

1 このドキュメントについて

検査会社様がCATSフォーマット対応のテストデータを作成する際、前回データから変更が生じた項目をチェックするものです。（初回作成時はチェック不要）

テストデータの変更項目、またはコンバータ改修の影響がある項目について「変更有無」の欄に「○」（変更あり）を入力してください。（「-」は変更なし）

※テストデータに表れないコンバータの改修を実施した場合、該当項目の備考欄にその旨記載をお願いします。

当資料で扱うCATSフォーマットのバージョンは下記です。

CATSフォーマット	v1.0.1
------------	--------

2 変更点チェックリスト

検査会社様入力欄

No	CATSフォーマット項目名	必須	配列の長さ	データ型	文字列正規表現	選択肢(enum, boolean)	変更有無	備考
1	metaData	○	-	オブジェクト	-	-	-	
2	schemaVersion	○	-	文字列	^[0-9¥¥.]+\$	-	-	
3	referenceGenome	○	-	オブジェクト	-	-	-	
4	name	○	-	文字列[選択式]	-	"GRCh37" "GRCh38"	-	
5	patch	○	-	文字列またはnull	^p[0-9]+\$	-	-	
6	descriptions		0～N	配列(文字列)	^.+\$	-	-	
7	configOptions		-	オブジェクト	-	-	-	
8	typeLabelsInterpretedAskbAmplification		1～4	配列(文字列[選択式])	-	"copyNumberAlterationType: amplification" "copyNumberAlterationType: gain" "copyNumberAlterationType: duplication" "rearrangementType: duplication"	-	
9	typeLabelsInterpretedAskbLoss		1～3	配列(文字列[選択式])	-	"copyNumberAlterationType: loss" "copyNumberAlterationType: deletion" "copyNumberAlterationType: homozygous deletion" "rearrangementType: deletion"	-	
10	typeLabelsInterpretedAskbGeneFusion		1～8	配列(文字列[選択式])	-	"rearrangementType: gene fusion" "rearrangementType: frameshift gene fusion" "rearrangementType: bidirectional gene fusion" "rearrangementType: duplication" "rearrangementType: deletion" "rearrangementType: inversion" "rearrangementType: truncation" "rearrangementType: other"	-	
11	hideAlleleFrequency		-	ブール型	-	true false	-	
12	hideCnaValue		-	ブール型	-	true false	-	
13	hideMsiValue		-	ブール型	-	true false	-	
14	hideTmbValue		-	ブール型	-	true false	-	
15	hideLohValue		-	ブール型	-	true false	-	
16	comments		1～N	配列(オブジェクト)	-	-	-	
17	itemIds	○	0～N	配列(文字列)	^.+\$	-	-	
18	contents	○	1～N	配列(文字列)	^.+\$	-	-	
19	testInfo	○	-	オブジェクト	-	-	-	
20	tesetId	○	-	文字列	^.+\$	-	-	
21	testType	○	-	文字列[選択式]	-	"tumor-only" "tumor and matched-normal" "tumor-only (cell-free)" "tumor (cell-free) and matched-normal"	-	
22	softwareName		-	文字列	^.+\$	-	-	
23	softwareVersion		-	文字列	^.+\$	-	-	
24	panelName	○	-	文字列	^.+\$	-	-	
25	panelVersion	○	-	文字列	^.+\$	-	-	
26	variants		-	オブジェクト	-	-	-	
27	shortVariants		1～N	配列(オブジェクト)	-	-	-	
28	itemId	○	-	文字列	^.+\$	-	-	
29	chromosome	○	-	文字列	^[a-zA-Z0-9_¥¥-]+\$	-	-	
30	position	○	-	整数	-	-	-	
31	referenceAllele	○	-	文字列	^[ACGTN]+\$	-	-	
32	alternateAllele	○	-	文字列	^[ACGTN¥¥*]+\$	-	-	
33	alternateAlleleFrequency	○	-	数値	-	最小値0、最大値1	-	
34	totalReadDepth		-	整数	-	最小値1	-	
35	alternateAlleleReadDepth		-	整数	-	最小値1	-	
36	variantType		-	文字列[選択式]	-	"SNV" "insertion" "deletion" "delins" "indel" "MNV"	-	

