

Schnittstellenspezifikation für geparte Output-Dateien

Input-Datei: Aufbau des von der Anwendung generierten 5-teiligen Dateinamens

<hg19|hg38>-<interne PatientenID>-<MGU-Bogenr>-<interne Mitarbeiter-ID>-<Zeitstempel>.vcf

Beispiel: hg19-3828678-2-3746-20220308095825.vcf

bis Juli 2022: Aufbau der Datenbank-Staging-Tabelle VCF_UPLOAD

Aus den Bestandteilen 2 bis 5 des Dateinamens und des "Parser-Outputs" wird diese Datenbanktabelle gefüllt:

Columns of VCF_UPLOAD					
Staging-Tabelle zur Zwischenspeicherung der zu einer Patientin hochgeladenen und geparten vcf-Daten; von hier aus werden die Daten in					
	Name	Type	Nullable	Default	Comments
1	MEMBER_ID	NUMBER			interne Pat-ID, Teil 2 des Dateinamens
2	BOGEN_NR	NUMBER			MGU-Bogenr, Teil 3 des Dateinamens
3	ERFMIT	NUMBER			interne Mitarbeiter-ID, Teil 4 des Dateinamens
4	ERFDAT	NUMBER			Zeitstempel des Uploads, Format YYYYMMDDHH24MISS, Teil 5 des Dateinamens
5	GEN2	VARCHAR2(20)	Y		Gen-Name bzw. "STATUS" bei Statuszeile
6	REFSEQ	VARCHAR2(20)	Y		Transkript/Referenzsequenz
7	HGVS_DNA	VARCHAR2(1000)	Y		HGVS-Nomenklatur DNA-Ebene bzw. bei Statuszeile "ok", "leer" oder Fehlermeldung
8	HGVS_PROT	VARCHAR2(1000)	Y		HGVS-Nomenklatur Protein-Ebene
9	ART	VARCHAR2(400)	Y		Effekt der Sequenzänderung
10	PATH	VARCHAR2(400)	Y		Pathogenität
11	CHROM	VARCHAR2(20)	Y		Chromosom (als Nummer bzw. X oder Y)
12	GPOS	NUMBER	Y		Genetische Positionsangabe auf dem Chromosom
13	REF	VARCHAR2(1000)	Y		Reference Sequence
14	ALT	VARCHAR2(1000)	Y		Alternate Sequence

Änderungswünsche für die neue Parserversion:

(1) HG19 / HG38 betreffend

Kleine Änderungen bei den 3 bisherigen Variablen:

- GPOS → POS_HG19
- REF → REF_HG19
- ALT → ALT_HG19

3 neue Variablen:

- POS_HG38
- REF_HG38
- ALT_HG38

(2) Neue Variable Zytogie

- Variablenname ZYGOT mit folgende Ausprägungen:

###	keine Angabe
0/1	heterozygot
1/1	homozygot
0/0	Wildtyp

Aufbau der neuen Datenbank-Staging-Tabelle VCF_UPLOAD

Columns of VCF_UPLOAD

Staging-Tabelle zur Zwischenspeicherung der zu einer Patientin hochgeladenen und geparsten vcf-Daten; von hier aus werden die Daten in

	Name	Type	Nullable	Default	Comments
1	MEMBER_ID	NUMBER			interne Pat-ID, Teil 2 des Dateinamens
2	BOGEN_NR	NUMBER			MGU-Bogennr, Teil 3 des Dateinamens
3	ERFMIT	NUMBER			interne Mitarbeiter-ID, Teil 4 des Dateinamens
4	ERFDAT	NUMBER			Zeitstempel des Uploads, Format YYYYMMDDHH24MISS, Teil 5 des Dateinamens
5	GEN2	VARCHAR2(20)	Y		Gen-Name bzw. "STATUS" bei Statuszeile
6	REFSEQ	VARCHAR2(20)	Y		Transkript/Referenzsequenz
7	HGVS_DNA	VARCHAR2(1000)	Y		HGVS-Nomenklatur DNA-Ebene bzw. bei Statuszeile "ok", "leer" oder Fehlermeldung
8	HGVS_PROT	VARCHAR2(1000)	Y		HGVS-Nomenklatur Protein-Ebene
9	ART	VARCHAR2(400)	Y		Effekt der Sequenzänderung
10	PATH	VARCHAR2(400)	Y		Pathogenität
11	CHROM	VARCHAR2(20)	Y		Chromosom (als Nummer bzw. X oder Y)
12	POS_HG19	NUMBER	Y		Genetische Positionsangabe auf dem Chromosom für hg19
13	REF_HG19	VARCHAR2(1000)	Y		Reference Sequence für hg19
14	ALT_HG19	VARCHAR2(1000)	Y		Alternate Sequence für hg19
15	POS_HG38	NUMBER	Y		Genetische Positionsangabe auf dem Chromosom für hg38
16	REF_HG38	VARCHAR2(1000)	Y		Reference Sequence für hg38
17	ALT_HG38	VARCHAR2(1000)	Y		Alternate Sequence für hg38
18	ZYGOT	VARCHAR2(20)	Y		Zygotie mit Ausprägungen 0/1, 1/1, 0/0, sonst ###