

EBiSC Information Sheet: Genomic data processing for catalogue search

Version 0.2

2022-08-18

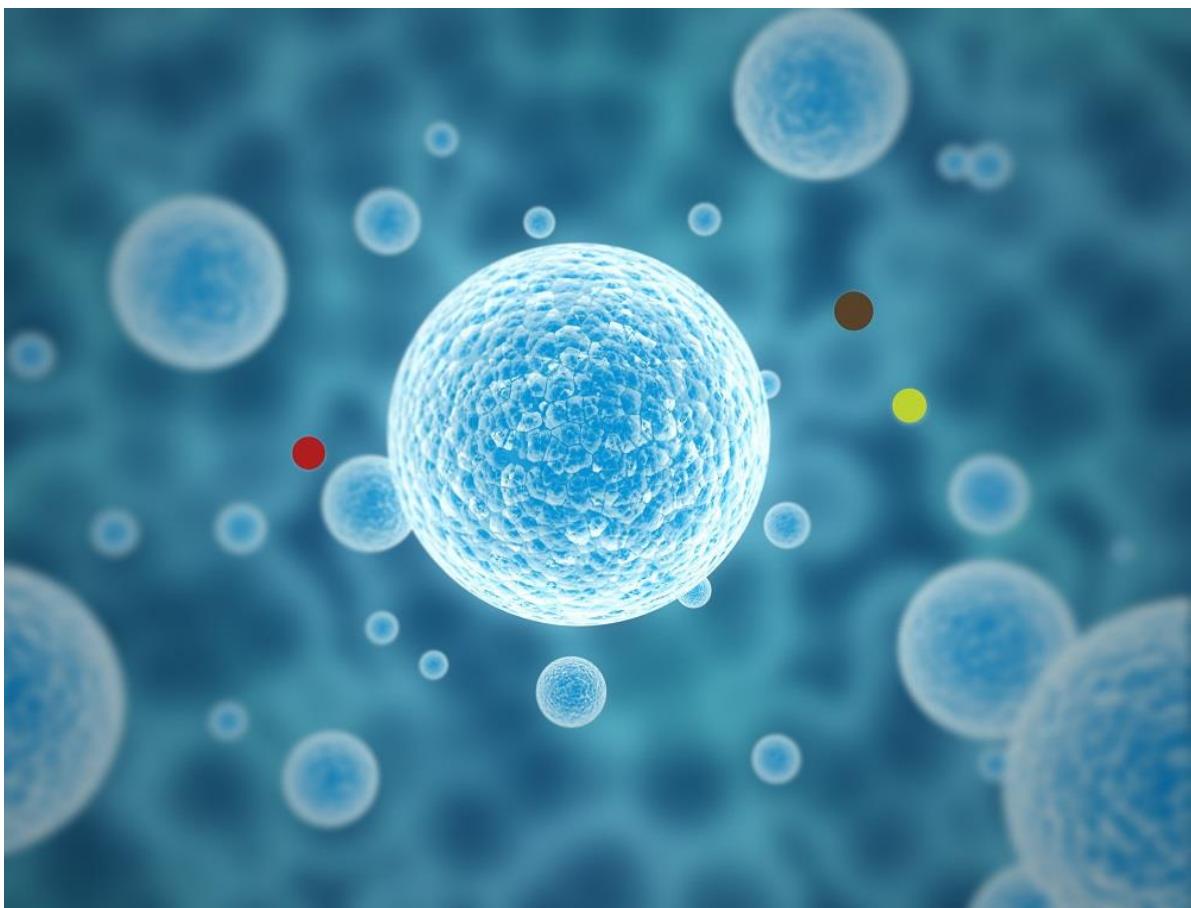


Table of Contents

1	Introduction	3
2	Prerequisites	4
3	Data Source.....	4
4	Workflow.....	5
5	Change History.....	6

1 Introduction

The EBiSC catalogue search page (<https://ebisc.org/search>) is testing new search functionality that will enable users to find the cell lines with mutations in disease-related genes.