No	CATSフォーマット項目名	必須	配列の長さ	データ型	文字列正規表現	選択肢(enum, boolean)	変更有無	備考
37	transcripts	○	1～N	配列(オブジェクト)	-	-	-	
38	transcriptId	○	-	文字列またはnull	^[^¥¥s]+\$	-	-	
39	transcriptDatabaseName	○	-	文字列[選択式]またはnull	-	"RefSeq" "Ensembl"	-	
40	transcriptDatabaseVersion		-	文字列またはnull	^.+\$	-	-	
41	geneSymbol	○	-	文字列またはnull	^[^¥¥s]+\$	-	-	
42	strand		-	文字列[選択式]またはnull	-	"+" "-"	-	
43	cdsChange	○	-	文字列またはnull	^.+\$	-	-	
44	aminoAcidsChange	○	-	文字列またはnull	^.+\$	-	-	
45	calculatedEffects		0～N	配列(文字列)	^.+\$	-	-	
46	testMethod	○	-	文字列[選択式]	-	"DNA-seq" "RNA-seq"	-	
47	variantOrigin		-	文字列[選択式]	-	"somatic" "germline" "likely somatic" "likely germline"	-	
48	reported	○	-	ブール型	-	true false	-	
49	copyNumberAlterations		1～N	配列(オブジェクト)	-	-	-	
50	itemId	○	-	文字列	^.+\$	-	-	
51	chromosome	○	-	文字列	^[a-zA-Z0-9_¥¥-]+\$	-	-	
52	startPosition		-	整数	-	-	-	
53	endPosition		-	整数	-	-	-	
54	copyNumberMetrics		0～N	配列(オブジェクト)	-	-	-	
55	value	○	-	数値	-	-	-	
56	unit	○	-	文字列[選択式]	-	"absolute copy number" "fold-change" "log2 fold-change" "fraction-of-gene"	-	
57	copyNumberAlterationType	○	-	文字列[選択式]	-	"amplification" "gain" "duplication" "loss" "deletion" "homozygous deletion" "neutral"	-	
58	transcripts	○	1～N	配列(オブジェクト)	-	-	-	
59	transcriptId		-	文字列	^[^¥¥s]+\$	-	-	
60	transcriptDatabaseName		-	文字列[選択式]	-	"RefSeq" "Ensembl"	-	
61	transcriptDatabaseVersion		-	文字列	^.+\$	-	-	
62	geneSymbol	○	-	文字列またはnull	^[^¥¥s]+\$	-	-	
63	strand		-	文字列[選択式]	-	"+" "-"	-	
64	cdsChange		-	文字列またはnull	^.+\$	-	-	
65	aminoAcidsChange		-	文字列またはnull	^.+\$	-	-	
66	calculatedEffects		0～N	配列(文字列)	^.+\$	-	-	
67	testMethod	○	-	文字列[選択式]	-	"DNA-seq" "RNA-seq"	-	
68	variantOrigin		-	文字列[選択式]	-	"somatic" "germline" "likely somatic" "likely germline"	-	
69	reported	○	-	ブール型	-	true false	-	
70	rearrangements		1～N	配列(オブジェクト)	-	-	-	
71	itemId	○	-	文字列	^.+\$	-	-	
72	breakends	○	2	配列(オブジェクト)	-	-	-	
73	chromosome	○	-	文字列	^[a-zA-Z0-9_¥¥-]+\$	-	-	
74	startPosition	○	-	整数	-	-	-	
75	endPosition	○	-	整数	-	-	-	
76	matePieceLocation		-	文字列[選択式]	-	"upstream" "downstream"	-	
77	transcripts	○	1～N	配列(オブジェクト)	-	-	-	
78	transcriptId		-	文字列	^[^¥¥s]+\$	-	-	
79	transcriptDatabaseName		-	文字列[選択式]	-	"RefSeq" "Ensembl"	-	
80	transcriptDatabaseVersion		-	文字列	^.+\$	-	-	
81	geneSymbol	○	-	文字列またはnull	^[^¥¥s]+\$	-	-	
82	strand		-	文字列[選択式]	-	"+" "-"	-	
83	cdsChange		-	文字列またはnull	^.+\$	-	-	
84	aminoAcidsChange		-	文字列またはnull	^.+\$	-	-	
85	calculatedEffects		0～N	配列(文字列)	^.+\$	-	-	
86	genePairs		0～N	配列(文字列)	^.+¥¥-.\$	-	-	
87	insertedSequence		-	文字列またはnull	^[ACGTN]+\$	-	-	
88	supportingReadCount		-	整数	-	-	-	
89	alternateAlleleFrequency		-	数値	-	-	-	
90	expressionLevelMetrics		0～N	配列(オブジェクト)	-	-	-	
91	value	○	-	数値	-	-	-	
92	unit	○	-	文字列[選択式]	-	"TPM" "FPKM" "FPM" "RPKM" "RPM"	-	
93	rearrangementNames		0～N	配列(文字列)	^.+\$	-	-	

