78/Un_JTFH01000013v1_decoy.regions does not exist
Un_JTFH01000013v1_decoy                         0                               22
GRCh38.78/2_GL582966v2_alt.regions does not exist
1_KI270709v1_random                         0                           11
GRCh38.78/HLA-B*53:11.regions does not exist
HLA-B*53:11                         0                       1
GRCh38.78/HLA-A*24:03:01.regions does not exist
HLA-A*24:03:01                          0                           14
GRCh38.78/Un_JTFH01001899v1_decoy.regions does not exist
Un_JTFH01001899v1_decoy                         0                               9
GRCh38.78/Un_JTFH01001336v1_decoy.regions does not exist
Un_JTFH01001336v1_decoy                         0                               5
1_KI270706v1_random                         0                           5
GRCh38.78/HLA-A*01:02.regions does not exist
GRCh38.78/HLA-DRB1*01:01:01.regions does not exist
HLA-DRB1*01:01:01                           0                       4
HLA-A*01:02                         0                       7
GRCh38.78/Un_JTFH01001168v1_decoy.regions does not exist
Un_JTFH01001168v1_decoy                         0                               1
Un_KI270507v1                           0                       7
GRCh38.78/HLA-B*40:06:01:01.regions does not exist
2_GL582966v2_alt                            0                           1
GRCh38.78/Un_JTFH01001255v1_decoy.regions does not exist
GRCh38.78/HLA-B*08:134.regions does not exist
GRCh38.78/HLA-B*08:79.regions does not exist
GRCh38.78/Un_JTFH01001101v1_decoy.regions does not exist
GRCh38.78/HLA-B*14:07N.regions does not exist
HLA-B*08:134                        0                       3
HLA-B*08:79                         0                       2
GRCh38.78/Un_JTFH01001219v1_decoy.regions does not exist
HLA-B*14:07N                        0                       3
Un_JTFH01001255v1_decoy                         0                               2
GRCh38.78/2_KI270893v1_alt.regions does not exist
GRCh38.78/3_JH636055v2_alt.regions does not exist
Un_JTFH01001101v1_decoy                         0                               4
Un_JTFH01001219v1_decoy                         0                               1
GRCh38.78/Un_KI270518v1.regions does not exist
GRCh38.78/HLA-B*51:02:01.regions does not exist
HLA-B*51:02:01                          0                           1
HLA-B*40:06:01:01                           0                       5
```

```
GRCh38.78/HLA-B*49:01:01.regions does not exist
HLA-B*49:01:01                          0                          5
2_KI270893v1_alt                          0                        2
GRCh38.78/HLA-DQA1*01:03:01:01.regions does not exist
HLA-DQA1*01:03:01:01                        0                        10
GRCh38.78/HLA-A*02:07:01.regions does not exist
GRCh38.78/HLA-B*08:19N.regions does not exist
3_JH636055v2_alt                          0                        15
HLA-B*08:19N                        0                          3
GRCh38.78/Un_JTFH01001553v1_decoy.regions does not exist
Un_JTFH01001553v1_decoy                         0                         1
HLA-A*02:07:01                        0                          8
GRCh38.78/22_KI270734v1_random.regions does not exist
GRCh38.78/HLA-B*13:01:01.regions does not exist
Un_KI270518v1                       0                          5
GRCh38.78/Un_KN707927v1_decoy.regions does not exist
HLA-B*13:01:01                          0                          2
Un_KN707927v1_decoy                           0                         1
GRCh38.78/HLA-A*02:05:01.regions does not exist
GRCh38.78/1_KI270710v1_random.regions does not exist
HLA-A*02:05:01                          0                          8
GRCh38.78/19_KI270914v1_alt.regions does not exist
GRCh38.78/HLA-A*01:09.regions does not exist
HLA-A*01:09                         0                          7
GRCh38.78/17_KI270910v1_alt.regions does not exist
GRCh38.78/Un_JTFH01001120v1_decoy.regions does not exist
GRCh38.78/19_KI270917v1_alt.regions does not exist
Un_JTFH01001120v1_decoy                         0                         1
GRCh38.78/6_GL000253v2_alt.regions does not exist
1_KI270710v1_random                           0                        2
GRCh38.78/Un_KN707971v1_decoy.regions does not exist
22_KI270734v1_random                          0                         5
19_KI270917v1_alt                         0                        108
19_KI270914v1_alt                         0                        97
Un_KN707971v1_decoy                           0                         71
6_GL000253v2_alt                        0                        795
GRCh38.78/HLA-B*59:01:01:02.regions does not exist
17_KI270910v1_alt                         0                        11
HLA-B*59:01:01:02                         0                        2
GRCh38.78/HLA-A*11:25.regions does not exist
GRCh38.78/Un_JTFH01001074v1_decoy.regions does not exist
GRCh38.78/Un_JTFH01000129v1_decoy.regions does not exist
Un_JTFH01001074v1_decoy                         0                         1
HLA-A*11:25                         0                          8
GRCh38.78/HLA-A*26:11N.regions does not exist
HLA-A*26:11N                          0                          2
GRCh38.78/9_GL383540v1_alt.regions does not exist
GRCh38.78/Un_JTFH01000148v1_decoy.regions does not exist
```