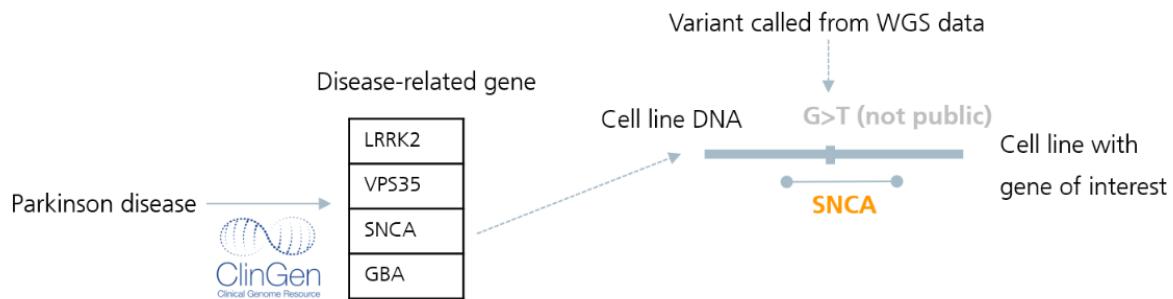


Figure 1. Overview of path from disease to cell line via disease-related genetic variants

2 Prerequisites

- 1) GATK (Genome Analysis Toolkit) v4.2.5.0: <https://github.com/broadinstitute/gatk>
- 2) SnpEff (Genetic variant annotation and functional effect prediction toolbox) version 5.0e: <https://pcingola.github.io/SnpEff/>
- 3) ClinVar vcf for GRCh37: https://ftp.ncbi.nlm.nih.gov/pub/clinvar/vcf_GRCh37/

3 Data source

The gene sequence variations from the VCF (Variant Call Format) data(.genotype.vcf.gz) of 127 HipSci lines, which have open data and were deposited in EBiSC, can be found at the HipSci resource website:
<https://www.hipsci.org/lines/#/lines?Banking%20availability%5B%5D=Banked%20at%20EBiSC&Assays%5B%5D=Whole%20genome%20sequencing>

A table of all links to each cell line's vcf data can be found in the Annex.

4 Workflow

To demonstrate how genes with variants called from WGS datasets (or any other high-throughput genomic data, such as SNP arrays or whole exome sequencing) could be made searchable in the EBiSC catalogue as a proof-of-principle exercise, openly accessible data from the HipSci project (<https://www.hipsci.org/>) was further processed as in Figure 2.

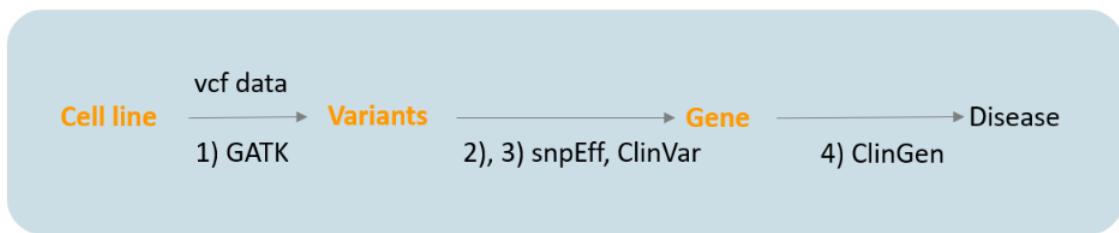


Figure 2. Workflow to obtain genes of interest for every cell lines.

- 1) Original VCF files are processed by GATK to capture the sites with variants.
- 2) The chromosomal locations of the genetic variants in the VCF data were annotated to genes using the snpEff software package (<https://pcingola.github.io/SnpEff/>).
- 3) Variants were filtered according to the following criteria: 1) The variant has an entry in ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>); 2) The variant is located in the coding region; 3) The impact of the variant is high according to snpEff impact prediction (https://pcingola.github.io/SnpEff/se_inputoutput/).
- 4) Genes affected by the filtered variants were recorded for 127 HipSci lines. Importantly, as a data protection measure, the precise gene variant was not recorded in the database. For example, with the variant NM_007262.5:c.105dupT, only the gene (PARK7) is recorded, not the exact description of the genetic variant. These genes were then annotated with disease (if extant) according to disease-gene associations from ClinGen (<https://clinicalgenome.org/>).

In principle, the workflow could be applied to any variant call file, regardless of the method used to generate the VCF data.