No	CATSフォーマット項目名	必須	配列の長さ	データ型	文字列正規表現	選択肢(enum, boolean)	変更有無	備考
94	rearrangementType	○	-	文字列[選択式]	-	"gene fusion" "frameshift gene fusion" "bidirectional gene fusion" "duplication" "deletion" "inversion" "truncation" "splice variant" "other"	-	
95	testMethod	○	-	文字列[選択式]	-	"DNA-seq" "RNA-seq"	-	
96	variantOrigin		-	文字列[選択式]	-	"somatic" "germline" "likely somatic" "likely germline"	-	
97	reported	○	-	ブール型	-	true false	-	
98	otherBiomarkers		0～N	配列(オブジェクト)	-	-	-	
99	itemId	○	-	文字列	^.+\$	-	-	
100	biomarkerType	○	-	文字列[選択式]	-	"TMB" "MSI" "LOH"	-	
101	biomarkerMetrics		0～N	配列(オブジェクト)	-	-	-	
102	value	○	-	数値	-	-	-	
103	unit	○	-	文字列	^.+\$	-	-	
104	state		-	文字列[選択式]またはnull	-	"high" "low" "intermediate" "stable"	-	
105	descriptions		0～N	配列(文字列)	^.+\$	-	-	
106	biomarkerOrigin		-	文字列[選択式]	-	"somatic" "germline" "likely somatic" "likely germline"	-	
107	reported	○	-	ブール型	-	true false	-	
108	compositeBiomarkers		0～N	配列(オブジェクト)	-	-	-	
109	itemId	○	-	文字列	^.+\$	-	-	
110	componentItemIds	○	2～N	配列(文字列)	^.+\$	-	-	
111	biomarkerNames	○	1～N	配列(文字列)	^.+\$	-	-	
112	descriptions		0～N	配列(文字列)	^.+\$	-	-	
113	reported	○	-	ブール型	-	true false	-	
114	sequencingSamples		1～N	配列(オブジェクト)	-	-	-	
115	itemId	○	-	文字列	^.+\$	-	-	
116	tumorOrNormal	○		文字列[選択式]	-	"tumor" "normal"	-	
117	testMethod	○		文字列[選択式]	-	"DNA-seq" "RNA-seq"	-	
118	duplicateReadsPercentage			数値	-	-	-	
119	mappedReadsPercentage			数値	-	-	-	
120	meanReadDepth			数値	-	-	-	
121	medianReadDepth			数値	-	-	-	
122	suspectedSampleStates		0～N	配列(文字列[選択式])	-	"contaminated" "deaminated" "fragmentated" "degraded"	-	