```
Un_JTFH01000129v1_decoy                          0                          3
Un_JTFH01000148v1_decoy                          0                          1
GRCh38.78/Un_JTFH01001870v1_decoy.regions does not exist
Un_JTFH01001870v1_decoy                          0                          1
GRCh38.78/HLA-B*07:156.regions does not exist
HLA-B*07:156                            0                           2
GRCh38.78/Un_JTFH01001014v1_decoy.regions does not exist
Un_JTFH01001014v1_decoy                          0                          1
GRCh38.78/HLA-DQB1*03:02:01.regions does not exist
HLA-DQB1*03:02:01                           0                           1
9_GL383540v1_alt                           0                           2
GRCh38.78/Un_JTFH01000546v1_decoy.regions does not exist
GRCh38.78/HLA-B*59:01:01:01.regions does not exist
GRCh38.78/9_GL383541v1_alt.regions does not exist
HLA-B*59:01:01:01                           0                           2
GRCh38.78/Un_JTFH01000230v1_decoy.regions does not exist
Un_JTFH01000230v1_decoy                          0                          1
GRCh38.78/Un_JTFH01000205v1_decoy.regions does not exist
Un_JTFH01000205v1_decoy                          0                          4
Un_JTFH01000546v1_decoy                          0                          1
GRCh38.78/HLA-A*31:46.regions does not exist
HLA-A*31:46                             0                           1
GRCh38.78/HLA-B*15:11:01.regions does not exist
HLA-B*15:11:01                            0                           4
9_GL383541v1_alt                            0                           6
GRCh38.78/6_GL000256v2_alt.regions does not exist
GRCh38.78/HLA-C*08:27.regions does not exist
HLA-C*08:27                             0                           2
GRCh38.78/Un_JTFH01000844v1_decoy.regions does not exist
GRCh38.78/22_KI270731v1_random.regions does not exist
GRCh38.78/Un_KN707660v1_decoy.regions does not exist
GRCh38.78/9_GL383542v1_alt.regions does not exist
GRCh38.78/HLA-A*11:05.regions does not exist
HLA-A*11:05                             0                           9
GRCh38.78/HLA-A*23:01:01.regions does not exist
HLA-A*23:01:01                            0                           13
22_KI270731v1_random                             0                           1
Un_KN707660v1_decoy                            0                           63
GRCh38.78/Un_JTFH01000124v1_decoy.regions does not exist
GRCh38.78/HLA-C*08:62.regions does not exist
Un_JTFH01000844v1_decoy                          0                          12
GRCh38.78/11_JH159136v1_alt.regions does not exist
HLA-C*08:62                             0                           2
6_GL000256v2_alt                            0                           773
GRCh38.78/Un_GL000224v1.regions does not exist
Un_JTFH01000124v1_decoy                          0                          1
9_GL383542v1_alt                           0                           5
GRCh38.78/HLA-A*25:01:01.regions does not exist
```

```
GRCh38.78/HLA-DQB1*02:02:01.regions does not exist
HLA-DQB1*02:02:01                          0                          3
HLA-A*25:01:01                           0                          2
GRCh38.78/Un_JTFH01001894v1_decoy.regions does not exist
GRCh38.78/HLA-C*08:41.regions does not exist
HLA-C*08:41                            0                        2
Un_JTFH01001894v1_decoy                          0                          1
GRCh38.78/6_KI270758v1_alt.regions does not exist
11_JH159136v1_alt                           0                         19
6_KI270758v1_alt                          0                         36
GRCh38.78/18_KI270911v1_alt.regions does not exist
GRCh38.78/HLA-A*26:01:01.regions does not exist
GRCh38.78/HLA-B*18:94N.regions does not exist
HLA-A*26:01:01                           0                          2
GRCh38.78/1_KI270892v1_alt.regions does not exist
GRCh38.78/Un_JTFH01001915v1_decoy.regions does not exist
Un_JTFH01001915v1_decoy                          0                          1
GRCh38.78/HLA-A*11:110.regions does not exist
GRCh38.78/6_GL000255v2_alt.regions does not exist
HLA-A*11:110                            0                         8
GRCh38.78/HLA-B*27:07:01.regions does not exist
HLA-B*27:07:01                            0                          5
HLA-B*18:94N                            0                         1
Un_GL000224v1                            0                         64
18_KI270911v1_alt                           0                          1
1_KI270892v1_alt                          0                          7
GRCh38.78/19_KI270916v1_alt.regions does not exist
GRCh38.78/Un_JTFH01001677v1_decoy.regions does not exist
Un_JTFH01001677v1_decoy                          0                          6
GRCh38.78/11_JH159137v1_alt.regions does not exist
GRCh38.78/HLA-A*68:02:02.regions does not exist
HLA-A*68:02:02                           0                          7
6_GL000255v2_alt                            0                          839
GRCh38.78/HLA-C*08:40.regions does not exist
HLA-C*08:40                            0                        2
GRCh38.78/6_GL000254v2_alt.regions does not exist
GRCh38.78/Un_JTFH01001783v1_decoy.regions does not exist
19_KI270916v1_alt                           0                          55
Un_JTFH01001783v1_decoy                          0                          1
GRCh38.78/HLA-B*55:02:01.regions does not exist
HLA-B*55:02:01                           0                          2
6_GL000254v2_alt                            0                          720
GRCh38.78/19_KI270915v1_alt.regions does not exist
11_JH159137v1_alt                           0                          14
GRCh38.78/HLA-A*33:07.regions does not exist
HLA-A*33:07                            0                        3
19_KI270915v1_alt                           0                          106
GRCh38.78/Un_JTFH01001888v1_decoy.regions does not exist
```

