DACC specification for variant submission

Variant information should be submitted in one of the two ways:

(1) *Tabular format* described below
(2) *VCF file* (GRCh38/GRCm39)

(1) - Tabular format would include required and optional fields:

REQUIRED fields:

(column 1) Genome reference (GRCh38 or GRCm39) (example GRCh38)

(column 2) Chromosome (RefSeq ID) type: string (example NC_0000012.2)

(column 3) Position (1-based coordinates) type:integer (example 234567)

(column 4) Reference allele (minimum 1 letter) type: string (example 1 A) (example 2 TTG)

(column 5) Alternative allele (could be "-" for deletion) type: string (example 1 G) (example 2 CC)

OPTIONAL fields:

(column 6) rsID type: string (example rs219)

(column 7) Variant type type: enum (SNV, MNV, INS, DEL, INDEL)

(column 8) Associated protein change type: string (example nonsense mutation)

Selection criteria

(column 9) Associated gene (HGNC) type: string

(column 10) Associated trait (DOID|EFO|HP|MONDO|OBA) type: string

(column 11) OMIM type: string

(column 12) Pubmed ref doi :/10.10.20.30)	type: string (example 1 PMID:123456) (example 2
(column 13) Other (Free Text)	type: string

(2) - VCF file specifications:

a. VCF file is using GRCh38 for human variants and GRCm39 for mouse variants. This information has to be included in the metainformation section at the header of the file.
b. If the variant is in dbSNP, the metainformation section should include information on te dbSNP version. The variant ID column should include the rsID, and DB flag should be included.