5 Annex

Table S1. Source links to variant call format files from open genomic datasets from HipSci

No.	Cell id	Download link
1	WTSLi041-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448214/HPSI0114i-bezi_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
2	WTSLi020-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448189/HPSI0114i-eipl_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
3	WTSLi020-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447849/HPSI0114i-eipl_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
4	WTSLi072-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448022/HPSI0114i-fikt_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
5	WTSLi019-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448198/HPSI0114i-iisa_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
6	WTSLi019-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447908/HPSI0114i-iisa_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
7	WTSLi074-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448028/HPSI0114i-jomx_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
8	WTSLi018-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447925/HPSI0114i-kolf_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
9	WTSLi018-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447889/HPSI0114i-kolf_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
10	WTSLi017-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448081/HPSI0114i-lexy_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
11	WTSLi017-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447860/HPSI0114i-lexy_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
12	WTSLi016-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448111/HPSI0114i-oevr_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
13	WTSLi015-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447912/HPSI0114i-rozh_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
14	WTSLi015-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448206/HPSI0114i-rozh_5.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
15	WTSLi024-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448064/HPSI0114i-vabj_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
16	WTSLi075-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447957/HPSI0114i-vass_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
17	WTSLi144-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ486/ERZ486063/HPSI0114i-wegi_1.wgs.gatk.haplotype_caller.20171212.genotypes.vcf.gz
18	WTSLi023-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448055/HPSI0114i-zapk_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
19	WTSLi023-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448083/HPSI0114i-zapk_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
20	WTSLi037-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448133/HPSI0114i-zoxy_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
21	WTSLi048-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448132/HPSI0214i-datg_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
22	WTSLi044-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448229/HPSI0214i-eiwy_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
23	WTSLi135-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447884/HPSI0214i-feec_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz

No.	Cell id	Download link
24	WTSLi014-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447959/HPSI0214i-heja_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
25	WTSLi014-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448188/HPSI0214i-heja_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
26	WTSLi045-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448234/HPSI0214i-heth_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
27	WTSLi076-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448161/HPSI0214i-kehc_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
28	WTSLi013-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447992/HPSI0214i-kucg_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
29	WTSLi030-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447865/HPSI0214i-pehm_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
30	WTSLi030-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447916/HPSI0214i-pehm_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
31	WTSLi168-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448196/HPSI0214i-poih_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
32	WTSLi168-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448117/HPSI0214i-poih_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
33	WTSLi143-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448215/HPSI0214i-rayr_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
34	WTSLi046-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447893/HPSI0214i-wibj_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
35	WTSLi046-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447910/HPSI0214i-wibj_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
36	WTSLi178-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448149/HPSI0215i-fawm_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
37	WTSLi178-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448084/HPSI0215i-fawm_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
38	WTSLi191-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448026/HPSI0215i-oilg_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
39	WTSLi031-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448072/HPSI0314i-bipt_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
40	WTSLi012-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448125/HPSI0314i-bubh_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
41	WTSLi012-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447866/HPSI0314i-bubh_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
42	WTSLi077-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448218/HPSI0314i-cuhk_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
43	WTSLi078-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448096/HPSI0314i-fafq_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
44	WTSLi026-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448006/HPSI0314i-hoik_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
45	WTSLi040-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447939/HPSI0314i-qaqx_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
46	WTSLi011-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448046/HPSI0314i-qonc_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
47	WTSLi073-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448211/HPSI0314i-sojd_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
48	WTSLi073-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448150/HPSI0314i-sojd_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz

No.	Cell id	Download link
49	WTSLi179-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448137/HPSI0314i-wigw_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
50	WTSLi010-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447901/HPSI0314i-xugn_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
51	WTSLi010-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447871/HPSI0314i-xugn_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
52	WTSLi138-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448104/HPSI0414i-mita_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
53	WTSLi173-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447891/HPSI0414i-naju_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
54	WTSLi183-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448047/HPSI0414i-oaqd_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
55	WTSLi183-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447882/HPSI0414i-oaqd_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
56	WTSLi185-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448168/HPSI0414i-seru_7.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
57	WTSLi177-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448144/HPSI0414i-xojn_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
58	WTSLi180-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447903/HPSI0514i-puie_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
59	WTSLi180-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447883/HPSI0514i-puie_5.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
60	WTSLi181-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447993/HPSI0514i-rutc_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
61	WTSLi049-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447966/HPSI0514i-wiii_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
62	WTSLi049-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447926/HPSI0514i-wiii_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
63	WTSLi182-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447955/HPSI0614i-kegd_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
64	WTSLi169-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448114/HPSI0614i-miaj_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
65	WTSLi169-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448102/HPSI0614i-miaj_6.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
66	WTSLi174-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448152/HPSI0614i-paab_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
67	WTSLi174-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447900/HPSI0614i-paab_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
68	WTSLi176-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447897/HPSI0714i-iudw_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
69	WTSLi176-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448230/HPSI0714i-iudw_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
70	WTSLi047-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448051/HPSI0814i-bokz_5.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
71	WTSLi047-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448228/HPSI0814i-bokz_6.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
72	WTSLi009-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448008/HPSI0913i-diku_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
73	WTSLi002-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447857/HPSI0913i-eika_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz

No.	Cell id	Download link
74	WTSLi071-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448000/HPSI0913i-lise_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
75	WTSLi071-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448023/HPSI0913i-lise_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
76	WTSLi027-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448171/HPSI0913i-oapg_5.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
77	WTSLi190-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448095/HPSI0914i-kajh_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
78	WTSLi190-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447895/HPSI0914i-kajh_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
79	WTSLi172-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448143/HPSI0914i-laey_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
80	WTSLi172-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447948/HPSI0914i-laey_6.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
81	WTSLi171-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447947/HPSI0914i-zerv_7.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
82	WTSLi171-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447914/HPSI0914i-zerv_8.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
83	WTSLi082-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448179/HPSI1013i-cups_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
84	WTSLi029-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448174/HPSI1013i-garx_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
85	WTSLi079-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448076/HPSI1013i-hiaf_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
86	WTSLi042-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447990/HPSI1013i-jufd_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
87	WTSLi042-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448224/HPSI1013i-jufd_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
88	WTSLi008-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448088/HPSI1013i-kuxp_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
89	WTSLi025-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448123/HPSI1013i-pamv_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
90	WTSLi025-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448009/HPSI1013i-pamv_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
91	WTSLi080-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448069/HPSI1013i-sebz_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
92	WTSLi187-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448175/HPSI1013i-sita_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
93	WTSLi039-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448099/HPSI1013i-wopl_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
94	WTSLi021-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447982/HPSI1013i-wuye_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
95	WTSLi021-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448203/HPSI1013i-wuye_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
96	WTSLi022-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447981/HPSI1013i-yemz_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
97	WTSLi022-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447853/HPSI1013i-yemz_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
98	WTSLi186-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448071/HPSI1014i-babz_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz

No.	Cell id	Download link
99	WTSLi188-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447984/HPSI1014i-kefb_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
100	WTSLi184-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447933/HPSI1014i-qayj_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
101	WTSLi184-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448160/HPSI1014i-qayj_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
102	WTSLi189-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447862/HPSI1014i-quls_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
103	WTSLi175-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448004/HPSI1014itixi_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
104	WTSLi050-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447964/HPSI1014itoss_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
105	WTSLi032-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448163/HPSI1113ibima_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
106	WTSLi007-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448034/HPSI1113idons_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
107	WTSLi085-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448056/HPSI1113ieofe_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
108	WTSLi006-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448050/HPSI1113ihayt_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
109	WTSLi006-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447954/HPSI1113ihayt_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
110	WTSLi038-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448219/HPSI1113ieki_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
111	WTSLi038-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448177/HPSI1113ieki_3.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
112	WTSLi139-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447934/HPSI1113ioaz_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
113	WTSLi005-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448044/HPSI1113ipodx_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
114	WTSLi004-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448057/HPSI1113iqolg_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
115	WTSLi004-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448170/HPSI1113iqolg_3.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
116	WTSLi081-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448107/HPSI1113iqorq_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
117	WTSLi081-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448093/HPSI1113iqorq_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
118	WTSLi051-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447915/HPSI1113iuofv_1.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
119	WTSLi052-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447918/HPSI1114iziyn_6.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
120	WTSLi028-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448089/HPSI1213ibabk_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
121	WTSLi003-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447950/HPSI1213ihehd_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz
122	WTSLi098-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448098/HPSI1213inekd_1.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
123	WTSLi103-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ448/ERZ448073/HPSI1213inusw_2.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz

No.	Cell id	Download link
124	WTSLi001-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447942/HPSI1213i-pahc_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
125	WTSLi170-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447946/HPSI1213i-tolg_4.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
126	WTSLi170-B	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447999/HPSI1213i-tolg_6.wgs.gatk.haplotype_caller.20161201.genotypes.vcf.gz
127	WTSLi083-A	ftp://ftp.sra.ebi.ac.uk/vol1/ERZ447/ERZ447977/HPSI1213i-xuja_2.wgs.gatk.haplotype_caller.20170425.genotypes.vcf.gz

6 Change History

Version	Valid from	Changes compared to previous version
0.1	17-Aug-2022	First draft
0.2	18-Aug-2022	Clarification: variants were not selected because they have clinical significance per se, they were selected because they were reported in ClinVar. Reported variants in ClinVar have diverse levels of confidence as to their medical relevance.