```
Un_JTFH01001888v1_decoy                         0                        18
22                        595                       1438                        Co
20                        602                       2044                        Co
13                        416                       1069                        Co
21                        436                       1690                        Co
X                        353                       500                        Comp
9                        1035                       1900                        Co
10                        1008                       2311                        C
18                        421                       660                        Com
19                        2073                       2152                        C
7                        1283                       2079                        Co
17                        1571                       2169                        C
16                        1057                       1841                        C
14                        932                       1533                        Co
6                        1442                       2287                        Co
11                        1776                       2121                        C
15                        841                       1442                        Co
12                        1343                       2011                        C
8                        877                       1463                        Com
3                        1459                       2220                        Co
4                        939                       2159                        Com
5                        1169                       1889                        Co
1                        2677                       3944                        Co
2                        1675                       2835                        Co

Merging temp files...
SIFT4G Annotation completed !
Output directory:NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10
End Time for parallel code: Mon Mar 27 11:59:37 PDT 2017
```

6.30 Question 16)

On Chromosome 17, how many variants are annotated? How many are unannotated?

6.31 Answer 16)

1571 annotated, 2169 unannotated

6.32 Question 17)

How many deleterious (not 'Low confidence') variants are found from these variants?

6.33 Answer 17)

```
In [35]: %%bash
         cat NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_
```

```
/NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_gt_
|grep 'DELETERIOUS'|grep -v 'Low confidence'|cut -f1,2,3,4 \
|sort|uniq|wc -l
```

```
1365
```

1365 deleterious variants.

6.34 Question 18)

How many genes have deleterious variants? Output the list of genes names into a file. Display the gene names.

6.35 Answer 18)

```
In [36]: %%bash
         cat NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_
         /NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_gt_
         |grep 'DELETERIOUS'|grep -v 'Low confidence'|cut -f7 \
         |sort|uniq \
         > 1KGenomesSample.SIFT4G.genes_with_deleterious_variants.txt
         wc -l 1KGenomesSample.SIFT4G.genes_with_deleterious_variants.txt
         cat 1KGenomesSample.SIFT4G.genes_with_deleterious_variants.txt
```

```
1154 1KGenomesSample.SIFT4G.genes_with_deleterious_variants.txt
A2ML1
ABCA12
ABCA4
ABCA9
ABCC11
ABCC4
ABCD4
ABHD11
ABO
AC244230.1
ACACB
ACAN
ACSM5
ACSS1
ACTRT2
ADAM15
ADAM18
ADAMDEC1
ADAMTS18
ADAMTSL3
ADO
ADORA3
ADPGK
```

AFMID
AHNAK
AK2
AKAP1
AKAP13
AKAP2
AKAP3
AKR1C2
ALB
ALDH1B1
ALG6
ALPK2
ALPK3
ALPL
AMACR
AMICA1
AMPD3
ANAPC10
ANGEL1
ANKK1
ANKRD30A
ANLN
ANO10
ANO2
ANO3
ANXA7
AOAH
AP1G2
AP2S1
APH1B
APOA1BP
APOB
APOBEC3H
APOL1
APOL4
AQP12A
ARHGAP9
ARHGEF28
ARHGEF37
ARMC9
ARRDC4
ASB16
ASB18
ASH1L
ASMTL
ASPSCR1
ATF7IP
ATP10A

ATP6V1C2
ATRX
ATXN1
AVPR2
B3GNT3
B4GALNT2
BAG3
BANK1
BARD1
BCL9
BMP2
BMP3
BNIPL
BPIFB2
BRIP1
BST1
BTD
BTNL2
BTNL8
BUD13
C14orf37
C14orf80
C18orf54
C1orf112
C1orf177
C1orf87
C1QTNF6
C1QTNF9B
C2orf61
C2orf73
C3
C3orf20
C3orf30
C4orf33
C6orf15
C6orf222
C7
C7orf31
C7orf57
C7orf72
C9orf43
CACNA1H
CACNA1S
CAGE1
CALCA
CALCOCO2
CAPN11
CAPN12

CAPN9
CARD14
CASC1
CASC5
CAV2
CCAR1
CCDC116
CCDC122
CCDC13
CCDC141
CCDC157
CCDC18
CCDC180
CCDC181
CCDC88C
CCDC89
CCHCR1
CCNA1
CCND3
CCNH
CCP110
CCT6B
CD109
CD163
CD200R1
CD207
CD27
CD5
CDAN1
CDC25C
CDH11
CDHR2
CDK11A
CDK11B
CDK15
CDK5RAP2
CEACAM21
CEACAM5
CENPQ
CEP192
CEP72
CEP89
CERKL
CES1
CETN1
CFAP53
CFAP69
CFAP74

CFB
CH17-3B23.1
CHFR
CHI3L1
CHIA
CHIT1
CHMP4A
CHPT1
CHRNA3
CHRNB4
CLCA2
CLCNKA
CLEC1A
CLEC4A
CLGN
CLIP1
CLUAP1
CMYA5
CNOT1
CNTNAP2
CNTNAP4
CNTRL
COASY
COBL
COL12A1
COL14A1
COL24A1
COL2A1
COL4A3
COL4A4
COL6A6
COL9A3
COMP
COQ7
CORO7
CORO7-PAM16
COX10
CPAMD8
CPN2
CPS1
CRYBG3
CRYGB
CSTA
CTNNA3
CTSB
CTU2
CUBN
CWH43

CYP11B1
CYP2A7
CYP4A22
CYP4B1
CYP4F12
CYP4F2
DACT1
DAPL1
DAW1
DCAF4
DCHS2
DCLRE1C
DDIAS
DDX39B
DDX4
DDX58
DDX60L
DEFB127
DENND1B
DENND1C
DHTKD1
DIP2B
DISC1
DLEC1
DLGAP2
DMP1
DMRT2
DNAAF2
DNAJA3
DNAJC16
DNHD1
DOCK10
DOCK6
DOCK8
DPYSL2
DRC7
DSG3
DUSP13
DUSP23
ECHDC3
EFCAB7
EFCC1
EIF2AK4
ELN
EMR1
ENAM
EPHA1
EPHA8

EPHX1
EPPK1
EPS8L1
ERCC4
ERCC5
ERO1LB
ESPL1
ESRRA
ETFA
ETFB
ETV2
EVC
EVC2
EXD3
EXPH5
EXTL1
F5
FAM151A
FAM153B
FAM154B
FAM175A
FAM188B
FAM26F
FAM47E
FAM47E-STBD1
FAM71C
FAM71F1
FAM71F2
FAM86B2
FASTKD5
FAT2
FBF1
FBLIM1
FBN3
FBXO36
FBXW10
FBXW8
FCN2
FERMT1
FHDC1
FHL5
FIGNL1
FLNB
FLNC
FLT4
FMN1
FMO2
FN1

FOS
FOXD4
FOXD4L1
FPR1
FRA10AC1
FRAS1
FREM2
FRG1B
FRG2B
FRMD4B
FRMPD2
FTSJ3
FUK
FUT10
FUT2
FUT3
FUT9
GAD2
GAK
GAL3ST1
GALC
GALNT16
GALNT8
GALNTL5
GALP
GBGT1
GBP3
GBP6
GCAT
GCNT1
GDF15
GDPD4
GEMIN4
GFAP
GGT6
GGTLC1
GIMAP6
GIMAP7
GIT2
GLUL
GNA12
GOLGA6L2
GOLGA6L9
GOLGB1
GON4L
GORAB
GPR108
GPR137C

GPR35
GPR45
GPR98
GPRIN1
GPRIN2
GPSM2
GRIK1
GRIN3A
GRIN3B
GSDMC
GSTA5
GTF3C1
GTPBP2
GUCA1C
H1FNT
HAP1
HEATR1
HEATR5A
HELB
HIBCH
HLA-C
HLA-DPB1
HLA-DQA1
HLA-DQB1
HLA-DQB2
HLA-DRB5
HLA-G
HMCN1
HMGA2
HMGXB4
HNRNPA1L2
HPS4
HRCT1
HRNR
HSD17B4
HSDL1
HSPA5
HSPA6
HTR3D
HUS1B
IBSP
IDNK
IFT88
IGFL1
IGLL5
IGSF10
IKBKAP
IL1F10

IL1RL1
IMPG1
INADL
INMT
INPP5B
INSR
IP6K1
IQCA1
IQCB1
IQCF6
IQGAP3
IQSEC1
ISM2
ITGA11
ITGA9
ITGAE
ITGB4
ITIH5
ITLN2
ITPR2
JAG1
KCNJ12
KCNJ18
KCNMB3
KDELR3
KDM3A
KDR
KIAA0753
KIAA1328
KIAA1377
KIAA1755
KIF20A
KIF2C
KIF5A
KIR2DL1
KIR2DL3
KIR2DL4
KIR3DL1
KIR3DL3
KLHDC1
KLHDC7A
KLRC2
KLRC3
KMT2C
KNG1
KRT13
KRT27
KRT3

KRT32
KRT33A
KRT37
KRT40
KRT6A
KRT78
KRT79
KRTAP10-1
KRTAP10-10
KRTAP10-11
KRTAP10-3
KRTAP10-4
KRTAP10-5
KRTAP11-1
KRTAP12-3
KRTAP15-1
KRTAP19-2
KRTAP26-1
KRTAP4-4
KRTAP5-10
KRTAP5-3
KRTAP5-7
KRTAP5-8
KRTAP5-9
KRTAP9-4
KRTAP9-8
LAG3
LAMA1
LAMA5
LAMB2
LARP1B
LCAT
LCE2D
LCE3D
LCE5A
LCTL
LENG9
LGALS3
LGALS8
LILRB1
LIMS1
LMO7
LPA
LRBA
LRP1B
LRP2
LRR1
LRRC25

LRRC34
LRRC43
LRRC48
LRRK2
LTA
LYAR
LYSMD4
MACC1
MACF1
MAGEC1
MAGEF1
MAN2B1
MAP2K3
MAP3K19
MAPKBP1
MARCH7
MASP2
MATN2
MBD1
MBLAC1
MC5R
MCEE
MCF2L2
MCM3AP
MCPH1
MDC1
MDGA1
MDGA2
MED13
MEGF10
MEGF11
MEP1A
MERTK
MFI2
MICA
MICALCL
MICB
MKI67
MLF1
MMS19
MMS22L
MOB3B
MPP3
MRGPRX4
MROH7
MROH7-TTC4
MRPL18
MS4A14

MS4A6E
MSI2
MST1
MTCH2
MTERF4
MTMR1
MTMR12
MTRR
MTUS2
MUC12
MUC15
MUC16
MUC3A
MUC4
MUC5AC
MUC6
MUS81
MXRA5
MYBPC1
MYBPC2
MYH11
MYH15
MYH4
MYH6
MYO1A
MYO5C
MYO6
MYOF
MYOM3
MYPN
NAAA
NAALADL2
NARFL
NAT2
NAV3
NBPF3
NCAPD2
NCAPG
NCF4
NCKAP5
NECAP1
NEFH
NEK11
NEMF
NFATC1
NFATC2IP
NFXL1
NID2

NIPSNAP3A
NKG2-E
NLRP12
NLRP13
NLRP14
NME4
NME8
NNT
NOD1
NOTCH3
NOX5
NPHP3
NPHS1
NPIPB15
NPIPB6
NPNT
NRG1
NT5C3B
NTMT1
NTSR1
NUDCD3
NUP133
NUP160
NUP210L
NUSAP1
NUTM1
NUTM2D
OBP2A
OBSCN
OBSL1
OC90
ODF3L1
OLFML1
OR10A2
OR10A6
OR10G2
OR10G3
OR10H1
OR10H3
OR10H5
OR10J1
OR10R2
OR10Z1
OR11G2
OR11H1
OR11H6
OR11L1
OR12D2

OR13C3
OR13C5
OR13C8
OR13D1
OR13J1
OR14C36
OR14I1
OR1A1
OR1D5
OR1E2
OR1I1
OR1L4
OR1L6
OR1Q1
OR1S2
OR2B11
OR2C1
OR2C3
OR2D3
OR2G2
OR2G3
OR2K2
OR2L3
OR2L8
OR2M5
OR2M7
OR2T12
OR2T29
OR2T3
OR2T33
OR2T34
OR2T7
OR2T8
OR3A1
OR4A16
OR4B1
OR4C11
OR4C46
OR4C6
OR4D2
OR4D6
OR4E2
OR4K1
OR4M1
OR4Q3
OR51B2
OR51B5
OR51B6

OR51G1
OR51G2
OR51I1
OR51I2
OR51M1
OR51Q1
OR51V1
OR52E2
OR52I2
OR52J3
OR52K2
OR52L1
OR52N1
OR52N2
OR52R1
OR52W1
OR56A5
OR56B1
OR5A1
OR5AU1
OR5D16
OR5H1
OR5H15
OR5H6
OR5L2
OR5M1
OR5R1
OR5T2
OR5V1
OR6B2
OR6C1
OR6C2
OR6C74
OR6K6
OR6M1
OR6N1
OR6S1
OR7A10
OR7C1
OR7G1
OR8B3
OR8D1
OR8G5
OR8J1
OR8K1
OR8S1
ORC1
OTOL1

OTOP1
OTOP2
OTOR
OVCH1
PABPC1
PABPC3
PADI4
PALM2-AKAP2
PAPPA
PARP10
PARP15
PARP4
PAX8
PCDHA9
PCDHB10
PCDHB7
PCDHGA4
PCDHGB6
PCNXL3
PCNXL4
PDE12
PDE2A
PDE4C
PDIA2
PDLIM5
PDLIM7
PECAM1
PER3
PFKFB2
PGM1
PGM2L1
PHLPP2
PHYHD1
PIGC
PIK3C2B
PJA2
PKD1
PKD1L1
PKD1L3
PKDREJ
PKHD1
PKHD1L1
PKN1
PLAUR
PLCL1
PLEKHG4B
PLXNA2
PM20D1

PNPLA3
PODXL
POLN
POLQ
PON1
PON2
POTEC
PPA2
PPEF1
PPEF2
PPM1F
PPP1R15A
PRAM1
PRAMEF1
PRAMEF12
PRB1
PRB2
PRB3
PRB4
PRDM7
PRG4
PRKAG2
PRMT6
PRMT7
PRR14
PRR4
PRRC2A
PRRC2C
PRSS48
PRSS50
PRUNE2
PSG5
PSG8
PSMD13
PSMF1
PTCH1
PTF1A
PTGER3
PTX4
PYGB
PZP
QRICH2
QSOX1
RAB11FIP1
RAD54L
RAET1E
RAI1
RAMP3

RASAL1
RBM18
RBM19
RBMX
RBMXL1
RET
RETNLB
RFPL1
RFPL2
RFX5
RFX7
RGL4
RGPD4
RHAG
RHBG
RHCG
RHOT2
RHPN1
RICTOR
RIN1
RIOK2
RMDN3
RNASEL
RNF157
RNF183
RNF213
RNF43
RP1
RP11-404P21.8
RP11-457D20.2
RP11-545J16.1
RP11-683L23.1
RP11-697E2.6
RP1L1
RPAP1
RPF1
RPTN
RREB1
RSPH4A
S100Z
S1PR3
SAA2
SCAF11
SCLT1
SCLY
SCYL2
SDCBP2
SEC14L3

SEC23B
SEC31B
SELE
SEMA4D
SEMA4G
SEMA6D
SEPN1
SEPT11
SERPINA5
SERPINA9
SERPINB10
SERPINB11
SERPINB8
SERPINI2
SETX
SH2D1B
SH2D4A
SH3TC1
SHROOM3
SIRPB1
SKAP2
SLC15A2
SLC15A4
SLC22A10
SLC22A2
SLC22A25
SLC22A4
SLC25A47
SLC26A6
SLC2A3
SLC2A9
SLC30A2
SLC45A2
SLC4A3
SLC4A4
SLC5A4
SLC6A18
SLC7A13
SLC7A14
SLC9A3R2
SLC9C1
SLCO1B7
SLCO6A1
SLFNL1
SLIT3
SLX4
SMC1B
SMC6

SMPDL3B
SMYD4
SNTB2
SPAG17
SPATA31A6
SPATA31E1
SPATA6
SPEF2
SPEG
SPIDR
SPINK5
SPINT4
SPNS3
SPTA1
SRRM2
SSX5
STAG3
STEAP2
STK11IP
STK16
STK31
STK36
STPG2
SULT1A2
SV2C
SVOPL
SYNC
SYNE1
SYNE2
SYNM
T
TAB3
TACC2
TAP2
TAPBPL
TAS2R19
TAS2R31
TAS2R4
TBC1D26
TBC1D28
TBC1D32
TBC1D9B
TBX10
TCEB3
TCOF1
TCP11L1
TDRD5
TEKT4

TEKT5
TGM4
TGOLN2
TGS1
THBS1
THOC1
THOP1
THSD7A
TIAM2
TLR2
TLR3
TMBIM1
TMC1
TMC3
TMEM106C
TMEM161B
TMEM171
TMEM173
TMEM176B
TMEM214
TMEM244
TMEM261
TMEM50A
TMEM71
TMF1
TMPRSS15
TMPRSS2
TNFRSF10D
TNIP2
TNK1
TNN
TNP2
TNS1
TNS3
TOX3
TP53BP1
TPPP2
TPSD1
TPSG1
TRAFD1
TRAPPC12
TRIB2
TRIM22
TRIM40
TRIM51
TRIM6
TRIM63
TRNT1

TRPA1
TRPV4
TSPAN8
TTC21A
TTC21B
TTC22
TTC30A
TTC30B
TTI2
TTLL4
TTLL8
TTN
TUBA3E
TUBB8
TYK2
TYR
TYW1B
UBR1
UBXN11
UCK1
UGT1A6
UGT2A1
UGT2B28
UGT2B4
UIMC1
UMODL1
UNC5C
UQCRHL
USP16
USP35
USP45
USP6
USP8
VARS2
VDR
VN1R4
VPS41
VRK2
VWA8
WBSCR27
WBSCR28
WDR49
WDR55
WDR66
WDR72
WDR90
WDR91
WFS1

WNK2
WNT10A
WNT9B
WRNIP1
WSCD2
WWC2
XRN1
YIF1A
YLPM1
ZAN
ZBBX
ZDHHC11
ZFP57
ZFP69B
ZNF154
ZNF155
ZNF177
ZNF180
ZNF19
ZNF208
ZNF214
ZNF221
ZNF223
ZNF229
ZNF230
ZNF239
ZNF28
ZNF283
ZNF285
ZNF286A
ZNF286B
ZNF30
ZNF404
ZNF415
ZNF443
ZNF45
ZNF462
ZNF474
ZNF493
ZNF501
ZNF502
ZNF534
ZNF543
ZNF549
ZNF559-ZNF177
ZNF571
ZNF573
ZNF594

```
ZNF596
ZNF598
ZNF607
ZNF611
ZNF654
ZNF667
ZNF679
ZNF682
ZNF705A
ZNF705G
ZNF708
ZNF728
ZNF736
ZNF737
ZNF761
ZNF778
ZNF79
ZNF799
ZNF835
ZNF845
ZNF85
ZNF852
ZNF880
ZNF99
ZNRF4
ZSCAN31
ZSCAN32
```

1154 genes. Gene names listed above.

6.36 Question 19)

What genes do Craig Venter, James Watson, and this 1000 Genomes sample All have deleterious variants in? How many genes is this?

6.37 Answer 19)

```
In [37]: %%bash
         join Venter_and_Watson.SIFT4G.genes_with_deleterious_variants.txt \
         1KGenomesSample.SIFT4G.genes_with_deleterious_variants.txt \
         > Venter_and_Watson_and_1KGenomesSample.SIFT4G.genes_with_deleterious_vari
         wc -l Venter_and_Watson_and_1KGenomesSample.SIFT4G.genes_with_deleterious_
         cat Venter_and_Watson_and_1KGenomesSample.SIFT4G.genes_with_deleterious_va

322 Venter_and_Watson_and_1KGenomesSample.SIFT4G.genes_with_deleterious_variants.tx
A2ML1
ACACB
```

ACAN
ADAMTSL3
AHNAK
AKAP13
AKR1C2
ALDH1B1
ALPK2
ALPK3
AMACR
AMPD3
ANKRD30A
APOA1BP
APOB
APOL4
ARHGEF37
ARMC9
ASB16
ATF7IP
BAG3
BMP2
BMP3
C14orf37
C1orf87
C2orf61
C2orf73
C4orf33
C7
C7orf31
C7orf57
C7orf72
CAPN9
CASC5
CCDC18
CCDC181
CD163
CDH11
CDHR2
CDK11A
CDK11B
CDK5RAP2
CENPQ
CFAP69
CFAP74
CHIA
CHIT1
COASY
COL2A1
COL4A3

COL4A4
COL6A6
COQ7
COX10
CRYBG3
CRYGB
CUBN
CWH43
DAPL1
DDX58
DDX60L
DHTKD1
DISC1
DLEC1
DLGAP2
DOCK6
DOCK8
DPYSL2
ECHDC3
EPPK1
EPS8L1
ERCC5
ESPL1
EVC2
FAM188B
FAT2
FBLIM1
FBN3
FBXW8
FHDC1
FPR1
FRAS1
FREM2
FRG1B
FRMD4B
FUT2
FUT3
FUT9
GALNTL5
GALP
GBP3
GBP6
GCAT
GGT6
GIMAP6
GOLGA6L2
GORAB
GPR137C

GPR98
GPRIN2
GRIN3A
GSTA5
HEATR5A
HIBCH
HMGXB4
HPS4
HRNR
HSDL1
HUS1B
IFT88
IKBKAP
INMT
IQGAP3
ITGA11
ITGA9
ITGAE
ITGB4
ITPR2
KCNJ12
KIAA0753
KIAA1755
KLHDC1
KLRC3
KRT13
KRT32
KRTAP10-1
KRTAP10-11
KRTAP10-3
KRTAP10-4
KRTAP10-5
KRTAP12-3
KRTAP9-4
LARP1B
LCE3D
LMO7
LRP2
LYSMD4
MAP2K3
MASP2
MBD1
MCF2L2
MCPH1
MEP1A
MKI67
MMS22L
MRGPRX4

MROH7
MROH7-TTC4
MS4A6E
MTCH2
MTMR1
MTRR
MTUS2
MUC12
MUC16
MUC4
MUC5AC
MUS81
MXRA5
MYH15
MYH4
MYOM3
MYPN
NAAA
NAALADL2
NCAPG
NCKAP5
NEK11
NEMF
NFATC1
NIPSNAP3A
NLRP13
NOTCH3
NPIPB15
NRG1
NT5C3B
NTMT1
NUP160
OBSCN
OBSL1
OR10A6
OR10H1
OR10J1
OR11G2
OR11H6
OR11L1
OR13J1
OR14C36
OR1I1
OR1L4
OR1L6
OR1Q1
OR2C1
OR2M7

OR2T12
OR2T7
OR4A16
OR4B1
OR4C11
OR4C46
OR4D6
OR4M1
OR51G1
OR51M1
OR51Q1
OR52E2
OR52J3
OR52N1
OR52W1
OR56B1
OR5AU1
OR5D16
OR5H15
OR5H6
OR5R1
OR6M1
OR7A10
OR8D1
OR8G5
OR8K1
OTOP2
OTOR
PADI4
PAPPA
PCDHB7
PER3
PIGC
PKD1L3
PKHD1L1
PLEKHG4B
PM20D1
PON2
PPA2
PPEF2
PRAMEF1
PRRC2C
PSMD13
PSMF1
PTGER3
RAB11FIP1
RASAL1
RBM19

RBMX
RFPL1
RFPL2
RHBG
RHOT2
RICTOR
RNF43
RP1
RP1L1
SCLT1
SEMA4D
SEMA4G
SEPN1
SIRPB1
SLC22A10
SLIT3
SMPDL3B
SMYD4
SPATA31A6
SPINK5
SPTA1
SSX5
STEAP2
STK36
SVOPL
SYNE1
SYNE2
TAS2R4
TEKT4
TGOLN2
TGS1
THOC1
TLR3
TMBIM1
TMEM244
TMPRSS15
TNK1
TNN
TNP2
TRIM22
TRIM51
TRNT1
TTC30B
TTI2
TTLL4
TTN
TUBB8
UCK1

```
UGT2B28
UNC5C
VRK2
WBSCR28
WDR49
WDR91
WNK2
ZAN
ZNF177
ZNF180
ZNF19
ZNF214
ZNF221
ZNF239
ZNF28
ZNF30
ZNF404
ZNF415
ZNF443
ZNF45
ZNF534
ZNF549
ZNF559-ZNF177
ZNF573
ZNF607
ZNF611
ZNF667
ZNF705A
ZNF728
ZNF737
ZNF880
ZNF99
```

Gene names provided above. 322 genes in common.

6.38 Question 20)

What is the lowest SIFT score of the deleterious variants?

6.39 Answer 20)

```
In [38]: %%bash
         cat NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_
         /NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_gt_
         |grep 'DELETERIOUS'|grep -v 'Low confidence' \
         |cut -f1,2,3,4,13 \
         |sort|uniq \
```

```
        |sort -k1,1 -k2,2n \
        |sort -k5,5n \
        |head
```

```
chr10       100506090       A       C       0.000
chr10       11755501        G       A       0.000
chr10       26219214        C       A       0.000
chr10       46549695        C       A       0.000
chr10       48086       G       A       0.000
chr10       73378933        C       T       0.000
chr10       97465888        G       A       0.000
chr1        11046609        T       C       0.000
chr11       108593482       T       C       0.000
chr11       26508237        C       T       0.000
```

0.0 is the lowest SIFT score.

6.40 Question 21)

What variants are annotated with the lowest SIFT score? Output the chromosome, coordinate, reference base, alternate base, gene name, reference amino acid, alternate amino acid, amino acid position, and sift score into a file. Display the first 10 lines of this file.

6.41 Answer 21)

```
In [39]: %%bash
         cat NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_
         /NA06984.alt_bwamem_GRCh38DH.20150826.CEU.exome.qual_gt_20.dp_gt_10.gq_gt_
         |cut -f1,2,3,4,7,10,11,12,13,17 \
         |grep '^CHROM\|DELETERIOUS'|grep -v 'Low confidence' \
         |awk '($9==0.0)||$1=="CHROM"' \
         > 1KGenomesSample.SIFT4G.sift_score_0.txt
         head -n10 1KGenomesSample.SIFT4G.sift_score_0.txt
```

```
CHROM       POS         REF_ALLELE      ALT_ALLELE      GENE_NAME       REF_AN
chr1        1956754       C       A       CFAP74      G       C       628
chr1        11046609        T       C       MASP2       D       G       120
chr1        17334004        G       C       PADI4       G       A       112
chr1        18483281        T       C       KLHDC7A       L       S       767
chr1        25342976        T       G       TMEM50A       W       G       37
chr1        26043403        G       T       SLC30A2       N       K       189
chr1        26043403        G       T       SLC30A2       N       K       140
chr1        54653861        C       T       MROH7       S       F       312
chr1        54653861        C       T       MROH7       S       F       312
```

7 4) References

1. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. Nat Protoc. 2016;11(1):1-9. (https://www.ncbi.nlm.nih.gov/pubmed/26633127)
2. Aken BL, Achuthan P, Akanni W, et al. Ensembl 2017. Nucleic Acids Res. 2017;45(D1):D635-D642. (https://www.ncbi.nlm.nih.gov/pubmed/27899575)
3. Auton A, Brooks LD, Durbin RM, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68-74. (https://www.ncbi.nlm.nih.gov/pubmed/26432245)
4. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics. 2009;25(16):2078-9. (https://www.ncbi.nlm.nih.gov/pubmed/19505943)
5. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics. 2011;27(21):2987-93. (https://www.ncbi.nlm.nih.gov/pubmed/21903627)